Department for Business, Energy & Industrial Strategy

Domestic National Energy Efficiency Data-Framework (NEED) Methodology

25 June 2020   **National Statistics**

This note outlines the methodology used for analyses using the domestic National Energy Efficiency Data-Framework (NEED) as published in June 2020.

NEED enables analysis of gas and electricity consumption by a range of property attributes including property type and property age. It also provides estimates of the changes in gas and electricity consumption following the installation of energy efficiency measures such as cavity wall insulation and loft insulation.

A detailed overview of the outputs from the quality assurance (QA) process used during the creation of NEED is published alongside this report in Annex A. It includes details of the strengths and weaknesses of both the data sources used in NEED and the address matching process used to link datasets.

The NEED data and methods are subject to continuous improvement and development. Sections 3 and 4 of this document gives an overview of the methodology including refinements. Changes to the methodology for estimating changes in consumption for "Impact of measures" analysis mean that a like for like comparison of estimates provided in the June 2020 publication cannot be made with previous years' publications where specified.

This note does not cover non-domestic NEED[1], although it uses some of the same data sources.

---

[1] More information on non-domestic NEED is available here:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/314725/non_domestic_need_framework.pdf

Contents

1. Background

Development of the framework

The National Energy Efficiency Data-Framework (NEED) project was set up by the Department for Business, Energy and Industrial Strategy (BEIS)[2]. NEED is a key element of the evidence base, supporting BEIS to:

- Develop, monitor and evaluate key policies including the Green Deal and Energy Company Obligation (ECO).

- Identify energy efficiency potential which sits outside the current policy framework.

- Develop a greater understanding of the drivers of energy consumption.

- Gain a deeper understanding of the impacts of energy efficiency measures for households.

The first results from NEED were published in June 2011[3]. The publication included summary statistics on consumption by property attributes and household characteristics, as well as an assessment of the change in households' gas consumption following the installation of retrofit insulation measures. Estimates of savings were produced using a difference in difference approach and covered cavity wall insulation, loft insulation and condensing boilers. Since then, further developments have been made to the methodology.

This note sets out the latest methodology used for the June 2020 NEED publication as well as a summary of developments since 2011.

All NEED publications can be found on the NEED webpage: https://www.gov.uk/government/collections/national-energy-efficiency-data-need-framework

The framework and data sources

NEED is a framework which combines data from existing sources (administrative and commercial) to provide insights into how energy is used and what the impact of energy efficiency measures are on gas and electricity consumption, for different types of properties and households.

Four main data sources are used to put NEED together, as shown in Figure 1.1 below:

- Energy Suppliers provide consumption data for gas and electricity through meter readings.

- Valuation Office Agency (VOA) supplies property attribute data.

---

[2] formerly the Department for Energy and Climate Change (DECC).
[3] http://webarchive.nationalarchives.gov.uk/20130109092117/http://decc.gov.uk/en/content/cms/statistics /energy_stats/en_effic_stats/need/need.aspx

- Experian supplies modelled household characteristics data.

- Energy Efficiency measure installations are collected within BEIS from obligated suppliers (through Ofgem), installations under the Feed in Tariff and Green Deal government schemes, Homes Energy Efficiency Data base (HEED is the historic energy efficiency measure data base) and other relevant data relating to energy efficiency data held in BEIS.

Figure 1.1: Structure of domestic NEED
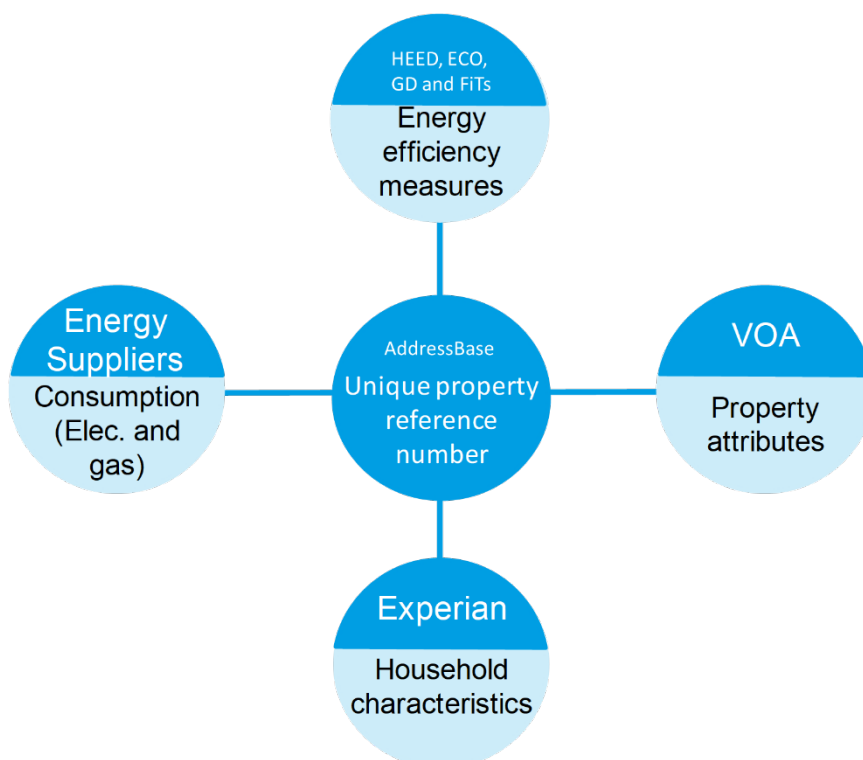
The data in these sources is linked together using the address information in each dataset, where a unique property reference number (UPRN) is assigned to each record.

Table 1.1 below summarizes the scope of data in NEED (domestic).

Table 1.1: Scope of Domestic NEED

In Scope	Out of Scope
Domestic properties	Non-Domestic/Commercial properties
Great Britain[4]	Northern Ireland
Metered gas and electricity[5]	Non-metered fuel use e.g. oil, coal, wood, auto-generated electricity

[4] Data for England and Wales is reported in the main NEED summary report and tables, whereas Scotland is reported in a separate section of the NEED report. There are also separate Scotland summary statistics.
[5] Properties that produce electricity on-site (e.g., solar panels) are included in the NEED data set. The annual kWh consumption for these properties reflects electricity imported from the grid. Any electricity generated and consumed on-site is not recorded by NEED.

| BEIS Energy Efficiency Measures data (e.g. ECO, HEED) | Measures installed outside of government schemes and DIY measures |

In addition to the four main data sources (outlined above), other indicators have been assigned to the properties based on their geographic locations. For example, an index of multiple deprivation and a fuel poverty indicator are assigned to each property based on the Lower Super Output Area (LSOA).

The most important sources for the current analysis of domestic data are summarised in Table 1.2 below.

More detailed information on all data sources which feed into NEED can be found here: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/720055/Annex_D_-_What_is_NEED.pdf

Quality assurance of all data sources is undertaken before they are used in NEED and throughout to ensure consistency. The validation and more information on the data sources have been published in the quality assurance Annex A of the NEED publication.

Table 1.2: Data in NEED

Category	Source	Description
Premises	Ordnance Survey AddressBase	Contains a unique identifier for each address in England and Wales which is matched to each of the data sources in NEED and then used to link data together.
Energy consumption	Energy suppliers and ECOES, Xoserve/Independent Gas Transporters	Gas and electricity consumption data for all domestic and non-domestic meters in GB, 2004-2018 and meter profile for electricity meters. Gas data are weather corrected.
Measures installed	Home Energy Efficiency Database (HEED), Green Deal (GD) schemes, Energy Company Obligation (ECO)[6], Feed-in Tariffs (FITs)	Information on energy efficiency measures installed through government schemes (including EEC, CERT and CESP and Green Deal Communities).
Property attributes	Valuation Office Agency (England and Wales), Scottish Assessor (Scotland)	VOA collects property attribute data to inform its function relating to business rates and council tax. The data include floor area, number of bedrooms, property type and property age. The Scottish Assessors Office perform a similar function to the VOA in Scotland and data covers property age, floor area and property type.
Household characteristics	Experian	Modelled data for household characteristics such as income and tenure.
Socio demographic	Ministry of Housing, Communities & Local Government, Welsh Assembly Government, Office for National Statistics, BEIS	Sub-regional flags such as index of multiple deprivation and fuel poverty levels.

---

[6] https://www.ofgem.gov.uk/environmental-programmes/energy-companies-obligation-eco/previous-energy-efficiency-schemes

Energy performance certificates	Ministry of Housing, Communities & Local Government	Energy efficiency rating of a domestic property

Gas meter data collection method

The summer of 2017 saw the implementation of new gas meter point management and settlement processes. For the old meter point management system, the average meter read period was 1st October – 30th September. Since the final year of the old meter point management system (2015 gas year) consistent dates have not been used, and therefore there have been gaps and periods of overlap. The gas consumption periods each year refers to are:

- Prior to 2014: October – September (same period as 2014 and 2015)
- 2014: October 2013 – September 2014
- 2015: October 2014 – September 2015
- 2016: mid July 2016 – mid July 2017
- 2017: mid June 2017 – mid June 2018
- 2018: mid May 2018 – mid May 2019

From here in this document the gas consumption periods will be referred to as the "gas period" for that year (e.g. the 2018 gas period refers to mid-May 2018 to mid-May 2019). All references to gas consumption years in this report signify the gas period for that year (e.g. "gas consumption 2015" refers to consumption of gas between October 2014 and September 2015). Note that one month of gas consumption for 2018 includes consumption which is also counted in 2017; this also occurred for the 2017 gas period (2019 publication).

It is expected that once the transition period has completed, future gas periods will not overlap or have gaps between them; they are expected to be the same 12 months every year. The impact of measures analysis adjusts for overlaps in consumption periods. Note that the consumption values for each year remain weather corrected (the process by which gas consumption values are adjusted to account for that period's weather).

2. Users and uses

Analysis using NEED has supported a number of BEIS policies:

- NEED has been used to understand the changes in consumption (and resulting impacts on energy bills) for households installing energy efficiency measures.

- The estimates from impact of measures analysis from NEED were used to inform "in use factors" for the Green Deal.

- NEED has also had a smaller, but still significant, part to play in a range of other BEIS policies, for example, the Renewable Heat Incentive and Fuel Poverty. Data on consumption by property attributes, including the distribution of households' consumption, has been used to help BEIS understand the likely under or over payment if payments for the renewable heat incentive were to be based solely on property attributes available in NEED.

- NEED has helped Fuel Poverty analysis to better understand the actual consumption for different types of properties and households, and therefore to better understand how policy options will impact on different households. Having this information enables BEIS to provide better value for money and understand the impacts of policy options better, for both BEIS and consumers.

- NEED has also helped BEIS understand where further research should be focused. It provides high level results which have highlighted a need for further investigation. For example, further investigation was carried out to understand why households which appear the same in physical property attributes use varying amounts of energy[7].

BEIS published an updated version of the anonymised record level NEED dataset in December 2019[8]. Users can access this and interrogate this data themselves. It is anticipated this will increase the use of NEED by enabling analysis by other individuals and organisations. NEED is also available to accredited researchers through the ONS Secure Research Service[9].

In addition, NEED published outputs are also used externally by a wide range of interested parties, including local authorities, academics and energy suppliers. These include:

- Committee on Climate Change use the outputs to inform reports, including recommendations to Government.

- Energy companies and academics used the outputs to validate and inform their own research and estimates.

- Energy suppliers use the NEED reports to act as an independent trusted source demonstrating the benefit of installing energy efficiency measures.

- Local authorities use the outputs to help with modelling housing stock and understanding the impact of installing energy efficiency measures.

---

[7] For more information go to:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/437099/Annex_C_Change_of_occupancy_analysis.pdf.
[8] To access these datasets go to: https://www.gov.uk/government/statistics/national-energy-efficiency-data-framework-need-anonymised-data-2019
[9] https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme

3. Current NEED Dataset

A detailed overview of the outputs from the quality assurance (QA) process used during the steps below is published alongside this report ("Annex A Comparison with other sources"). It includes details of the strengths and weaknesses of each of the data sources used in NEED.

A significant change to NEED in 2019 was the inclusion of all properties contained in the VOA Council Tax Database. This has been made possible through an enhanced data sharing agreement with the VOA. The NEED dataset compiled in 2020 uses all properties contained on the 2019 VOA council tax database, where the property is assessed to have valid gas or electricity consumption in 2018.

Prior to NEED in 2019 a stratified random sample of approximately one in five records was selected from the complete property attribute dataset held by VOA. Resulting in a sample containing approximately 4 million records.

The final NEED dataset (combining the UPRNs from both gas and electric data) for 2020 contains 23.8 million records. This is 90 per cent of the full VOA Council Tax Database. The remaining 10 per cent of records are excluded as either a valid address match cannot be made or the property is categorised as an annex, caravan or cluster house. Once the data are matched to other sources, further records were removed because of invalid[10] or missing consumption values. For the current NEED dataset, 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. As a comparison with electric meters, of the 21.9m gas meters matched to NEED, 83.3 per cent had a valid gas value. 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 linking to other sources.

The 10 per cent of records removed through property address matching is not evenly distributed as there were more records that could not be matched for flats (as these can be harder to match to addresses) and consequently proportionately more records were lost in cities and metropolitan areas than other areas of England and Wales.

Steps to create the NEED dataset

1. A database of all UPRNs that have been linked to an electric or gas meter is the starting point for creating NEED.
2. Link gas and electric consumption data (2005 to current year) and calculate valid consumption and valid gas or electric consumption identifiers for each UPRN.
3. Add property attribute data from Experian, the VOA, the Scottish Assessor, ONS and EPC.
4. Filter only for UPRNs that are contained in the VOA Council Tax and Scottish Assessor Databases.
5. Only UPRNs that have valid gas or electric consumption identifiers are included in the consumption table data and IoM analysis.

---

[10] Meter point consumption is regarded as valid if the annual figure is not an estimate from previous years and the range of consumption is between 100kWh and 25,000kWh (electric) or 50,000kWh (gas).

6. Energy efficiency data (e.g., Gas safe, ECO) is added when IoM analysis is undertaken.

4. Impact of measures methodology

A difference in difference approach is used to estimate the impact of installing retrofit energy efficiency measures on the amount of gas required to heat a home. This approach has been used, with refinements, since the first publication of the NEED impact of measures analysis in 2011.

In the 2019 publication, the method was developed to improve the quality of outputs and speed of processing. This change was prompted by the recognised need for multiple iterations in the analysis, BEIS' increased access to data and the need to understand causes of differences between years. Further details of the changes made in the 2019 publication are set out in Annex C[11] to that release.

A few minor changes have been implemented for the 2020 publication. This section sets out the full methodology, highlighting the changes, for the most recent analysis (measures installed between July 2017 and May 2018).

Overview of Difference in Difference Approach

A difference in difference approach for impact of measures works by comparing the gas consumption in properties before and after an energy efficiency measure has been installed (or electricity consumption before and after Solar PV has been installed). The analysis explores the change in consumption over the same period for similar properties which have not had a measure installed.

To do this, an intervention and a comparator group are created for each energy efficiency measure being considered. The **intervention group** contains properties which have received the energy efficiency measure, and no other measure[12] (including smart meters)[13] in the year before, during or after installation as far as recorded in NEED.

The intervention group cannot be analysed in isolation since a change in gas consumption could be due to other factors, such as a change in energy prices, energy efficiency awareness or changes in household size or occupants. Therefore, a **comparator group** is created which contains similar properties, but have no record of an energy efficiency measure[12] installed in the year before, after or during the year of installation as defined by the intervention group. This helps to control for some of the other factors which may have resulted in gas consumption, providing a more accurate estimate of the impact of the energy efficiency measure.

---

[11] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/812417/Annex_C_Impact_of_Measures_Method_Update.pdf

[12] Energy efficiency measures recorded as being installed in HEED, GD, EO: cavity wall insulation, solid wall insulation, loft insulation (including roof and room in roof insulation), condensing boiler, double glazing, smart meters and draught proofing. Additionally they also have no installations under the FiTs and RHI schemes, or a boiler registration by Gas Safe. These properties may have had a measure installed which has not been recorded, such as DIY loft insulation

[13] The impact of installing a smart meter on energy consumption is currently unknown however is assumed to exist. Smart meters are therefore treated like other energy efficiency measures listed in footnote 12. As a result, evidence of installation of either a gas or electric smart meter in the year before, after or during the year of installation defined by the intervention group results in that property being removed from both the comparator and intervention group. This is done regardless of whether the analysis is observing changes in gas or electric consumption. In the 2019 publication, BEIS only held electric smart meter data and therefore some properties included in the comparator and intervention group may have had a gas smart meter installed.

A savings estimate is produced for each individual property in the intervention group. This is done by matching each property in the intervention group to a property with similar attributes in the comparator group and looking at the difference in the change in gas consumption between the two properties. The typical difference (median), the distribution of savings and the mean can be derived from the saving estimates across all properties in the intervention group. The estimates of savings are given both as percentage and in kWh.

An illustration of the change in consumption of installing the energy efficiency measures is shown in Figure 4.1, where 'C' represents the energy saving. This saving is equivalent to 'A minus B' on the diagram, where A is the change in consumption for the intervention group before and after the installation of the measure; and B is the difference in consumption for the comparator group.
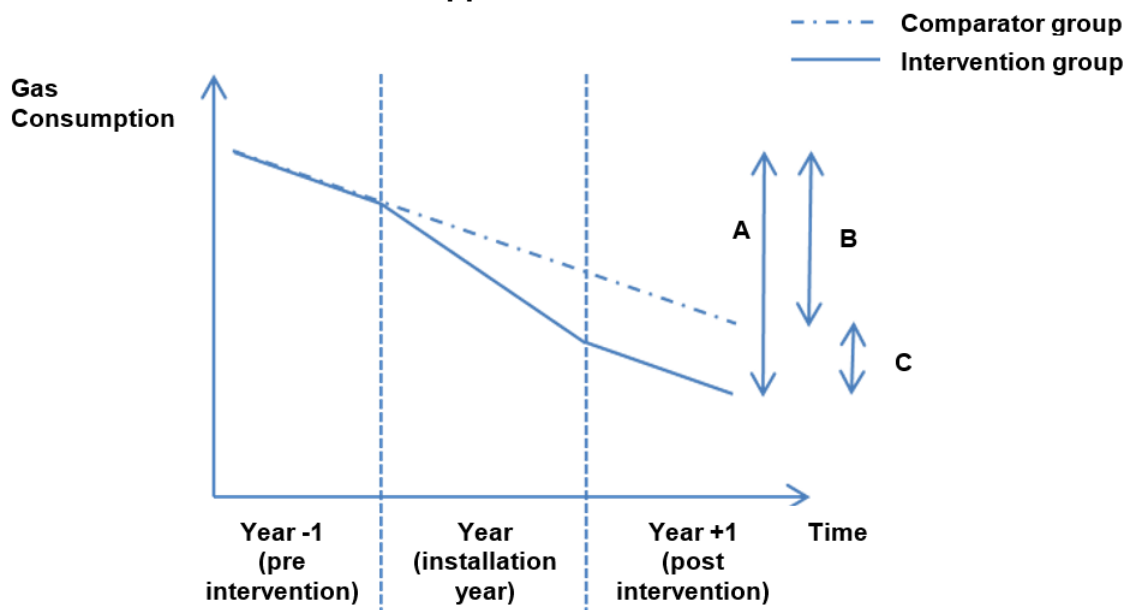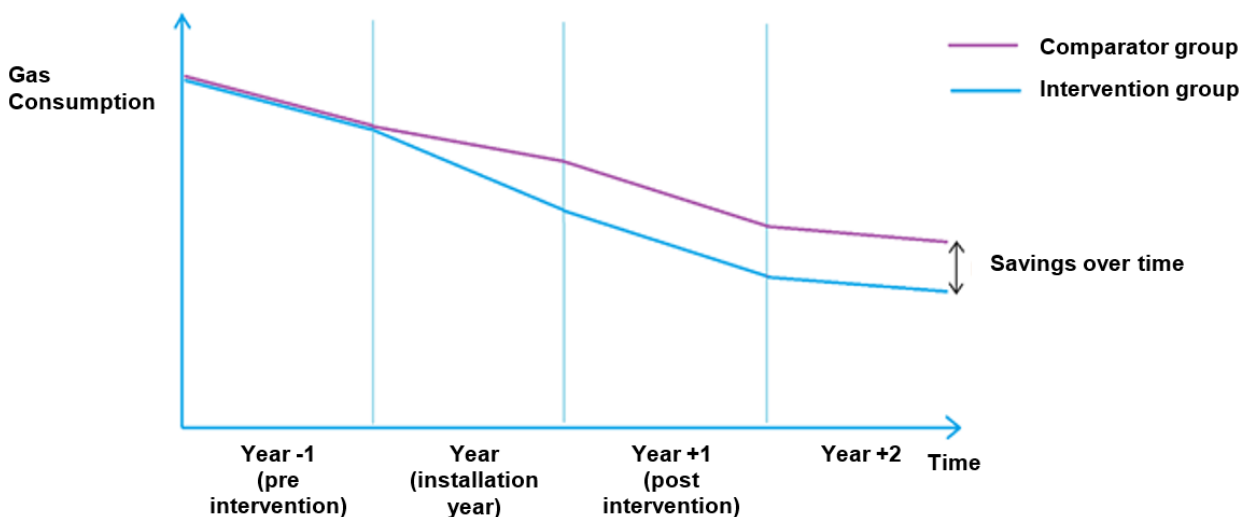
Figure 4.1: Difference in difference approach

Figure 4.2 below shows the difference in difference approach to finding savings for multiple years after installation ("Savings in years following installation").

Figure 4.2: Difference in difference approach for savings in years following installation

The rest of this chapter provides details on:

1. How the intervention and comparator groups are formed.
2. How the intervention and control groups are paired.
3. The difference in difference process which is used to calculate the savings estimates for each property in the intervention group for the given installation.
4. The weighting applied to the savings estimates for individual properties to derive summary statistics representative of the whole housing stock rather than just the stock which had the measures installed in the period considered.

Intervention group

A new intervention group is created for each energy efficiency measure considered every year. These properties must meet the following conditions:

- Have no other recorded energy efficiency measure recorded (on HEED, GD, ECO, FiTs or Gas Safe) in the year before, during or after installation[14].

- Have a valid gas consumption (between 2,500 and 50,000 kWh[15]) the year before and after installation – since these are the years gas consumption will be analysed to investigate the differences.

- For the analysis of solar PV installations, have a valid electricity consumption (between 500 and 25,000 kWh[16]) the year before and after installation.

Properties which meet these criteria are also excluded if:

- Gas consumption is suspected to be an estimated reading that has either been carried over from either of the previous two years, or has been assigned by suppliers based on typical consumption for certain household types, sizes etc. See Page 12 for more information on the process.

- The change in consumption between the year before and the year before that, or the change in consumption between the year after and the year of installation has increased by 50 per cent or more, or decreased by 80 per cent or more. Changes in consumption outside of this range would suggest something other than the installation of the measure is causing this change, such as a change in occupants or their circumstances.

- The property type is a flat. In a large number of cases insufficient address details were available to identify which flat within a building had the energy efficiency measure

---

[14] See footnote 15, these properties may have had a measure installed, which has not been recorded by BEIS. Note that the Gas Safe records that BEIS holds for is incomplete, containing records of boiler installations only up to mid-September 2018. Therefore properties that had a boiler installed after this period may be incorrectly selected in the control group or intervention group for other measures. Analysis of the impact of this based on previous years' results indicates this has little impact on the savings estimates for boilers, likely due to the large sample size.

[15] This range was taken following analysis of the sample and comparison with the range of consumption for properties with gas heating in the English Housing Survey, see November 2012 Annex B section 3.2:
www.gov.uk/government/uploads/system/uploads/attachment_data/file/65971/6871-need-report-annex-b.pdf.

[16] Prior to 2018, electricity consumption below 2,500 kWh has been excluded from the impact of measures analysis for Solar PV. In 2018 this exclusion was changed to only exclude electricity consumption below 500 kWh, due to many households having an electricity consumption below 2,500 kWh. In 2019 the minimum valid electricity consumption value was 100kWh. Following a review this was deemed too low and so for the 2020 publication 500kWh has been reinstated as the minimum valid electricity consumption value. This change in methodology for Solar PV means that direct comparisons with previous NEED impact of measures analysis should be made with caution. As electricity consumption in most households is far below 25,000 kWh, the relatively small number of properties with consumption greater than this have been excluded from this analysis.

installed and therefore it is not possible to accurately identify which gas meter relates to the flat which received the measure.

- Properties built post-1999 have specific installation standards regarding loft and wall insulation. Therefore, post-1999 properties have also been excluded from the cavity wall, solid wall and loft insulation analyses within this section[17].

- If a household has 'unknown' property features when matched to VOA and Experian data, it is removed from the sample. This is to ensure that any breakdowns by property feature will be consistent with the overall total. The only exception is where the energy efficiency rating band[18] is unknown.

Table 4.1 below summarises the conditions that had to be met for a property record to be included in the impact of measures analysis for the NEED 2020 publication.

Table 4.1: Summary of intervention group conditions

Variable	Condition
Date of installation[19]	Energy efficiency measure recorded as being installed between mid-July 2017 and mid-May 2018 and Solar PV recorded as being installed between 31st January 2017 to 30th January 2018 (inclusive).
Energy efficiency measures	No record of any other measure available through government schemes or recorded in NEED in the periods before, during or after installation.
Consumption	Gas consumption in 2016, 2017 and 2018 between 2,500 kWh and 50,000 kWh (excluding estimated readings). Electricity consumption in 2016, 2017 and 2018 between 100 kWh and 25,000 kWh (excluding estimated readings).
Change in consumption	Change in consumption between 2015 and 2016 is between -80 per cent and + 50 per cent, and the change in consumption between 2017 and 2018 is between -80 per cent and +50 per cent.
Property type	Flats are excluded due to insufficient address details being available to identify which flat in a block received the energy efficiency measure.
Property Features	If a household has 'unknown' property features when matched to VOA and Experian data it is removed from the sample. This is to ensure that any breakdowns by property feature will be consistent with the overall total. The only exception is where the energy efficiency rating band rating is unknown.

---

[17] In 2017 and 2018, newer properties were investigated as studies have shown that properties post 1999 have not necessarily had cavity wall and or loft insulation installed.

[18] This can be found on the Energy Performance Certificate (EPC) for a property, ranging from A-G. Note not all properties have an EPC.

[19] As is shown in the table, the range of dates from which installations are eligible for inclusion in impacts of measures are defined by the gas and electricity consumption periods for the year before and after. For the 2020 publication, the range of dates from which installations eligible for inclusion in impact of measures (2017) is defined as the day after the 2016 consumption year finishes to the day before the 2018 consumption year starts.

Applying these conditions reduces the number of properties included in the analysis, but it is considered necessary to make the analysis as robust as possible. For example, excluding properties known to have had other energy efficiency measures installed over the period in question, allows us to better isolate the impact of the energy efficiency measure being considered.

Comparator group

As the analysis is carried out retrospectively it is not possible to create a comparator group in advance of measures being installed. A separate comparator group is created for each measure for each intervention year from the properties in the NEED dataset.

All the conditions applied to the intervention group are also applied to the comparator group, with the exception of the comparator group having no recorded energy efficiency measure installations in NEED over the period in question (the year before, during or after the given installation year).

The comparator group is both the same size as the intervention group and contains properties that have similar property features[20]. Once the comparator group has been selected, each property is matched to an equivalent property in the intervention group (i.e. one that has similar property features). To ensure that there is no loss of properties (owing to pairing not being possible in all cases), while pairing by as many features as possible, an algorithm is used which pairs by as many of the below characteristics as possible while meeting the constraint that all the intervention group properties must be paired. The characteristics, in order of the importance the algorithm assigns to them, are:

- Gas consumption band

- Type of property

- Age band of property

- Number of adults living in the property

- Energy efficiency rating band (from EPC)

- Region

This process is referred to "elastic pairing", as the characteristic which properties are paired by changes depending on the distribution of characteristics in the intervention group.

Removing properties with estimated consumption

For properties without a meter reading in a particular year, a reading for that meter is generally estimated in the meter data received by BEIS, and recorded in NEED, rather than being left blank. As actual consumption is not recorded for these properties, they should not be included

---

[20] In a small number of cases there is no property with the same combination of characteristics in the group of properties with no measures installed (i.e. the group the comparator group is selected from), this results in a small number of properties being excluded from the matched pairs analysis for each energy efficiency measure.

in either intervention or comparator groups. There are two scenarios in which meter reads are estimated:

1) When a meter has had consumption recorded in the past, the consumption figure recorded will be carried forward to future years as the estimate. Properties with these estimates are removed by finding properties where annual consumption is identical to either of the previous two years. This filter was introduced in the NEED 2018 publication.

2) When a meter is new it is assigned an estimated consumption figure based on its profile. This can be detected by looking for a large number of properties with identical annual consumption. In practice this is done by rounding the consumption of all properties to the nearest kWh, counting the number of properties by each kWh, ordering the counts by kWh, and looking for changes in counts between sequential kWh values of over 300 per cent. For example, if 50 properties have a consumption of 12,235 kWh and 345 properties have a consumption of 12,236, all 345 properties with a consumption of 12,236 kWh would be removed from the intervention and comparator groups. Such a change would be a 590 per cent increase in the frequency of consumption, above the 300 per cent threshold, meaning that it is very likely that most properties with gas consumption of 12,236 kWh in fact have estimated readings. For some years this finds 10 – 15 rounded kWh figures each year and for others it finds none. While it is a beneficial filter, is it likely that lower frequency estimated consumption figures remain in the analysis e.g. from smaller new build developments. This filter was introduced in the NEED 2019 publication.

Calculate the relative change in consumption for each pair

After each property in the intervention group is paired with a property from the comparator group, the savings resulting from the given energy efficiency measure are calculated as detailed below:

$$Intervention\ property\ change\ in\ consumption$$
$$= Intervention\ property\ consumption\ year\ after\ installation\ year$$
$$- Intervention\ property\ consuption\ year\ before\ installation\ year$$

$$Comparator\ property\ change\ in\ consumption$$
$$= Paired\ comparator\ property\ consumption\ year\ after\ installation\ year$$
$$- Paired\ comparator\ property\ consumption\ year\ before\ installation\ year$$

$$Estimated\ savings\ from\ the\ given\ energy\ efficiency\ measure$$
$$= -(Intervention\ property\ change\ in\ consumption$$
$$- Comparator\ property\ change\ in\ consumption)$$

The range of differences in change in consumption between paired properties vary considerably. Figure 4.3 shows an example of the differences in changes in gas consumption between the intervention and comparator pairs (unweighted), illustrating the large variation in consumption changes. This variation is due to wide range of events which take place and remain unaccounted for (e.g. change in occupants of property).

Figure 4.3: Difference of change in gas consumption between intervention and comparator properties after installation of solid wall insulation in 2015
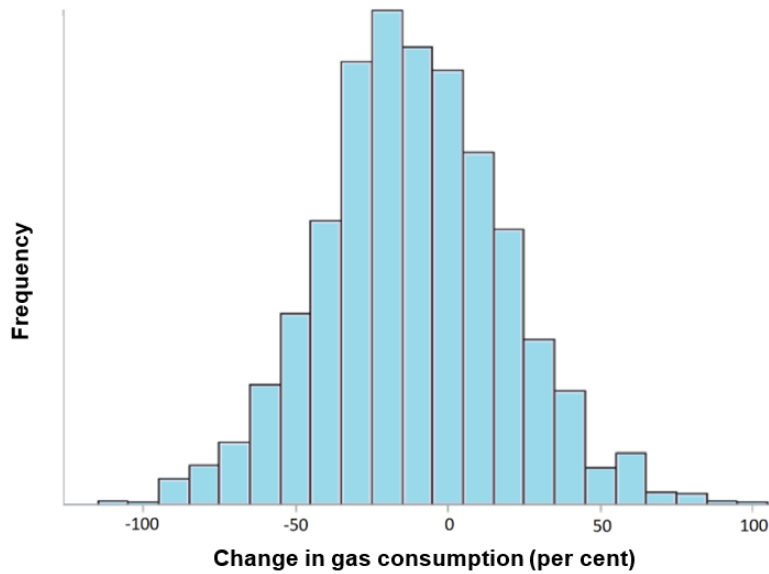
Removing outliers

Differences outside of a 95 per cent window[21] are considered to be the result of an inappropriate pairing and the filter prevents these pairs from being included. This is done by removing the largest and smallest 2.5% of differences[22]. To illustrate this Figure 4.4 below shows how the values to filter on are determined using this filter for boilers installed in 2016.

Figure 4.4: Filtering boundaries for paired properties using 95 per cent window for boilers installed in 2016

---
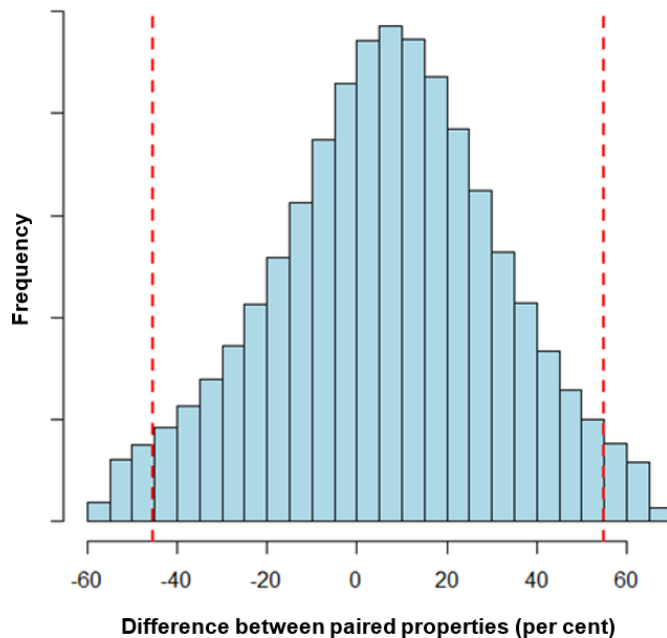
[21] The 95 per cent window is achieved by removing pairings with savings below the 2.5th percentile and above the 97.5th percentile. The central 95 per cent of the data is retained for analysis.
[22] In 2019 the removal of outliers was done by identifying the high density region, set at 95 per cent, which selected the 95 per cent of the data which is most densely populated.

Weighting

Since the November 2013 NEED publication, weighting has been applied to the results to mitigate the impact of a biased housing stock in the intervention group. The analysis covers a representative sample of households which have received a measure; however, this may not be representative of the full housing stock. The weighting provides a better understanding of the impact of the measure for the general housing stock for England and Wales.

To ensure that the majority of properties were used in the analysis while weighting by as many features as possible, an algorithm was created which weighted by as many of the below characteristics as possible while meeting the constraint that over 99 per cent[23] of the national housing stock (with flats excluded) be represented. The characteristics, in order of the importance the algorithm assigns to them, are identical to those used for pairing:

- Gas consumption band

- Type of property

- Age band of property

- Number of adults living in the property

- Energy efficiency rating band (from EPC)

- Region

This is similar to the elastic pairing process, however in this case the process is referred to as "elastic weighting".

The weighting factor weights each record in the intervention group based on its frequency - relative to how often it should appear if this group had the same distribution as the total housing stock. For example, if a property type is more common in the intervention group than the housing stock as a whole, then the weighting factor acts to reduce its overall contribution, while if the property type is less frequent in the group than is the case in the housing stock, then weighting factor acts to increase its contribution.

The weighting factor is calculated using the following formula:

$$\text{Weighting factor for record n } (w_n) = \frac{Housing\ stock\ _g}{Total\ housing\ stock} \times \frac{1}{Sample\ stock_g} = A \times B$$

Where: *n*   *is the record number in the sample*
      g   *is the group number (e.g. group 1 = built pre-1919, detached, with 1 to 2 bedrooms)*
  housing_g   *is the number of properties in group g in the total housing stock (VOA)*
  sample_g   *is the number of properties in group g from the intervention population (full NEED data set)*

---

[23] In the 2019 publication, the threshold of 83 per cent was used. This was amended for the 2020 publication to ensure the weighting the results better represents the housing stock.

Part A is equivalent to the proportion of the total housing stock a group accounts for (i.e. the number of records in group g in the housing stock divided by the total housing stock).

A is multiplied by B; where B the reciprocal of the number of properties in the intervention group in the same group (i.e. one over the number of properties in the intervention group in group g).

It is worth noting that the sum of the weights for the sample always totals one.

The percentage saving from each measure can then be calculated by summing the weighted percentage saving for each record, to give a percentage saving for the population:

$$Population \ \% \ saving = \sum_{n}\left(\% \ age \ saving \ for \ record \ n \ \times w_{n}\right)$$

Historically, the weighting had the least impact on estimates of savings from installation of loft insulation and boilers, suggesting that these two measures have been installed in a range of properties most reflective of the housing stock.

Impact of measures in algorithmic form

The method is presented in algorithmic form below.

Select all properties in England and Wales in year n as basis for comparator group, then apply filters:
1) Remove where the consumption value is estimated
2) Remove where the change in consumption between the year before and the year before that, or the change in consumption between the year after and the year of installation, is either over 80 per cent reduction or 50 per cent increase
3) Remove where consumption value is not between 2,500 – 50,000 kWh for gas, or 500[24] – 25,000 kWh for electricity
4) Remove where one or more measures were installed in the years before, during or after the period in question
5) Remove flats
Create interventions group by selecting all properties in NEED with an installation of each measure during the period in question.

Create intervention groups for combinations of measures by finding properties with multiple interventions during the period in question.

Apply filters to intervention groups:
1) Remove where the consumption value is estimated
2) Remove where the change in consumption between the year before and the year before that, or the change in consumption between the year after and the year of installation, is either over 80 per cent reduction or 50 per cent increase
3) Remove where consumption value is not between 2,500 – 50,000 kWh for gas, or 500[24] – 25,000 kWh for electricity
4) Remove where one or more measures were installed in the year before or the year after

---

[24] Note that in the 2019 publication the range of 100-25,000kWh was given for valid electricity consumption. This has been reviewed and changed to 500-25,000kWh, in line with the 2018 publication.

5) Remove flats

6) If assessing solid wall insulation, cavity wall insulation or loft insulation[25], remove where the property was built after 1999 (due to changes in building regulations for properties built after 1999)

Calculate banded gas or electricity consumption for properties in comparator and intervention groups.

Assign energy efficiency rating band (from EPC) to comparator and intervention groups by UPRN.

Apply elastic pairing process to find optimal characteristics to pair by.

Apply elastic weighting process to find optimal characteristics to weight by.

The following section is carried out with 50 iterations:

> Randomise the order of the comparator properties prior to joining the properties
> Join the intervention and comparator properties
> Calculate the relative changes in consumption
> Remove outliers
> Create summary statistics (End of iterations)

Calculate averages of summary statistics across all iterations.

Variations in estimated savings between years

The impact of measures analysis has produced estimated savings for each measure, using methodology which is not fully consistent with previous years. It is therefore important to consider each estimated saving as indicative rather than precise. There are a number of factors that are likely to contribute to variations in estimated savings:

Method and Data

- Whilst the fundamental difference in difference approach for analysis of the impact of measures has remained consistent since the creation of NEED, refinements in methodology have been made over time. The sensitivity of the estimates to these changes has not been fully assessed and therefore variation seen in estimates may be a result of methodological changes. As a result, comparisons of estimated savings between results published in different years should be done with caution.

- The number of measures installed differs year on year. In installation periods where fewer measures have been installed, the intervention and comparator group sample size will likely be smaller as a result. Estimates based on small sample sizes are likely to be less reliable.

- Increasing prevalence of measures outside of government schemes. This means that properties included in the intervention and comparator group may have had energy

---

[25] In the 2019 publication properties were removed if built after 1999 for solid wall insulation and cavity wall insulation. In line with building regulations information, properties built after 1999 have also been removed for loft insulation. Prior to 2019 publication, this exclusion was only applied to cavity wall insulation and loft insulation.

efficiency improvements made that have not been identified in the NEED data. All other things being equal, this would lead to a decrease in the savings found using NEED over time.

Unknown Installation Information

- The quality of installations may vary between years.

- The average size of the install may vary between years (e.g. larger solar panels can generate more electricity).

- The attributes of the installation are unknown, for example the number of walls in the property covered by wall insulation and the thickness of loft insulation.

- The brand or subtype of measure may vary between years. For example, while cavity wall insulation is considered to be a single class of intervention, in reality there are several types of cavity fill (notably bead and mineral wool), which may have different impacts [26].

Unknown Household Information

- The results may be different for early adopters of novel measures because this self-selecting treated population may consume energy in a different pattern to general consumers.

- Any variation between the treated populations which is not known from the data. This could include age of residents and the number of children in the home.

- Changes in energy consumption behaviour which follow the installation of an energy efficiency measure and may also vary over time and between different types of household. An example is when a household chooses to heat their home to a higher temperature following installation of a measure; this is a known phenomenon called comfort taking.

BEIS seeks to continuously improve the data and methods used for analysis in the NEED publication, along with developing outputs presented that best suit user needs. The methodology therefore may be updated in future years as further methodological improvements are made. Note this may result in publishing revised estimates for previous years. Users are encouraged to provide comments and feedback and should be sent to: energyefficiency.stats@beis.gov.uk

---

[26] For more information go to:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/853563/Gas_consumption_savings_from_bead_and_mineral_wool_cavity_wall_insulation.pdf.