

## Creating socio-economic data for small areas: methods and outcomes

Canberra, 19 September 2008

- 9.15 am Registration and coffee
- 9.30 am Welcome and introduction  
 Chair, Professor Ann Harding  
 Director, National Centre for Social and Economic  
 Modelling (NATSEM), University of Canberra
- 9.40am Susan Linacre (Deputy Australian Statistician,  
 Australian Bureau of Statistics)  
*The importance of small area socio-economic data*

### Statistical approaches

- 10 – 11 am **John Glover** (Public Health Information Development  
 Unit, University of Adelaide)  
*Developing and reporting small area estimates to  
 inform policy and planning for prevention and health  
 promotion activities*
- Raymond Chambers** (Centre for Statistical and Survey  
 Methodology, University of Wollongong), Nikos  
 Tzavidis (University of Manchester) and Nicola Salvati  
 (University of Pisa)  
*Poverty mapping: integrating geographically weighted  
 regression with M-quantile modelling*
- 11 – 11.20 am Morning tea
- 11.20 am – 12.20 pm **Daniel Elazar**, Noel Hansen, Xin Lei, Janice Wooton  
 and Matthew Hardy (Australian Bureau of Statistics)  
*Small Area Estimation of Labour Force: A  
 Methodological Feasibility Study*  
 Pramod Adhikari (Australian Bureau of Statistics)  
*Local area estimates of literacy: comparing small area  
 estimation methods*
- 12.20 – 1.20 pm Lunch

## **Spatial microsimulation approaches**

- 1.20 – 2.20 pm                      **Mark Birkin**, Belinda Wu (School of Geography, University of Leeds), Haibo Chen (Institute for Transport Studies, University of Leeds) and Jinyi Guo (Northern Jiaotang University)  
*Traffic planning and the evaluation of transport sustainability with dynamic spatial microsimulation*
- Kimberly Edwards**, Graham Clarke (School of Geography, University of Leeds) and David Forman (Cancer Epidemiology Group, University of Leeds)  
*Validation of SimObesity: a spatial microsimulation model to synthesise small area estimates for obesogenic environments and obesity*
- 2.20 – 2.45pm                      Afternoon tea
- 2.45 – 3.45 pm                      **Robert Tanton**, Justine McNamara, Riyana Miranti and Ann Harding (NATSEM, University of Canberra)  
*Poverty at the local level: an application of SpatialMSM*
- Sharyn Lymer**, Laurie Brown and Ann Harding (NATSEM, University of Canberra)  
*Small area estimates of the need for care among the aged population*
- 3.45 – 4.15pm                      *Panel on: Future Opportunities, Collaborations and Directions: Where to From Here?*  
Professor Raymond Chambers (University of Wollongong), John Glover (University of Adelaide); Professor Robert Stimson (CR-SURF, University of Queensland), Dr Mark Birkin (University of Leeds); Dr Kimberly Edwards (University of Leeds); Robert Tanton (University of Canberra).
- 4.15 – 5 pm                      Drinks and collaborative canapés!

(Presenters are in **bold**. Other authors may not be at the workshop.)

**Workshop Venue:**

University of Canberra Innovation Centre  
Building 22, Level B  
Room 22B19  
University of Canberra  
ACT 2601

**Cost:**

\$396 per person (including GST).  
\$198 full time student (including GST).  
\$198 half day (including GST).

Cost to ARCRNSISS members: up to 10 members of ARCRNSISS will be entitled to complimentary workshop registration and the payment of their travel and accommodation costs.

**RSVP:**

Lucie Marshall

Email: [reception@natsem.canberra.edu.au](mailto:reception@natsem.canberra.edu.au)

Phone: 02 6201 2780

Registration form can be downloaded from: [www.natsem.canberra.edu.au](http://www.natsem.canberra.edu.au)

**Funding:**

NATSEM would like to acknowledge funding support from The Australian Research Council Research Network on Spatially Integrated Social Science (ARCRNSISS), NATSEM, and the Australian Research Council 'Opportunity and Disadvantage' Discovery grant (DP0664429).

Abstracts received to date:

***The importance of small area socio-economic data***

Susan Linacre (Deputy Australian Statistician, ABS)

Susan Linacre has a long-standing interest in small area socio-economic data, having been joint author of a 1976 paper on 'Techniques for the Estimation of Small Area Characteristics'. In recent years the ABS has increasingly focussed on this area of data requirements. This presentation will cover the drivers for small area socio-economic information and the issues created in satisfying the need, while also providing an overview of ABS activity in the area.

***Developing and reporting small area estimates to inform policy and planning for prevention and health promotion activities***

John Glover (Public Health Information Development Unit, University of Adelaide)

Data as to the extent of morbidity (illness or disease), disability and risk factors in the community have generally not been available at the small area level, apart from proxy measures such as hospital admissions, or incidence of cancers. Policy makers, service providers, health practitioners and others are becoming increasingly aware of the value of understanding geographic variations in the prevalence rates of chronic diseases and their determinants. The presentation of estimates of chronic disease and associated risk factors at the small area level, along with indicators of socioeconomic status, goes far towards meeting this need.

***Poverty mapping: integrating geographically weighted regression with M-quantile modelling***

Professor Raymond Chambers (Centre for Statistical and Survey Methodology, University of Wollongong), Nikos Tzavidis (University of Manchester) and Nicola Salvati (University of Pisa)

Over the last decade there has been growing demand for estimates of population characteristics at small area level. Unfortunately, cost constraints in the design of sample surveys lead to small sample sizes within these areas. Small area methods are designed to tackle this small sample size problem. The most popular class of models for small area estimation is linear mixed models, which include random area effects to account for between area variation. However, such models depend on strong distributional assumptions, require a formal specification of the random part of the model and do not easily allow for outlier robust inference. An alternative approach to small area estimation, based on the use of M-quantile models, has been proposed in Chambers and Tzavidis (Biometrika, 2006). Unlike traditional random effects models, M-quantile models do not depend on strong distributional assumption and automatically provide outlier robust inference.

The first half of this presentation details recent research that has used linear M-quantile models for deriving small area estimates of poverty and inequality. This methodology improves upon traditional poverty mapping methods in that it enables estimation of the distribution of the study variable within the small areas of interest.

An application to data from the 2002 Living Standards Measurement Survey (LSMS) in Albania will be discussed. A key advantage of M-quantile models is that they are easily generalised to non-linear situations. Consequently, in the second half of this talk I will present an extension of this approach that uses Geographically Weighted Regression (GWR) to specify the M-quantile model structure, thus allowing the approach to capitalise on spatially correlated data. This can be contrasted with the use of linear mixed models with Simultaneous Autoregressive Regression (SAR) random effects for this purpose. An application to predicting average levels of the Acid Neutralizing Capacity at 8-digit Hydrologic Unit Code level in the Northeast states of the U.S.A. will be presented and prospects for application of this method in poverty mapping will be discussed.

***Small Area Estimation of Labour Force: A Methodological Feasibility Study***

Daniel Elazar, Noel Hansen, Xin Lei, Janice Wooton and Matthew Hardy (Australian Bureau of Statistics)

This presentation will cover current work at the ABS on the feasibility of producing an annual series of SAEs of Labour Force that is of sufficient quality to meet users' needs. We will describe the models being used, the auxiliary data we have available, the methodology for estimating SAEs and their associated MSEs. We will also describe the methods for assessing the goodness of fit of the models and measures of quality of the resulting SAEs. A discussion of the methodology used will be presented, together with an assessment of how these diagnostics can be used to better inform users of the quality and fitness for purpose of the SAEs. We will then briefly cover areas for further possible work.

***Local area estimates of literacy: comparing small area estimation methods***

Pramod Adhikari (Australian Bureau of Statistics)

The Adult Literacy and Lifeskills Survey (ALLS) 2006 collected information from about 9,000 adults measuring different dimensions of literacy in Australia. Whilst ALLS has the potential to provide an evidence base for policy development, the relatively small sample size means that it is not possible to obtain reliable estimates for regional or local areas from the sample data alone. In order to ascertain the feasibility of developing spatially-based evidence to inform policy alternatives, the paper utilises a range of Small Area Estimation (SAE) methods to predict literacy rates for regional or local government areas. The SAE methods range from a simple method that uses regression modelling at an aggregate area level to a computationally intensive and statistically rigorous method. We will demonstrate results obtained from these two methods: (i) an area level synthetic model; and (ii) an empirical best linear unbiased predictor method using nested regression model. We will contrast the literacy levels at small area with another ABS area level disadvantage measure (namely Index of Relative Socio-economic Disadvantage) and show how literacy levels for an area can be used as yet another measure of area level disadvantage. The paper will provide examples of how small area estimates can be used to visualise gradients in literacy rates using mapping techniques and discuss the potential that SAE methods have for informing public policy.

***Traffic planning and the evaluation of transport sustainability with dynamic spatial microsimulation***

Mark Birkin, Belinda Wu (School of Geography, University of Leeds) Haibo Chen (Institute for Transport Studies, University of Leeds) and Jinyi Guo (Northern Jiaotang University)

The authors have made substantial progress with Moses, a dynamic spatial microsimulation model which aims to generate long-term demographic projections for small geographical areas. This paper explores the usefulness of Moses outputs for two applications in urban transportation. In the first case, we consider the development of a traffic planning model to evaluate the efficiency of the road network against a detailed social and demographic baseline from the Moses simulation. In the second scenario, transport sustainability is considered with regard to changing geodemographics of the Leeds population.

***Validation of SimObesity: a spatial microsimulation model to synthesise small area estimates for obesogenic environments and obesity***

Dr Kimberly Edwards, James Thomas (School of Medicine, University of Leeds)

This paper explores the notoriously difficult issue of validation of spatial microsimulation models. This is because normally it is not possible to directly validate the output variables themselves, as the reason they are being simulated is because these data do not otherwise exist. Thus often validation/calibration is undertaken by comparing the suitably aggregated simulated (individual level) data with the known data for the constraint variables, or for variables not used in the simulation (i.e. not input or output variables). This paper addresses some of these issues and suggests some novel alternatives.

Firstly, SimObesity, a deterministic reweighting spatial microsimulation model, was used to generate a synthetic population of individuals for Leeds with a multitude of attributes, including obesogenic environment and behavioural variables, from the 2001 Census, the Health Survey for England 2002, and the Food and Expenditure Survey 2005. Validation was undertaken by means of linear regression and equal variance t-tests using known and simulated data for the constraint variables aggregated to ward level. This confirmed the micro-level simulations were robust. A key aspect to this validation process for the output variables was that they were correlated with the constraint variables.

SimObesity is also being used to simulate obesity data for adults in Northern England, using the 2001 Census and the Health Survey for England 2006. The validation of this model will extend beyond the previous methodology. Data (including full postcode) on cancer incidence for eight cancers known to be associated with obesity (tumour sites: endometrium, colorectal, prostate, post-menopausal breast, kidney, gall-bladder, pancreas, oesophagus) has been obtained from the Northern and Yorkshire Cancer Registry. Accordingly we would expect to find 'hot spots' of cancer incidence in the same locations as 'hot spots' of obesity prevalence. This is a novel way to address the difficult issue of validating spatial microsimulation models.

***Poverty at the local level: an application of SpatialMSM***

Robert Tanton, Justine McNamara, Riyana Miranti and Ann Harding (NATSEM, University of Canberra)

The paper provides new estimates of poverty rates at the statistical local area level for all of Australia. The estimates are based on the latest version of NATSEM's spatialMSM model, which takes the unit record files from the 2002-03 and 2003-04 ABS Surveys of Income and Housing and reweights them to 2006 Census small area data. This technique creates a synthetic household database, with information about the characteristics of households living within each small area. This paper canvasses the techniques used to create the household database, the challenges faced when validating the output and issues encountered in the implementation of the methodology.

***Small area estimates of the need for care among the aged population using the CAREMOD model***

Sharyn Lymer, Laurie Brown and Ann Harding (NATSEM, University of Canberra)

To plan the effective delivery of aged care services to those still living in their own homes, governments and businesses require reliable small area estimates of the extent of disability, the availability of informal carers and the financial resources of older citizens. The spatial microsimulation model, CAREMOD, was developed to explore the issues of need for aged care at a small area level in New South Wales. It combines the 1998 ABS survey of Ageing, Disability and Carers and the 2001 Census small area data. In this paper, the model has been extended to provide projections of potential care needs in small areas of New South Wales in 2021. The projections consider the effects of population ageing and the changing age-sex structure of small areas by incorporating ABS SLA age-sex medium scenario projections for the next 20 years through a static ageing process of the microdata.

## About the presenters

### **Pramod Adhikari**

Pramod Adhikari is currently working at the National Centre for Education and Training Statistics, Australian Bureau of Statistics, Central Office, Canberra. Pramod has been with the ABS for nearly four years. Prior to ABS Pramod worked at the Department of Health and Ageing. Pramod has been with the Australian Public Service for nearly 15 years and has worked with AIHW, ATSI and DEETYA. He also spent some time at WESTAT, Maryland. Pramod has a PhD from University of New South Wales. His main areas of interest are survey sampling, socioeconomic indexes and small area estimates.

### **Mark Birkin**

Mark Birkin is a Senior Lecturer in the School of Geography, University of Leeds, UK. Previously he was Director of the Institute for Interdisciplinary Informatics at the University of Leeds (2001-2005) and Managing Director of GMAP Limited (1996-2001).

Mark has long-standing research interests in geographical information systems (GIS), spatial microsimulation, geodemographics, and 'what if?' predictive modelling. He is currently the Principal Investigator of Moses, a research node of the UK e-social science programme ([www.ncess.ac.uk](http://www.ncess.ac.uk) <<http://www.ncess.ac.uk>> ). Moses aims to construct a dynamic individual level simulation of UK cities and regions into the medium-term future. Mark is the co-editor of *Applied Spatial Analysis and Policy*, and a member of the editorial board of *Transactions in GIS*.

### **Ray Chambers**

Ray Chambers is Professor of Statistical Methodology at the Centre for Statistical and Survey Methodology, University of Wollongong. Prior to moving to Wollongong in March 2006, he was Leverhulme Professor of Social Statistics and Director of the Southampton Statistical Sciences Research Institute at the University of Southampton. He has worked in the Australian Bureau of Statistics, the Australian Bureau of Agricultural and Resource Economics and the Australian National University. He is a member of the International Statistical Institute and a Fellow of the American Statistical Association.

### **Kimberly Edwards**

Dr Kimberly Edwards (formerly Procter) is a lecturer in epidemiology at the University of Leeds (UK), based in the Cancer Epidemiology Group, Division of Epidemiology and Biostatistics, School of Medicine. Her research interests include: spatial epidemiology; geography of disease (particularly obesity and cancer); GIS and applied spatial modelling; small area estimation; spatial microsimulation modelling.

**Daniel Elazar**

Daniel Elazar has led the Small Area Estimation team in the Analytical Services Unit (ASU) of the ABS for the past five years. During this time the Small Area Estimation team has been responsible for producing a Small Area Estimation Practice Manual and work on the production of small area estimates (SAE) of disability. Currently the team is working on a feasibility study of SAEs of labour force status. Prior to moving to ASU, Mr Elazar worked in the Statistical Services Branch of the ABS on survey designs, sample selection and estimation and on methodologies for better measuring and improving the quality of survey estimates. Daniel Elazar has a degree in mathematics from UNSW and a Post-Graduate Diploma in Statistics from the ANU.

**John Glover**

John Glover is director of PHIDU (the Public Health Information Development Unit) at the University of Adelaide. His interest is in presenting data to describe and monitor social inequality and its impact on the health of population groups. The major work in this area has been the Social Health Atlases of Australia, which are unique in providing information about the social determinants of health and health outcomes at a local, state/territory and national level, presented in ways that are useful for policy development, as well as being accessible to consumers and other community advocates who may have limited skills in handling statistical information presented in more traditional ways. The PHIDU website [www.publichealth.gov.au](http://www.publichealth.gov.au) provides access to the atlases and other publications, as well as to interactive software to monitor inequality (currently limited to SA data: national data to be added in the coming months).

**Ann Harding**

Ann joined NATSEM as the inaugural Director in January 1993 and is Professor of Applied Economics and Social Policy at the University of Canberra. During the past 15 years Ann has led the development of highly sophisticated microsimulation models and databases, so that policy makers can gain much better information about the likely distributional impact of current and proposed policies. More recently Ann has steered microsimulation modelling in Australia beyond its traditional 'tax and social security' focus to such new areas as health, aged care, housing and regional issues, with the goal of extending sophisticated quantitative decision-support tools to policy makers in these areas.

Ann has published widely on income inequality, poverty, and the distributional impact of government programs, and is a prolific contributor to public policy debate in Australia, with her work typically being cited every week in the media. In 1996 she was elected a fellow of the Academy of the Social Sciences in Australia and in 2003 was elected President of the International Microsimulation Association. She serves on a range of government and academic boards and advisory committees. Prior to joining NATSEM Ann worked on policy and research within the Federal departments of Social Security, Treasury and Health, Housing and Community Services.

**Susan Linacre**

Susan is currently Deputy Australian Statistician, Social Statistics Group, having taken up leadership of the Population Statistics Group in 2002. Prior to this Susan spent a period in the UK, working as head of the Directorate of Methodology and IT in the Office for National Statistics. Susan spent the early part of her career in the ABS as a survey statistician working across both the economic and social fields.

Susan is an elected member of the International Statistical Institute (ISI) and a current member of the ISI Council. She has been elected President Elect of the International Association of Survey Statisticians (2007-2009). In her current position, Susan is responsible for the social statistics produced by the ABS. This includes a comprehensive program of household surveys and data collections covering areas such as health, family and community, housing, education and training, income, recreation, and crime and justice.

**Sharyn Lymer**

Sharyn joined NATSEM in 2004 from NSW Health. She has a strong background in statistics, particularly in the health arena. She is interested in health and health modelling especially inequality and access issues. At NATSEM Sharyn has been the lead modeller in developing the microsimulation model of the Australian Medicare system. She has also been heavily involved in small area microsimulation in the need for aged care in Australia. Currently, she is pursuing her PhD, looking at the projected costs of health care in an ageing Australia over the next 50 years, using dynamic microsimulation.

**Justine McNamara**

Justine joined NATSEM in June 2004, and over this period has presented and published work focused on regional differences in advantage and disadvantage across Australia, child care, social exclusion and poverty. She is currently working with the regional modelling team on a number of projects focused on social and economic analyses of poverty and exclusion from a spatial perspective, particularly focused on children. Other current projects include work focused on academic women's career development and the gender wage gap for young Australians. Justine's PhD related to long term poverty among older American women, and she has published in American journals on this and related topics.

**Robert Tanton**

Robert joined NATSEM in 2005, and has led the team working on poverty and income inequality. Robert also started leading the research on spatial microsimulation in 2006, and now leads the Social and Spatial Inclusion area at NATSEM, which conducts research into spatial poverty and disadvantage using spatial microsimulation techniques. Robert's research priorities are socio-economic indexes, spatial disadvantage and social inclusion, which covers housing, poverty and homelessness. Robert has a Masters in Statistics from the ANU, and worked in the ABS on small area estimates of crime and the 2001 SEIFA indexes before joining NATSEM.