Harris, R. J. (2020). Featured Graphic: Visualising the scales of ethnic diversity in London using a multilevel entropy index. *Environment and Planning A*. https://doi.org/10.1177/0308518X20904728

Peer reviewed version

Visualising the scales of ethnic diversity in London using a multilevel entropy index

Richard Harris,

School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 1SS.
rich.harris@bris.ac.uk

(Abstract) A challenge emerging from the literature on measuring and modelling segregation at multiple scales is how to map the results. This Featured Graphic uses a method based on the CMY colour model, allowing each of the C, M and Y values to represent one of the micro-, meso- and macro-levels of analysis separately, then combining them in an overall map. It presents the method using a new index, a multilevel index of diversity, with a case study looking at the ethnic diversity of state school pupils living in London's neighbourhoods.

Key words: ethnic diversity, entropy index, multilevel, segregation, London

In recent literature about how to measure segregation, methods have been developed that treat segregation not just as a numeric quantity (the amount of segregation measured) but also in terms of a spatial pattern: the geographical scales at which population groups are clustered together or apart (see, for example, the special editions of two journals: Harris & Johnston, 2018 and Piekut et al., 2019).

Some of these methods employ a three-level approach, casting segregation in terms of 'micro', 'meso' and 'macro' effects (Manley et al., 2015). An attraction of this, but not one previously employed, is to map each level to a base colour, providing a palette from which the various scales of segregation are painted. Here the approach is demonstrated with a multilevel index of ethnic diversity, a CMY(K) colour scheme and a case study mapping the ethnic diversity of state school pupils in neighbourhoods within London based on information extracted from the National Pupil Database.[1]

The multilevel index is based on the commonly used entropy index (Theil and Finezza, 1971; Theil, 1972), which is

$$E = -\sum_{1}^{n_g} p_g \log(p_g)$$

where the summation is over n_g ethnic categories and where p_g is the proportion of all pupils in a neighbourhood that are of ethnic group g. From this, a three-level model of LSOAs, Wards and LAs[2], may be formed as,

$$f(E_{ijk}) = \beta_0 + \epsilon_{ijk}$$

where ϵ_{ijk} is the (residual) difference between the measured value of diversity for each LSOA (E_{ijk}) and the constant, β_0 (which is the average diversity score across all LSOAs); and where

$$\epsilon_{ijk} = C_i + M_j + Y_k$$

[1] https://www.gov.uk/government/collections/national-pupil-database

[2] LSOAs (Lower Level Super Output Areas) are the second tier of the English and Welsh Census geography. They nest, hierarchically, into Wards and Local Authorities (LAs).

decomposes the residuals into three levels – those at the LSOA level (C_i), those at the Ward level (M_j) and those at the LAD level (Y_k). It is the estimates of C, M and Y that are used to colour the map.[3]

The results are shown in the graphic. In terms of the overall pattern of diversity, the greatest differences are between local authorities (LAs): 46.9 per cent of the variation in the spatial pattern is at the LA scale – local authorities such as Waltham Forest (numbered 31 on the map), Redbridge (26), Barking and Dagenham (1), Newham (25) and Hackney (12) have the greatest ethnic diversity; Bromley (5), Bexley (3) and Havering (16) are amongst those with the least. Nevertheless, there are variations within LAs both at the Ward and LSOA level. For example, there is a Ward within Bromley with a diversity that is greater than expected for that LA (it is shaded dark magenta in the top-right panel of the map) and there are LSOAs within Redbridge and also within Barking and Dagenham that have the highest overall levels of ethnic diversity of any in London. Those spatial clusters of greatest ethnic diversity do not stop at the boundaries of LAs but 'over-spill' into adjoining areas, including a part of Havering. In short, the geographical scales of diversity vary across the map, with 'pockets' of diversity adding to more broad-scale trends.

A tutorial and code to reproduce the graphic (with minor variations) is available at https://rpubs.com/profrichharris/MLentropy

References

Harris, R. & Johnston, R. (2018). Measuring and modelling segregation – New concepts, new methods and new data. *Environment and Planning B*, 45 (6), pp. 999–1002.

Manley, D., Johnston, R., Jones, K. and Owen, D. (2015). Macro-, Meso- and Microscale Segregation: Modeling Changing Ethnic Residential Patterns in Auckland, New Zealand, 2001–2013. Annals of the Association of American Geographers, 105 (5), pp. 951 – 967.

Piekut, A., Pryce, G. & van Gent, W. (2019). Segregation in the Twenty First Century: Processes, Complexities and Future Directions. *Environment and Planning B*, 110 (3), pp. 225–380.

Theil, H. (1972). *Statistical decomposition analysis*. North-Holland Publishing Company, Amsterdam.

Thiel, H. and Finezza, A. J. (1971). A note on the measurement of racial integration of schools by means of informational concepts. *Journal of Mathematical Sociology* 1 (2) pp. 187 – 94.

[3] Prior to fitting the model, a Box-Cox transformation is applied to the entropy scores, E_{ijk}, yielding $f(E_{ijk})$ in the equation above.