Conference theme: International Migration and Migrant Populations

Paper Title (provisional): The World in Names: New Geographies of Ethnic Minorities Revealed through Family Name and Personal Name Analysis.

Authors:
Mateos, Pablo
Webber, Richard
Longley, Paul

Department of Geography
University College London
CASA, 1-19 Torrington Place
London, WC1E 7HB, England
 www.casa.ucl.ac.uk
p.mateos@ucl.ac.uk
Tel +44 (0)20 75306300

Abstract:

Understanding of the nature and detailed composition of ethnic groups remains key to a vast swathe of social science, to medicine and to our understanding of human biology. In practice, however, much research remains hamstrung by the quality and availability of ethnicity classifications, and consequent shortcomings in our ability meaningfully to subdivide populations. Ethnicity information is not available in most large scale public sector data series, and such data as are routinely available are based upon diverse, vague or over-generalised classifications. Common consequences are a lack of consistency between different studies, incompleteness of coverage, and impediments to longitudinal analysis. More fundamentally still, the various socioeconomic, demographic, and cultural correlates of ethnicity cannot be identified accurately.

In the UK, these problems are very evident with respect to geographically disaggregate Census of Population data. Census Output Areas are equivalent to approximately 150 people (Output Area) for which data are available for 16 ethnic groups, subject to a minimum disclosure threshold of 3 persons per Output Area per ethnic group. In practice, this threshold reduces the typical number of groups available in most areas to a maximum of six. A further (obvious) constraint is that Census data are usually collected only every ten years, making it impossible to chart the rapidly changing ethnic geography of many parts of the UK.

An alternative methodology has been developed by the authors to ascribe population ethnicity using family name and personal name analysis, at very fine geographical levels (typically unit postcode, accounting for c. 16 individuals) and using a very detailed typology of 100 ethnic groups. This classification can be continuously updated using a basic name and address register. This paper presents the results of its successful application in London.

The method uses a tool to assign an individual's personal name and family name to one of 100 Cultural Ethnic and Linguistic groups (CEL), which are weighted according to name scores, in order to assign the most probable CEL allocated to each individual. This tool has been applied to the UK Electoral Roll as well as to several health registers in London. The accuracy of the method has been evaluated using separate datasets where the self-reported ethnicity of individuals has been recorded.

The outcome of the research is an improved methodology for classifying population registers, as well as small areas (typically postcode units), into cultural, ethnic and linguistic

groups, that makes possible the creation of much more detailed, frequently updated, representations of the ethnic kaleidoscope of UK cities. The results of the pilot study are used to illustrate the diverse ethnic geography of London in 2004, and the detailed nature of spatial segregation operating at neighbourhood scales. These have hitherto been hidden by the coarser spatial and temporal granularity of Census data.