Geodemographic Output Area Classifications for London, 2001-2011

Gale, C.G., Longley, P.A.

University College London, Department of Geography, Gower Street, London, WC1E 6BT.
Tel: +44 (0)20 7679 0510 Fax: +44 (0)20 7679 0565
Email: c.gale.10@ucl.ac.uk, p.longley@ucl.ac.uk

ABSTRACT

London is a dynamic global city, yet the most recent 'open' public geodemographic classification is Vickers and Rees' Output Area Classification of 2001 Census data. In London the population allocation between the classification's seven Supergroups is uneven: 77% of areas are assigned to just two groups. This paper describes preparations for the 2011 Output Area Classification, that will address this problem, and which may accommodate temporal updating through use of government open data. The methodology will also allow regional and bespoke geodemographic classifications to be created, based on the same methodological concepts as the national classification.

KEYWORDS: Geodemographics, Census, Open Data, OAC, London

1. Introduction

It is intended that the Office for National Statistics (ONS) sponsored 2011 Output Area Classification (2011 OAC) will provide a new open-source geodemographic classification of the United Kingdom. Defined as "small area classifications that provide summary indicators of the social, economic and demographic characteristics of neighbourhoods" (Adnan *et. al.* 2010) geodemographic classifications have been widely used both in the commercial and public sectors for strategic resource planning and allocation, along with tactical marketing (Shelton *et. al.* 2006). The data used to create geodemographic classifications are traditionally derived from national Census of the Population data; although they can also be obtained from a range of other sources, such as commercial surveys (Experian 2010) or freely available public data. The 2011 OAC methodology is currently under construction and will be applied to small area data from the 2011 Census of Population in the United Kingdom when they become available in late-2012 (ONS 2011a). There is also the possibility that the classification will be augmented with open data sources.

2. The 2001 Output Area Classification

Creating the 2011 OAC requires an understanding of how the current classification using, at present, the most up-to-date Census data from 2001 – the 2001 Output Area Classification (2001 OAC) – functions. The 2001 OAC (Vickers and Rees 2007) assigns each of the Output Areas (OAs) of United Kingdom to one of seven Supergroups, 21 Groups and 52 Subgroups in a three tiered hierarchical classification (Vickers *et. al.* 2005). Each group in each tier has a unique socioeconomic composition; the names and descriptions of each group reflect this. The specification and estimation of the 2001 OAC is of fundamental importance when accessing how any new classification might be devised and implemented. London is a global city (Sassen 2001) which has a different make-up to anywhere else in the United Kingdom (CACI 2009). It provides a good starting point to investigate a range of classification issues, and identify problems that might be addressed using a new classification methodology. London's economic, political, cultural and infrastructural characteristics set it apart from the rest of the United Kingdom and to a large extent the rest of the world. Petersen *et. al.* (2010)

described London as having a "special settlement status" within the United Kingdom. Despite this well documented uniqueness, London is brought together with the rest of the United Kingdom in a multitude of geodemographic classifications. The data used to create the 2001 OAC is over a decade old, and it is important that the 2011 OAC methodology not only provides a newer classification of the United Kingdom, but also accommodates the changes in social, economic and demographic structure that have characterised the most recent inter-censal period.

3. London and the 2001 Output Area Classification

Examining the assignments of the seven Supergroups within London gives a good indicator as to how useful and representative the deployment of a United Kingdom national classification in London can be. The United Kingdom has 223,060 OAs covering every square mile of the total land mass. London accounts for 24,140 of these, 10.8% of the total. Table 1 shows the assignments of the seven Supergroups of OAC across the United Kingdom (with and without London) and London.

Supergroup Name/ Geographical Boundary	United Kingdom	United Kingdom without London	London
Blue Collar Communities	16.1% (35837)	17.7% (35231)	2.5% (606)
City Living	7.5% (16637)	5.8% (11463)	21.4% (5174)
Countryside	12.4% (27681)	13.9% (27660)	0.1% (21)
Prospering Suburbs	21.2% (47250)	22.9% (45468)	7.4% (1782)
Constrained by Circumstances	14.9% (33165)	16.4% (32573)	2.5% (592)
Typical Traits	18.3% (40769)	19.3% (38339)	10.1% (2430)
Multicultural	9.7% (21721)	4.1% (8186)	56.1% (13535)

Table 1: OAC Supergroup percentages (counts in brackets)

Table 1 identifies the high proportion of OAs in London classed as Multicultural, over 56% of the total. This Supergroup is so dominant in London, that only 37.7% (8186) of the total OAs designated Multicultural are found elsewhere in the United Kingdom. Figure 1 displays the uneven distributions in cluster assignment in London when compared to the United Kingdom, with and without London, data sets. At a national level (either with or without London included) the distribution of the cluster assignments is more uniform, resulting in the smoother lines seen in Figure 1. This is an indication of the potential inadequacies of a national classification such as the 2001 OAC at representing the diversity of neighbourhood circumstances that characterise London.

This problem is caused by the diverse characteristics of London being accommodated within a single national classification; in important respects the United Kingdom is set apart from the prevailing characteristics of its capital city. The numerical size of the Multicultural Supergroup attests to this fact, while the cluster profiles of Vickers *et. al.* (2005) identifies the average numbers of individuals described as 'Indian, Pakistani and Bangladeshi', 'Black' and 'Born Outside UK' as key to defining this Supergroup. In a sense we can think of London as handcuffed to the rest of the United Kingdom in the quest to devise a 'one size fits all' classification, a decision that has hindered the effectiveness of the 2001 OAC. Undoubtedly the 2011 OAC will need to be a national classification that includes London within it, but it is desirable to address the dominance of the multicultural category seen in
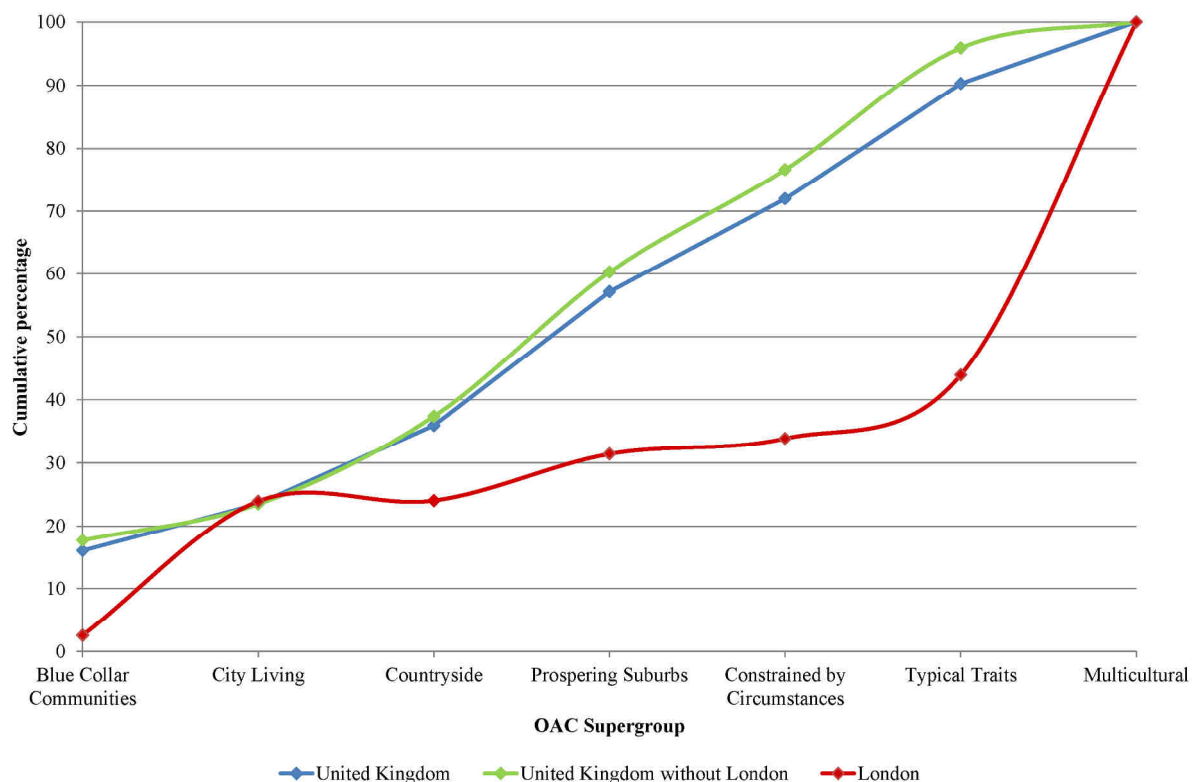
Figure 1: OAC Supergroup cumulative percentages for the United Kingdom
(with and without London) and London

London with the 2001 OAC. The London Output Area Classification (LOAC) created by Peterson *et. al.* (2010) identifies one logical response, to create a regional classification to co-exist alongside a national classification. LOAC, which combines a London only dataset with the 2001 OAC methodology, promotes the concept of creating both national and regional classifications using an identical methodological approach. This suggests that the methodology used to create the 2011 OAC should therefore be able to accommodate regional classifications from the start. In essence the creation of a national 2011 OAC, similar to the 2001 OAC, is just the beginning of the process rather than the end as regional and other bespoke classifications can then follow.

4. Change since the 2001 Census

Indicators of the changes that have taken place across the United Kingdom since 2001 can be devised in order to evaluate the changes that might be desirable in the 2011 OAC. Some (Sleight 2004) have suggested this change does not matter as certain areas will always be dominated by certain types of people and as people move out similar people move in. Longley *et. al.* (2011), by creating a regional geography of Britain using surnames, agree partially with Sleight's assessment. They found that a large proportion of the British population has remained settled for at least the past 600 years with the possible exception of urban conurbations, such as London. Since 2001 London has undergone large population changes, increasing in population size by 9.1 percent from 2001 to 2010 (ONS 2011b). Figure 2 – using unpublished mid-2010 OA population estimates – displays the change in population per OA between 2001 and 2010, with each OA having its size modified to represent the population count for each in 2010 to give a better indication on how densely populated inner London is compared to outer London. Clearly change has happened to London's population structure, not only in the increase in the total population but how many people live in each OA.

Output Area Percentage Change

	-100.00 - -26.43
	-26.42 - -5.79
	-5.78 - 0.00
	0.01 - 23.43
	23.44 - 46.72
	46.73 - 86.46
	86.47 - 158.26
	158.27 - 294.44
	294.45 - 611.30
	611.31 - 1228.86

Figure 2: London Population Change 2001 to 2010 Cartogram

It is reasonable to assume that the dramatic changes in the population sizes of London neighbourhoods have been accompanied by changes in the other characteristics captured by the 2001 OAC. Given the magnitude of local changes in size, it is also reasonable to anticipate changes in population composition as well, and to envisage that this will manifest in a geodemographic classification. Geodemographic classification is an inherently dynamic procedure, as acknowledged in the changing group compositions over time of commercial classifications. In practice the nature of these changes may require a regional rather than a national classification to accommodate these changes, and this is one consideration to bear in mind when re-engineering the 2001 OAC for 2011 Census data.

5. Conclusion

The results of a consultation exercise carried out in late 2011 confirm that there is a clear need for the 2011 OAC. This need comes from the desire to have an updated version – albeit using a different methodology – of the 2001 OAC to cater for the observed changes seen in the population during the last decade. An additional consideration is the facility to create regional and bespoke classifications, particularly in the light of evidence of London's very distinctive geodemographic structure. The view of Petersen *et. al.* (2010) that London has a "special settlement status" within the United Kingdom led to the release of LOAC using the same Vickers and Rees methodology, although it can be argued that this was not the best strategy for designing an open source geodemographic classification for London. These issues can be addressed more strategically if the 2011 OAC has a clear release strategy, and a flexible open reproducible methodology that allows for the creation of regional classifications. These

issues need to be addressed at the planning stage, as do decisions regarding all potential outputs, as this will influence the design of the classification and therefore the overall effectiveness of the 2011 OAC as a tool. This will allow the 2011 OAC to build on the success of the 2001 OAC while accommodating the changes in regional and national population structure that have taken place since 2001.

6. References

Adnan, M., Longley, P., Singleton, A. and Brunson, C. (2010) Towards Real-Time Geodemographics: Clustering Algorithm Performance for Large Multidimensional Spatial Databases. Transactions in GIS, 14(3), 283-297.

CACI (2009) Proud to be different – London found to be nothing like rest of UK. http://www.caci.co.uk/395.aspx (Accessed 20th May 2011)

Experian (2010) Optimise the value of your customers and locations, now and in the future. http://www.experian.co.uk/assets/business-strategies/brochures/mosaic-uk-2009-brochure-jun10.pdf (Accessed 20th November 2011)

Longley, P.A., Cheshire, J.A. and Mateos, P. (2011) Creating a regional geography of Britain through the spatial analysis of surnames. Geoforum 42, 506-516.

ONS (2011a) 2011 Census Outputs. http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/producing-and-delivering-data/2011-census-outputs/index.html (Accessed 20th November 2011)

ONS (2011b) Methodology Note on production of Super Output Area Population Estimates. http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/pop-ests/methodology-note-on-production-of-super-output-area-population-estimates.pdf (Accessed 20th November 2011)

Petersen, J., Gibin, M., Longley, P., Mateos, P., Atkinson, P. and Ashby, D. (2010) Geodemographics as a tool for targeting neighbourhoods in public health campaigns. J Geogr Syst 13(2), 173-192.

Sassen, S. (2001) The Global City: New York, London, Tokyo. Woodstock: Princeton University Press

Shelton, N, Birkin, M. and Dorking, D. (2006) Where not to live: a geo-demographic classification of mortality for England and Wales, 1981 – 2000, Health and Place. 12(4).

Sleight, P. (2004) Targeting customers. How to use geodemographics and lifestyle data in your business. WARC, Henley-on-Thames.

Vickers, D., Rees, P. and Birkin, M. (2005) Creating the National Classification of Census Output Areas: data, methods and results. Working Paper 05/2. School of Geography, University of Leeds, Leeds. (Available from: http://www.geog.leeds.ac.uk/wpapers/05-2.pdf)

Vickers, D. and Rees, P. (2007) Creating the UK National Statistics 2001 output areas classification. J R Stat Soc 170(2), 379-403.

7. Biography

Chris Gale *is a second year PhD student at University College London, funded by studentships from UCL and the Office for National Statistics. He is working towards creating better area classifications for the 2011 Census, specifically a 2011 version of the Output Area Classification, with particular focus on new modes of dissemination that better utilise web technologies and new advances in GIS and geodemographics.*

Professor Paul Longley *is Professor of Geographic Information Science at the Department of Geography, University College London.*