

Patterns of spatial dependence and heterogeneity and the scale of urban deprivation

Carolina Tobón[†] and Paul Longley

Department of Geography and Centre for Advanced Spatial Analysis (CASA)
University College London (UCL)

Abstract

The scale and pace of change in urban systems is without historic precedent. Today's increasingly affluent populations have ever more diverse lifestyles, and it is increasingly untenable to think of intra urban social patterning by analogy to an inert mosaic of internally homogeneous statistical reporting zones (Johnston, 1999). Small area spatial differentiation in physical and social conditions thus remains an important and developing focus for policy concern: most cities are restricted in spatial extent by planning policy, and so experience complex processes of change in neighbourhood composition over even quite short time periods (Harris, 1999). Appropriate allocation of public resources within and between urban areas requires that social conditions be represented in ways that are open to scrutiny (Gordon and Pantazis, 1997), while there is increasing realisation that 'if government has no settled and adequate measure of poverty, then it cannot reliably assess how its policies are contributing to reducing poverty' (White, 2002). Such measures need to be generalisable, transparent, pertinent, up-to-date and safe to use (Campbell, 1999). Taken together, one of the greatest challenges to policy-relevant urban geography is to keep pace with the fission of many urban lifestyles and effectively measure and monitor pertinent social conditions at fine spatial scales.

In recent years our ability to measure and monitor the morphology and extent of urban areas has improved considerably. Yet there has been no commensurate improvement in our ability to measure or infer urban social conditions. The barriers to creation of pertinent, robust and defensible small area measures are well known: confidentiality strictures dictate the maximum resolution at which census variables may be differentiated (Rees and Martin, 2002); population censuses are carried out only every ten years (UK researchers will still be using 1991 Census data well into 2003); resource constraints on public sector surveys limit the sampling interval and hence increase zone size for sample estimates (e.g. Dale and Teague, 2002); non-response is an increasing problem in most national settings (see the discussion of the 'missing millions' in the 1991 Census: Champion *et al.*, 1996); and the remit of many public sector surveys is restricted to what are deemed acceptable intrusions into private lives (and in the UK this precludes any income question in population censuses). The consequence for quantitative urban geography is representation of cities as crude mosaics of rough-hewn tiles, coloured according to very imperfect surrogate attribute measures, coincident with reporting zones that may be arbitrary, and remaining inert for the duration of successive decennial census periods.

[†] Corresponding Author. Phone: (+44) 20 7679 4260, Fax: (+44) 7813 2843, Email: c.tobon@ucl.ac.uk

We address some of these issues in the context of the debate about the intra-urban geography of hardship and social exclusion. Low income fundamentally restricts the abilities of people to participate actively in society (Harris and Longley, 2002), yet reliable, up-to-date income measures at fine spatial scales are rarely available from conventional sources. As a consequence, Lee (1999) reflects that many indicators of deprivation are reliant upon data sources that are out of date and/or entail use of crude surrogate measures. Some measures bear little clear correspondence with hardship at all – for example, even following re-engineering of socio-economic classifications (Rose and Pevalin, 2002), diversity in terms and conditions of employment make occupational classifications an increasingly opaque indicator of household circumstances. Other widely-used indicators are spatially variable in their operation. For example: low car ownership is a poor bellwether of socio-economic conditions in transit rich metropolitan centres; lacking freezers has different connotations in densely-populated urban areas where residents are almost invariably close to a shop (which will often stay open late); and socio-economic descriptors of occupation can have different connotations in different settings (e.g. the conditions of employment and remuneration of personal assistants, accountants and solicitors in the English provinces compared to their City of London counterparts). One response that has been better received (Gordon *et al.*, 2000) is to use a ‘democratic majority’ to identify the goods and services that, by common consent, define social exclusion; to then infer the relationship between these goods and services and census variables in a nationally representative sample survey; and to then infer the detailed geography of hardship through the geography of the census variables. This is essentially a return to the ‘budget standards’ approaches pioneered in Victorian Britain (White, 2002) and represents progress of sorts (Senior *et al.*, 2000), although it also entails problematic ecological inferences and insensitivity to spatial setting. Specifically, the indicators deemed to define deprivation are not allowed to vary spatially, even though the goal of subsequent linkage to the UK Census through surrogate indicators is identification of spatial variation. Thus the specification of deprivation is aspatial, and it is only when parameter estimates are used predictively that spatial variation is assumed to occur. Can a model of deprivation that is mis-specified in this way provide a reliable indication of geographic variation in the phenomenon of interest? And can ‘top down’ inferences from a national survey to local scales be used reliably to infer characteristics of small areas?

The broader issue concerns the scale and extent of ‘pockets’ of hardship and the scale ranges at which difference is deemed manifests. The problems are compounded if each of the range of surrogate measures used to specify a concept operates at different scales. Taken together, it remains unclear whether meaningful indicators of social conditions can ever be adequately specified, or whether generalised representations can be sufficiently sensitive to place. For example, at what scale range does the impact of predictor variables hold constant, and what are the possible distorting effects on the scale of hardship?

In many countries (e.g. the UK) no small area income measures are collected at all, and this forces reliance upon commercial sources. Yet, the use of such data in academic research is not without considerable problems. In the same spirit as Gordon and Pantazis (1995) we thus think it necessary to retain some linkage to population census data – but in a way which is much more sensitive to spatial context. A critical issue is thus to understand the scales at which both income, and the variables that are

used to predict it, vary (see also Rees, 1998; Harris and Longley, 2002). In our empirical analysis we examine a number of the facets to this issue through a detailed empirical case study, which begins to assess the contribution of new data sources to building improved representations of income distributions. We take a lifestyles survey that enables us to create small area estimates of household incomes for a study of the City of Bristol, UK. We use this analysis to begin to identify the relative priorities of better data, better specification of spatial dependence to draw inferences across space, and better understanding of local heterogeneity. The case study is also used to evaluate, in an inductive way, the scale at which it seems meaningful to draw inferences about hardship in our study area.

The ongoing revolution in the provision and handling of socio-economic information is improving the supply of geographic data, yet a challenge to researchers is to better understand the sources and operation of biases in data collection. Hence, we use analysis techniques that are sensitive to context to begin to assess the inherent differences and spatial mismatches between conventional indicators of income from the UK Census and direct income measures from a lifestyles survey. We believe that extending the interests of urban geographers from census analysis towards work with direct, timely, spatially disaggregate indicators is key to developing the data foundations to a new, data rich and relevant urban geography. The issues of using lifestyles data are fraught with problems, not least because some users of such surveys have little interest in the vagaries of spatial heterogeneity and non-stationarity. However, just as retailers use such data to supplement conventional geodemographic indicators, there is a need to tease out the relationship between these new pertinent measures and conventional indicators.

The problems of inference in urban geography are not just of technique, but of generalised, timely, spatially disaggregate data. Our study has begun to address the advantages of using new sources of digital data for estimating models that allow statistical inference and a better understanding of their processes in space. The range of spatial models permitted the identification of spatial patterns of interest beyond averaged relations between dependent and explanatory variables. It may appear as intuitive to argue that variables can better explain the variability in the level of income in some EDs rather than others, or in statistical terms, that the significance of their relation to the dependant variable can vary across space. Yet this confirmation of local heterogeneity and spatial dependence can also be interpreted as providing evidence of local misspecification of multifaceted concepts such as hardship. Use of a range of techniques points both towards interesting local spatial patterns and at problems in the specification of the models, both of which encourage iterative refinement and better understanding of local heterogeneity. This also underscores the importance of including spatial effects when they exist as a means to account for local variability, and the complexity of specifying its manifestations at finer scales.

References

Campbell H. (1999). Institutional consequences of the use of GIS. In *Geographical Information Systems: Principles, Techniques, Management and Applications* (2nd Edn.) Eds P A Longley, M F Goodchild, D J Maguire, D W Rhind. New York: Wiley: 621-31

Champion, T., C. Wong, A. Rooke, D. Dorling, M. Coombes and C. Brunson (1996). *The Population of Britain in the 1990s: a Social and Economic Atlas*. Oxford: Oxford University Press.

Dale, A. and A. Teague (2002). Microdata from the census: samples of anonymised records. In *The Census Data System*. Eds. P. Rees, D. Martin, P. Williamson. Chichester: Wiley: 203-12

Gordon D. and C. Pantazis (1995). *Breadline Britain in the 1990s*. York: Joseph Rowntree Foundation.

Harris, R. J. (1999). Geodemographics and the Analysis of Urban Lifestyles. School of Geographical Sciences: University of Bristol, 383pp. (Unpublished).

Harris R. J. and P. A. Longley (2002). Creating Small Area Measures of Urban Deprivation. *Environment and Planning A* 34: 1073-93

Johnston R. J. (1999). Geography and GIS. In *Geographical Information Systems: Principles, Techniques, Management and Applications* (2nd Edn.) Eds P A Longley, M F Goodchild, D J Maguire, D W Rhind. New York: Wiley: 39-47

Lee P. (1999). Where Are the Deprived? Measuring Deprivation in Cities and Regions. In *Statistics In Society: the Arithmetic Of Politics* Eds D Dorling, S Simpson London: Arnold: 172-180

Rees, P. and D. Martin (2002). The debate about census geography. In *The Census Data System*. Eds. P. Rees, D. Martin, P. Williamson. Chichester: Wiley: 27-36

Rose, D. and D. Pevalin (2002). A Researcher's Guide to the National Statistics Socio-Economic Classification. London: Sage

White, M. (2002). Clearer poverty definition 'vital'. *The Guardian* 27 August: 2.