Estimating the Prevalence of Shared Accommodation across the UK from Big Data

Ananya Samuel[*1], Guy Lansley[†1] and Rory Coulter[‡1]

[1]Department of Geography, UCL

January 29, 2019

Summary

This paper introduces a new means of measuring the proportion of shared accommodation within UK neighbourhoods using linked administrative and consumer data. An address-level multi-person household indicator (MHI) was produced for individual years spanning 1997 to 2016. Crucially, this new indicator enables fine-grained spatial analysis of trends in house-sharing outside of decennial census years. This abstract discusses the purpose and derivation methodology of the MHI, before illustrating how it can be used and outlining some future research and policy applications.

KEYWORDS: Shared housing, household composition, neighbourhood change, administrative data, consumer data

1. Introduction

Relatively little is known about the prevalence of shared accommodation across the UK. Comprehensive census data are collected infrequently and local authority registrations of Houses in Multiple Occupation (HMOs) are patchy, partial in coverage, and difficult to access. This paper aims to enrich the data landscape on shared accommodation by presenting a Multi-person Household Indicator (MHI) built from routinely collected and annually refreshed linked administrative and consumer data about the adult population. The indicator could provide insights into changing household structures, neighbourhood dynamics and housing policy effects. It also contributes to a growing body of quantitative work that harnesses big data to supplement government surveys and the census (Lansley and Cheshire, 2018). This paper will provide an account of this dataset and how it can be applied, using the City of Bristol as an illustrative case study.

2. Background

Unfortunately, there is no agreed definition of shared accommodation. Shared housing usually encompasses dwellings in which there are two or more unrelated family units or single individuals (Ahrentzen, 2003). Such accommodation typically consists of private bedrooms with communal facilities and spaces for which sharers are jointly responsible (Hemmens *et al.*, 1996). There are two broad types of sharing described in the literature, both of which recognise young people to be a key demographic (Hemmens *et al.*, 1996). Firstly, *choice-based shares* by young professionals and students who are not yet in a position to enter the housing market. Secondly, *constraint-based* 'stranger shares' that are typically linked to poor housing conditions, poor health and poverty (Bush and Shinn, 2017; Dehavenon, 1996).

In the UK, limited housing construction combined with population growth and an uneven labour market

[*] ananya.samuel.16@ucl.ac.uk
[†] g.lansley@ucl.ac.uk
[‡] r.coulter@ucl.ac.uk

has led to housing pressures in larger cities. In the South East especially, this has driven up house prices and contributed to affordability problems in owner-occupied and rental markets. High costs have pushed adults into living in shared accommodation as combining incomes enhances purchasing power. In many neighbourhoods the dwelling stock has barely changed, yet the population has increased substantially due to the decline of traditional family households and the rise of house conversions and shared accommodation. Indeed, much of the densification of British cities has occurred within pre-existing buildings rather than through the construction of high-rise blocks (Barker, 2008). This has fundamentally changed the fabric of many neighbourhoods, for instance as an increase in the supply of shared accommodation often coincides with gentrification (Cameron, 2003).

While shared accommodation can provide suitable short-term housing for some people, for others they are costly, exploitative and sometimes dangerous. Whilst there are regulations on housing quality these are poorly enforced and private rents are completely deregulated. Although this means there is a pressing need to understand the geography of shared housing, high levels of mobility within the sector makes sharers hard to reach and they are typically under-enumerated in official datasets. Therefore, this study aims to present a means to identify concentrations of multi-person households using linked big data resources in order to assist local authorities with service delivery.

3. Defining a Multi-person Household

The MHI was built from the Linked Consumer Register (LCR), an address-level register of adults which records the duration of their occupancy across a 20-year period. The register was formed from the linkage of public versions of the electoral register and various anonymised consumer datasets. The database is representative of the vast majority of the adult population and records annual changes from 1997 until 2016 (Lansley et al, 2018). While the consumer data mitigate data biases in the component electoral registers, it is important to recognise that those in shared accommodation are still likely to be under-enumerated at their correct addresses. Thus, we anticipate that we may slightly under-enumerate multi-person households.

For this analysis, two parameters were used to define a multi-person household. These were addresses that:

- Contain at least 3 adults
- AND none of whom share a surname[§]

We assume that firstly, each address refers to a single household unit. It is not currently possible to match the Office for National Statistics (ONS) 'shared facilities' definition of households using administrative and consumer datasets, thus like the administrative census project, we consider occupied addresses as our unit of analysis (ONS, 2017). Secondly, we assume that surnames distinguish family identity. Of course, there will be exceptions such as wives that retained their maiden surnames or unrelated occupants who coincidentally share a surname.

The data were aggregated to the LSOA level and steps were taken to ensure the outputs were not disclosive. The MHI indicator represents the proportion of identified multi-person households out of the total number of occupied addresses per LSOA for each year.

Encouragingly, testing the 2011 MHI against 2011 Census data on multi-person households revealed a strong correlation between the two measures (r= 0.891). This is despite the fact that the Census defines multi-person households more broadly to be essentially any household with at least two unrelated

[§] This rule is relaxed for addresses with more than 20 surnames, which are all classed as multi-person households.

adults. This includes households where a family live with an unrelated adult and also multi-family households. However, it is feasible for analysis on the raw LCR to also identify these household types.

4. Outputs

To illustrate how the MHI can be used, **Figure 1** shows the proportion of multi-person households within LSOAs across the City of Bristol. The highest rates of multi-person households are situated in cosmopolitan neighbourhoods near the university campuses. The two maps indicate that over the 10-year period there has been an increase in the proportion of multi-person households throughout the central parts of the city, even in areas not traditionally occupied by students. This suggests that a growing proportion of the population in Bristol are living in shared accommodation. Indeed, the 2011 Census confirms that Bristol has a large young professional population which are mainly concentrated in neighbourhoods proximal to the city centre. The increase of multi-person households to the north east was perhaps caused by the expansion of the University of the West of England.

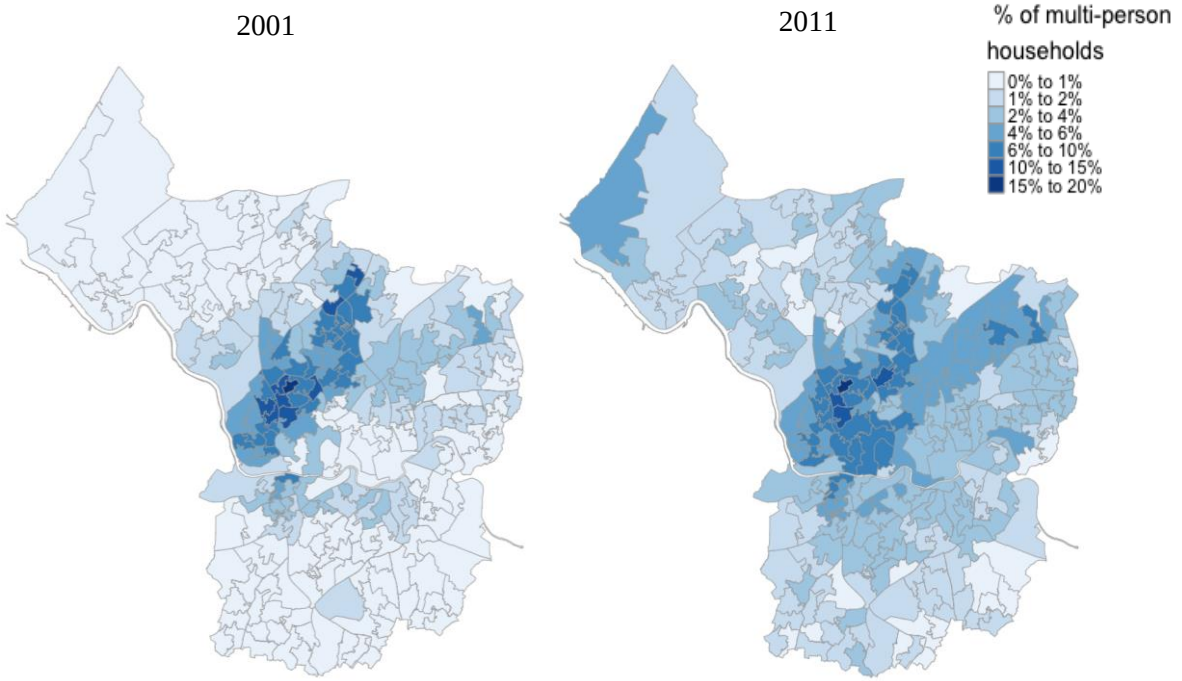

Figure 1 Multi-person Households in Bristol

It is likely that areas with high rates of shared accommodation also on average have high proportions of students and higher proportions of households in the private rental sector. To explore this, a regression model was constructed using the proportions of students and private rented households in the 2011 Census as independent variables and the 2011 MHI as the dependent. The results are shown in **Table 1**. While both variables are significant predictors, the distribution of MHI is still independent to a large degree. Given the robustness of the methods used to determine the index and the high association with equivalent data during census years, this demonstrates there is utility in a bespoke product specifically designed to identify shared accommodation.

Table 1 Results of linear regression of MHI on covariates in Bristol

	Estimate	P-Value
Intercept	0.00437	0.0035
Private renters	0.0917	0.00
Students	0.0858	0.00
Adjusted R-square	0.720	

5. Conclusions

The Multi-person Household Indicator (MHI) demonstrates how linked Big Data can be harnessed to examine local-scale population and residential change at more regular refresh cycles than with traditional datasets. Although we have presented a robust indicator, the definitions and granularity of the MHI can be tailored further to test specific hypotheses or evaluate particular policy interventions (e.g. the housing market impacts of housing benefit reforms).

6. Acknowledgements

This work is funded by the UK ESRC Consumer Data Research Centre (CDRC) grant reference ES/L011840/1.

Biographies

Ananya Samuel is an intern at the UK Consumer Data Research Centre.

Guy Lansley is a Research Associate at the UK Consumer Data Research Centre and the Department of Geography, University College London (UCL).

Rory Coulter is a Lecturer in Quantitative Human Geography at the Department of Geography, UCL.

References

Ahrentzen, S. (2003). Double indemnity or double delight? The health consequences of shared housing and "doubling up". *Journal of Social Issues*, *59*(3); 547-568.

Barker, K. (2008). Planning policy, planning practice, and housing supply. *Oxford Review of Economic Policy*, *24*(1); 34-49.

Bush, H. and Shinn, M. (2017). Families' Experiences of Doubling Up After Homelessness. *Cityscape*, *19*(3); p.331-356.

Cameron, S. (2003). Gentrification, housing redifferentiation and urban regeneration: 'going for growth' in Newcastle upon Tyne. *Urban Studies*, *40*(12); pp. 2367-2382.

Dehavenon, A.L. (1996). Doubling-up and New York City's policies for sheltering homeless families. *There's no place like home: Anthropological perspectives on housing and homelessness in the United States*, Greenwood, Westport. pp. 51-66.

Hemmens, G. and Charles Hoch (1996). Shared housing in low income households. In: *Under one roof: Issues and innovations in shared housing*, SUNY Press, New York. pp. 17-32.

Hemmens, G.C., Hoch, C. and Carp, J. eds. (1996). Introduction. In: *Under one roof: Issues and innovations in shared housing*. SUNY Press, New York. pp. 1-16

Lansley, G. and Cheshire, J. (2018). Challenges to representing the population from new forms of consumer data. *Geography Compass*, e12374.

Lansley, G., Li, W. and Longley, P., 2018. Modelling Small Area Level Population Change from Administrative and Consumer Data. *Geographical Information Science Research UK (GISRUK)*.

Office for National Statistics. (2017) Research Outputs: Occupied address (household) estimates by size, 2011 - Office for National Statistics. Online: https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusresearchoutputs/householdsandfamilies/researchoutputsoccupiedaddresshouseholdestimatesbysize2011 (Accessed 02.01.2019)