Integrating Public Domain Data to Construct Community Profiles

*by Ken Reed, Betsy Blunsdon,
Nicola McNeil
& Steven McEachern **

Introduction

In recent times there has been an increase in the availability of and requirement for public domain data that can be used to construct community level profiles in order to study contextual level phenomena. In this paper, we will review the developments that have given rise to this requirement, and explain how we have constructed a database of community characteristics at a postcode level. We shall then review applications of this type of data, and the opportunities for both aggregation and extension.

The demand for community level data has grown following developments in academic research and also in policy related work. Academics in many fields, including the government, are now examining aspects of the social world at multiple levels and therefore need accurate data at these levels (Sampson 1988; Griffen 1997; Raudenbush and Sampson, 1999; Sampson and Raudenbush 1999).

There has also been an increased demand for community level data within local government areas in order to better understand and address local issues, and to be used for evidence based planning. The approach we will outline provides an opportunity to do this, particularly when integrating a variety of data sources and utilising multiple methods of data collection.

The need to measure 'context'

Researchers in many fields are now studying social phenomena at multiple levels of analysis. It has been previously recognised that to understand the social world in detail requires a multi-level approach, yet until recently this has not been possible. New theoretical and empirical developments in the area of multi-level modelling make it possible to measure 'contexts' (such as 'environment', 'organisation', 'community' or 'neighbourhood') in addition to 'individuals'. Hierarchical or 'multi-level' modelling techniques estimate the effects of variables at different hierarchical levels. These techniques were developed primarily in educational research, to estimate separate parameters for the effects of the school, teacher and pupil on a student's academic achievement (Goldstein 1995; Singer 1997). The technique of multi-level modelling involves estimating higher-level effects from the data, then using these higher-level estimates as constants in the analysis of units at lower levels of analysis. This enables effects at lower levels to be determined while taking into account the tendency of (lower-level) individuals in similar (higher-level) contexts to exhibit similar responses (Goldstein 1995; Bryk and Raudenbush 1992).

These recent developments in multi-level research have seen major advances in many fields, including Sociology, Psychology, Social Work, Organisational Theory and Criminology. For example, understanding aspects of social capital, such as the tendency to volunteer, involves understanding both aspects of individuals and the characteristics of their communities. Communities can differ in terms of density of social relationships, infrastructure characteristics (such as number of business premises versus residential, parkland, number of schools), homogeneity, heritage, community safety and structural aspects (such as age and gender distribution and household composition). Similarly, individuals may differ in terms of demographics, personality, attitudes, values, behaviours or motivation.

It is not surprising given these developments that there has been growth in the number of multi-level and ecological studies in a broad range of fields. Two examples of such research designs are the Project on Human Development in Chicago Neighborhoods and the Los Angeles Family and Neighborhood Survey (LA FANS). Both projects attempt to disentangle the intricate effects of individual, family and community factors on the human development of children and juveniles, by undertaking a longitudinal investigation of individuals, communities and individuals within their community.

The aim of the Los Angeles Family and Neighborhood Survey[1] is twofold – firstly to study the impact of neighbourhood, family and peers on children's development, and secondly to study how and why families choose to live in certain neighbourhoods. The focus of this study is to provide in-depth understanding of these relationships to inform policy development, particularly welfare policy.

The Chicago Neighborhoods Project[2] aims to reveal the causes and implications of deviancy amongst residents of several Chicago neighbourhoods. Over an 8-year period, researchers will gather official data about neighbourhood characteristics, including the economic, social, political, business and cultural structures within the neighbourhood. Information will also be collected through a series of co-ordinated longitudinal studies that trace 7,000 randomly selected children through to young adulthood, looking at personal characteristics, changing circumstances and the exhibition of delinquent behaviours.

In Australia, there has been no large scale, systematic study that attempts to describe the characteristics of geographic regions in this manner. The main difficulty for Australian researchers attempting this type of undertaking is that there was no consensus amongst government agencies about how the geographic element of this data would be collected. For example, the Business Register identifies the geographic location of businesses by the postcode they are located in, while Federal and State Government election results are recorded by electoral district and crime statistics are recorded by local government area. In an attempt to address this inconsistency, the Australian Bureau of Statistics has recently revised a classification scheme that will enable the production of statistics that can be integrated. Statistical units, such as households and businesses, are assigned to a geographical area, with the data collected from these statistical units then compiled into pre-defined geographic aggregations that, subject to confidentiality restrictions, are then available for publication.

With multi-level research comes the need to measure characteristics of communities and neighbourhoods with much more precision than has been required in the past. This provides the opportunity to integrate a higher number of community and neighbourhood indicators into comprehensive data sets of characteristics.

Sources of data at different levels
The data we have sourced and integrated has been collected at the postcode level. That is, the data can be treated as roughly representative of geographical communities. The postcode has both advantages and disadvantages as a unit of analysis. On the one hand it is an administrative boundary that may not exactly reflect the social ecology of the residence; on the other hand, it offers an advantage in that there are many different types of data available at the postcode level.

The type of data available includes: demographic data (in Australia this is available for 1996 and 2001, both years in which there was a census), data on the incidence and perceptions of crime, the number and locations of licensed premises, the number and location of churches and religious institutions (as well as other cultural institutions), the number of schools, the number and type of recreational facilities, social services (such as law courts, legal centres, community centres) and the number and location of businesses, government offices and agencies. A detailed list of the type and source of data is available in Table 1

Aggregation and extensions
Community data collected at the postcode level can be aggregated into local council areas or regions and into state level areas. This type of dataset also provides the opportunity for extension through integration with other survey data and qualitative techniques. By constructing a sampling frame for household or individual surveys from electronic phone directories, there is also the opportunity to identify neighbourhoods within postcode levels (by street location, for example). Telephone survey data can also be supplemented at the level of the respondent through more detailed personal interviews, following the analysis of the survey data.

Contextual and situational data can be further extended through systematic social observation (SSO) and ethnographic fieldwork. Mapping the variation in community characteristics provides a rich information source for case selection in ethnographic fieldwork. Information about the characteristics of communities and neighbourhoods can also be integrated with the detailed descriptions of ethnographies and social observations in order to further understand complex social situations and interactions. An example of this type of extension is found in Sampson and Raudenbush (1999), a study of social disorder in Chicago. They combine social observation, census data, police records and a survey of residents to test a theory of structural constraints with local collective efficacy in order to understand the sources and consequences of social and physical disorder in urban areas.

Sourcing, collating and integrating community level data over time also provides enormous potential for identifying macro-level structural change. Our database enables us to aggregate the data of local government areas for 1996 and 2001. This enables the analyses of variation between these areas at each point in time, and reveals changes within council areas over the two periods. This has important policy applications for community capacity building, community regeneration and social capital.

Applications
It is widely accepted that an understanding of the social world requires theories at multiple levels to achieve complete explanation. While a 'complete' explanation is a *direction* for progress rather than an attainable goal, this type of data set will provide much information about macro-level contexts such as environment, community and neighbourhood. It also makes possible analyses of the landscapes in which many social issues arise. The following examples illustrate how this data might be

applied for multi-level research.

Our first example is a study of the propensity to volunteer in community-based fire fighting. The study will take into account the effects of the individual, life-situation and community on tendency to volunteer. People may choose to become volunteer fire fighters for a variety of personal reasons, including feeling part of a community and a sense of obligation, or for instrumental reasons, such as maintaining property security and developing skills that enhance employment prospects. Analyses at the individual level may capture such differences in perception, but not the effects associated with influences derived from living within different types of community. Community level data such as that described will give us insight into the relationship between types of community characteristics and the tendency to volunteer. This contributes to theoretical development and policy and planning issues involving volunteer activity and community development.

A second example is the study of the effects of social relationships and networks on personal outcomes, in this case health. Like the problems described thus far, a central issue in this investigation is distinguishing between the effects at the individual level from effects at the community level. For example, if we compared two communities, one characterised by dense social networks and the other by a high amount of individualism, we cannot determine whether it is living in that type of community or having many social relationships that produces individual outcomes, including better health. To this end, community level data provides the opportunity to study the interaction effects of individual and community characteristics on personal health.

In these cases we treated postcodes as roughly representing geographical communities. We employed a sample design capable of allowing statistical inferences that estimate the effects at both the postcode level and again at the individual level. As referred to previously this offers both advantages and disadvantages. The community level data that we have described is available at the postcode level while individual data can be aggregated from sources, such as the Census, to provide measures of affluence, demographic structure and residential stability.

Conclusion

Recent developments in multi-level theory and policy and planning requirements make it timely to collect and integrate data from a number of sources to better understand the context in which social action occurs. The data set that we describe is constructed by collecting data from a number of different sources, almost all of which is available at postcode level. This data can then be aggregated and extended to a higher level of analysis, such as the region or the state, or disaggregated into neighbourhoods or streets, if combined with further data collection methods. This

extended data set, or community profile, can then be used to identify and pinpoint the effects of social change across a range of social levels, from a state wide macro-level down to the micro-level of a particular street if need be.

These applications stem directly from the recent abundance of publicly available data and the development of effective ways in which the community profile may be used. Academic researchers, policy workers and local or state governments may immediately benefit from the adoption of this method of social inquiry. At present, the only limiting factor for this method is the potential incompatibility of the various data sets available. However, as seen by the recent actions of the Australian Bureau of Statistics, this is a problem that has been recognised and is being addressed.

In closing, it can only be expected that the kind of multi-level data collection and analysis that we have presented today, and that can be seen in the Los Angeles Family and Neighborhood Survey and the project on Human Development in Chicago Neighborhoods projects, will become an established and invaluable part of social research and planning.

References

Bryk, A. and Raudenbush, S. (1992) *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, California: Sage.

Goldstein, H. (1995) *Multilevel statistical models*. London: E Arnold, Halsted Press.

Griffen, M (1997) "Interaction between individuals and situations: Using HLM Procedures to Estimate Reciprocal Relationships." *Journal of Management* 23 (6):759-773.

Raudenbush, S. and Sampson, R. J. (1999) "'Ecometrics': Toward a Science of Assessing Ecological Settings, with Application to the Systematic Social Observation of Neighborhoods." *Sociological Methodology* 29:1-41.

Raudenbush, S. and Sampson, R. J. (1999) "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods (1)" The American Journal of Sociology, 105: 603-636.

Sampson, R. J. (1988) "Local Friendship Ties and Community Attachment in Mass Society: A Multi-Level Systemic Model." *American Sociological Review* 53:766-769.

Singer, J. D. (1997) "Using SAS PROC MIXED to fit multilevel models, hierarchical models and individual growth models.*" Journal of Educational and Behavioral Statistics* 23: 323-355.

Footnotes

Table 1: Data type and source

CATEGORY	DATA TYPE	DATA SOURCE
Demographics	Census	Australian Bureau of Statistics
Business	Victorian Business Registry	Australian Bureau of Statistics
Crime	Victorian Crime Statistics	Australian Bureau of Statistics
Education	State Schools – primary, secondary and technical	Victorian Government: Department of Education & Training website
	Special Developmental Schools	
	Special Schools	
	Language Schools	
	Special Education Unites	
	Private School	Victorian Independent Schools Directory
	Victorian University	Federal Government: Department of Education: Science & Technology website
	Victorian TAFEs	Street Directory Index
	Adult Migration Education	Yellow Pages
	Kindergartens	Yellow Pages
Deaths	Coroners Court data	Monash University
	Cemeteries	Street Directory Index
	Funeral Directors	Yellow Pages
Religious Centres	Religious Centres by Denomination	Street Directory Index
	Masonic Centres	Street Directory Index
Recreation	Reserves/parks & leisure centres	
	List of all athletics clubs	
	Bicycle tracks	
	BMX tracks	
	Boat Clubs	
	Boat Launching Sites	
	Bowling Clubs	
	Indoor Bowling Clubs	
	Cinemas	Yellow Pages
	Croquet Clubs	Street Directory Index
	Golf Clubs – Public & Private	Victorian Golf Association website
	Golf Driving Ranges	Street Directory Index
	Gun Clubs	Victorian Government: Department of Tourism, Sport & the Commonwealth Games
	Indoor Cricket	Street Directory Index
	Life Saving Clubs	Victorian Government: Department of Tourism, Sport & the Commonwealth Games
	Motor Cycle Tracks	Street Directory Index
	Night sports	Yellow Pages
	Racing tracks (horse and greyhound)	Street Directory Index
	Riding Schools and Clubs (horse)	
	Scout Parks	
	Public Squash Courts	
	Public Skating Rinks	
	Public Swimming Pools	Street Directory Index
	Public Tennis Courts	
	Liquor Licences	Victorian Government: Department of Human Services website
	Gaming Licences	Victorian Government: Department of Human Services website

Services	Ambulance Services	Street Directory Index
	Drug Advice	Yellow pages
	Law Courts	Victorian Government: Department of Human Services website
	Nursing Homes	Street Directory Index
	Police Stations	Street Directory Index
	Post Offices	Australia Post Website
	RSL Clubs	Returned Services League website
	Railway stations	Street Directory Index
	Designated refuge areas	Street Directory Index
	State Emergency Services	Victorian Government: Department of Infrastructure website
	Animal Hospitals	Street Directory Index
	Veterinarians	Yellow Pages
Cultural	Art Galleries	Yellow Pages
	Theatres	Yellow Pages
	Libraries	Street Directory Index
	Museums	Street Directory Index
Holiday	Caravan Parks	Street Directory Index
	Hotels and Motels	Street Directory Index
Government Services	Centrelink	Victorian Government: Department of Employment website
	Consulates	Street Directory Index
	State Emergency Services	Victorian Government: Department of Infrastructure website
Community Services	Community Corrections Centres	Street Directory Index
	Community Health Centres	Victorian Government: Department of Human Services website
	Halls	Street Directory Index
	Health and Community Services	Victorian Government: Department of Human Services website
	Hospice and Palliative Care	
	Hospitals – public	
	Hospitals – private	Street Directory Index
	Hostels – migrant	Street Directory Index
	Hostels – youth	
	Maternal and Child Health Centres	Street Directory Index
	Municipal Offices	Street Directory Index
	Neighbourhood Houses	Local Government websites
	Retirement Villages	Yellow Pages
	Child Care Centers	Yellow Pages