Local area estimation of expenditure profiles and consumer attitudes

William H.M. James[1], Nik Lomax[1]

[1]School of Geography and Leeds Institute for Data Analytics, University of Leeds, Woodhouse Lane, Leeds, West Yorkshire, LS2 9JT, UK

29th January 2019

Summary

Examining patterns of expenditure and consumer attitudes is important for research into environmental impact, health, obesity and other 'lifestyle diseases'. In this research paper, we present new GIS datasets of household expenditure and consumer attitudes across Great Britain. The technique of spatial microsimulation was used to estimate household expenditure of 103 categories of food and drink for each Local Authority District of Great Britain for the years 2008 – 2016. Additional datasets relating to consumer attitudes were generated using microsimulation, allowing for the interplay between consumer attitudes and their expenditure profiles to be investigated.

KEYWORDS: Spatial Microsimulation, Household Expenditure, Consumer Attitudes

1. Background

In common with the rest of the world, the UK is experiencing major shifts in dietary patterns, as evidenced by changing patterns of food and drink expenditure (Kearney, 2010). Consumer attitudes play a major role in product choices, with views on environmental sustainability, health and ethics all influencing the expenditure profile of an individual (Kriwy and Mecking, 2012). Understanding the local scale patterns of expenditure and consumer attitudes is therefore critical for assessing market stability and for forecasting future trends. Examining patterns of expenditure and attitudes is also important for research into environmental impact, health, obesity and other 'lifestyle diseases'. Despite its importance, there has been relatively little research into the spatial patterns of expenditure and associated consumer attitudes in the UK.

In this research paper, we present new GIS datasets of household expenditure and consumer attitudes across Great Britain. All datasets are at the Local Authority District level (n = 380) allowing for spatial variation to be assessed. The technique of spatial microsimulation was used to estimate household expenditure of 103 categories of food and drink for each local authority district of Great Britain for the years 2008 – 2016. Additional datasets relating to consumer attitudes (supermarket shopping behaviours, attitudes to sustainability, environmental issues and food) were also generated using microsimulation, allowing the relationships between consumer attitudes and expenditure to be assessed.

2. Methods

This study employs the technique of Spatial Microsimulation in the form of Iterative Proportional Fitting (IPF), implemented within the R programming environment (https://www.r-project.org). IPF works by adjusting a large array of weights - rows corresponding to individuals and columns corresponding to the geographic zones (e.g. Local Authority Districts) - iteratively, to maximise the fit between simulated and known (e.g. census/survey) data. The mathematics of IPF are covered by

Fienberg (1970), a guide to implementation is provided in Lomax and Norman (2016) whilst the code used here for implementing IPF in R was developed by Lovelace and Dumont (2016).

2.1. Expenditure microsimulation

For the years 2008 – 2016, spatial microsimulation was used to generate estimates of expenditure, as shown in Figure 1. Firstly, a baseline population structured by age and sex was generated for each Local Authority District. This was based on data from the Office for National Statistics mid-year population estimates (ONS, 2018), with residents living in communal establishments (ONS, 2011) removed to ensure that the baseline was compatible with the individual level expenditure data.

Additional demographic constraints (ethnicity, student status, unemployment status, income, household characteristics) were taken from the Census (ONS, 2011b), Annual Population Survey (ONS, 2018a), ONS Estimates of Unemployment (ONS, 2018c) and the Annual Survey of Hours and Earnings (ONS, 2018b). These were all standardised to the baseline population, as required for microsimulation.

Microdata were taken from the Living Cost and Food Survey (LCF), covering approximately 12,000 respondents from 6,000 households each year. The LCF comprises an expenditure diary detailing purchases over a two-week period and an interview covering socio-demographic characteristics, income and regular items of household expenditure. The diary provides a record of expenditure grouped by category, based on The Classification of Individual Consumption by Purpose (COICOP) coding framework. This allows for the detailed recording of expenditure items (e.g. code 01.1.2.5.1 = sausages). The framework also allows easy aggregation to higher levels (e.g. 01.1.2.5 = Dried, salted or smoked meat and edible meat offal; 01.1.2 = Meat; and 01.1 = Food). There are a total of 103 food, drink and tobacco categories covered by the COICOP coding framework. For the years 2008 – 2016 these microdata were combined with the constraint variables using microsimulation to estimate expenditure at the local authority level. One of the intermediate outputs of this process is a 'synthetic population' – a list of all simulated individuals in each Local Authority District with associated demographics. The process for producing this synthetic population is outlined in Figure 1.
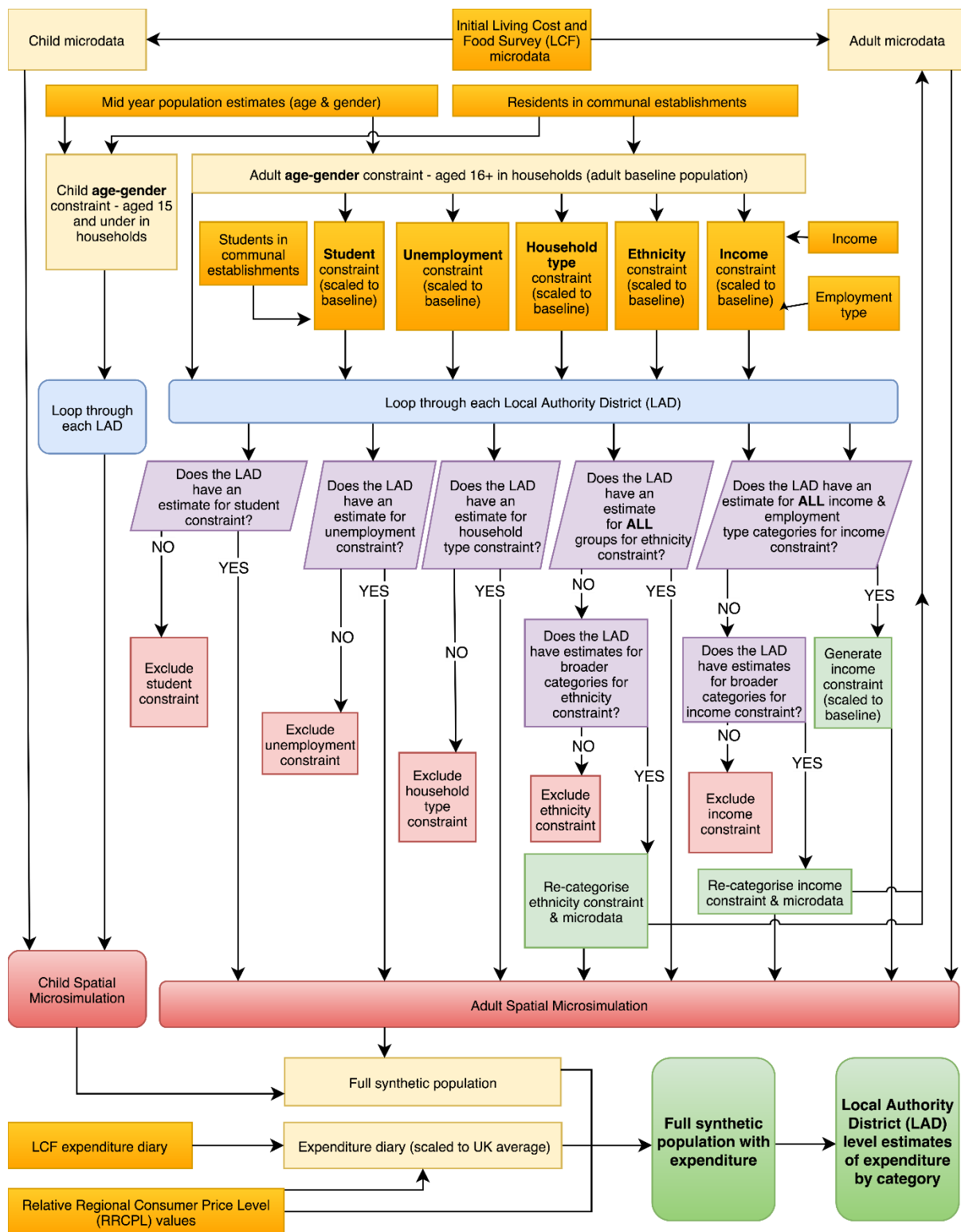
Child microdata

Initial Living Cost and Food Survey (LCF) microdata

Adult microdata

Mid year population estimates (age & gender)

Residents in communal establishments

Child **age-gender** constraint - aged 15 and under in households

Adult **age-gender** constraint - aged 16+ in households (adult baseline population)

Students in communal establishments

Student constraint (scaled to baseline)

Unemployment constraint (scaled to baseline)

Household type constraint (scaled to baseline)

Ethnicity constraint (scaled to baseline)

Income constraint (scaled to baseline)

Income

Employment type

Loop through each LAD

Loop through each Local Authority District (LAD)

Does the LAD have an estimate for student constraint?

Does the LAD have an estimate for unemployment constraint?

Does the LAD have an estimate for household type constraint?

Does the LAD have an estimate for **ALL** groups for ethnicity constraint?

Does the LAD have an estimate for **ALL** income & employment type categories for income constraint?

NO

YES

YES

YES

NO

YES

NO

YES

Exclude student constraint

NO

Exclude unemployment constraint

NO

Exclude household type constraint

Does the LAD have estimates for broader categories for ethnicity constraint?

Does the LAD have estimates for broader categories for income constraint?

Generate income constraint (scaled to baseline)

NO

Exclude ethnicity constraint

NO

Exclude income constraint

YES

YES

Re-categorise ethnicity constraint & microdata

Re-categorise income constraint & microdata

Child Spatial Microsimulation

Adult Spatial Microsimulation

Full synthetic population

LCF expenditure diary

Expenditure diary (scaled to UK average)

Relative Regional Consumer Price Level (RRCPL) values

Full synthetic population with expenditure

Local Authority District (LAD) level estimates of expenditure by category

Figure 1 Schematic diagram of the expenditure estimation framework

2.2. Consumer attitudes microsimulation – YouGov Survey

For the year 2016, consumer attitudes for each Local Authority District were estimated using YouGov survey data, Living Cost and Food Survey (LCF) microdata and the synthetic population generated in section 2.1, as shown in Figure 2. The YouGov survey data was from 2015 - 2016 and records responses to questions relating to supermarket shopping behaviour, attitudes to sustainability, environmental issues, food-related behaviours and commuting. Additional demographic information regarding age, sex and socio-demographic class were also available as part of the survey. Using the common demographic variables between the YouGov survey and from the LCF (age, sex, social grade), survey responses were re-weighted to best fit the LCF. These weights were subsequently applied to the 2016 synthetic population (generated from the previous microsimulation procedure) to generate an estimate of survey response for each Local Authority District.

Figure 2 Schematic diagram of the consumer attitudes framework
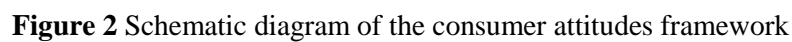
3. Results

Geographical datasets were created by joining the results of the microsimulation with Local Authority GIS boundaries. For visualisation and integration with other datasets, the GIS vector shapefiles were converted to a spatial grid of data cells in a similar manner to other studies (e.g. James *et al.*, 2018). Figure 3 shows examples of selected expenditure datasets for the year 2012. Cumulative categories are generated by summing the appropriate individual COICOP categories (e.g. all food and drink: Figure 3a, alcoholic drinks: Figure 3b, tobacco and cigarettes: Figure 3c) whilst individual COICOP categories can also be mapped (e.g. bacon and ham purchased for household supplies: Figure 3d). Figure 4 shows examples of selected consumer attitude datasets for the year 2016, relating to attitudes to 'healthy' foods (Figure 4a) and environmental attitudes (Figure 4b).
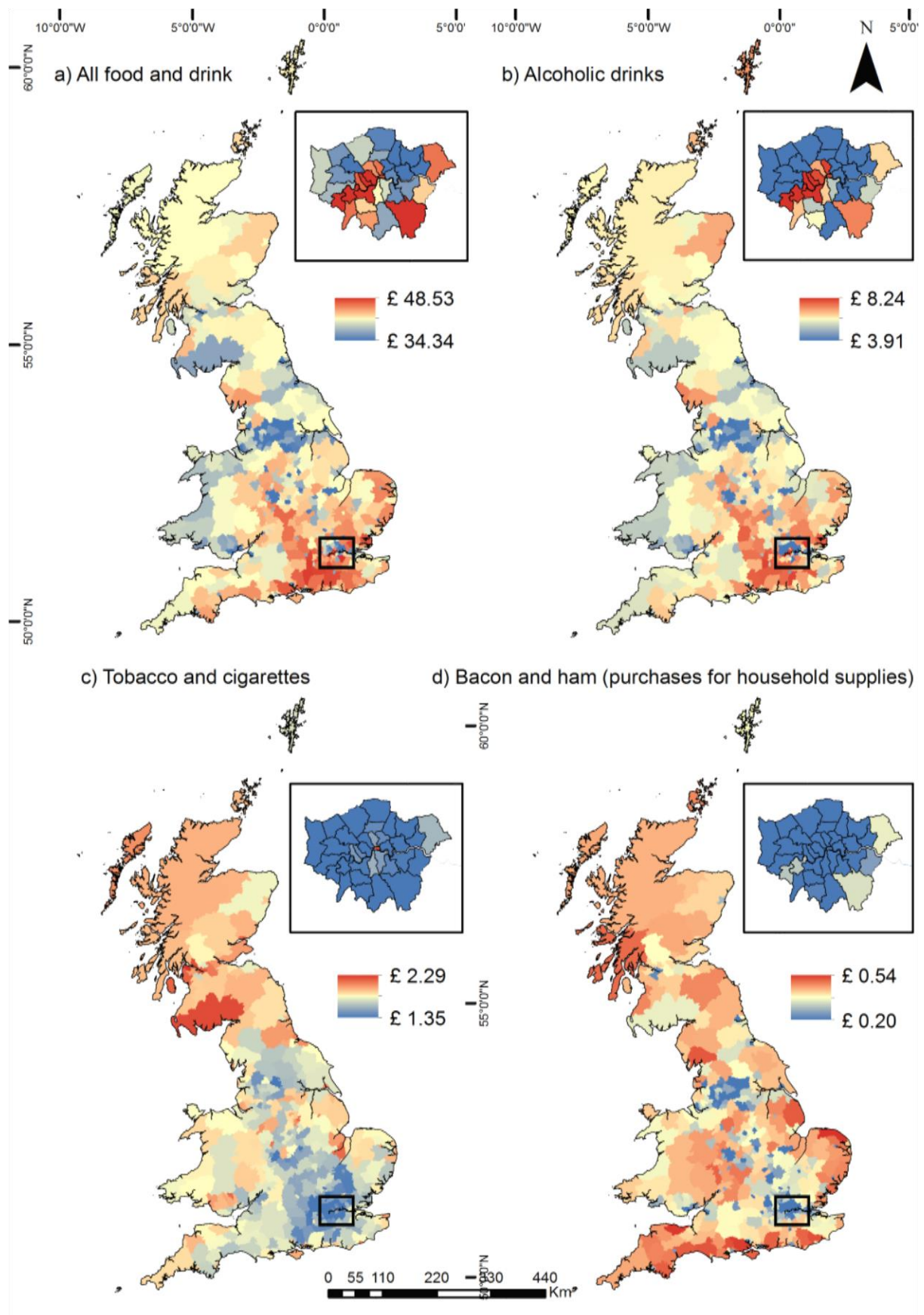
Figure 3 Estimated average weekly expenditure per person for 2012
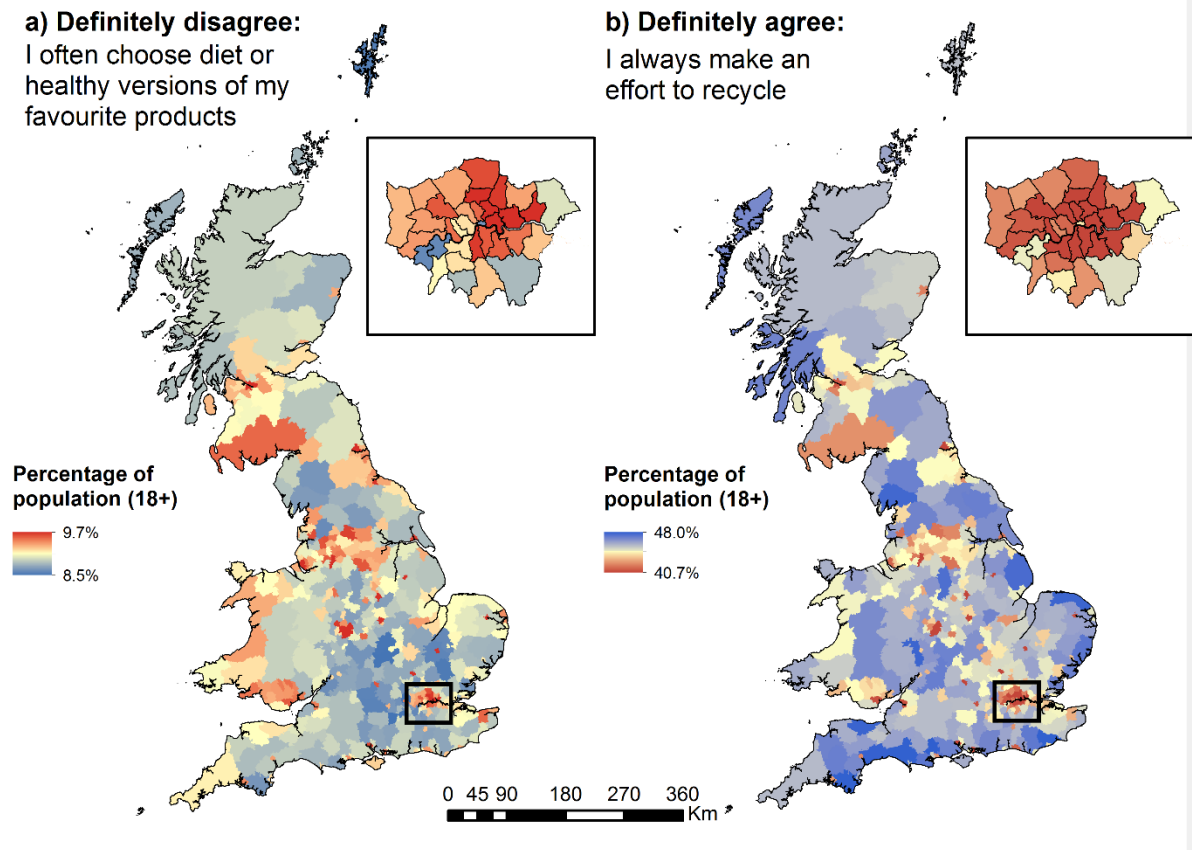
Figure 4 Estimated responses to questions related to consumer attitudes (2016)

4. Conclusions and further work

This paper presents new GIS datasets of household expenditure and consumer attitudes across Great Britain. We believe spatial microsimulation techniques of the type described in this paper hold great potential benefits for a range of disciplines including economics, retail geography and public health. Whilst this study focusses on Great Britain, the framework here could be applied to any location with the appropriate data sources. Whilst analysis of these datasets will help improve our knowledge of the current relationships between consumer attitudes and expenditure profiles, they can also form the basis for future projections under different scenarios.

Acknowledgements

Reference List

Fienberg, S. E. (1970) 'An iterative procedure for estimation in contingency tables', *The Annals of*

Mathematical Statistics, 41, pp. 907–917.

James, W. H. M. *et al.* (2018) 'Gridded birth and pregnancy datasets for Africa, Latin America and the Caribbean', *Scientific data*, 5, p. 180090.

Kearney, J. (2010) 'Food consumption trends and drivers', *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365, pp. 2793–2807.

Kriwy, P. and Mecking, R. (2012) 'Health and environmental consciousness, costs of behaviour and the purchase of organic food', *International Journal of Consumer Studies*. Wiley Online Library, 36(1), pp. 30–37.

Lomax, N. and Norman, P. (2016) 'Estimating population attribute values in a table:"get me started in" iterative proportional fitting', *The Professional Geographer*, 68, pp. 451–461.

Lovelace, R. and Dumont, M. (2016) *Spatial microsimulation with R*. CRC Press.

Office for National Statistics (2011a) 'Census 2011 Table LC1105EW - Residence type by sex by age'.

Office for National Statistics (2011b) 'Census 2011 Table QS110UK - Adult lifestage (alternative adult definition)'.

Office for National Statistics (2018a) 'Annual Population Survey'.

Office for National Statistics (2018b) 'Annual Survey of Hours and Earnings'.

Office for National Statistics (2018c) 'Model-based estimates of unemployment'.

Office for National Statistics (2018d) 'Population estimates - local authority based by single year of age'.

Biographies

William James is a research fellow at the School of Geography and Leeds Institute for Data Analytics, University of Leeds. His work focusses on consumer data analysis and applied GIS as part of the PigSustain project.

Nik Lomax is a University Academic Fellow in Data Analytics, based in the School of Geography, University of Leeds. Nik's research interests are in population estimation and projection. Nik is a Co-Investigator of the ESRC funded Consumer Data Research Centre and is a funded Turing Fellow.