Ernestina Coast
The household in surveys: a focus on fuzziness

Original citation:
© 2010 ESRC

This version available at: http://eprints.lse.ac.uk/36679/
Available in LSE Research Online: July 2011

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (http://eprints.lse.ac.uk) of the LSE Research Online website.

This document is the author’s final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

http://eprints.lse.ac.uk
Social scientists may use biomarkers not only as measures of outcomes, but also as predictors of behavior. We are beginning to understand how behavior is shaped by biology, including genetic and epigenetic factors. Twin studies suggest an important heritable component to many domains of social and economic behavior, from willingness to cooperate in laboratory “trust games”, to risk aversion and even religious and political attitudes. Biological differences interact with, and are mediated by, social and economic circumstances and institutions. It is the role of social scientists to study these processes.

An important part of this developing field is the inclusion of biomarkers in longitudinal studies of population ageing and other large panel surveys. The English Longitudinal Study of Ageing (ELSA) collects a wide range of biomarkers, which include anthropometrics, blood pressure, and lung function; glucose and c-reactive protein; and grip strength and walking speed. Understanding Society plans to collect a similarly broad range of biomarkers.

While biomarker data offer many opportunities for social scientists, there are also significant challenges to be overcome. A major challenge is to determine what to collect and how. The collection of biomarkers may be costly and intrusive, and it is exceedingly difficult to anticipate the research hypotheses we will want to test, even in the medium term. We must also work out protocols for the storage and release of data and biological material, and deal appropriately with the important ethical issues arising from the collection of such data. Despite these difficulties, the potential rewards are too large to ignore.

Thomas Crossley is Reader in Economics at the University of Cambridge and a Programme Director at the Institute for Fiscal Studies in London. In April 2010 the Centre for Microdata Methods and Practice at the IFS organized a workshop on Biomarkers in Social Science Research on behalf of the ESRC National Centre for Research Methods (NCRM).
Analysing the spatio-temporal distribution of crime

MADE (Multi-Agency Data Exchange) is a data warehouse tool for datasets that are relevant to crime and disorder in Lancashire. The MADE project was established in 2001 to help people living in Lancashire to make better-informed decisions about community safety issues in their neighbourhood and “to create a common collation and dissemination facility which will improve the speed and reliability of multi-agency information exchange for crime and disorder strategies and other multi-agency policies throughout Lancashire”. As a contribution to this, researchers from Lancaster-Warwick-Stirling node of NCRM have conducted a descriptive analysis of the spatio-temporal distribution of crime in Lancashire. This note summarises the statistical model and the results of the analysis.

The data used cover the period April 2003 to March 2009 inclusive. Information on each recorded crime consists of its location, time and type of crime, classified as criminal damage (51% of crimes), serious acquisitive crime (30%) and other wounding crime (19%). To preserve confidentiality, location is aggregated into lower super-output areas (LSOA’s). In a mixed urban-rural county such as Lancashire, LSOA’s are highly variable in geographical size to achieve relatively constant population size (average 1,500 per LSOA). The population size and relevant covariates for each LSOA were extracted from the 2001 national census. The time of reporting of each crime is recorded to the nearest minute, but the time of occurrence is not recorded. For this reason, and to eliminate time-of-day effects, our analysis uses a time-resolution of one day.

Our methodological approach is to derive the properties of the resulting spatio-temporal count data from an underlying spatially and temporally continuous point process model. To estimate the “normal” pattern of incident crimes we consider a multiplicative decomposition of the local (in space and time) intensity of crimes into three components. The first two correspond to spatially averaged temporal variation and temporally averaged spatial variation, and are modelled by log-linear regressions. The third component corresponds to residual spatio-temporal variation and is modelled as a latent stochastic process, whose covariance structure describes any spatio-temporal variation not captured by the available covariates.

Temporal variation of crime

The regression components of the model reveal strong and statistically significant day-of-week effects for crime-types: criminal damage and other wounding. For criminal damage the effects are highest for Mondays, followed by Sundays and Saturdays. On other week-days, crime-rates are between 20% and 30% lower than on Sundays. For other wounding, the difference between week-days and weekends is greater. For serious acquisitive crime, day-of-week effects and seasonal effects are smaller and only marginally statistically significant. All three crime types show seasonal variation in incidence and a decreasing trend over the study-period that we model as log-quadratic. Seasonal effects are strongest for criminal damage, whilst the decreasing trend is most pronounced for serious acquisitive crime.

Spatial variation of crime

The available spatial covariates are the density of licensed premises, the deprivation rates for income and employment, and domains of the index of multiple deprivation corresponding to health and disability, education skills and training, barriers to housing and services, and living environment. The effect of density of licensed premises is highly significant for all three types of crime, whereas the significance of the effects of the individual domains of deprivation are not consistent over the three categories of crime.

Crime rates are generally higher in and near urban areas. Rates in the other wounding category show a strong peak at each city centre. For the criminal damage category, the gradation in incidence from rural to urban areas is less strong, and for the serious acquisitive crime category even less so. For serious acquisitive crime, incidence also appears to be generally higher in the southern part of the county than in the northern part. Direct examination of animations of the data indicates that the urban areas
have consistently high rates, whereas departures from zero in the rural areas can be characterised more as sporadic events.

The pattern of crime rates varies considerably over the 14 districts in Lancashire. We consider the results relating to criminal damage in three representative districts: Blackpool, Preston and Lancaster. These districts vary in size, overall crime rates, decreasing trends and seasonal pattern. All three show a highly significant association between local crime rates and density of licensed premises.

Examining associations between deprivation effects and crime rates we conclude that health, barriers to housing and services, and living environment deprivation effects are significant for Blackpool; income, employment, education and health effects are significant for Preston; income, employment, health, and barriers to housing and services effects are significant for Lancaster.

**Spatio-temporal variation of crime**

The results relating to the stochastic model for residual spatio-temporal variation also suggest that the covariance functions for each of the three crime categories are separable, with marginal correlations following exponential decay with increasing separation in time and space. The residual spatio-temporal correlation is strongest for serious acquisitive crime and weakest for other wounding. We conjecture that this might be explainable by the spatial pattern of land-use types affecting the rates of acquisitive crimes, whereas the other wounding category includes crimes that are the result of violent interactions between individuals, irrespective of their location.

In future work, we intend to use the descriptive model described in this note as the basis for developing a real-time surveillance system to detect emergent clustering of criminal activity and inform decisions on short-term tactical responses to these emergent clusters.

For further details please see http://www.lancs.ac.uk/staff/diggle/MADE

---

**Lancashire MADE project**

MADE is short for Multi-Agency Data Exchange. The project team at the Lancashire County Council collect and process data on behalf of the agencies responsible for working towards a Lancashire that is a safe place to grow, live and work.

The 1998 Crime & Disorder Act recognised that by working together, local authorities, police, fire and health services could tackle crime and disorder more effectively and set up Crime and Disorder Reduction Partnerships (CDRP) in each district. In 2006 the Police and Justice Act required CDRPs to produce an annual strategic assessment and community safety plan. In addition, the agencies who work at a county level are also obliged to produce a community safety agreement, detailing where value is added by working together to support the CDRPs.

The local community safety plans and the county community safety agreement are used to direct resources from the partnership into the areas where they will be most effective. Data from across the partnership is needed to inform strategic assessments and monitor the progress of interventions. Each partner has its own computer systems and data collection techniques, therefore the Multi-Agency Data Exchange (MADE) project was set up to provide a central data collection and processing function. Partners can download spreadsheets containing ward profiles or query the online database.

There was also a demand for unrestricted access to some of the data collected by MADE. Therefore, a public facing Lancashire MADE Public website was developed to publish selected data from the MADE partners for the benefit of the general public. The aim is to ensure that people in Lancashire can make an informed decision about community safety issues in their neighbourhood and to provide a resource for community groups and neighbourhood watch schemes.

For more information about the MADE project please see their website http://www.saferlancashire.co.uk/statistics
School, family, neighbourhood: Which is the most important to a child’s education?

George Leckie, Rebecca Pillinger, Jennifer Jenkins and Jon Rasbash, LEMMA2 node, University of Bristol

Children grow up in complex social environments. The influences on a child’s development are many, and many layered. The school they go to, the family around them, and the neighbourhood they grow up in are often highlighted as particularly important for children’s educational achievements.

Our knowledge of the impact of schools, neighbourhoods and families comes from studies that look at each of these influences separately and therefore little is known about their relative importance. For example, school effectiveness studies attempt to measure the importance of schools on children’s academic progress, but nearly always ignore the role of family. Family research using siblings, on the other hand, attempts to quantify the importance of genetic and environmental family influences, but nearly always ignores the role of the wider shared environment of schools and neighbourhoods.

Knowledge of the relative effects of schools, neighbourhoods and families is needed to help inform decisions about the allocation of government resources to programmes and policies that will support children’s learning.

We follow a cohort of half a million school children through secondary schooling. We use cross-classified multilevel models to estimate the impact of the different influences of children’s complex social environments on academic progress during secondary schooling.

Our final model decomposes the variance in children’s progress into effects attributable to six influences: Local Education Authorities (LEAs), secondary schools, neighbourhoods, primary schools, families and children. This is an improvement on previous school effectiveness studies because it differentiates, for the first time, between the effects of the family and of the pupil; and it is an improvement on previous family research because it estimates the influence of the family-shared environment separately from those of the wider shared environments of schools and areas. Thus, we unite school effectiveness studies and family research under a single framework and therefore disentangle, for the first time, the various influences on children’s progress.

Our key finding is that families – in their dual roles of providing genes and parenting – account for 40% of the overall variance between children in their academic progress during secondary schooling; the wider shared environments of primary school (9%), secondary school (10%), neighbourhood (2%) and LEA (1%) account for 22%. The remaining 38% is attributed to the children themselves.

From a school effectiveness point of view it is interesting that the family effect is just as important as the child effect. Indeed, around half of what had been thought of in school effectiveness research as child variation is really due to family level factors.

However, this does not necessarily imply that interventions to improve child academic outcomes are best implemented at the family level. While primary and secondary schools appear less important than families, they still make reasonably substantial contributions; and it is at these levels that many educational policies will be most effective. It is much easier for governments to improve schools than to improve families.

From the point of view of family studies, the finding that 22% of the variance in progress during secondary school is attributable to the non-familial shared environment is both new and important. Indeed, studies that ignore the wider shared environment will misattribute this to families, making them appear more important than they actually are. It is perhaps encouraging that the neighbourhood effect is so small. It suggests that a bright child from a deprived neighbourhood can still do well if the schools and the parenting are good.

The largest variance component is the family effect – the genetic and environmental influences that parents provide. We were not able to distinguish between the “nature” and the “nurture” effects of the family because we did not observe which of the twins in our data were identical and which were not. Obviously this is a question that we would like to investigate. We see the next stage of this work as the integration of a genetically sensitive design with the cross-classified family, school and neighbourhood design on which this study was based.

References


This research is published in the Journal of the Royal Statistical Society, Series A. It is also published in the RSS Significance magazine and has been selected as the best JRSSA journal article in 2010. It will be presented in the “Best of the Society Journal Papers” session at the 2010 RSS conference.
The household in surveys: a focus on fuzziness

Ernestina Coast, The London School of Economics and Political Science

‘Household’ is a word that is simultaneously part of common language yet is also a technical term used in surveys. Household surveys are the mainstay of micro-level data for developing countries, providing data for more than half of the Millennium Development Goal indicators. In high income countries, in addition to census and routine statistics, micro-level data collection includes increasingly complex longitudinal household surveys.

Household surveys require a household to be identified, and the household is seen to be one of the building blocks of human societies, despite being a contested and problematic concept across disciplines, time and space. It is both a crisply defined concept in survey manuals and a fuzzy concept from the perspective of those people that live in households.

Our research used multiple methods (document review, key informant interviews, in-depth interviews) for a pilot case study of Tanzania. It established how the concept of the household is defined and used in household surveys in sub-Saharan Africa, and the implications of these definitions for analyses and policy-making.

Methodological implications

Survey professionals, who use the household as a technical concept, consider it as an analytic unit which just exists as a statistical artefact. The household has a crisp definition, is clearly bounded and has unambiguous membership criteria, reflected in the instructions given to interviewers in training. These clear boundaries are challenged during survey data collection in a way that means that respondents and interviewers have to negotiate the locally understood and lived basic social unit into this definition. We do not reject the household as being too contrived; to do so would be counter-productive and rob us of a useful concept that is articulated in everyday and non-technical language. However, collecting data on people’s lives, we have to recognise that humans live in groups within which there are social, and often biological, relationships where different types of individuals support each other.

These groups are fundamental building blocks of society; they do have their own independent characteristics and recognising them renders data collection more practical and subsequent analysis more sensible.

We contend that the very complexity and flexibility of social organisation that is understood by this minimal group as it is lived, could be better acknowledged by those that commission, produce and use these data. The very concept ‘household’ is ambiguous which needs to be situated in its context, and be qualified. The household is potentially very useful analytically, provided that its constraints, boundaries and parameters are placed centre stage, not just for survey respondents and data collectors, but also for people that consume, use and analyse data produced by household surveys. We do not suggest that households should be excluded from comparative research across and through cultural, geographic, and linguistic borders. However harmonising household definitions does not simplify the analytic task. Generic ‘one size fits all’ household definitions may do more damage than good if what we really want to understand and compare is how people actually live together.

In reality, the household is polymorphous - it is made up of individuals who have allegiances and connections, to varying degrees and types, with other households. Household membership can be blurred and indistinct, across both space and time and whereas for many survey professionals, defining who belongs to a household appears relatively straightforward, our research demonstrates that actually identifying the group who make joint provision for food and other essentials for living is frequently very difficult. Reducing the household down to a simple de facto residential group with limited criteria for membership loses important dimensions of the household as a fundamental social unit.

Household surveys in low income countries

How can household surveys in low income countries be adapted to capture better the population they purport to represent? Routinely collecting information on both de facto and de jure household members whilst having a long reference period of at least a year would be a simple first step. Household surveys need to collect data in more flexible ways that allow better configurations of individuals, including moving beyond simply listing relationship to household head, to documenting detailed relationships and nuclei within the household. Surveys and censuses in higher income countries are increasingly flexible in their approach to household membership. For example, the new French continuous census allows multiple household membership. Where possible it is critical to avoid assumptions of crisp boundaries to households, and instead to allow multiple membership of households. Ultimately, this approach is likely to give a much more realistic picture of household structure and access to resources and potential wellbeing.

Future research

A forthcoming research project develops these findings and aims to understand the implications of harmonizing definitions of the ‘household’ for survey data to represent the realities of intergenerational relationships in Anglophone and Francophone settings in Europe and Africa.

For further information please see the project website http://www.householdsurvey.info/ or contact Dr Ernestina Coast at e.coast@lse.ac.uk
Collaborative analysis of international microdata resources: The Pathfinder projects

Peter Elias, ESRC Strategic Advisor (Data Resources), University of Warwick

In a review of the work of the Economic and Social Research Council (ESRC), the House of Commons Science and Technology Committee (House of Commons, HC13, 2004) commended moves taken by the ESRC to broaden its research agenda, moving it towards the funding of issues of global importance.

The ESRC responded by commissioning a Review of International Data Needs and Resources in 2005. The Review made a series of recommendations. Foremost among these recommendations was the need to develop knowledge about and improve access to data on the emerging economies, particularly Brazil, South Africa, India and China.

To carry forward these recommendations four scoping studies were commissioned, one in each of these countries, to review and document microdata resources available for research purposes. A synthesis of the findings from these studies was published in 2006.

These studies highlight the wealth of microdata available for secondary data analysis in these countries in areas of health, education, housing, labour markets, crime and population.

The synthesis report argues persuasively for an increased sharing of such microdata across the global social research community.

However, the scoping studies also identified some stumbling blocks that may arise in negotiating ways to share such microdata effectively across national borders, namely: language, the absence of centralised data repositories, confidentiality and legal/cultural issues. To address these issues, whilst meeting the aim of the ESRC in promoting international research networks, sharing research knowledge and expertise and widening research opportunities for UK researchers, plans were laid to stimulate secondary data analysis projects using the microdata resources revealed in the scoping studies. These projects, which have subsequently become known as the ‘Pathfinder’ projects, aim to:

• build or strengthen research networks, joining social scientists in the four countries concerned (Brazil, India, China and South Africa) with UK counterparts;
• help build capacity in quantitative methods in the social sciences within these countries; and
• familiarise UK social scientists with a range of data resources not widely recognised as available for research in these countries.

Commissioning of projects has now been undertaken in Brazil and India. Twelve projects were funded from among 25 submitted proposals (please see table below for details).

Further reading

The Call for Proposals for projects in South Africa and China

The call was published at the end of June 2010 with a closing date of 12th October 2010. See www.esrc.ac.uk/ESRCInfoCentre/opportunities/current_funding_opportunities/Pathfinders.aspx

Pathfinder projects in India, Brazil and UK

India/Brazil/UK projects

• Inequalities in access to health care in Brazil and India: closing the gap for the poorest poor.
• Employment and Occupational Outcomes: The Impact of Widening Access to Education in Brazil and India.
• Urbanisation and Spatial Inequalities in Health in Brazil and India.
• Assessing the Impact of Higher Education Expansion on Economic Restructuring, Occupational Change and Access to Opportunities in Brazil and India

Brazil/UK projects

• Brazilian Manufacturing in the face of Chinese Competition: Economic Restructuring, Competitiveness and Employment.
• Exchanging Data and Skills on Place Inequality: A UK-Brazilian Collaboration.
• Gender and Ethnic Bias in Sentencing - A study using Brazilian Micro-data.
• The PBR-5 Network for the Interdisciplinary Analysis of the Brazilian Economic Restructuring Process.
• Social Inequality and Women’s response to domestic violence: A UK-Brazil Research Partnership.
• Firm Behaviour and the Introduction of New Products: Evidence Using a Firm Micro Data Set in Brazil

India/UK projects

• Economic Reforms and the Evolution of Productivity in Indian Manufacturing Firms.
• Women’s Autonomy and the Nutritional Status of children in India.

Pathfinder projects in India, Brazil and UK

...
Reflecting on the 4th ESRC Research Methods Festival

Professor Dharmendra Pratap Singh, Tata Institute of Social Sciences, Mumbai, India

It was a great pleasure for me to attend the ESRC Research Methods Festival in Oxford. I would like to congratulate NCRM for organising such a great event with many experts and participants from different countries.

I was very fortunate to receive a bursary from RCUK India office to attend the festival and feel that the aim of the bursary scheme was fulfilled. During the festival I made many new friends and had the opportunity to discuss and share my research interests with them.

The festival programme offered many simultaneous sessions on different research topics and I chose to attend seven sessions that were of most interest to me. I found that the sessions were very well planned and highlighted new techniques and future prospects in social science research methods. Personally I benefitted a lot from finding out about new techniques and applications. I am hoping to use this newly gained knowledge in my research work as well as to share this knowledge to our students at Masters and Ph.D. level.

With regard to the venue, the accommodation and food facility at St. Catherine College was excellent. Services such as internet, social programme and volunteers were so helpful that I personally did not face any problems during my entire stay. If I would give a grade to the festival I would give it 4.8 points out of 5.

The Centre for Research Methodology, where I hold a professorship, is unique among Indian universities as no other university offers a full course on research methodology. We have been running courses in Research Methodology since 1952. Our centre covers all aspects of research methodology from various theories of social science research to qualitative and quantitative aspects of research. The Tata Institute of Social Sciences was founded in 1936. The university was recently judged by the university Grant Commission New Delhi, as one of five best universities in India. Organising a similar research methods event in India is most desirable. I would be especially interested in inviting participants from many developing countries and encourage them to share their experiences.

Lastly, let me thank ESRC and NCRM for holding such a grand research methods festival.

4th ESRC Research Methods Festival took place on 5-8 July 2010 at St Catherine’s College, Oxford.

The event attracted over 800 delegates, with more than 10% of delegates from countries across the globe, from Finland to India.

Delegates came from across the social sciences, with sociology, education, social policy, health, social statistics, economics and psychology each accounting for over 5% of the people present, and geography, political science, and management also well-represented.

In terms of sectors, the majority of delegates came from higher education, but nearly 20% came from government, social or market research.

Regarding the representation of researchers at different career stages, 22% of delegates were research students, and 47 of these exhibited posters about their work which generated a great deal of interest.

The backbone of the Festival was the programme of 67 sessions which ranged from introductory ‘What Is?’ sessions on 27 different topics to presentations about developments at the cutting edge of more advanced topics such as multi-level modelling. There were also workshops and masterclasses, and sessions devoted to considering the future of publishing and teaching research methods, and to the vexed question of the ‘impact’ of social science research.

Many of the presentations held at the 4th ESRC Research Methods Festival are available through the festival web pages in http://www.ncrm.ac.uk/TandE/other/RMF2010/

The 5th ESRC Research Methods Festival will be held on 2-5 July 2012 at St Catherine’s College, Oxford.

Picture: Festival organisers were lucky with the sunny weather. Here the festival delegates are enjoying themselves at one of the early evening receptions.
Vital Signs 2 conference: Engaging Research Imaginations

Registration is now open for the forthcoming 3 day conference Vital Signs 2: Engaging Research Imaginations. The conference takes place on 7-9 September 2010 at the University of Manchester.

Vital Signs 2 is an international, interdisciplinary conference organised by Realities, part of the ESRC National Centre for Research Methods and is based in the Morgan Centre at the University of Manchester.

How can we engage our research imaginations and rise to the challenge of generating knowledge this is vital and resonates with complex and multi-dimensional lived realities? Vital Signs 2 will provide a major forum for lively and engaged discussion of these issues.

Keynote speakers are:
- David Inglis (Cultural Sociology, University of Aberdeen)
- Nigel Rapport (Anthropology/Philosophy, St Andrews University)
- Gillian Rose (Geography/Visual Culture, Open University)
- Jennifer Mason (Sociology, University of Manchester)

To view the full programme and register, please visit the conference website http://www.manchester.ac.uk/realities/events/vitalsigns/

Everyone wishing to attend must register as a conference delegate.

Fees: £160 standard; £80 concessionary