2022 Heat Network Consumer and Operator Survey

Technical Report

Kantar Public

April 2023

Contents

Survey

Overview of survey methodology

Heat Network domestic consumers

The survey was carried out using a mixed-mode, self-completion approach known as Addressed Based Online Surveying (ABOS). Fieldwork took place between 22 March and 12 July 2022. The target populations were a sample of households with a heat network as their energy or water source and a matched sample of comparison households (without a heat network).

Selected households in both the heat network consumer and comparison samples were sent a letter inviting them to participate in the survey online. Two reminders were sent to those who did respond to the invitation. The first reminder included a paper copy of the 23-page questionnaire. To maximise response rates, respondents were sent a £5 gift voucher on completion of the survey. This was to maximise the response rate and minimise the risk of non-response bias.

The average interview length online was approximately 25 minutes.

Heat Network non-domestic consumers

The initial research plan involved conducting a survey of 1,500 non-domestic heat network consumers (and a matched comparison sample of 500 non-domestic energy consumers). Kantar Public issued survey invitation letters to approximately one-third of non-domestic heat network consumers in their sample (8,326 out of 24,000). This sample batch sample was expected to achieve c. 650 surveys with non-domestic heat network consumers. However, the initial response was substantially poorer than expected. Following a reminder invitation to participate, the response levels stayed low. The total number of surveys returned was 102. Following checks on the data there were only 32 complete surveys with non-domestic consumers on heat networks. Kantar Public and BEIS agreed that issuing the remainder of sample to non-domestic consumers would not achieve a large enough sample to provide robust analysis and that the non-domestic survey would not continue. All analysis of heat network consumers in the main report and this technical annex relate to domestic heat network consumers.

Heat Network operators

Kantar Public conducted a Computer Assisted Telephone Interview (CATI) survey with 130 heat network providers. CATI was chosen as it typically returns higher

response rates than other modes for surveys of organisations and makes it possible to screen for the correct person within the organisation to answer the questions. Operators were sent advance letters or emails to notify them of the research and provide an opportunity to opt out. Fieldwork took place between 18 March and 30 May.

Sampling

Building a sample frame of Heat Networks

An address-level sample frame was produced using the Heat Metering and Billing regulations (HMBR) database shared by BEIS and publicly available Energy Performance Certificate (EPC) databases. A summary of the process is provided below.

The HMBR database included the postcode of the energy centre of each registered heat network and the number of domestic and non-domestic consumers served by each centre. When processing the HMBR database we have followed the procedures developed for the 2017 survey[1].

The HMBR database was firstly aggregated by the energy centre postcode so that there was one record per energy centre postcode. This step was necessary to build an address-level file. This is because the addresses closest to each energy centre were selected as being likely to be on a heat network. If there are duplicate energy centre postcodes, the same addresses would be selected multiple times. In addition, there was a low likelihood that two separate heat networks would have boilers in exactly the same postcode (as full postcodes only typically cover a small geographical area).

To begin with the HMBR database that Kantar Public received had 8,640 rows. Following the aggregation by energy centre postcode described above there were 7,442 unique energy centre postcode entries (Heat Networks). After this aggregation, there were 136 rows in the aggregated database which contained information from more than one row in the original database.

During the aggregation process Kantar Public retained all information from the original HMBR extract. Where there were duplicate energy centre postcodes, the data was aggregated into a single record. This involved adding up the number of customers, and the number of buildings supplier from each of the original rows to

[1] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/665448/HNCS_-_Technical_Report_-_FINAL.pdf

calculate an overall total. In addition, all information was kept on the building supplier, technology and energy source used.

It should be noted that the HMBR database contained 109 duplicate energy centre postcodes which had different heat network supplier names, of these 26 were the same name with slightly different spelling. For the remaining 83 duplicates with different supplier names, a count of the number of supplier names listed was produced. The supplier name with the most customers was taken as the supplier name for the aggregated file. It is worth noting that there is some uncertainty around the supplier name in the HMBR file as it is possible that networks have been registered through agents, management companies or developers.

Kantar Public then converted the HMBR data into an address-level file by using the residential Postcode Address File[2] (the Royal Mail database of delivery points) to select the closest neighbouring postcodes to each energy centre, until the number of addresses selected was a close match to the number of domestic dwellings listed as on the heat network in the HMBR dataset (after the earlier aggregation by energy centre postcode). It should be noted that this approach is imperfect and the address-level file may include properties which are adjacent to heat networks but not actually heated by them.

The total number of addresses identified through the HMBR database was 381,538. This is a relatively close match to the 322,181 residential customers listed on the HMBR database. However, there is still a risk that many of the domestic address listed as being on a heat network may not actually be on a heat network.

EPC

The publicly available EPC databases for England and Wales and Scotland were used to identify Heat Network addresses[3]. The EPC database was filtered to only include addresses that had a Heat Network as a main heating source, this filter was based on the main source of heating description. Any address which had a community scheme or community heat pump listed as the main source of heating was identified as being on a heat network. It should be noted that the EPC databases do not comprehensively cover all addresses. They only include buildings that have been sold, rented, or constructed in the last ten years or so.

Therefore, Kantar Public used the EPC data in conjunction with the publicly available dataset Royal Mail Postcode Address File (PAF). These were used to identify other properties that are not listed in the EPC but that are likely to be on a heat network

[2] This excludes large mail users and any records flagged as a business

[3] EPC data for England and Wales was taken from the domestic Energy Performance of Buildings register (https://epc.opendatacommunities.org/domestic/search), EPC data for Scotland was taken from the Scottish Energy Performance Certificate Register (https://statistics.gov.scot/data/domestic-energy-performance-certificates)

because of their conjunction to buildings known to be on a heat network. The Royal Mail PAF database was used to obtain address counts for each postcode. The EPC was then used to produce counts of the number of properties heated by a Heat Network in each postcode. Where the number of addresses identified through EPC was greater than 50% of the number of addresses listed in the postcode in the PAF, Kantar Public made the assumption that all other addresses in that postcode were also likely to also be on the same heat network.

Combining the data

Finally, the HMBR and EPC data was combined and de-duplicated to ensure that addresses could only appear once in the final sample frame. In total, Kantar Public identified 1,012,049 domestic addresses that were likely to be on a Heat Network. Of these 187,156 (19%) were identified from the HMBR database, 194,382 (19%) from both the HMBR and EPC and 630,511 (62%) just from the EPC.

Risk of inaccurately identifying addresses as being on a heat network

As noted previously, the way in which the address-level sample frame was created means that there is a risk that some addresses may not actually be on a heat network.

From the addresses identified through the HMBR, 193,635 (51%) were identified within an energy centre postcode listed in HMBR. The remaining 187,903 (49%) were identified from postcodes adjacent to an energy centre postcode. There is a risk that some of these addresses are adjacent to heat networks but not actually heated by them.

Of the addresses identified through the EPC, 492,444 (78.1%) were identified as being on a heat network – it should be very likely that these are on a heat network as these addresses have a heat network listed as their main source of heating[4]. 138,067 (21.9%) addresses were identified from a postcode where more than 50% of addresses were on a heat network[5].

As the 2017 survey used this methodological approach for sampling successfully, the approach was repeated with projected minimal risk.

Overall sample size

The target sample size for domestic consumers on heat networks proposed was 4,800.

[4] A Community scheme is listed as their main form of heating

[5] In some cases, there were postcodes where the number of addresses listed as being on a Heat Network exceeded the number of addresses listed in PAF. Where this was the case all PAF addresses in that postcode have been included.

However, as noted above, there was a degree of uncertainty associated with the sample building process. As for the 2017 study, it was likely that some of the addresses sampled as being on a heat network, will report that they are not on a heat network when participating in the survey.

As such, the sample design needed to take into account likely eligibility rates. The starting sample design aimed for 5,300 complete surveys with heat network consumer sample. This meant that even if c.10% of addresses that participate reported not being on a heat network then the target of 4,800 would be achieved.

It should be noted that the interview data was still used if people reported not being on a heat network; they were included in the comparison sample. This was consistent with the approach used in 2017.

Sample design

Kantar Public derived an explicit stratification variable based on HMBR registration, nation, scheme type and Heat Trust registration. The table below shows the population profile and the proposed sample design if response rate assumptions were met.

Table 1: Domestic consumers sample design

Stratum	**Population**		**Proposed sample design**	
	N	**%**	**N completes targeted**	**N to sample (assuming 16% RR)6**
HMBR registered, LA scheme type, in Scotland, not registered with the Heat Trust	4,854	0.5%	c.162	c.1,012
HMBR registered, LA scheme type, not in Scotland, registered with the Heat Trust	308	0.0%	c.12	c.75
HMBR registered, LA scheme type, not in Scotland, not registered with the Heat Trust	62,123	6.1%	c.928	c.5,801

[6] We have assumed a 16% response as outlined in the ITT

HMBR registered, Private scheme type, in Scotland, registered with the Heat Trust	1,824	0.2%	c.61	c.380
HMBR registered, Private scheme type, in Scotland, not registered with the Heat Trust	5,384	0.5%	c.180	c.1123
HMBR registered, Private scheme type, not in Scotland, registered with the Heat Trust	20,988	2.1%	c.819	c.5,120
HMBR registered, Private scheme type, not in Scotland, not registered with the Heat Trust	124,609	12.3%	c.695	c.4,342
HMBR registered, Social scheme type, in Scotland, not registered with the Heat Trust	11,149	1.1%	c.372	c.2,325
HMBR registered, Social scheme type, not in Scotland, registered with the Heat Trust	1,338	0.1%	c.52	c.326
HMBR registered, Social scheme type, not in Scotland, not registered with the Heat Trust	148,961	14.7%	c.831	c.5,191
Not HMBR registered, in Scotland, registered with the Heat Trust	94	0.0%	c.3	c.20
Not HMBR registered, in Scotland, not registered with the Heat Trust	21,391	2.1%	c.352	c.2,202
Not HMBR registered, not in Scotland, registered with the Heat Trust	982	0.1%	c.38	c.240

Not HMBR registered, not in Scotland, not registered with the Heat Trust	608,044	60.1%	c795	c.4,967
Total	1,012,049	100.0%	c.5,300	c.33,125

This design aimed to achieve the following effective sample sizes (after design weighting has been applied to compensate for the fact sampling probabilities vary between strata):

Table 2: Target sample size by stratum

	Likely Number of interviews[7]	**Estimated effective sample size[8]**	**Estimated Design Effect**
Scheme type			
Local Authority Scheme type	c.1,098	c.1,045	1.05
Private Scheme type	c.1,765	c.1,006	1.75
Social Scheme type	c.1,252	c.958	1.31
Not HMBR Registered & Scheme type unknown	c.1,185	c.848	1.40
HMBR Registered			
HMBR registered	c.4,115	c.2,659	1.55
Not HMBR Registered	c.1,185	c.848	1.40
Scottish Heat Networks			
Scotland	c.1,130	c.1,000	1.13

[7] These are achieved completes and do not account for ineligibility. The likely number of completes and effective sample size are likely to be lower once this has been accounted for.
[8] Non-response weighting is likely to reduce these figures a bit further.

England / Wales	c.4,170	c.1,794	2.32
Heat Trust Registered			
Heat Trust	c.1,002	c.1,000	1.00
Not Heat Trust	c.4,298	c.1,861	2.31
Overall	c.5,300	c.2,456	2.71

Selecting the sample

Within each stratum Kantar Public sorted the sample frame by a range of variables:

- Region
- Indices of Multiple Deprivation (IMD)
- Output Area Classification (OAC)
- Proportion of residential buildings in LSOA/dataZone which are flats (Based on published VOA and Scottish Assessor statistics)
- Postcode

The sample was then systematically selected within each stratum, so it was representative of the population for these characteristics.

The sampling probability for each stratum was calculated as the number of cases issued divided by the number of cases in the sample frame.

Domestic consumers comparison sample

Given that the population of heat network consumers differs from the general population, Kantar Public drew a matched comparison sample as similar as possible to the population of heat network consumers. This allows the research to more confidently attribute differences between the two groups to membership of a heat network, rather than other unrelated differences between the two populations.

The target sample size proposed for the domestic comparison sample in the original specification was 1,700. To take into account overlap in sample files, where some households sampled as having a heat network reported that they did not have a heat network (and were used in the final comparison sample), Kantar Public designed the

matched comparison sample with a target of c.1,200 completes. In total, there were 1,733 comparison respondents included in the final data..

For the matched comparison sample, Kantar Public used a similar approach to that successfully used for the 2017 study. This used statistical modelling (a binary logistic regression) to identify areas (Census Output Areas) that are not on heat networks but that are the most similar to the areas in which heat networks are located. Kantar Public included the following variables in the matching algorithm:

- Region
- Deprivation (IMD)
- Area Characteristics (OAC11, Rural/Urban classification)
- Housing stock data (types of building, council tax bands[9], building age[10])
- Census data (e.g., tenure, average HH size[11], average number of rooms[12]).

Once Kantar Public identified the output areas to draw the comparison sample from, sample was drawn from the PAF. Prior to making a systematic random selection, addresses in the selected output areas were sorted by the same characteristics as used for the domestic consumer sample:

- Region
- Indices of Multiple Deprivation (IMD)
- Output Area Classification (OAC)
- Proportion of residential buildings in LSOA/dataZone which are flats (Based on published VOA and Scottish Accessor statistics)
- Postcode

Operators Survey

Kantar Public aggregated the supplier names listed in the HMBR file to produce a file containing unique Heat Network suppliers. This file is based on the 7,400 unique Heat Networks identified in the HMBR file following the aggregation to one row per energy centre postcode.

The aggregation primarily used supplier name, with some manual review to adjust for misspelt or extremely similar names with matching supplier addresses. It should

[9] VOA statistics were sourced for England and Wales LSOA's, Scottish Assessors statistics were sourced for Scottish DataZones
[10] For England and Wales only, as equivalent data for Scotland at an appropriate granularity was not available.
[11] England and Wales only
[12] England and Wales only

be noted that in 68 instances Kantar Public made the assumption that two networks were operated by the same supplier even though the address provided for the supplied differed. This is likely to be because some suppliers had more than one office.

The unique supplier file produced contained 1,323 unique suppliers. In order to maximise the likelihood of a successful tele-match, Kantar Public used a number of external suppliers to tele-match all combinations of supplier names and addresses in the HMBR file. In total there were 1,543 combinations of supplier name and address that were tele-matched. Following tele-matching, the contact names, email address and telephone numbers were deduplicated. This returned the file to be unique by supplier name. Where multiple contact details were matched with different supplier addresses, both phone numbers were appended to the sample. The table below shows the breakdown of tele-matching and deduplication

Table 3: Operator sample telematching

	Count
No telephone number found	737
Telephone number identified	792
Phone number provided by BEIS	14
Removed in deduplication	172
Usable	634

In order to achieve as many of the 300 telephone interviews as originally required (when working under the assumption that the sample would have 1,500 contactable suppliers), all 634 suppliers identified were issued.

The table below shows the population profile of the HMBR Operators sample and the profile of the issued sample.

Table 4: Operator sample

	Population		**Issued**	
	N	**%**	**N**	**%**

Country				
England	1,170	88%	571	90%
Scotland	134	10%	51	8%
Wales	19	1%	12	2%
Number of networks supplied				
1	910	69%	356	56%
2-4	205	16%	111	18%
5-9	79	6%	58	9%
10+	129	10%	109	17%
Scheme Type*				
Local Authority	69	5%	53	8%
Private	992	75%	376	59%
Social	260	20%	205	32%

* This was derived as described in the domestic consumer survey section above.

Questionnaire development

The survey questionnaires were developed through a mixture of desk research, stakeholder workshops and cognitive testing.

For both domestic consumers and operators surveys, the starting point was reviewing and agreeing the key research questions. For consumers, Kantar Public mapped the research questions to the 2017 questionnaire to see where existing questions needed changing or where additional questions were needed.

Kantar Public and BEIS held two questionnaire workshops to agree which areas were priorities to develop new questions. From this, an iterative approach was taken to develop a draft questionnaire for cognitive testing.

The primary purpose of cognitive testing is to examine how questions perform when asked of survey respondents. That is, if respondents understand the questions as intended, and if they can provide accurate and consistent answers. For the operators' questionnaire, this was achieved through in-depth, semi-structured interviews with a small number of heat network operators.

The objectives of cognitively testing the questionnaire were to:

- Explore understanding of question wording and phraseology for both the question as a whole and any key words and phrases it might contain. We were particularly interested in how respondents understood the descriptions we used of the services they had interacted with (such as ''district heat network').

- Understand what decision processes the respondent uses in coming to an answer

- Test overall feelings about the questionnaire

- Make recommendations on how the questions can be improved or refined

Cognitive interviews were conducted with five with domestic heat network consumers and one non-domestic heat network consumer.

Domestic consumers were identified through the researcher's personal networks. Participant characteristics are shown in Table 5 below.

Table 5: Domestic consumer cognitive testing sample

Gender	
Male	0
Female	5
Age	
18-24	0
25-34	4
35-44	1
45+	0
Ethnicity	
White	4

BAME	1
Region	
London	5
Rest of the UK	0

Property	
Flat or Maisonette	5
Terrace house	0
Terrace bungalow	0
Semi-detached house / end of terrace house	0
Semi-detached bungalow / end of terrace bungalow	0
Detached house	0
Detached bungalow	0
Tenure	
Own home outright	0
Buying it with the help of a mortgage or loan	0
Part own and part rent (shared ownership)	3
Rent from a council or local authority	2
Rent from a housing association, housing co-operative, charitable trust or registered social landlord	0
Rent from a private landlord or letting agency	0
Rent from someone else	0
Live rent-free in another person’s property	0

Heat network type	
Communal	5
District	0

The non-domestic consumer was identified from an address-level sample frame produced using the Heat Metering and Billings Regulations (HMBR) data shared by BEIS and the publicly available non-domestic EPC databases. Interviews lasted approximately one hour and were conducted face to face or online via Zoom.

Recruitment with non-domestic consumers was challenging and only one cognitive interview was conducted with a non-domestic consumer. This respondent was a facilities manager in a building. They were employed by an organisation which was the acting landlord for the building owner.

Following the cognitive testing, Kantar Public provided a report with recommended revisions.

The full domestic consumer and operators questionnaires are in Annex C and D of this report. A summary of the content of each is below.

Domestic consumers questionnaire structure

- Property characteristics
- Satisfaction with heating
- Type of heating and insulation in property
- Heating problems
- Outages
- Underheating
- Overheating
- Complaints
- Information received about heating in property
- Billing
- Heating payments
- Amount paid for energy
- Amount paid for standing charges
- Whether the price of energy has changed

- Views on the price paid for energy
- Is it in line with expectations
- Whether struggling to pay energy bills
- Demographics

Operators questionnaire structure

- Organisation type and characteristics
- Types of networks operated
- Heating sources used
- Attitudes towards low carbon sources
- Billing
- Format and frequency
- What is included within bills
- How often tariffs are reviewed
- Customer survey and complaints
- Vulnerable customer support
- Whether experienced heating outages or planned interruptions
- Knowledge of rights, regulations and powers

Survey fieldwork

Heat Network consumers

The survey was carried out using a mixed-mode, self-completion approach known as Addressed Based Online Surveying (ABOS). Fieldwork took place between 22 March and 12 July 2022. Consumers were initially sent an invitation to complete the survey online. Those who had not responded were sent a reminder letter (including a paper questionnaire) asking them to either complete the survey online or complete a paper questionnaire and send back to Kantar Public. A £5 incentive was provided upon completion of the survey.

Kantar Public carried out verification of all responses received to ensure they were eligible as completed surveys. The initial aim was for 6,500 responses (with a response rate of 16%). After verification there were 3,977 complete surveys. In total, 2,214 verified surveys were completed online and 1,853 by postal questionnaire.

In total, 2,244 valid responses from Heat Network consumers and 1,733 valid comparison group responses. The final response rate was 10%.

Heat Network operators

Kantar Public conducted a Computer Assisted Telephone Interview (CATI) survey with 130 heat network providers. Operators were sent advance letters or emails to notify them of the research and provide an opportunity to opt out. Fieldwork took place between 18 March and 30 May. The average interview length was approximately 31 ½ minutes. The final response rate was 25%.

Weighting

Domestic Consumer survey

Identifying heat network households

In total, 37,000 addresses were sampled from the sample frame of addresses likely to be on a heat network. However, as noted in the sampling section above, not all addresses on the sample frame were part of a heat network. In identifying heat network and non-heat network households Kantar Public were attempting to minimise the amount of misclassification between heat network and non-heat network households.

Heat network households could be identified in two ways:

- Sample Frame data: Kantar Public used the EPC to identify addresses that had a heat network as their main source of heating.
- Respondent data: In the questionnaire, respondents were asked if their household was part of a heat network (survey questions NETWORK/WATERSOURCE)

Households were classified as being on a heat network if:

- The household was identified as being on a heat network in the EPC data (even if the respondent was unaware that they are on a heat network).
- The household was identified as being on a heat network in the respondent data and was identified as being on a heat network in the HMBR sample frame (but not the EPC data).
- The household was selected as part of the comparison sample but identified as being on a heat network in the respondent data and no EPC certificate could be found for the property.

Households were classified as not being on a heat network if:

- The household indicated they are not on a heat network (and an EPC certificate was available for the address and indicated an alternative main heat source)
- The household was not identified as being on a heat network in the sample or in the respondent data

In total 2,242 households were classified as being on a heat network. The table below shows the breakdown of households indicating being on a heat network in the respondent data, and the final heat network classification by sample source.

Table 6: Heat network allocation

Sample source	Respondent data		Heat network classification		Total
	On a heat network	Not on a heat network	On a heat network	Not on a heat network	
HMBR Only	321	827	321	827	1148
	28%	72%	28%	72%	
HMBR and EPC	1028	223	1251	0	1251
	82%	18%	100%	0%	
EPC Only	480	158	638	0	638
	75%	25%	100%	0%	
Comparison sample	42	896	32	906	938
	4%	96%	3%	97%	
Total	1871	2104	2242	1733	
	47%	53%	56%	44%	

42 comparison sample households indicated they were on a heat network in the respondent data. These addresses were searched against the domestic EPC dataset. Certificates were found for 10 of these. These certificates indicated that the household's main heat was supplied by a source other than a heat network. These households were classified as not being on a heat network. EPC certificates could not be found for 32 households, these were classified as being on a heat network.

The heat network classification rate varied across sample sources, with all households identified through the EPC classified as being on a heat network, while 28% of households identified solely through the HMBR register were classified as being on a heat network.

To account for this varying rate, all households identified in the sample as being on a heat network were included in the weighting process, even if they were identified in the respondent data as not being on a heat network. These ineligible households were then removed from the heat network sample and moved to the comparison

sample. These households were then included in the comparison sample weighting process. The 32 comparison sample households which were classified as being on a heat network were assigned a modelled heat network weight.

Weighting heat network consumers

Weighting was required to compensate for variation in sampling fractions and systematic differences in response rate between population sub-groups. To compensate for variation in sampling fractions, a design weight was calculated based on the sampling probability of each heat network consumer address. This varied by stratum due to the disproportionate sample design[13].

There were no comprehensive population statistics available for heat network consumers which could be used as weighting targets to account for systematic differences in response rate between population sub-groups.

For all households sampled as being on a heat network Kantar Public calculated non-response weights using the data available on the sample frame. This was consistent with the approach used for the previous heat networks study in 2017. Kantar Public ran a logistic regression model to predict the estimated probability of any given sampled household taking part in the survey. The non-response weight was then calculated by inversing these estimated response probabilities. Consumers which were estimated to have been less likely to take part were given larger weights, so that they were not under-represented.

As there were no comprehensive population statistics, the weighting was limited to variables which are included on or can be appended to the sample frame.

The non-response weight was calculated using the data available on the heat network sample frame. This included:

- Number of dwellings in the network
- Region
- Index of Multiple Deprivation (IMD)
- Output Area Classifications (OAC)
- Urban/Rural classification
- Area level information matched from VOA/SSA:
- The percentage of dwellings in the area which are flats

[13] The sample design disproportionately sampled addresses to achieve a c.1,000 interviews in Scotland, c.1,000 interviews with consumers in Heat Trust registered heat networks and to achieve roughly even number of addresses by scheme type (whether operated by a Local Authority, Private operator or social operator).

- The age of dwellings in the area[14]
- The percentage of dwellings in the area in each council tax band[15]

Separate models for England and Wales, and Scotland were used to generate this weight. This was to account for differences in available area level statistics.

The weight for the sample identified heat network households was a product of the non-response weight and the design weight.

The table below shows the population profile, the unweighted profile and the weighted profile the sample identified heat network households.

Table 7: Domestic consumers weighted profile

	Population profile %	**Unweighted profile %**	**Sample identified heat network households weighted profile %**
Scheme type			
Local Authority Scheme type	7%	21%	7%
Private Scheme type	15%	34%	15%
Social Scheme type	16%	24%	16%
HMBR registered	38%	79%	38%
Not HMBR registered	62%	21%	62%
Region/Country			
North East	4%	3%	3%
North West	8%	5%	8%

[14] For England and Wales only, as equivalent data for Scotland at an appropriate granularity was not available.
[15] VOA statistics were sourced for England and Wales LSOA's, Scottish Assessors statistics were sourced for Scottish DataZones

Yorkshire and the Humber	5%	5%	5%
East Midlands	4%	2%	4%
West Midlands	5%	3%	4%
East	5%	4%	5%
London	48%	43%	50%
South East	9%	8%	10%
South West	6%	5%	6%
Wales	2%	1%	2%
Scotland	4%	21%	4%
England / Wales	96%	79%	96%
Heat Trust Registered			
Heat Trust	3%	19%	2%
Not Heat Trust	97%	81%	98%
Mean Absolute Error From population profile	-	18%	0%

Due to the varying heat network classification rate across the sample sources, all households identified in the sample as being on a heat network were included in the weighting process up until this point, even if they were identified in the respondent data as not being on a heat network.

Ineligible households were then assigned to the comparison sample, due to uncertainty of their network status.

The table below shows the population profile, the unweighted profile and the weighted profile after excluding the ineligible heat network sampled households.

Table 8: Domestic consumers weighted profile – excluding ineligible households

	Population profile %	Unweighted profile %	Heat Network sample weighted profile % (excl. ineligible households)
Scheme type			
Local Authority Scheme type	7%	21%	5%
Private Scheme type	15%	34%	11%
Social Scheme type	16%	24%	12%
HMBR registered	38%	79%	28%
Not HMBR registered	62%	21%	72%
Region/Country			
North East	4%	3%	2%
North West	8%	5%	7%
Yorkshire and the Humber	5%	5%	5%
East Midlands	4%	2%	4%
West Midlands	5%	3%	4%
East	5%	4%	5%
London	48%	43%	52%
South East	9%	8%	9%
South West	6%	5%	6%
Wales	2%	1%	2%

Scotland	4%	21%	4%
England / Wales	96%	79%	96%
Heat Trust Registered			
Heat Trust	3%	19%	2%
Not Heat Trust	97%	81%	98%
Mean Absolute Error From population profile	-	18%	2%

A weight was produced for the 32 comparison sampled households which were identified as being on a heat network. To do this, Kantar Public ran a regression model to predict the value of the heat network weight based on questionnaire and sample variables. These variables included:

- Type of property (TypeA)
- Housing tenure (TENUREA)
- When the property was built (WHENBUILT)
- Index of Multiple Deprivation (IMD)
- Output Area Classifications (OAC)
- Urban/Rural classification

The predicted values from this model were used as the heat network weight for the 32 comparison sample households that were identified as being on a heat network.

The table below shows the population margins and the final heat network weighted households.

Table 9: Domestic consumers weighted profile – comparison sample profile targets

	Population profile %	Heat Network weighted profile
Scheme type		
Local Authority Scheme type	7%	5%
Private Scheme type	15%	10%
Social Scheme type	16%	12%
HMBR registered	38%	27%
Not HMBR registered	62%	73%
Region/Country		
North East	4%	2%
North West	8%	8%
Yorkshire and the Humber	5%	5%
East Midlands	4%	4%
West Midlands	5%	4%
East	5%	5%
London	48%	52%
South East	9%	9%
South West	6%	6%
Wales	2%	2%
Scotland	4%	4%
England / Wales (total)	96%	96%

Heat Trust Registered		
Heat Trust	3%	2%
Not Heat Trust	97%	98%
Mean Absolute Error from population profile	-	3%

The mean overall design effect of the heat network households at the overall level has been estimated as: 2.84[16].

Table 10: Domestic consumers weighted profile – heat networks effective sample size

	Design effect	Unweighted Sample size	Effective sample size
Heat network households	2.84	2,242	790

Weighting the comparison sample

For the non-heat network comparison sample, Kantar Public constructed weights using Propensity Score Matching (PSM). This ensured the profile of the comparison sample closely matched that of the heat network consumer sample across a range of characteristics.

Kantar Public ran a logistic regression model to predict the likelihood of a case being on a heat network. A range of predictor variables were included in the model:

- Region (Sample frame)
- Urban/Rural classification (Sample frame)
- IMD (Sample frame)

[16] Where the design effect from weighting = 1 + cov(W)^2) – where cov(W) is the coefficient of variation of the weights.

- Output Area Classifications
- Area level information matched from VOA/SSA:
- The percentage of dwellings in the area in each council tax band[17]
- Type of property (TypeA)
- Housing tenure (TENUREA)
- When the property was built (WHENBUILT)

This model calculated for all households, a predicted probability (or 'propensity score') of being on a heat network, based on the model above.

Using these predicted probabilities, we ran a Propensity Score Matching (PSM) algorithm which assigned a weight to each case in the comparison sample. The effect of this weight was to ensure that the profile for the comparison sample closely matched that of the weighted heat network sample for the predictor variables included in the logistic regression model. The purpose of this was to more confidently attribute differences in survey responses to actual differences in experiences between heat network households and households not on heat networks, rather than differences in characteristics or demographic profile.

The table below shows the weighted heat network margins, and the unweighted and weighted comparison sample margins.

Table 11: Domestic consumers weighted profile – comparison sample profile

	Heat network weighted profile **%**	**Unweighted Comparison profile** **%**	**Weighted Comparison profile** **%**
Region			
North East	2%	5%	4%
North West	8%	7%	8%
Yorkshire and the Humber	5%	6%	6%
East Midlands	4%	4%	4%
West Midlands	4%	4%	4%

[17] VOA statistics were sourced for England and Wales LSOA's, Scottish Assessors statistics were sourced for Scottish DataZones

East of England	5%	5%	5%
London	52%	37%	48%
South East	9%	10%	8%
South West	6%	7%	6%
Scotland	4%	13%	5%
Wales	2%	2%	2%
IMD Quintile			
1	32%	31%	32%
2	33%	28%	31%
3	18%	20%	21%
4	12%	14%	11%
5	6%	7%	4%
Urban/rural classification			
Urban	91%	81%	91%
Rural	9%	19%	9%
Property Type			
Flat or Maisonette	91%	57%	90%
Terrace house/bungalow	3%	15%	3%
Semi-detached house/bungalow	2%	18%	2%
Detached house/bungalow	2%	8%	2%
Other/refused	3%	2%	3%
Housing Tenure			

Own outright	13%	28%	14%
Buying with mortgage/loan	14%	18%	11%
Shared ownership	5%	2%	2%
Rent from council/local authority	22%	18%	22%
Rent from HA	22%	14%	18%
Rent from private landlord	22%	19%	32%
Rent Other/free	1%	1%	1%
Prefer not to say/refused	1%	1%	1%
Age of property			
Before 1960	12%	32%	17%
1960-1999	25%	32%	24%
2000-2009	6%	11%	23%
2010 or more recently	39%	9%	18%
Don't know	16%	15%	18%
Prefer not to say/refused	2%	1%	1%
Mean absolute error From HN weighted profile	-	7%	1%

The mean overall design effect of the combined weight at the overall level has been estimated as: 2.52[18].

[18] Where the design effect from weighting = 1 + cov(W)^2) – where cov(W) is the coefficient of variation of the weights.

Table 12: Domestic consumers weighted profile – overall effective sample size

	Design effect	Unweighted Sample size	Effective sample size
Overall	2.52	3,975	1,578
Heat network households	2.84	2,242	790
Comparison households	2.11	1,733	823

Weighting the operators survey

Kantar Public calculated a non-response weight using the data available on the sample frame, this included:

- Strata (Scheme Type, Country, Heat Trust registration)
- Region
- Number of networks supplied in OPSS register
- Number of customers supplied
- Number of buildings supplied
- Whether heating, cooling or hot water supplier
- Energy source Electricity/Natural Gas/Other
- Technology type Boiler/Chiller/CHP

As all of the usable operator sample derived from the HMBR database was selected and issued for fieldwork the sampling fractions did not vary and no explicit design weight was produced. However, as the tele-matching, and response rates varied between strata, a strata variable was included in the non-response model to control for this variation.

The table below shows the HMBR operator population margins and the unweighted and weighted sample margins.

Table 13: Operators weighted sample profile

Variable	HMBR operator	Unweighted profile	Weighted profile

	Population profile %	%	%
Strata			
HMBR registered, LA scheme type, in Scotland, not registered with the Heat Trust	1%	2%	1%
HMBR registered, LA scheme type, not in Scotland, not registered with the Heat Trust	5%	12%	5%
HMBR registered, Private scheme type, in Scotland, not registered with the Heat Trust	2%	5%	2%
HMBR registered, Private scheme type, not in Scotland, not registered with the Heat Trust	18%	32%	18%
HMBR registered, Social scheme type, in Scotland, not registered with the Heat Trust	5%	4%	5%
HMBR registered, Social scheme type, not in Scotland, not registered with the Heat Trust	69%	45%	69%
Scheme type			
LA scheme type	5%	14%	5%
Social scheme type	20%	37%	20%
Private scheme type	75%	49%	75%
Country			
England	89%	85%	91%
Scotland	10%	11%	8%

Wales	1%	4%	1%
Number of networks supplied			
1	69%	46%	69%
2-4	16%	24%	16%
5-9	6%	13%	6%
10+	10%	17%	10%
Mean absolute error from population profile	-	10%	0%

The mean overall design effect for analysis of HMBR heat network Operators has been estimated as: 1.75[19].

The table below shows the design effects and effective sample size.

Table 14: Operators effective sample size

	Design effect	Unweighted sample size	Effective sample size
Overall	1.75	130	74

Statistical significance

Results from all surveys are attempts to estimate “true values” in a wider population; and therefore all survey statistics come with an associated margin of error within which the “true” population measure is expected to lie. As such, all differences quoted in the main survey report have been tested to for statistical significance; that is, the difference between two compared values are significant even after we have accounted for the margins of error.

[19] Where the design effect from weighting = 1 + cov(W)^2) – where cov(W) is the coefficient of variation of the weights.

Unless otherwise specified, all commentary in the main report focuses exclusively on differences that are statistically significant at a 95% confidence level[20]. In basic terms, this means that if the survey was conducted 100 times, a finding of the same nature would be found in at least 95 cases. In a few exceptional circumstances, survey findings which were not statistically significant have been presented where their inclusion was important for context or was consistent with a wider trend. In every case any non-statistically significant results are clearly stated as such.

Applying weights to data as described above, while tending to make the quoted figures more representative of the population of interest, has the effect of reducing the effective sample size of the data. As such the effective base size, which is used in any statistical testing, is smaller than the unweighted base size and this has the effect of increasing the confidence intervals around the survey estimates. This effect has been taken into account when determining whether or not differences between survey estimates described throughout the report are statistically significant. Therefore, while the base sizes reported throughout this report are the actual base sizes, the statistical analysis is based on the effective base.

Regression modelling

Logistic regression modelling was used to analyse the key drivers of overall satisfaction, satisfaction with the level of control and the perceived fairness of cost. This approach estimates the influence of a single factor on the outcome variables while keeping all other variables in the model fixed.

The core purposes of the models were to:

- Investigate whether or not the selected influencers were significantly associated with the outcome variables (after controlling for other variables), and
- Understand the direction and magnitude of these significant associations.

The table below includes details of all variables included in the models, how they were coded and how we dealt with missing data.

Where there were low levels (<10%) of missing values, Kantar Public combined the missing values with another category. Missing values were combined with the category that was most similar to not having an opinion. While the approach of combining missing values with another category in a statistical model carries the risk of introducing bias, loss of information, and increased variance, these risks are relatively low in this case as the proportion of missing values is low.

[20] Differences across sub-groups were tested using chi square tests.

Where there were moderate levels (10% - <15%) of missing data, Kantar Public imputed the missing values using other survey variables that were likely to be related. For example, Kantar Public used regression modelling to estimate the number of bedrooms (where missing) based on other characteristics, such as the number of people in the household, the property type and the age of the property.

The variable for the type of heat network had missing values for 25% of cases. This was because data on the type of heat network was not available for cases identified from the EPC or sampled as part of the comparison sample but later identified as being on a heat network. As this is a distinct category and could have missing data imputed with an acceptable level of certainty these cases were included in a distinct category.

Some variables could not be included in the models due to very high levels of missing data (<30%)[21].

A summary of variables and recoding of missing data is in the table below.

Table 15: Summary of variables in regression modelling

Variable/Variable name	Categories *reference category	How we dealt with missing data	Number of imputed cases22
Outcome variable			
Overall satisfaction (Q16_1 - SATISFACTION_1)	1 - Very satisfied/Satisfied 0 - Neither/dissatisfied/very dissatisfied*	DK/blank combined with 'Neither/dissatisfied/very dissatisfied'	N/A

[21] We considered including the variable PayAmount (Thinking about the last payment you made for heating and hot water, what was the total amount paid?) as an influencing variable, however 52% of cases have missing values at this variable. With this level of missing data, an imputation model was unlikely to work very well. We could not including the missing data as a distinct category either, as the completed data may bias the coefficients. For example, if we included categories, the £1-£100 group would only represent those that provided a valid answer to PayAmount, and not those that did not know their bill amount or declined to answer.

[22] Includes just those variables with missing data

Satisfaction with the level of control (Q16_2 - SATISFACTION_2)	1 - Very satisfied/Satisfied 0 - Neither/dissatisfied/very dissatisfied*	DK/blank combined with 'Neither/dissatisfied/very dissatisfied'	N/A
Fairness of cost (Q65 - FAIR)	1 – Very fair/fair 0 – Not very/Not at all* fair	DK/blank combined with 'very fair/fair'	N/A
Influencer variables			
Underheating (Q28 – COLD)	0 – Not experienced underheating 1 – Experienced any underheating*	DK/blank combined with 'Not experiencing overheating'	N/A
Overheating (Q31 – WARM)	0 – Not experienced overheating 1 – Experienced any overheating*	DK/blank combined with 'Not experiencing overheating'	N/A
Satisfaction with amount of information on the bill (Q48 – INFOSAT)	1 – Very/fairly satisfied 2 – Not very/fairly satisfied*	DK/blank combined with 'Not very/fairly satisfied'	N/A
Complaints (Q35 – COMPLAINT)	1 – Very/fairly satisfied 2 – Not very/fairly satisfied	DK/blank combined with 'no reason to complain'	N/A

	3 – No reason to complain*		
Fairness of cost (Q65)	1 – Very fair 2 – Fair 3 – Not at all fair 4 – Not very*	DK/blank combined with 'very fair/fair'	N/A
Financially struggling (Q63-STRUGGLE)	1 –Agree/Strongly agree 2 – Neither 3 –Strongly disagree/disagree *	DK/blank combined with 'disagree/strongly disagree'	N/A
Control variables			
Heat network operator (sample data/ Q46 - BILLPROV[23])	1 – Private 2 – Local Authority 3 – Housing Association*	DK/blank imputed by modelling	290
Type of heat network (Sample variable)	1 – District 2 – Missing/EPC case 3 –Communal *	Distinct category	N/A
Heat Trust property (sample data)	1 – Heat Trust registered 2 – Not Heat Trust registered*	Blank combined with 'Not Heat Trust registered'	N/A

[23] For EPC/comparison sample cases, Heat network operator has been derived from Q46 - BILLPROV

Property age (Q7 - WHENBUILT)	1 – pre-1960 2 – 1960-1999 3 – 2000-2009 4 – Post 2010*	DK/blank imputed by modelling	365
Property type (Q2 - TYPEA)	1 – Flat or maisonette 2 – All other types*	DK/blank combined with 'all other types'	N/A
Vulnerable people in household (Q85 - ADDNEEDS, Q86 - INJURY, Q87 - SUPPORT)	1 – Vulnerable person present 2 – no vulnerable person*	DK/blank combined with 'no vulnerable person'	N/A
Older people in household (Q13 - AGE)	1 – No person 65 or older in household 2 – Person 65 or older in household*	DK/blank combined with 'no person 65 or older'	N/A
Children in household (Q13 - AGE)	1 – Children (<18) in household 2 – No children (<18) in household*	DK/blank combined with 'no child (<18)'	N/A
Number of people in household (Q11 - HOUSEHOLD)	1 – 1-person household* 2 – 2-person household 3 – 3 or more in household	DK/blank imputed by modelling	16

Receiving separate bill (Q56 - SPLITBILL)	1 – Separate bill 2 – Not separate bill*	DK/blank combined with 'no/other'	N/A
Level of information on the bill (Q51 - BILLINFO)	1 – Too little information 0 - Too much / About right *	DK/blank combined with 'too much/about right'	N/A

Testing for multicollinearity

Prior to finalising the model specification, Kantar Public tested for multicollinearity. Multicollinearity is when two or more influencer variables in a model are highly correlated with each other. Including all of these variables in a model can mean that significant associations are not identified or that coefficients can be in the wrong direction or of an implausible magnitude.

Kantar Public calculated Variance Inflation Factors (VIF) to identify whether multicollinearity could be an issue. This identified three variables with a VIF of over 2[24], this indicated that they were moderately correlated with other variables in the model:

- Satisfaction with the amount of information on the bill (VIF=3.6)
- Satisfaction with the clarity of information on the bill (VIF=3.6)
- Property size (VIF=2.2).

Kantar Public excluded satisfaction with the clarity of the information on the bill from the models. This reduced the VIF for satisfaction with amount of information on the bill to below 2, so this variable was kept in the models. Kantar Public also excluded property size from the model as it was fairly strongly correlated with household size and slightly correlated with whether there were children present in the household.

[24] The variance for these coefficients is at least twice as large as would be expected if there was no multicollinearity

Regression model outputs

Overall satisfaction

The table below shows the output from the logistic regression model for overall satisfaction, including the confidence intervals associated with the odds ratios.

Table 16: Overall satisfaction model outputs

	Beta	Std. err.	P value	Odds ratio	Confidence interval	
					Lower limit	Upper limit
Influencer Variables						
Satisfaction with level of control						
'Very Satisfied' (n=517) vs. 'Not satisfied/Neither/no answer' (n=833)	3.113	0.361	0.000	22.488	11.078	45.650
'Fairly Satisfied' (n=894) vs. 'Not satisfied/Neither/no answer'(n=833)	2.254	0.247	0.000	9.523	5.869	15.451
Perceives cost as Fair						
'Very fair' (n=233) vs 'not at all fair/no answer' (n=561)	0.890	0.499	0.074	2.436	0.916	6.482
'Fair' (n=915) vs 'not at all fair/no answer' (n=561)	0.733	0.259	0.005	2.082	1.252	3.462

'Not very fair' (n=545) vs 'not at all fair/no answer' (n=561)	0.393	0.266	0.141	1.481	0.878	2.497
Satisfaction with the amount of information received						
Did not receive (n=821) vs. 'Neither/dissatisfied/very dissatisfied/no answer' (n=570)	1.009	0.320	0.002	2.743	1.463	5.142
'Very satisfied'/'satisfied' (n=853) vs. 'Neither/dissatisfied/very dissatisfied/no answer' (n=570)	1.151	0.322	0.000	3.160	1.681	5.940
Not experienced under-heating (n=1493) vs. no under-heating/no answer(n=819)	0.930	0.222	0.000	2.534	1.641	3.915
Satisfaction with handling of complaint						
'Very satisfied/satisfied' (n=233) vs Did not complain (n=1683)	0.474	0.374	0.205	1.606	0.772	3.343
'Neither/dissatisfied/very dissatisfied' (n=328) vs Did not complain (n=1683)	-1.002	0.281	0.000	0.367	0.212	0.637
Not experienced overheating (n=1425) vs. over-heating/no answer (n=819)	-0.282	0.210	0.179	0.754	0.500	1.138

Heat network operator						
Private (n=1164) vs. Housing association (n=598)	0.084	0.261	0.747	1.088	0.652	1.814
Local authority (n=482) vs. Housing association (n=598)	0.321	0.331	0.331	1.379	0.721	2.637
Control Variables						
Heat Network Type						
District (n=439) vs Communal (n=1239)	-0.233	0.319	0.466	0.792	0.424	1.482
Unknown (n=566) vs Communal (n=1239)	0.260	0.240	0.279	1.297	0.810	2.078
Receiving separate bill for heating and hot water (n=922) vs not Receiving separate bill (n=1322)	0.256	0.222	0.248	1.292	0.836	1.995
Vulnerable people in household (n=817) vs no vulnerable people in household (n=1427)	0.021	0.245	0.930	1.022	0.632	1.652
Age of property						
Pre-1960 (n=309) vs. 2010 or later (n=863)	0.089	0.325	0.785	1.093	0.578	2.067
1960-1999 (n=875) vs. 2010 or later (n=863)	-0.010	0.266	0.969	0.990	0.587	1.667
2000-2009 (n=197) vs. 2010 or later (n=863)	0.203	0.471	0.667	1.225	0.486	3.084
Not financially struggling						
'Strongly agree/Agree' (n=1016) vs. 'Strongly	-0.501	0.272	0.066	0.606	0.355	1.034

disagree/disagree/DK/no answer' (n=525)						
'Neither' (n=703) vs. 'Strongly disagree/disagree/DK/no answer' (n=525)	0.082	0.293	0.780	1.085	0.611	1.926
Not registered with Heat Trust (n=1807) vs Heat Trust registered (n=437)	0.147	0.217	0.500	1.158	0.756	1.773
No children in household (n=1961) vs children in household (n=283)	0.480	0.395	0.224	1.616	0.745	3.505
No people aged 65 or above in household (n=1458) vs people aged 65+ in household (n=786)	-0.439	0.234	0.061	0.645	0.407	1.020
Household size						
1 person (n=1253) vs. 3 or more (n=368)	-0.171	0.370	0.645	0.843	0.408	1.742
2 people (n=623) vs. 3 or more (n=368)	0.338	0.398	0.396	1.402	0.642	3.063
Property type: flat (n=1979) vs any other (n=265)	-0.217	0.374	0.561	0.805	0.386	1.676
Information on bill 'Too little' (n=357) vs 'About right/Too much' (n=1887)	0.563	0.368	0.126	1.756	0.853	3.614
Intercept	-1.893	0.734	0.010	0.151	0.036	0.636

The nagelkerke pseudo-R^2 for this model was 0.493.[25]
Base HN consumers (2,444)

Satisfaction with overall control

The table below shows the output from the logistic regression model for satisfaction with the level of control, including the confidence intervals associated with the odds ratios.

Table 17: Satisfaction with level of control model outputs

	Beta	Std. err.	P value	Odds ratio	Confidence interval	
					Lower limit	Upper limit
Influencer Variables						
Perceives cost as Fair						
'Very fair' (n=233) vs 'not at all fair/no answer' (n=561)	1.052	0.324	0.001	2.863	1.518	5.400
'Fair' (n=915) vs 'not at all fair/no answer' (n=561)	0.897	0.215	0.000	2.452	1.608	3.738
'Not very fair' (n=545) vs 'not at all fair/no answer' (n=561)	0.543	0.222	0.014	1.721	1.114	2.660
Satisfaction with the amount of information received						
Did not receive (n=821) vs. 'Neither/dissatisfied/very dissatisfied/no answer' (n=570)	-0.273	0.232	0.239	0.761	0.483	1.199

[25] Using the Nagelkerke approach, the analysis calculated pseudo R-squared metrics for the models. Effectively, this provides a quantification of the outcome variability that is explained by the models. However, we note that the usefulness of pseudo R-squared metrics is open to debate amongst data users, with concerns being raised regarding the extent to which these are intuitively interpretable in relation to non-linear outcomes (such as the binary outcomes modelled in this study). For a brief review of pseudo R-squared metrics, see: Tabachnick, B. G; & Fidell, L. S. (2007). Using Multivariate Statistics. Boston: Pearson / Allyn & Bacon.

'Very satisfied'/'satisfied' (n=853) vs. 'Neither/dissatisfied/very dissatisfied/no answer' (n=570)	0.312	0.222	0.160	1.367	0.884	2.113
Not experienced under-heating (n=1493) vs. no under-heating/no answer(n=819)	0.029	0.183	0.876	1.029	0.719	1.472
Satisfaction with handling of complaint						
'Very satisfied/satisfied' (n=233) vs Did not complain (n=1683)	-0.153	0.257	0.551	0.858	0.518	1.420
'Neither/dissatisfied/very dissatisfied' (n=328) vs Did not complain (n=1683)	-1.175	0.250	0.000	0.309	0.189	0.504
Not experienced overheating (n=1425) vs. over-heating/no answer (n=819)	0.240	0.169	0.156	1.272	0.912	1.773
Heat network operator						
Private (n=1164) vs. Housing association (n=598)	-0.076	0.201	0.706	0.927	0.624	1.376
Local authority (n=482) vs. Housing association (n=598)	0.088	0.250	0.725	1.092	0.669	1.784
Control Variables						
Heat Network Type						
District (n=439) vs Communal (n=1239)	0.239	0.227	0.292	1.270	0.814	1.980

Unknown (n=566) vs Communal (n=1239)	-0.185	0.191	0.333	0.831	0.572	1.208
Receiving separate bill for heating and hot water (n=922) vs not Receiving separate bill (n=1322)	0.083	0.179	0.643	1.087	0.764	1.545
Vulnerable people in household (n=817) vs no vulnerable people in household (n=1427)	-0.208	0.174	0.232	0.812	0.577	1.143
Age of property						
Pre-1960 (n=309) vs. 2010 or later (n=863)	-0.756	0.248	0.002	0.470	0.289	0.764
1960-1999 (n=875) vs. 2010 or later (n=863)	-0.676	0.224	0.003	0.508	0.328	0.789
2000-2009 (n=197) vs. 2010 or later (n=863)	-0.603	0.330	0.068	0.547	0.286	1.045
Not financially struggling						
'Strongly agree/Agree' (n=1016) vs. 'Strongly disagree/disagree/DK/no answer' (n=525)	0.083	0.219	0.704	1.087	0.707	1.670
'Neither' (n=703) vs. 'Strongly disagree/disagree/DK/no answer' (n=525)	-0.250	0.217	0.250	0.779	0.509	1.193
Not registered with Heat Trust (n=1807) vs Heat Trust registered (n=437)	0.037	0.171	0.831	1.037	0.741	1.451
No children in household (n=1961) vs children in household (n=283)	0.193	0.375	0.607	1.213	0.581	2.534

No people aged 65 or above in household (n=1458) vs people aged 65+ in household (n=786)	0.230	0.184	0.212	1.259	0.877	1.806
Household size						
1 person (n=1253) vs. 3 or more (n=368)	-0.088	0.343	0.798	0.916	0.467	1.795
2 people (n=623) vs. 3 or more (n=368)	0.058	0.345	0.866	1.060	0.539	2.085
Property type: flat (n=1979) vs any other (n=265)	-0.107	0.248	0.665	0.898	0.553	1.460
Information on bill 'Too little' (n=357) vs 'About right/Too much' (n=1887)	-0.200	0.242	0.409	0.819	0.510	1.316
Intercept	0.322	0.555	0.562	1.380	0.465	4.094

The nagelkerke pseudo-R^2 for this model was 0.183.[26]

Perceived fairness of cost

The table below shows the output from the logistic regression model for Perceived fairness of cost, including the confidence intervals associated with the odds ratios.

Table 18: Perceived fairness of cost model outputs

	Beta	Std. err.	P value	Odds ratio	Confidence interval	
					Lower limit	Upper limit

[26] Using the Nagelkerke approach, the analysis calculated pseudo R-squared metrics for the models. Effectively, this provides a quantification of the outcome variability that is explained by the models. However, we note that the usefulness of pseudo R-squared metrics is open to debate amongst data users, with concerns being raised regarding the extent to which these are intuitively interpretable in relation to non-linear outcomes (such as the binary outcomes modelled in this study). For a brief review of pseudo R-squared metrics, see: Tabachnick, B. G; & Fidell, L. S. (2007). Using Multivariate Statistics. Boston: Pearson / Allyn & Bacon.

Influencer variables						
Satisfaction with the amount of information received						
Did not receive (n=821) vs. 'Neither/dissatisfied/very dissatisfied/no answer' (n=570)	0.728	0.248	0.003	2.072	1.273	3.373
'Very satisfied'/'satisfied' (n=853) vs. 'Neither/dissatisfied/very dissatisfied/no answer' (n=570)	1.297	0.230	0.000	3.658	2.328	5.748
Not experienced under-heating (n=1493) vs. no under-heating/no answer(n=819)	0.474	0.198	0.016	1.607	1.091	2.368
Satisfaction with handling of complaint						
'Very satisfied/satisfied' (n=233) vs Did not complain (n=1683)	-0.409	0.282	0.148	0.665	0.382	1.155
'Neither/dissatisfied/very dissatisfied' (n=328) vs Did not complain (n=1683)	-1.064	0.304	0.000	0.345	0.190	0.626
Not experienced overheating (n=1425) vs. over-heating/no answer (n=819)	0.293	0.176	0.096	1.340	0.949	1.892
Heat network operator						
Private (n=1164) vs. Housing association (n=598)	-0.440	0.214	0.040	0.644	0.423	0.981

Local authority (n=482) vs. Housing association (n=598)	-0.072	0.277	0.795	0.931	0.541	1.601
Control Variables						
Heat Network Type						
District (n=439) vs Communal (n=1239)	0.173	0.252	0.493	1.189	0.725	1.949
Unknown (n=566) vs Communal (n=1239)	-0.713	0.197	0.000	0.490	0.333	0.721
Receiving separate bill for heating and hot water (n=922) vs not Receiving separate bill (n=1322)	0.541	0.192	0.005	1.717	1.179	2.501
Vulnerable people in household (n=817) vs no vulnerable people in household (n=1427)	0.497	0.199	0.012	1.644	1.114	2.427
Age of property						
Pre-1960 (n=309) vs. 2010 or later (n=863)	-0.202	0.284	0.477	0.817	0.468	1.426
1960-1999 (n=875) vs. 2010 or later (n=863)	0.033	0.240	0.890	1.034	0.646	1.655
2000-2009 (n=197) vs. 2010 or later (n=863)	-0.438	0.314	0.163	0.646	0.349	1.194
Not financially struggling						
'Strongly agree/Agree' (n=1016) vs. 'Strongly disagree/disagree/DK/no answer' (n=525)	-1.435	0.229	0.000	0.238	0.152	0.373
'Neither' (n=703) vs. 'Strongly	-0.419	0.232	0.071	0.658	0.417	1.037

disagree/disagree/DK/no answer' (n=525)						
Not registered with Heat Trust (n=1807) vs Heat Trust registered (n=437)	0.426	0.186	0.022	1.532	1.064	2.206
No children in household (n=1961) vs children in household (n=283)	-0.243	0.358	0.497	0.785	0.389	1.582
No people aged 65 or above in household (n=1458) vs people aged 65+ in household (n=786)	-0.699	0.194	0.000	0.497	0.339	0.728
Household size						
1 person (n=1253) vs. 3 or more (n=368)	-0.047	0.335	0.888	0.954	0.494	1.840
2 people (n=623) vs. 3 or more (n=368)	-0.375	0.343	0.275	0.687	0.350	1.347
Property type: flat (n=1979) vs any other (n=265)	0.134	0.279	0.632	1.143	0.662	1.975
Information on bill 'Too little' (n=357) vs 'About right/Too much' (n=1887)	-0.700	0.260	0.007	0.497	0.298	0.827
Intercept	0.559	0.563	0.321	1.748	0.580	5.273

The nagelkerke pseudo-R^2 for this model was 0.396.[27]

[27] Using the Nagelkerke approach, the analysis calculated pseudo R-squared metrics for the models. Effectively, this provides a quantification of the outcome variability that is explained by the models. However, we note that the usefulness of pseudo R-squared metrics is open to debate amongst data users, with concerns being raised regarding the extent to which these are intuitively interpretable in relation to non-linear outcomes (such as the binary outcomes modelled in this study). For a brief review of pseudo R-squared metrics, see: Tabachnick, B. G; & Fidell, L. S. (2007). Using Multivariate Statistics. Boston: Pearson / Allyn & Bacon.

Additional regression models to guide our analytical approach

We fitted additional logistic regression models to determine which characteristics were most strongly associated with six key survey outcomes:

- Overall satisfaction
- Satisfaction with the level of control over heating
- Experience of over-heating
- Perceived fairness of cost
- Perceptions of the level of information on bills (Too little information)
- Whether consumers had made a complaint (or had reason to make a complaint)

The same predictor variables used in the key driver models were included for each additional model.

These models were not designed to provide substantive inferences regarding these key outcomes. Instead, the purpose of these models was to identify characteristics which were strongly associated with most of the key outcomes listed above. This analysis would then inform the choice of which sub-groups within the population of heat network households would be referenced in the descriptive report of findings.

Table 19: Additional regression model outputs

Sig at p<0.01 level	Sig at p<0.05 level	Overall satisfaction (satisfied)			Control satisfaction (satisfied)			Experienced Overheating			Perceives cost as very/not fair			Too little information on bill			No reason to complain		
Predictor		Odds Ratio	95% Confidence Interval Lower	Upper	Odds Ratio	95% Confidence Interval Lower	Upper	Odds Ratio	95% Confidence Interval Lower	Upper	Odds Ratio	95% Confidence Interval Lower	Upper	Odds Ratio	95% Confidence Interval Lower	Upper	Odds Ratio	95% Confidence Interval Lower	Upper
Financially struggling: Agree Vs Disagree/Neither/DK	.00 Struggling vs. 1.00 not struggling/neither	0.5	0.36	0.69	0.91	0.67	1.24	0.99	0.73	1.35	3.77	2.72	5.24	2.01	1.36	2.98	0.54	0.39	0.74
Not receiving a separate bill vs Receiving separate bill	1.00 Not separate vs. 2.00 Separate bill	1.39	0.97	1.97	1.03	0.75	1.41	1.05	0.76	1.45	0.59	0.42	0.83	0.61	0.4	0.94	1.3	0.93	1.83
District Vs Communal	1.00 District vs. 3.00 Communal	0.72	0.46	1.14	1.16	0.75	1.78	0.91	0.6	1.38	1.07	0.65	1.75	1.78	1.04	3.05	0.77	0.48	1.21
EPC/Missing Vs Communal	2.00 EPC/Missing vs. 3.00 Communal	1.07	0.72	1.57	0.87	0.61	1.23	0.74	0.52	1.05	1.77	1.25	2.52	0.92	0.58	1.44	1.66	1.14	2.41
Private vs housing association	1.00 Private vs. 3.00 Housing Association	0.96	0.62	1.47	0.95	0.66	1.37	1.02	0.71	1.47	1.33	0.9	1.98	1.16	0.69	1.95	1.26	0.86	1.84
Local authority vs housing association	2.00 Local Authority vs. 3.00 Housing Association	1.23	0.72	2.12	1.1	0.7	1.74	1.25	0.77	2.04	1.03	0.61	1.74	0.85	0.44	1.65	1	0.6	1.65
Not heat trust vs Heat Trust	1.00 Not Heat Trust registered vs. 2.00 Heat Trust registered	1.14	0.83	1.59	1	0.73	1.37	0.98	0.73	1.32	0.73	0.53	1.01	1.6	1.1	2.32	1.32	0.96	1.8
Pre-1960 vs. 2010 or later	1.00 pre-1960 vs. 4.00 2010 and later	0.59	0.35	1	0.42	0.26	0.68	1	0.61	1.63	1.5	0.91	2.46	1.6	0.85	2.99	0.69	0.41	1.16
1960-1999 vs. 2010 or later	2.00 1960-1999 vs. 4.00 2010 and later	0.62	0.39	0.98	0.48	0.32	0.71	0.74	0.48	1.12	1.01	0.64	1.6	0.9	0.51	1.61	0.84	0.54	1.3
2000-2009 vs. 2010 or later	3.00 2000-2009 vs. 4.00 2010 and later	0.69	0.36	1.32	0.48	0.27	0.86	1.04	0.61	1.76	1.67	0.91	3.06	1.73	0.87	3.42	1.11	0.61	2.02
No vulnerable person in hh vs	1.00 No vs. 2.00 Yes	1.19	0.84	1.68	1.25	0.91	1.74	0.81	0.58	1.13	1.47	1.03	2.1	0.98	0.64	1.5	1.93	1.35	2.76

vulnerable person in hh																			
No Older person in HH vs older person in HH	1.00 No person 65 or older vs. 2.00 Person 65 or older in household	0.65	0.44	0.97	1	0.71	1.41	2.38	1.65	3.44	2.39	1.65	3.45	1.42	0.86	2.35	0.55	0.37	0.8
No children in HH vs children in HH	1.00 No children vs. 2.00 Children present	1.34	0.7	2.56	1.04	0.54	2.03	0.94	0.51	1.72	1.18	0.6	2.32	1.24	0.6	2.57	0.77	0.38	1.55
1 person vs. 3 or more	1.00 1 vs. 3.00 3 or more	1.02	0.5	2.08	1.19	0.6	2.33	0.95	0.49	1.83	0.96	0.47	1.98	0.67	0.3	1.5	1.65	0.79	3.43
2 people vs. 3 or more	2.00 2 vs. 3.00 3 or more	1.41	0.72	2.76	1.37	0.72	2.63	1.23	0.65	2.32	1.3	0.65	2.61	0.74	0.36	1.55	1.66	0.83	3.33
0 or 1 bedroom vs. 3 or more	1.00 1 bedroom or studio vs. 3.00 3 bedrooms or more	1.47	0.81	2.67	1.06	0.62	1.83	0.89	0.5	1.57	0.85	0.47	1.53	0.95	0.46	1.97	1.52	0.86	2.7
2 bedrooms vs. 3 or more	2.00 2 vs. 3.00 3 bedrooms or more	1.62	0.91	2.9	1.04	0.62	1.74	0.99	0.58	1.68	0.87	0.51	1.5	0.84	0.44	1.61	1.07	0.63	1.81
Flat vs Other types	1.00 Flat or maisonette vs. 2.00 All other types	0.69	0.38	1.26	0.85	0.5	1.43	2.51	1.43	4.43	1.06	0.63	1.76	1.38	0.69	2.77	0.73	0.42	1.27

Interviews

Overview of interview methodology

To complement the survey findings, qualitative interviews were conducted between the 15th of November 2022 and 30th of January 2023 with 50 domestic consumers and 18 operators across England and Scotland. The interviews aimed to provide a more in-depth understanding of consumers and operators based on survey findings, and to address emerging policy topics not covered by the survey such as the cost-of-living crisis and decarbonised energy production.

Consumers

For consumers, the topic guide was built around the following topics (for more information see Annex E:

- Background and perception of heat networks, including low carbon energy production
- Billing & loss of heating
- Perceptions of affordability, including cost-of-living crisis
- Satisfaction with heat networks

Consumers were recruited from 765 heat network consumers who completed the survey and consented to be recontacted. Interviews were conducted 30 to 45 minutes in length via Zoom and telephone. Interviews were recorded in full using Zoom software. Respondents were sent a £30 gift voucher on completion of the interview.

Operators

For operators, the topic guide was built around the following topics (for more information see Annex F):

- Organisation background and heat network specifications
- Heat network operations such as maintenance responsibility or planned changes to primary energy sources
- Operator-consumer relationships including billing, consumer service and cost-of-living crisis
- Investment decision-making, including barriers to investment
- Regulatory frameworks & standards, including outages planning
- Step-in rights and risks of insolvency

Operators were recruited from the 33 surveyed who consented to be recontacted, the HMBR dataset as well as a DESNZ and Scottish Government operator outreach. Operator Interviews were conducted of up to 45 minutes in length via Zoom and telephone. Interviews were recorded in full using Zoom software. Operators were a difficult-to-reach audience, partly due to the presence of gatekeepers, major variations in profile and challenges identifying the right role to speak to. The insights are based on 18 interviews with operators across England and Scotland. The discussion guide was tested with first respondents which led to minor adjustments to the initial guide. Respondents were sent a £50 gift voucher on completion of the interview.

Qualitative analysis

A framework analysis approach was used to synthesise the responses and experiences of both audiences. This involved constructing a thematic framework against which qualitative data was synthesised and then mapped to identify features and patterns by different profiles. Microsoft Excel was used to write up each interview and for analysis. The consumer sample size (50 interviewed consumers with a target of 50) and spread of desired profiling criteria allowed for detailed case analysis through this criteria (including vulnerable/non-vulnerable, district/communal heat network, satisfied/dissatisfied etc.). - based on set criteria contained within the sample For operators, due to the smaller sample (18 interviewed operators with a target of 50) and this being a less familiar audience, the framework was built around the range of profiles identified through the interviews(including non-operators, public entities etc.

Qualitative sample frame

Consumers

Domestic Consumers on a Heat Network (target 50 interviews)*			
Respondent group		**Minimum target**	**Recruited (50)**
Region	Scotland	Monitor	15
	England	Monitor	35
Heat network type	Communal	20	23
	District	20	27
Satisfaction with heating and hot water system	Satisfied	30	34
	Dissatisfied	10	16
Vulnerable consumer	Vulnerable	15	15
	Not vulnerable	30	35
Tenure	Owns home	10	17
	Privately rents	10	14
	Socially rents	10	19
Heat network scheme type	Private	Monitor	25
	Local authority	Monitor	10
	Housing association	Monitor	15
Metering	Whether meter for individual property	Monitor	33
	Whether meter for entire building	Monitor	17
Age of scheme	Installed within the last year	Monitor	1
	Installed 1-5 years ago	Monitor	11

	Installed more than 5 years ago	Monitor	38
*Note. respondents to be classified within groups per responses to quantitative survey			

Operators

Heat network operator (target 50 interviews)			
Respondent group		**Minimum target**	**Recruited (18)**
Region	Scotland	Monitor	7
	England	Monitor	10
	Both	Monitor	1
Heat network type	Communal	30	3
	District	10	13
Number of heat networks responsible for	1-4	10	2
	Between 5 and 9	10	11
	10 or more	5	4
Whether Operator has had to change energy supplier	Yes	5	2
	No	30	10
Age of heat network	Already installed in property	5	6
	Installed in the last 5 years	5	2
	Installed 5 or more years ago	10	9

Annex A: Consumer questionnaire

See the Technical Report Annexes.

Annex B Operator questionnaire

See the Technical Report Annexes.

Annex C Consumer survey invitation letter

See the Technical Report Annexes.

Annex D Operator survey invitation letter

See the Technical Report Annexes.

Annex E Consumer interview topic guide

See the Technical Report Annexes.

Annex F Operator interview topic guide

See the Technical Report Annexes.

This publication is available from: www.gov.uk/government/publications/heat-network-consumer-and-operator-survey-2022

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