Smart Meter Data and Public Interest Issues – The National Perspective
Discussion Paper 1

Annex A – Existing data

Simon Elam
University College London
s.elam@ucl.ac.uk

Foreword

Much discussion about smart energy data is understandably focused on its likely use in consumer feedback, in enabling the future smart grid, and in the potential for commercial applications, but it could have wider uses. In summer 2015, TEDDINET[*], the Centre for Sustainable Energy and Sustainability First launched a joint 'research challenge' to understand how future household smart-meter energy data might be deployed to serve the public interest. We jointly commissioned two university researchers to consider these questions and for each to write a short discussion paper, one on the national perspective and the other on sub-national issues, for publication and debate.

The sponsors have provided assistance and guidance, and financial support. But the papers and the views expressed in them are those of the researchers concerned.

Following an invited workshop with interested parties, the papers will be published on the TEDDINET, CSE and Sustainability First websites.

S. Elam (2016) Smart Meter Data and Public Interest Issues – The National Perspective Discussion Paper 1.
Annex A – Existing data

Annex A - Existing Data

Summary

This appendix provides additional technical detail on datasets discussed in section 3 of Discussion Paper 1. Section 3 provides analysis of the strengths and weaknesses of existing energy datasets alongside commentary on the potential benefits of smart meter data.

Over 100 datasets were identified as being potentially relevant to this report. These datasets either contain energy consumption data on domestic buildings in the UK or contain data that that is related to energy consumption in domestic buildings. Of these, the most relevant have been selected for a more detailed review in this appendix. An implicit consideration for selecting the most relevant datasets is how closely it aligns with smart meter data that will become available once the mass rollout of smart meter installations is under way from 2016 to 2020. The key criteria for selecting datasets include:

- Relevancy - contain energy consumption data or data that is relevant to energy consumption in domestic buildings.
- Availability – data should be publically available or available to a wide range of organizations that might act in the public interest with limited restrictions on use.
- Granularity – micro-data (e.g. records based on individual dwellings) is favoured over data aggregated to large geographic areas.
- Coverage and number of records – large-scale, nationally representative datasets with a broad range of data variables will be prioritized over, for example, small trials focused on a specific subject
- Recency – More recent data is favoured over older data.
- Knowledge and judgement of the author.

The more detailed review of datasets below will inevitably omit datasets considered to be important or vital by some. As a rule of thumb, datasets included in this technical annexe contain records based on individual domestic buildings and thus aggregated statistics have not been included, although exceptions have been made for the Digest of UK Energy Statistics (DUKES) and Energy Consumption in the UK (ECUK) as these datasets provide key baseline statistics widely used by the energy sector. Additionally the Department of Energy and Climate Change's (DECC) sub-national statistics have been included on the basis of providing energy consumption data at relatively high resolution e.g. Lower Level Super Output Area (or postcode for 2013 data).

Equally there will be many micro-datasets that will have valuable public interest use for organizations working in specific areas that have been excluded from this review e.g. field trials on heat pumps. Thus inclusion or exclusion of a dataset from the list of datasets reviewed in more detail below should not be seen as a reflection of its value as a data resource for public interest or other uses.

Summary of strengths and weaknesses of existing data

Dataset	Key Positives	Key Negatives
Sub-national energy statistics	• Near complete coverage of electricity and gas meters in GB	• Aggregated statistics • Very limited set of contextual variables • Annual energy data only
NEED	• Near complete coverage of electricity and gas meters in GB • Dwelling-level data • A moderate set of contextual variables (e.g. building characteristics, socio-demographics).	• Annual energy data only • Access to the full dataset is highly restricted beyond DECC. • Accessible data from 4 million homes is de-identified thus further data linking is not possible, although analysis is possible. • Uncertainties limit its utility for making finely balanced decisions.
Low Carbon London	• Dwelling-level data • High temporal resolution energy data (e.g. half-hourly)	• Very limited set of contextual variables • Limited sample (5,567 households) • Accessible data is de-identified thus further data linking is not possible • One-time dataset
Energy Demand Research Project	• Dwelling-level data • High temporal resolution energy data (e.g. half-hourly)	• Very limited set of project-specific contextual variables • Limited sample (~11,000 domestic participants) • Accessible data is de-identified thus further data linking is not possible • One-time dataset
Customer Led Network Revolution	• Dwelling-level data • High temporal resolution energy data (e.g. half-hourly)	• Limited set of contextual variables • Limited sample (~14,621 households) • Accessible data is de-identified thus further data linking is not possible • One-time dataset
Energy Performance Certificate (EPC) data	• Dwelling-level data • Standard methodology for collecting data (RdSAP)	• Energy consumption is modelled and annual. • Some concerns with quality of EPC data[1]

---

[1] DCLG. EPCs: notes and definitions - Detailed guidance - GOV.UK [Internet]. 2014 Available from: https://www.gov.uk/guidance/energy-performance-of-buildings-certificates-notes-and-definitions#data-quality

S. Elam (2016) Smart Meter Data and Public Interest Issues – The National Perspective Discussion Paper 1.
Annex A – Existing data

	• Large set of contextual variables (building characteristics) • Large dataset (> 9m buildings)	• Access to bulk data is restricted to "approved recipients" and the financial cost of purchasing data can be prohibitive.
Energy Follow Up Survey	• Dwelling-level data • Sub-dwelling energy consumption for some participants • High temporal resolution energy data (e.g. half-hourly) for some participants. • Large set of contextual variables across the socio-technical spectrum	• Small (but representative) sample (2,616 homes) • Accessible data is de-identified thus further data linking is not possible • Repeat but cross-sectional dataset
Home Electricity Use Study	• Dwelling-level data • Sub-dwelling energy consumption for some participants • High temporal resolution energy data (e.g. half-hourly). • Moderate set of contextual variables across the socio-technical spectrum	• Small sample (251 homes, owner-occupiers) • Accessible data is de-identified thus further data linking is not possible • One-time dataset
English Housing Survey (EHS) / Scottish Household Survey (SHS) / National Survey for Wales (NSW)	• Dwelling-level data • Nationally representative samples • Large set of contextual variables across the socio-technical spectrum	• Limited energy consumption data – annual (from bills) where available. • Accessible data is de-identified thus further data linking is not possible • Repeat but cross-sectional dataset

Existing UK domestic energy datasets relevant to this report

Dataset	GB sub-national statistics – Electricity and Gas
Description	DECC's sub-national estimates of annual electricity and gas based on meter point data from England, Wales and Scotland. Data is aggregated to Region, Local Authority, Middle Layer Super Output Area (MSOA), Lower Level Super Output Area (LSOA), and for 2013 to postcode. Economy7 consumption is included in the electricity data and estimates of off-grid households are included in the gas data.
Timespan	Since 2005. Most recent data available is 2014.
Population	Domestic and non-domestic buildings in the UK
Number of records/units	Aggregated statistics based on over 27 million domestic electricity meters and over 23 million domestic gas meters.
Data availability	Publically available via DECC's website: https://www.gov.uk/government/collections/sub-national-electricity-consumption-data https://www.gov.uk/government/collections/sub-national-gas-consumption-data
Energy consumption data	Total and average (mean and median) gas and electricity consumption
Other data	Number of meters

Database	National Energy Efficiency Data-framework (NEED)
Description	A national dataset of energy use and energy efficiency in domestic and non-domestic buildings in England, Wales and Scotland. The data framework matches (via an address-spine) gas and electricity consumption data with records of energy efficiency measures from HEED and other sources, building characteristics from VOA and EPC, and householder characteristics from Experian. Not all data sources (e.g. VOA and Experian) are accessible to non-government users. More information - https://www.gov.uk/government/collections/national-energy-efficiency-data-need-framework
Timespan	HEED - 1995-2012 Energy consumption – annual since 2004 EPC – since 2007
Population	Domestic and non-domestic buildings in the UK
Number of records/units.	Over 25 million homes and approximately 2 million non-domestic buildings.
Data availability	An anonymized public use dataset is available via DECC (50,000 homes) with an end-user license version (4 million homes) available via the UK Data Service.
Energy consumption data	Annualised gas and electricity consumption from 2004. Gas is weather corrected.
Other data	Retrofit energy efficiency measures from HEED e.g. loft insulation, cavity wall / solid wall insulation, boilers etc. Property characteristics e.g. dwelling age, type, floor area, no. of bedrooms from VOA or EPC.

Dataset	Low Carbon London project
Description	Electricity consumption readings at half-hourly intervals from 5,567 London households that took part in Low Carbon London project. One group of participants were put on dynamic Time of Use tariffs while the other group remained on a flat rate tariff.
Timespan	November 2011 to February 2014
Population	London households
Number of records/units	5,567 households. 167 million meter reads.
Data availability	Publically available via the London Data Store: http://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households
Energy consumption data	Half-hourly electricity consumption.
Other data	Acorn group (modelled socio-demographic variable). Tariff. Date and Time of meter read.

Dataset	Customer Led Network Revolution
Description	"The Customer-Led Network Revolution was one of the largest smart grid demonstration projects ever undertaken in the UK. This anonymised disaggregated dataset (i.e. data on a per customers basis) from CLNR trials with domestic and SME customers. This data comprises of electricity consumption and PV generation for individual customers, and allied information about the customer, the tariff and the technology being studied." (CLNR)
Timespan	2011-2014
Population	13,000 GB electricity customers
Number of records/units	Approximately 11,000 domestic and 2,000 non-domestic participants.
Data availability	Publically available via a Creative Commons license - http://www.networkrevolution.co.uk/resources/project-data/
Energy consumption data	Basic profiling of domestic smart meter customers; Basic profiling of small and medium sized enterprise (SME) customers; Enhanced profiling of domestic smart meter customers; Enhanced profiling of domestic customers with air source heat pumps; Enhanced profiling of domestic customers with solar photovoltaics (PV); Enhanced profiling of domestic customers with electric vehicles (EVs); Domestic smart meter customers on time of use tariffs; Domestic solar PV customers with automatic in-premises balancing for hot water charging; Domestic solar PV customers using in-home displays for manual in-premises balancing.
Other data	

Dataset	Energy Demand Research Project
Description	"The Energy Demand Research Project (EDRP) was a suite of large scale trials across Great Britain involving over 50,000 households. The aim was to understand how consumers react to improved information about their energy consumption over the long term." (Ofgem)
Timespan	2007-2010
Population	GB households
Number of records/units	Anonymised dataset - 14,621 households; 246 million gas meter readings; 413 million electricity meter readings
Data availability	An anonymised dataset is available via the UK Data Service.
Energy consumption data	Electricity and gas consumption.
Other data	

Database	Energy Performance Certificates (EPC)
Description	Energy Performance Certificates are required whenever a property is leased or sold in the UK. A survey is conducted by a qualified Domestic Energy Assessor using Reduced SAP (RdSAP) methodology
Timespan	August 2007-present
Population	UK buildings (domestic and non-domestic)
Number of records/units.	Over 9 million buildings
Data availability	• England/Wales/NI – The EPC Register is "owned" by DCLG and operated by Landmark. • Scotland - The EPC Register is "owned" by the Scottish Government and operated by the Energy Saving Trust (Scotland). Access to disaggregated data in England and Wales is possible but restricted to "Approved Recipients". Currently payment is required to obtain bulk EPC data.
Energy consumption data	Modelled annual energy consumption.
Other data	Energy Efficiency Rating, Environmental Impact Rating, Recommendations to improve energy efficiency. Building characteristics e.g. dwelling age, type, size, wall construction etc. Presence of energy efficiency measures e.g. loft insulation, cavity wall / solid wall insulation. Main heating system, main heating fuel, boiler type, secondary heating etc. Heating controls

Dataset	Energy Follow-up Survey (EFUS)
Description	"The main aim of the 2011 Energy Follow-Up Survey (EFUS) was to collect new data on domestic energy use, in order to update the current modelling assumptions about how energy is used in the home, and to inform energy efficiency policy. The 2011 EFUS consisted of a follow-up interview survey of a sub-set of households first visited as part of the 2010/2011 English Housing Survey (EHS)".
Timespan	Surveys conducted between December 2010 and April 2011
Population	England
Number of records	2,616 homes. 4 levels of survey: a) Interview Survey (2,616 homes): b) Temperature monitoring survey (a sub-sample of 823 homes): c) Meter readings (a sub-sample of 1,345 homes): d) Electricity profiling (a sub-sample of 79 homes): Finally, a small number of properties had electricity profiling equipment installed to examine patterns of lighting, appliance and electrical cooking use. This equipment collected data at rapid intervals on the profile of electricity use in the home.
Data availability	Initial reports prepared by BRE and published by DECC - https://www.gov.uk/government/publications/energy-follow-up-survey-efus-2011 De-identified data available via the UK Data Service.
Energy Consumption data	Meter readings from a sub-sample of 1,345 homes. Matched to meter readings in EHS. Also a sub-sample of 79 homes had electrical profiling equipment installed collecting high-frequency data on appliance and lighting usage.
Other data	Socio-demographics and building demographics from EHS. Temperature data from Temperature loggers in subsample of 823 homes. External temp from met office weather stations (MIDAS). Questions on usage patterns/behaviours for a wide variety of appliances and lighting. Heating patterns and thermal comfort.

Dataset	Household Electricity Use Study
Description	"An in-depth Household Electricity Use Study aimed to cover the electricity usage of a representative sample of English owner-occupier homes. This data offers an unparalleled source of very detailed electricity profiles" Reports and further information available via https://www.gov.uk/government/publications/early-findings-demand-side-management
Timespan	2010-11
Population	Households in England, mainly owner occupied homes
Number of records/units	251 households. Approximately 250 million records of monitored data.
Data availability	Anonymized micro-data available via UKDS - https://discover.ukdataservice.ac.uk/catalogue/?sn=7874
Energy consumption data	Detailed monitoring of electricity use split by appliance and many other variables
Other data	House type, size (m3), property age etc. A range of environmental and energy use attitude and behaviour data. Electrical v non-electrical heating/HW

Dataset	English Housing Survey (EHS)
Description	"The English Housing Survey (EHS) began in 2008 and is a continuous national survey that collects information about people's housing circumstances and the condition and energy efficiency of housing in England. The EHS brings together two previous survey series into a single fieldwork operation: the English House Condition Survey (EHCS) and the Survey of English Housing (SEH). The EHS covers all housing tenures and provides valuable information and evidence to inform the development and monitoring of government housing policies." (UKDS)
Timespan	Since 2008. Previously EHCS and SEH
Population	England
Number of records	Approximately 17,000 households annually.
Data availability	Available via the UK Data Service - https://discover.ukdataservice.ac.uk/series/?sn=200010
Energy consumption data	Electricity and gas bill data collected where permission obtained.
Other data	Socio-demographics, a comprehensive set of building demographics, including (but not limited to): type of dwelling; Detailed tenure; Work done to the home (rotating); Dwelling age, size (floor area), wall insulation, windows (including glazing type), no of storeys, construction type, no of bedrooms, improvement potential, plots and gardens, Main heating system, main heating fuel, secondary heating, water heating system, boiler type, etc.

Dataset	National Survey for Wales / Living in Wales
Description	The Living in Wales survey closed in 2008 and was replaced by the National Survey for Wales which ran from 2012 to 2015. "The Living in Wales (LIW) survey series, which ran from 2004-2008, was the main general source of statistical information about households and the condition of homes in Wales. The LIW survey had two separate but linked components: the Household Survey and the Property Survey. The Property Survey was conducted in 2004 and 2008 and comprised an internal and external assessment of the property which was completed by a qualified surveyor. "
Timespan	2004-2008 and 2012-2015
Population	Private households in Wales
Number of records/units	2008 – Household survey - 12,811 addresses 2008 – Physical survey - 2,741 properties
Data availability	Data available via UK Data Service
Energy consumption data	SAP modelled energy consumption
Other data	Property age, size, condition, insulation, construction type. Heating system and controllers. Household composition, ethnicity, religion, employment. Fuel poverty, SAP energy efficiency, Energy saving measures

Dataset	Scottish Household Survey / Scottish House Conditions Survey
Description	The Scottish House Conditions Survey ran from 1999 to 2011 and was then incorporated into the Scottish Household Survey in 2012 "The SHS is designed to provide reliable and up-to-date information on the composition, characteristics, attitudes and behaviour of private households and individuals, both nationally and at a sub-national level and to examine the physical condition of Scotland's homes. It covers a wide range of topics to allow links to be made between different policy areas." (UKDS)
Timespan	SHCS - 1999-2011, SHS since 2012
Population	Scottish dwellings
Number of records/units	2012 - 10,642 cases
Data availability	Data available via UK Data Archive – https://discover.ukdataservice.ac.uk/series/?sn=2000048
Energy consumption data	Heating costs
Other data	Basic description of the dwelling; the material used in construction; the state of repair of the internal, external and common parts (where relevant); insulation. Heating system. Household composition; tenure, previous moves and intention to move; employment status and income. Satisfaction with the home and the local environment;

Dataset	Digest of UK Energy Statistics (DUKES)
Description	A comprehensive set of aggregated energy statistics covering both energy production and consumption in the UK. Includes tables, charts and commentary on energy statistics. Trends over 20-30 years are provided for some statistics.
Timespan	Most statistics cover the previous 5 years while some statistics have data from the 1970s
Population	UK
Number of records/units	N/A
Data availability	Publically available via https://www.gov.uk/government/collections/digest-of-uk-energy-statistics-dukes
Energy consumption data	Energy consumption by sector; also covers non-gas and electricity data
Other data	

Dataset	Energy Consumption in the UK (ECUK)
Description	Aggregated statistics covering energy consumption in the UK which is split into 5 chapters. Each chapter has a series of data tables in an Excel file and an accompanying fact sheet: Chapter 1: Overall energy consumption in the UK Chapter 2: Transport sector energy consumption Chapter 3: Domestic sector energy consumption Chapter 4: Industrial sector energy consumption Chapter 5: Services sector energy consumption
Timespan	1970s to present
Population	UK
Number of records/units	N/A
Data availability	Publically available via https://www.gov.uk/government/collections/energy-consumption-in-the-uk
Energy consumption data	The domestic energy consumption chapter covers topics such as energy consumption by fuel type, average energy consumption, electricity use by appliance type etc.
Other data	

References and useful data resources

Several data repositories and registries have been utilised in writing this report. These, and others, should be investigated to get a more complete picture of existing energy data.

1. UK Data Archive - https://www.ukdataservice.ac.uk/
2. DECC Energy Statistics - https://www.gov.uk/government/organisations/department-of-energy-climate-change/about/statistics
3. Centre for Sustainable Energy Open Data - https://www.cse.org.uk/projects/view/1259#Energy_consumption_data__domestic
4. UKERC Energy Data Centre - http://ukedc.rl.ac.uk/
5. Centre for Energy Epidemiology Data Asset Register (CEEDAR) - http://www.energy-epidemiology.info/data
6. LUKES Feasibility Study - http://www.energy-epidemiology.info/research-projects/research-projects/lukes/lukes-overview