Computers, Environment and Urban Systems 63 (2017) 58–67

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Electricity consumption and household characteristics: Implications for census-taking in a smart metered future

Ben Anderson [a,*], Sharon Lin [a], Andy Newing [b], AbuBakr Bahaj [a], Patrick James [a]

[a] Sustainable Energy Research Group, Energy & Climate Change Division, University of Southampton, Southampton, UK
[b] School of Geography, University of Leeds, Leeds, United Kingdom

ARTICLE INFO

Article history:
Received 16 October 2015
Received in revised form 19 April 2016
Accepted 15 June 2016
Available online 1 July 2016

Keywords:
Census
Smart meter
Transactional data
Big data
Households

ABSTRACT

This paper assesses the feasibility of determining key household characteristics based on temporal load profiles of household electricity demand. It is known that household characteristics, behaviours and routines drive a number of features of household electricity loads in ways which are currently not fully understood. The roll out of domestic smart meters in the UK and elsewhere could enable better understanding through the collection of high temporal resolution electricity monitoring data at the household level. Such data affords tremendous potential to invert the established relationship between household characteristics and temporal load profiles. Rather than use household characteristics as a predictor of loads, observed electricity load profiles, or indicators based on them, could instead be used to impute household characteristics. These micro level imputed characteristics could then be aggregated at the small area level to produce 'census-like' small area indicators. This work briefly reviews the nature of current and future census taking in the UK before outlining the household characteristics that are to be found in the UK census and which are also known to influence electricity load profiles. It then presents descriptive analysis of a large scale smart meter-like dataset of half-hourly domestic electricity consumption before reviewing the correlation between household attributes and electricity load profiles. The paper then reports the results of multilevel model-based analysis of these relationships. The work concludes that a number of household characteristics of the kind to be found in UK census-derived small area statistics may be predicted from particular load profile indicators. A discussion of the steps required to test and validate this approach and the wider implications for census taking is also provided.

1. Energy monitoring for a 'Smart Census'

Area based population statistics in the United Kingdom (UK) have historically been derived from the decadal census of housing and population. In addition to basic demographic statistics, the socio-economic information collected is used to produce robust small area estimates of a range of characteristics for every neighbourhood. Representing 'a definitive snapshot of the nation' (Calder & Teague, 2013) this data provides a backbone for commercial, academic and social research as well as policy analysis, a decadal 're-grouping' and 're-basing' of all small area population projections statistics (Norman, 2013) and, crucially, national and local resource allocation (Eurostat, 2011; Norman, 2013). Nonetheless, the UK census has also faced criticism as a costly and frequently outdated source of population statistics, with a time lag of at least two years between data collection and reporting (Dugmore, Furness, Leventhal, & Moy, 2011b).

Currently considered approaches for the future provision of population statistics include decennial census-taking, more frequent social surveys or administrative (Government held) data linkage and aggregation (ONS, 2013). In contrast, this work explores the possibility of deriving small area estimates of traditional socio-economic indicators from 'digital trace' or transactional data collected by utility (or other) services as part of normal service provision. As a number of recent authors have noted large-scale geo-coded transactional datasets, such as those collected in the retail, telecommunications, finance and utilities sectors could offer opportunities to supplement census based small area statistics by supporting the delivery of area-based population statistics, and generating novel indicators at a neighbourhood level (Deville et al., 2014; Dugmore et al., 2011b; Struijs, Braaksma, & Daas, 2014). For the United Kingdom Statistics Authority, via its executive office the Office for National Statistics (ONS) in England and Wales, the use of commercial data to support census taking may therefore help address census users' requests for more frequent and timely reporting of census-type statistics in the intercensal periods.

* Corresponding author.
E-mail addresses: b.anderson@soton.ac.uk (B. Anderson), X.Lin@soton.ac.uk (S. Lin), A.Newing@leeds.ac.uk (A. Newing), A.S.Bahaj@soton.ac.uk (A. Bahaj), P.A.James@soton.ac.uk (P. James).

http://dx.doi.org/10.1016/j.compenvurbsys.2016.06.003

Recent related work suggests that commercial 'big data' could both support near real time census taking and also provide unique insights into household or individual behaviours (Carroll, Lyons, & Denny, 2014; Claxton, Reades, & Anderson, 2012; Deville et al., 2014; Douglass, Meyer, Ram, Rideout, & Song, 2015; Dugmore et al., 2011b; Pucci, Manfredini, & Tagliolato, 2015). In this work we consider household level data held by a range of utility companies before focusing in particular on smart meter derived electricity consumption data. Compared to a number of other forms of potentially useful 'big data', a grid-connected electricity supply is almost universally available in the UK, almost universally connected to domestic dwellings and metering of consumption is mandatory. Furthermore the planned universal roll-out of electricity smart meters collecting at least half-hourly consumption data (DECC, 2013) means that consideration of the value of suitably anonymised and aggregated smart meter data in the production of official statistics is now timely.

The use of this kind of data for market segmentation and other electricity related services has been noted in the literature (McKenna, Richardson, & Thomson, 2012) and was noted by Dugmore et al. (Dugmore et al., 2011b) in the context of future census data collection. However, as far as we are aware only one published study has investigated its potential in the development of official and/or small area statistics (Carroll et al., 2014). A growing literature suggests that household level electricity load data, collected via smart metering, could provide considerable opportunities to infer household characteristics (Beckel, Sadamori, & Santini, 2012; Newing, et al., 2015; Struijs et al., 2014) The link between household characteristics and household energy consumption is long established and the literature recognises that household characteristics will give rise to different load profiles and subsequent demand on the electricity supply network (e.g. see McLoughlin, Duffy, & Conlon, 2013 for a summary). Consequently, the energy sector uses household or area based indicators of household composition and characteristics to predict electricity 'demand' in order to manage networks and target interventions designed to reduce or time-shift peak loads (e.g. see Elexon, 2013; Hamidi, Li, & Robinson, 2009; Wright & Firth, 2007).

The purpose of this work is to explore the value of inverting this approach to assess the feasibility of using observed high temporal resolution electricity consumption data to infer household characteristics as a first step in the aggregation of household characteristics to form 'normal' area level population statistics. It should be emphasised therefore that the overall objective is not to characterise or 'profile' individual households, rather we seek to aggregate inferred household characteristics to develop area based 'neighbourhood' indicators similar to or in combination with Census estimates or other appropriate datasets.

This work briefly reviews the future provision of area based statistics in the UK, recognising the opportunities to enhance or supplement the census taking process with digital trace data. It then considers the extent to which digital trace data from the commercial sector could represent a novel tool to generate census type small-area statistics, before focusing on the use high resolution electricity consumption monitoring data collected via smart metering. Based on preliminary analyses of a 'smart meter-like' dataset the research highlights the potential value of the approach and then discusses significant challenges and concludes by setting out a research programme which could systematically test the value of the approach.

2. Future provision of area based population statistics in the UK

As a consistent and robust source of small area population statistics, the United Kingdom census is used to allocate billions of pounds of government and commercial investment at the local level. It represents a fundamental tool for market research, policy making, commercial decision making, resource allocation and for academic research (ONS, 2013; Watson, 2009). Estimates of population counts by age and sex are a key census output, yet the detailed attribute information related to households and their usual residents - determining characteristics such as ethnic composition, education, socio-economic status, religion and employment – offer greatest value to the academic and commercial sector.

Census data are not made available at the individual household level but are published as non-disclosive aggregated counts within a hierarchy of 'output zones' or areas. These are built from unit postcodes, designed for the release of aggregate population statistics and represent small areas ranging from Output Areas (OAs – typically containing around 125 households) through to local authority districts (LADs) or Unitary Authorities (UAs). The former represents an important analytical unit for resource allocation and policy making at the local level, especially within the commercial sector (Dugmore, 2013; ONS, 2014a). It is this combination of universal geographic coverage at the small area level coupled with detailed attribute data that represents a major strength of the census (House of Commons Treasury Committee, 2008).

However, inevitably increasing costs, difficulties of ensuring full response, concerns over the decadal reporting cycle and the two year time-lag between census-taking and the delivery of initial outputs has given rise to a search for alternatives (Dugmore et al., 2011b). This work has been conducted by the ONS 'Beyond 2011' programme (ONS, 2014a) and, together with subsequent reviews of international census taking practice (see for example Dugmore, Furness, Leventhal, and Moy (2011a); and Martin (2006)), has highlighted a variety of approaches to collecting area based statistics including the use of governmental administrative sources (e.g. Netherlands and Denmark) or a rolling census (France). However the work also showed that a number of options under consideration by 'Beyond 2011', particularly those driven by administrative data, were unable to provide the level of socio-economic attribute data that many census users rely upon for commercial analysis, policy making and resource allocation (Calder & Teague, 2013; ONS, 2014a). Additionally, concerns have been raised over the likely success and practicalities of a census based on an administrative or register based system given the lack of a population register within the UK (Skinner, Hollis, & Murphy, 2013).

Based on the recommendations of the Beyond 2011 program (ONS, 2014a) on, extensive user consultation (ONS, 2014b) and an independent review (Skinner et al., 2013), the UK Statistics Authority recommended to parliament that a 'traditional' decadal census should be carried out in 2021 (Dilnot, 2014). They also noted that this should be primarily carried out online and that the considerable potential of utilising administrative data and larger scale household surveys as a supplement to census based statistics should be developed further (Dilnot, 2014).

Whilst recognising that data held by commercial organisations may offer more cost effective or timely reporting (Dugmore et al., 2011b), this avenue has received far less attention and discussion has tended to refer only to 'customer information' recorded in customer service databases and/or retail transaction data. As far as we are aware, commercial data does not currently feature within the national statistical census taking or population statistics of any nation. As Struijs et al. (2014) note such data could be used to provide substantial additional data over and above basic address listings.

3. Smart meters for a Smart Census

The nascent roll-out of domestic electricity smart meters in a number of major markets including the US, China, Brazil, India and Japan (Deloitte, 2011) and the UK (DECC, 2012) provides an opportunity for the exploration of precisely the scenario described above.

In the UK, smart meters incorporate communication infrastructure allowing them to transmit near real-time energy usage data to in home display units (IHDs), to energy demand service operators selected by the customer and to a centralised data retrieval service to extract half-hourly data from all smart meters for use by energy suppliers (billing and fraud prevention); network operators (network management)