Spatial Microsimulation and Levelling Up: The Importance of Fine Scale Food Consumption Data

Dr William James*[1] and Prof Nik Lomax†[1]

[1]School of Geography, University of Leeds

GISRUK 2023

Summary

There are considerable spatial inequalities in food consumption which are associated with health issues such as obesity, cardiovascular disease and cancer. These issues need to be tackled as part of the levelling up agenda to improve health outcomes in deprived areas. Whilst traditional data sources are insufficient to capture the fine scale variability of food consumption, Spatial Microsimulation can be used to estimate consumption at the local level. In this study, we show how individual food diary data can be combined with Census data to derive fine scale estimates. This will enable the targeting of resources, funding and support.

KEYWORDS: Spatial Microsimulation, food consumption, health.

1 Background

Levelling up is a term used to to describe the ambition for addressing longstanding spatial inequalities and for specifically tackling the problems of left-behind places (Tomaney and Pike, 2020). Socio-economic inequalities persist at a range of spatial scales, although there is growing recognition that focusing on broad-scale trends can mask important fine-scale variations (e.g. Victora et al., 2017). Furthermore, knowledge of fine scale variations allows the geographically targeted deployment of resources, funding and policy implementations where they are most critically needed. The utility and success of such targeted approaches is explored in James et al. (2022).

There are existing datasets of fine-scale socio-economic inequalities in England, namely the Index of Multiple Deprivation (IMD). The IMD is constructed of seven distinct domains (income, employment, health, education, crime, housing and living environment) which are combined to provide a measure of relative deprivation for every Lower-layer Super Output Area in England (i.e. 32,844 distinct areas). Whilst useful for assessing relative deprivation at the fine spatial scale, the definition of deprivation is relatively broad (constructed from the seven indicators) which may not be suitable for informing policy, which is often focused on a specific issue. For example, policies for

*w.h.m.james@leeds.ac.uk
†N.M.Lomax@leeds.ac.uk

tacking inequalities in food and drink consumption require specialised data related to consumption patterns, which may not correspond directly to the broader measures of deprivation. Fine spatial scale data relating to a specific issue (e.g. food consumption) are seldom available, limiting the scope of geographically targeted policy intervention.

Spatial Microsimulation is a method for assimilating existing fine spatial scale data (e.g. common socio-economic variables from the Census) with specialised a-spatial data (e.g. food consumption data collected via a sample survey). This allows estimates of the specialised data to be constructed at the fine spatial scale. This paper discusses how Spatial Microsimulation can be used to map fine scale (LSOA level) food consumption.

1.1 Socio-economic patterns of food and drink consumption (a-spatial analysis)

The National Diet and Nutrition Survey (National Institute of Health Research Biomedical Research Centre, 2020) is a representative survey which provides information on socio-economic patterns of food and drink consumption across the UK. For example, Figure 1 demonstrates the relationship between selected socio-economic variables (collected as part of the NDNS) and processed red meat (bacon, ham etc.) consumption. Figure 1a shows consumption is significantly higher for individuals classified as white ethnic group compared to other groups. This is potentially due to cultural and/or religious reasons for which ethnic group is a proxy. Figure 1b shows that red meat consumption initially increases with age (likely due the overall smaller portion sizes of children vs adults), before remaining relatively stable for older age groups. Males consume significantly more red meat in all the adult age groups, perhaps due to differing portion sizes and general dietary preferences.

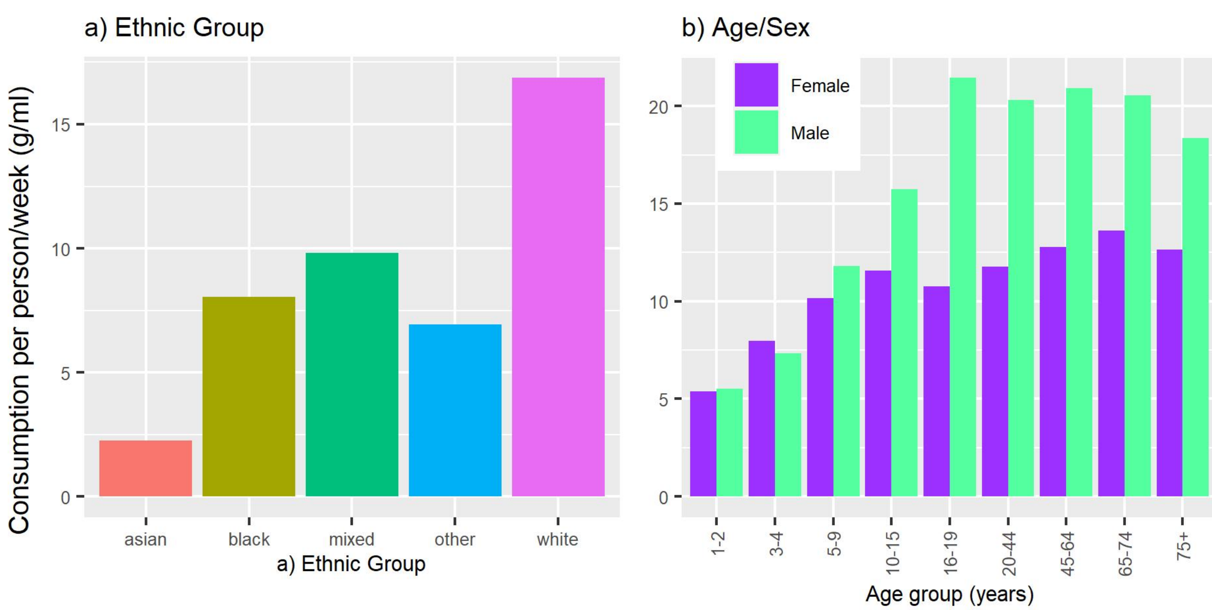

Figure 1: Socio-economic variation of food and drink consumption.

2 Methods

It is widely known that food consumption patterns vary by socio-economic variables (as shown in section 1.1 for processed red meat). These same variables show distinctive spatial patterns in the UK. For example, certain areas may have a higher proportion of elderly people or certain ethnic group than others. These spatial variations are known at the granual level, with the UK Census providing detailed breakdowns at a variety of scales. In this study, we focus on the Lower-layer Super Output Area (LSOA) level (i.e. 32,844 distinct areas).

In this study, we use Spatial Microsimulation to integrate the detailed (but a-spatial) individual level food consumption data (as discussed in section 1.1) with the fine scale UK Census data based on the common socio-economic variables. Specifically, we use Spatial Microsimulation in the form of Iterative Proportional Fitting (IPF), implemented within the R programming language. IPF works by iteratively adjusting an array of weights - rows corresponding to individuals and columns corresponding to the geographic zones (in this case LSOA's), to maximise the fit between simulated and known (in this case Census) data. The mathematics of IPF are described 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 Dumont and Lovelace (2016).

As mentioned in section 1.1, consumption data is sourced from the National Diet and Nutrition Survey (NDNS), a representative survey of approximately 1,000 individuals each year. The NDNS comprises of a 'food-diary' for each individual which provides detailed information on the quantities of each specific food product consumed (e.g. pork sausages, wholemeal bread etc.). A range of aggregated categories are also calculated as part of the survey (e.g. processed red meat) which includes items from composite meals (e.g. the ham from a ham and pineapple pizza). Importantly, the NDNS also includes data relating to the socio-economic characteristics of each individual (e.g. age/sex, ethnicity, highest qualification gained etc.), allowing cross-matching with corresponding Census variables in order for Spatial Microsimulation to be employed. Variables are chosen based on their known relationships with food consumption and inclusion in both the NDNS/Census. Variables need to be aggregated/binned to ensure matching between the NDNS and Census datasets, with the full list of variables shown in Table 2. A full discussion of variable choice/alignment for Spatial Microsimulation is provided by Lovelace et al. (2017).

This study currently uses combined NDNS data from 2008 - 2019 and 2011 Census data (i.e. it provides a temporally composite estimate which does not detail temporal consumption trends). These trends could be investigated by using a subset of the NDNS data and alternative Census data (e.g. 2021 counts).

3 Results

Figure 2 shows estimated processed red meat consumption (grams per person per day) for every LSOA area in England (i.e. 32,844 distinct areas), with the inset showing a detailed view for the Leeds/Bradford region. A clear urban/rural pattern is visible, with consumption significantly lower (minimum of 2.7 grams per person per day) in major cities such as London and higher in rural regions such as the Yorkshire Dales (maximum 17 grams per person per day). Within the Leeds/Bradford

Table 1: Linking variables

Variable	**Category/Bin**	**Census 2011 Table**
Age/Sex	1-2, 3-4, 5-9, 10-15, 16-19, 20-44, 45-64, 65-74, 75+ (cross tabulated by Male/Female)	LC1117EW
Ethnic Group	Asian, Black, Mixed, Other, White	KS201EW
Marital Status	Single, Non-single	KS103EW
Employment (Ages 16-74)	In employment, Not working	KS601EW to KS603EW
Student/living arrangement	Student (living alone), Student (living with parents), Student (Halls of residence/other), Non-student	LC4411EW
Highest Qualification (ages 16 - 64)	Degree, A-level/GCSE/Other, No Qualifications	LC5102EW
Tenure	Owned/mortgaged accommodation, Rented/other accommodation	QS403EW
Health (self reported)	Very bad, Bad, Fair, Good, Very Good	QS302EW

region, the lowest consumption is found in the eastern suburbs of Bradford (LSOA code E01010621: point A on the inset map) whilst the highest consumption (16.8 grams per person per day) is found in the village of Allerton Bywater to the South-East of Leeds (LSOA code E01011307: point B on the inset map).

By examining the socio-demographic characteristics of these two LSOAs (and relating these to the a-spatial consumption patterns outlined in section 1.1), we can start to understand why these areas have extreme consumption values and how the process of Spatial Microsimulation generated these estimates. For each LSOA, Figure 3 shows selected socio-economic variables compared to the national average. For the LSOA with the lowest consumption (code E01010621: point A on the inset map), there are a greater proportion of young people than the national average (and a corresponding lower proportion of elderly people). This partly explains the low estimate of processed red meat consumption, as we know that young people consume less than older age groups (as detailed in section 1.1). Importantly, this LSOA has a far greater proportion of individuals classed as asian ethnic group than the national average and a much lower proportion of those classed as white ethnic group. Individuals classed as asian ethnic group are known to consume significantly less processed red meat than their white counterparts (as shown in Figure 1), leading to the low estimate for the LSOA. Conversely, the age structure of LSOA code E01011307 is close to the national average and is over-represented by individuals classified as white ethnicity, producing the high estimate of processed red meat consumption. It should be noted the microsimulation process considers all the variables listed in Table 1 (not just age/sex and ethnicity), with the interaction of all these variables contributing to the final estimates.

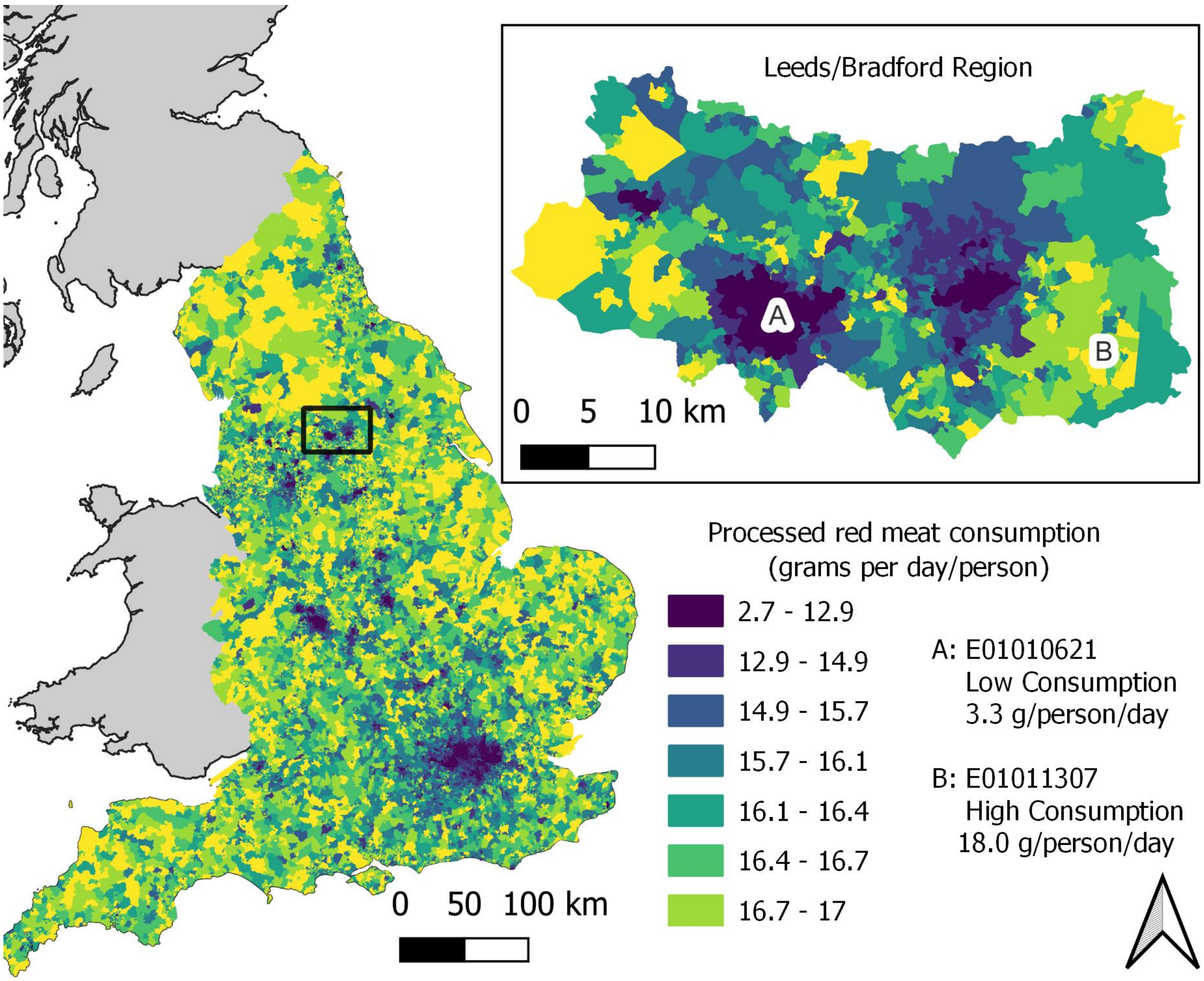

Figure 2: LSOA level consumption estimates of processed red meat for England

4 Discussion and Conclusions

This study shows the potential of microsimulation for generating fine scale (LSOA level) estimates of food consumption. Further work is needed to validate the model (namely by bench-marking against alternative independent data-sets) but the initial results are plausible and highly promising. Whilst the framework shown here is applied to a specific food group (processed red meat) and is temporally composite, the same process can be used for any food item/group and subset of data.

Patterns of food consumption are highly linked to health issues/outcomes (e.g. obesity, cardio-vascular disease, cancer, alcohol harm) with an increasing awareness that targeted policies are needed to tackle these issues (James et al., 2022). The levelling up agenda will require datasets such as these to identify specific areas which could benefit from specialist support and funding.

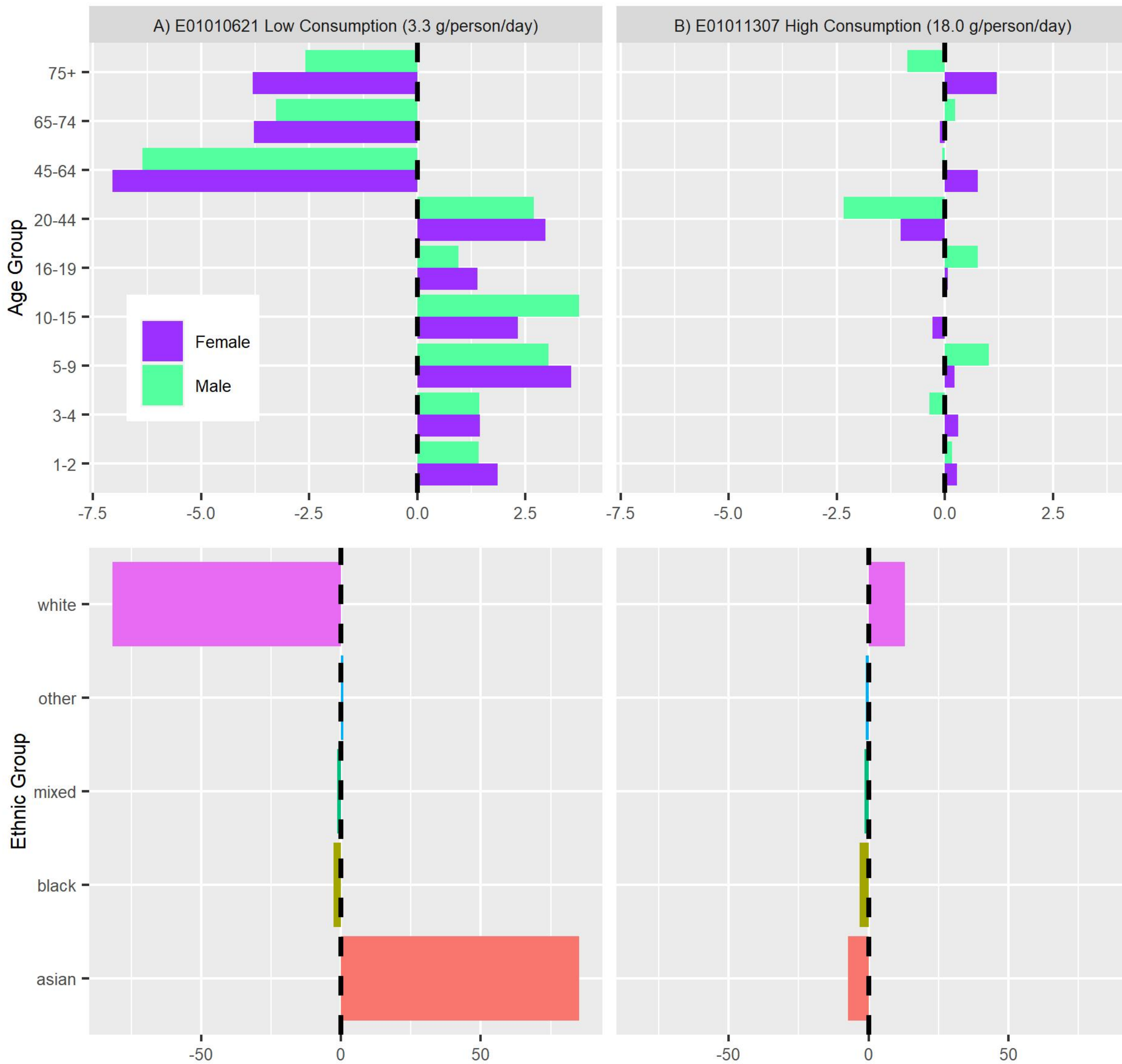

Figure 3: Variation between national average and LSOA specific socio-demographic variables

5 Biography

William James is a lecturer at the School of Geography, University of Leeds. His work focusses on consumer data analysis and applied GIS.

Nik Lomax is a Professor of Population Geography, based in the School of Geography, University of Leeds. Nik's research interests are in population estimation and projection. Nik is also a fellow at the Alan Turing Institute for Data Science and Artificial Intelligence and co-Director of the ESRC

funded Consumer Data Research Centre.

References

Dumont, M. & Lovelace, R. (2016). *Spatial Microsimulation with R.* CRC Press.

Fienberg, S. E. (1970). An iterative procedure for estimation in contingency tables. *The Annals of Mathematical Statistics*, *41*(3), 907–917.

James, W. H., Lomax, N., Birkin, M., & Collins, L. M. (2022). Targeted policy intervention for reducing red meat consumption: conflicts and trade-offs. *BMC nutrition*, *8*(1), 1–18.

Lomax, N. & Norman, P. (2016). Estimating population attribute values in a table: "get me started in" iterative proportional fitting. *The Professional Geographer*, *68*(3), 451–461.

Lovelace, R., Dumont, M., Ellison, R., & Založnik, M. (2017). *Spatial microsimulation with R.* Chapman and Hall/CRC.

National Institute of Health Research Biomedical Research Centre (2020). National Diet and Nutrition Survey Years 1-9, 2008-2017 `https://ukdataservice.ac.uk`.

Tomaney, J. & Pike, A. (2020). Levelling up? *The Political Quarterly*, *91*(1), 43–48.

Victora, C. G., Barros, A. J., França, G. V., da Silva, I. C., Carvajal-Velez, L., & Amouzou, A. (2017). The contribution of poor and rural populations to national trends in reproductive, maternal, newborn, and child health coverage: analyses of cross-sectional surveys from 64 countries. *The Lancet Global Health*, *5*(4), e402–e407.