Scottish Social Survey Data, Past, Present and Future – Does Scotland Need its Own Data Strategy?

Vernon Gayle, Christopher Playford and Paul Lambert

Introduction

Throughout the UK an increasing number of large-scale multiple-purpose social surveys are collected and made accessible to the social science research community. These data resources provide rigorously and systematically collected data that is representative and suitably robust to enable reliable and valid generalisations. There is also an increasing amount of social survey data collecting measures relating to social and economic life in Scotland. These include United Kingdom (or Great Britain) based surveys with a Scottish component, and surveys focussed on Scotland.

Post-devolution, with the emergence of new forms of governance and new institutional arrangements, the political desire to research life in Scotland has been brought into sharper resolution. Social and economic life in Scotland shares both similarities and differences with life south of the border. Large-scale social surveys are important because they provide an obvious source of empirical data from which these characteristics can be correctly identified (rather than simply being assumed). Both UK and Scottish social surveys offer exciting opportunities to document and understand living in Scotland in the 21st Century.

The Economic and Social Research Council (ESRC) have recently devised a UK national strategy for data resources for research in the social sciences (ESRC 2007). 'This National Data Strategy aims to provide a coherent framework for the development and maintenance of a robust data infrastructure, ensuring that relevant and timely data are available to inform and address future research priorities' (ESRC 2007 p.4). The strategy emphasises the importance of creating better resources for the study of complex research questions. Hitherto, a comparable strategy document on survey data in Scotland has not been produced.

In this paper we proceed by outlining some of the existing Scottish survey data resources. Our aim is to provide the reader with an overview of the composition of the existing survey data portfolio. We will also outline emerging data resources, and then highlight the benefits of Scotland having its own data strategy.

Survey Data Resources

Past

As described below, there has been a rapid and substantial expansion in the volume of survey data resources with extended coverage of Scotland, which can be traced to the period of the mid to late 1990's. However, prior to the last decade, Scotland already benefited from an extended history of participation in major social surveys.

Firstly, large scale surveys of Britain or the UK conventionally included a proportional sample from the Scottish population (albeit from south of the Caledonian canal), meaning that major annual resources such as the General Household Surveys (1971 onwards); the Labour Force Surveys (1975 onwards); and the New Earnings Survey (1970 onwards) could be fruitfully used for the analysis of the Scottish population. Nevertheless, many other influential social surveys from this period, whilst including coverage of the Scottish population, did not generally include sufficient cases that Scottish focussed analyses were easily achieved (see for example the British Social Attitudes Surveys 1983 onward; the British Election Studies 1963 onwards; and the British Household Panel Survey 1991-1999). It is of particular note that this latter situation has historically been the case for cross-national social survey projects (e.g. the Eurobarometer 1974 onwards; International Social Survey Project 1985 onwards; the European and World Values Surveys 1981 onwards), and continues to be observed for these ongoing, and many new, cross-national resources (e.g. the European Social Survey 2001 onwards).

Secondly, a history of Scottish focussed major surveys can also be identified. The 1974/5 Scottish Mobility Study represents a famous example of an influential Scottish focussed survey (see Iannelli and Paterson 2006). Scottish sub-samples were occasionally added to major survey resources such as in the example of the 1979 Scottish Election Study and the Scottish boost to the 1974 British Election Study. Two Scottish locations were included within the influential Social Change and Economic Life Initiative surveys conducted in 1986.

In addition to social surveys there are several long-running projects in the medical and health areas that include social science data. A notable example is the West of Scotland Twenty-07 Study: Health in the Community[24], which was established in 1986. Three generations born around 1932, 1952 and 1972 have been followed form 1987 (i.e. the study will be age twenty in 07!). Other examples include the Walker Project[25] based in Dundee, which is a longitudinal study of 48,000 children born between 1952 and 1966 (aged 36-55 in 2002) and their families. Two other examples are the Aberdeen Children of the 1950s Cohort Study[26] and the Lothian Birth Cohort Studies[27].

Nevertheless, these older survey resources oriented to Scotland are often eclipsed by the considerable expansion of Scottish surveys over the last decade. Moreover, there have arguably been some negative experiences in the historical collection and preservation of Scottish survey data. Whilst we do not consider that it is particularly useful to dwell on mistakes of the past, we do argue that one example illustrates the shortcoming of not having a clearly articulated data strategy. The original Scottish Longitudinal Study, which was based on Census data, was discontinued soon after the 1981 Census, when the Rayner Review of the Government Statistical Service[28] concluded that it was not cost effective. A major factor was the small sample size (one percent, approximately 50,000 members) which supposedly limited its potential for analysis. By contrast the data series for England and Wales continued. The Office of National Statistics Longitudinal Study (LS)[29] is based on linked individual decennial Census records from the 1971, 1981, 1991 and 2001 for individuals in England and Wales (and is linked to other data). The LS is an important data resource because it supports a wide variety of analyses of a diverse range of social science topics (e.g. births, deaths, ageing, fertility, health, cancer, households, occupations and transport). The scale of the LS (approximately one percent of the population) renders it an important large-scale social science data source. The temporal aspect of the LS supports analyses of trends over time. The data structure is based on linked individual records and therefore supports analyses of micro-level social change, and equally importantly social stability, over time.

[24] http://www.sphsu.mrc.ac.uk/studies/2007_study/ .
[25] http://www.ingentaconnect.com/content/bsc/ppe/2004/00000018/00000004/art00010 .

[26] http://www.abdn.ac.uk/childrenofthe1950s/ .
[27] http://www.psy.ed.ac.uk/research/lbc/LBC.html .
[28] See Rayner Report (1980).
[29] http://www.statistics.gov.uk/services/Longitudinal.asp .

In our view the discontinuation of the original Scottish Longitudinal Study has left a gap in Scottish data portfolio which has been disadvantageous and has restricted potential analyses. This situation is suboptimal and we conjecture that with a clearly articulated data strategy this would be far less likely to occur. On a brighter note, as we will discuss below, work to re-establish a Scottish Longitudinal Study has recent been successful (although linking to original records has not been possible).

Present

Currently survey data is collected across a broad spectrum of social science topics in Scotland. We consider that it is a useful heuristic device to make the distinction between UK (or British) survey datasets that contain Scottish data and Scottish datasets. In this section we will outline some of the UK surveys before moving on to Scottish datasets.

Williamson (2008) is an excellent guide to the Scottish large-scale government surveys supported by the Economic and Social Data Service (ESDS), although some other surveys are also covered. It is organised by themes (employment and labour market, health, crime, identity and social capital, and demography). In addition to concise details on various surveys, the guide also provides summary information of sample size and Scottish geography and succinct information on sample weighting. In particular, the appendix provides a useful table comparing information contained in Scottish datasets.

A number of other helpful resources exist. Hinds, Sproston and Taylor (2001) provide useful information in their report of a scoping study of longitudinal research needs carried out on behalf of the Scottish Executive. Bell, Jack and Wright (2004) provide a useful overview of micro-social science data suitable for Scottish policy research. In our view it would be very constructive and highly valuable if the Scottish Government funded a full review project that documented the entire portfolio of Scottish social survey resources.

- The British Household Panel Survey
[http://www.iser.essex.ac.uk/ulsc/bhps/]

The British Household Panel Survey (BHPS) is a multi-purpose survey based at Essex University. It began as a representative study of 5,500 British Households (approximately 10,000 adult individuals). All individuals from these households are re-interviewed annually (this is the panel). A new sweep of data collection begins in the autumn. Each

sweep of data collection is called a wave, and each wave is labelled alphabetically (e.g. wave A is the 1991 data and wave N is the 2004 data). The data are updated and released annually. The BHPS is also known as 'Living in Britain'.

A major development in 1999 (wave I, the ninth wave of survey data collection) was the recruitment of two additional samples to the BHPS, one in Scotland and another in Wales (a further additional sample from Northern Ireland was added to the BHPS in 2001). The Scottish sample, also known as 'Living in Scotland', is often referred to as the Scottish Booster. There were two main aims of the extensions. The first aim was to increase the relatively small Scottish and Welsh sample sizes in order to permit independent analysis of the two territories. The original sample size for Scotland was correctly proportional within Britain, but generally considered too small for independent analysis. Laurie and Wright (2000) report that the Scottish sample was between 452 and 568 households between 1991 and 1998 (p.337). The second aim was to facilitate analysis of the two countries compared with England in order to assess the impacts of the substantial public policy changes which may be expected to follow from devolution.

A clear account of the Living in Scotland survey is provided in Laurie and Wright (2000). At the current time data are available from 1999 to 2005 for the Scottish booster. In 1999 1,458 household were surveyed and 2,405 adults gave a full interview. In 2005 1,274 household were surveyed and 1,977 adults gave a full interview. Ermisch and Wright (2005) present a number of empirical analyses in the areas of families and households, inequalities, labour market issues and social and political behaviour using early waves of the Scottish data.

- The Millennium Cohort Survey
[www.cls.ioe.ac.uk/mcs]

The Millennium Cohort Survey (MCS) is a major survey data development. Since the 1940s Britain has an established tradition of collecting large scale birth cohort data. After the 1970 cohort a gap in the sequence of data collection emerged (Martin *et al.* 2006). The MCS is the first large-scale national birth cohort study to be undertaken for thirty years. As the title indicates it is a study of babies born at the millennium.

The survey for the first sweep took place between June 2001 and January 2003, gathering information from the parents of 18,819 babies born in the UK (of which 2,370 were born in Scotland). The

second sweep took place when the babies were around 3 years of age and the third sweep went into the field when the children were starting primary school. The study is an orthodox birth cohort study design and it presents an exceptional research opportunity to study the all-important first years of life for children growing up in the 21st Century. These data have the potentially to inform research that may move us closer to resolving many of the questions regarding the long-term impact of the early years upon outcomes in later life.

At the current time the data support research into the early stages of the lifecourse. Research into the health and well being of infants and parents is particularly well supported by the data, as are social research topics such as poverty, social exclusion and social capital. Future sweeps of the study will collect more information on emerging aspects of the lifecourse, such as advantage and disadvantage in education, health, employment, parenting, family life, technology, gender roles etc.

We can also expect that in future, the MCS will follow the earlier UK birth cohort studies and will become a study with data across the lifecourse. We consider that the MCS data are most likely to approach their optimal analytical potential if sample members are followed from birth to death. The reality of a birth to death cohort is that it is implausible that any of the original (or early) data collectors will be working with the data in the later stages of the cohort's lifecourse. Therefore appropriate forward planning is fundamental and suitable stewardship and data curation are vital. Such actions are too critical to be left to ad hoc arrangements. This is one obvious area of the data portfolio in which longer term and joined-up strategic planning are imperative, for both Britain and Scotland.

- The Labour Force Survey
[http://www.esds.ac.uk/government/lfs/]

The Labour Force Survey (LFS) is the main source of information on employment and the labour market. It is carried out by the Office of National Statistics (ONS) and used by many government departments to monitor social and economic policies.

Since its inception in 1975 the LFS has experienced several reviews to its sample design. The current format incorporates an overlapping panel design, leading to a structure that supports different types of secondary analysis. The LFS can be used for cross-sectional analyses (e.g. for calculating an unemployment rate). The data can be used for analysing trends by comparing points in time with cross-sectional

data (for example unemployment in Sept 2000, 2001 and 2002). The panel element of the LFS can also be exploited for full-scale longitudinal data analyses. As a large scale survey, the LFS always included considerable numbers of cases from Scotland, but of particular note is that from March 2003 there was a Local Scottish LFS enhancement that increased the overall target population to 23,000 households with 50,000 individuals (40,000 aged over 16). In the Quarterly LFS (January to March 2006) there were 10,738 people usually resident in Scotland in the sample (Williamson 2008 p.12).

The UK is required by European Union Regulations to carry out an annual Labour Force Survey and therefore we do not foresee changes to the LFS in the immediate future. The uniquely Scottish aspects of the survey are particularly important for Scottish economic analyses, but also for comparative economic and social research. A Scottish data strategy could be particularly important in ensuring that suitable equivalent, harmonized and comparative measures are collected in other datasets which allow them to appropriately dovetail with LFS data resources.

- The UK Census
[http://www.statistics.gov.uk/census/]
[http://www.gro-scotland.gov.uk/census]

The Census is an important resource in the UK data infrastructure. The Scottish Census is planned and carried out by the General Register Office for Scotland, and in England and Wales by the Office for National Statistics (ONS). Census data is legally restricted and to aid access a large catalogue of Census related products and outputs are produced and made available for social science research. In Scotland, the SCROL online service[30] provides a particularly popular source of access to results from the 2001 census.

The Samples of Anonymised Records (SARs) is a Census output product which makes a unique and valuable contribution to the social science data infrastructure. The SARs are samples of individual (and household) records from the 1991 and 2001 Censuses (with identifying information removed to protect confidentiality). The SARs provide very large datasets resembling social survey data that contain a wealth of information for social scientists and policy makers. The sample size is much larger than many surveys, and this can facilitate analyses of some small groups and sub-national geographical analyses. The SARs allow data analysis to be undertaken in a similar manner to most other large-scale datasets. The SARs cover a range of

[30] http://www.scrol.gov.uk/scrol/common/home.jsp

topics including housing, education, health, transport, employment and ethnicity, but in practice there are far fewer variables than in most large-scale surveys.

Some Thoughts on the Multiple-Purpose Surveys

We employ the term 'multiple-purpose' to describe the general nature of the large-scale datasets that have been outlined in the section above. At the current time many large scale multiple-purpose datasets appear to be secure in the near future, however their funding is never guaranteed. It is worth remembering that the funding of these resources frequently comes from different sources. The curation of the multiple-purpose data resources can be complex. The responsibility and stewardship of these data resources is currently distributed across a number of stakeholders in the social science community (these are often located in England). The multiple-purpose datasets are general infrastructural data resources that support a range of analyses across a number of social science disciplines. Given their significance to the portfolio of Scottish survey data, we consider that, taken together, these issues point to the need for a Scottish data strategy in which these multiple-purpose surveys are given prominent acknowledgement.

Scottish Datasets

Since the mid- to late 1990's Scotland has been well served with large-scale data resources. The following short list of social surveys provides some insight into the broad coverage. The Scottish Health Survey (initiated in 1998)[31], Scottish Crime Survey (1993)[32], Scottish Social Attitudes Survey (1999)[33], Scottish School Leavers Survey (1998)[34], Scottish Household Survey (1999)[35] and the Scottish House Conditions Survey (1991)[36]. All of these surveys are recurrent over time, and some have a longitudinal panel component (e.g. the Scottish House Conditions Survey). Some are comparable with non-Scottish datasets (e.g. the Scottish School Leavers Survey is generally comparable to the Youth Cohort Study of England and Wales[37]). In addition to these Scotland-wide datasets, some geographically focused

[31] http://www.scotland.gov.uk/Topics/Statistics/Browse/Health/scottish-health-survey .
[32] http://www.esds.ac.uk/government/scs/ .
[33] http://www.natcen.ac.uk/natcen/pages/or_socialattitudes.htm#ssa .
[34] http://www.statistics.gov.uk/STATBASE/Source.asp?vlnk=1323&More=Y .
[35] http://www.scotland.gov.uk/Topics/Statistics/16002 .
[36] http://www.scotland.gov.uk/Topics/Statistics/SHCS .
[37] http://www.statistics.gov.uk/STATBASE/Source.asp?vlnk=668&More=Y .

datasets also exist. A notable example is the Edinburgh Study of Youth Transitions and Crime (initiated in 1998)[38].

The expansion in survey data resources collected in Scotland offers a new depth of social survey information on many particular aspects of contemporary social life. The expansion in data also has parallels with many other countries. In many smaller European nations, as well as in many developing countries, there has been a rapid increase in the availability of new social survey datasets. Such developments appear to offer exciting opportunities for comparative survey data analysis which were not possible in previous generations. However, an important lesson from comparative survey research concerns the benefits of tightly coordinated data collection and data curation strategies in order to maximise the possibilities for cross-national comparability (e.g. Harkness, van de Vijver and Mohler 2003; Hoffmeyer-Zlotnick and Wolf 2003). We therefore argue that an important component of a Scottish data strategy should be an effort to pay great attention to issues of comparability with other national data collection standards.

Future

Having outlined some presently available data resources we will now turn to future (or emerging) resources. We will first outline two major UK survey data sources and then two new Scottish based social surveys. These resources are likely to make a substantial contribution to the social science data infrastructure.

- The UK Household Longitudinal Study
[http://www.iser.essex.ac.uk/ukhls]

The ESRC recently received funding for a new household study from the Office of Science and Innovation's (OSI) Large Scale Facilities Programme. The UK Household Longitudinal Study (UKHLS) is a major study (approximately £15.5M) and will be based on a large sample of households. The UKHLS will be the largest study of its type anywhere in the world. The study is an exciting development and will be a valuable asset in the social science data infrastructure. The planned sample size is 40,000 households and the study will incorporate one hundred percent of the existing British Household Panel Survey.

[38] http://www.law.ed.ac.uk/cls/esytc/aboutthestudy.htm .

The structure of the UKHLS sample and sub-samples will be complex, but the design will facilitate effective Scottish analyses and analyses that make within UK comparisons. The UKHLS will include a number of innovative features, most notably, a special (boosted) sample of key ethnic minority groups and interviews from all household members aged ten and above. The UKHLS will also contain an 'innovation panel' of households to better facilitate methodological research. In addition there are plans for the collection of biomedical measures and samples, as well as the linking of administrative and other (e.g. geographical) data. The exact title of the study has not been decided upon but it may become known as Understanding Society (US).

- The Integrated Household Survey (Formerly the Continuous Population Survey)
[http://www.statistics.gov.uk/CCI/nugget.asp?ID=936&Pos=&ColRank=1&Rank=192]

The ONS Integrated Household Survey (IHS) is currently being developed. The motivation is for a redesign of the existing ONS continuous household surveys into a single module-based survey. The surveys that are proposed for integration are the Labour Force Survey (LFS) and all associated boosts, the General Household Survey (GHS), the Expenditure and Food Survey (EFS), and the Omnibus Survey.

The proposed integrated survey is a response to the perceived need for better National information on key social and economic variables, and aims to meet the increasing demand for regional and sub-regional information. The ONS considers that increasing demands cannot be met within the current survey arrangements because the existing surveys have reached their limit in terms of both survey length and respondents' burden. The data from the existing surveys are not easily pooled as their designs are different. In addition there are practical and financial efficiency gains associated with combining data collection into a single redesigned survey.

- The Scottish Longitudinal Study
http://www.lscs.ac.uk/sls/

The Scottish Longitudinal Study (SLS) is an interesting new development that has very recently been launched. The SLS is a large-scale linkage study which has been created by using data available from the Scottish Census and administrative data. The SLS is a replica of the England and Wales Longitudinal Study (LS) (mentioned above). The SLS has an approximate sample size of 274,000 and

therefore can be regarded as the biggest Scottish social science dataset.

Members of the study have been identified from the 1991 census and information for these individuals are linked to the 2001 census and other datasets, including vital events and health information. The SLS provides a new data source for studying social life in Scotland especially in the areas of demography, housing, family life, marriage, fertility, social inequality and employment. A clear and concise introduction is available from Hattersley and Boyle (2007).

It is plausible that the SLS will facilitate comparative research, and it is particularly suited to analyses that also use the LS. In future a possible development would be a Great British (or even UK) Longitudinal Study, based on linked Census records. This would be a major data construction exercise (particularly as these two major datasets are similar but not identical). We assert that this is the type of infrastructural development that requires in-depth strategic planning.

- Growing Up in Scotland
[http://www.crfr.ac.uk/gus/]

The Growing Up in Scotland Study (GUS) is a new longitudinal research project aimed at tracking the lives of a cohort of Scottish children from the early years, through childhood and potentially beyond. Funded by the Scottish Executive Education Department, its principal aim is to provide information to support policy-making, but it is also intended to be a broader data resource.

The survey is based on two cohorts of children. The first cohort was aged approximately 10 months at the time of first interview and the second aged approximately 34 months. These two groups are referred to as the birth cohort and the toddler cohort respectively. There is an approximate sample size of 8,000 overall. The first sweep of the survey was carried out between April 2004 and March 2005. At present, the intention is to launch a new birth cohort (BC2) in 2009.

Research Capacity

It is a paradox that both Scotland and UK overall are now rich in survey datasets, but this has not been mirrored by research outputs. At the current time some survey datasets remain under-analysed (and some components verge on the unanalysed) and it is safe to conclude that Scottish survey data have not yet reached their full analytical

potential. For a small country Scotland has a good social science tradition, a relatively healthy academic sector and a large proportion of young people in higher education. Scotland has some skilled and talented survey analysts, but they are spread thinly across the academic and non-academic research sectors. At the present time there is a shortage of young and junior researchers with survey data analysis skills.

Scottish based projects such as the Practical Exemplars and Survey Analysis Project[39] and the Longitudinal Data Analysis for Social Science Researchers Project[40], both funded by the ESRC, have provided skills training in survey analysis. The Scottish Social Survey Network[41], also funded by the ESRC, provides a focussed network and advanced training for Scottish survey researchers. In addition the Scottish based ESRC National Centre for e-Social Science Node[42] will provide resources for survey data management. These projects are likely to have some impact but we argue that in the longer term a step-change needs to be effected to increase research capacity in Scotland.

Conclusions

We can clearly identify a number of advantages to Scotland having its own data strategy, but it is difficult to identify disadvantages. Our prime message is that Scotland needs its own data strategy given the diverse range of datasets (some of which have been briefly outlined above). This would minimise the risk of duplication of data resources; assist in standardisation and harmonisation of data when necessary; and protect against discontinuities in critical data resources (such as the example of the original Scottish Longitudinal Study).

As a nation it is imperative that social scientists are able to document and understand living in Scotland in the 21st Century. It is essential that appropriate survey data resources are available to study both social changes and social stability. Social and economic life in Scotland shares both similarities and differences with life south of the border. It is important that these characteristics can be correctly identified and this requires appropriate social survey data.

[39] http://www2.napier.ac.uk/depts/fhls/peas/index.htm .
[40] http://www.longitudinal.stir.ac.uk/ .
[41] http://www.longitudinal.stir.ac.uk/surveynetwork/ .
[42] http://www.dames.org.uk/ .

Current data resources are largely supported through UK based institutions such as the Economic and Social Data Service (ESDS)[43] and the Data Archive[44]. These institutions provide world-class data curation and state of the art access to survey data resources. In a more devolved, or independent, Scotland an alternative set of data management, curation, storage and access arrangements would probably be politically desirable. At present it does not appear that any thought has been directed towards this issue, and we envisage that this would be part of the remit of a Scottish Data Strategy.

Although we are making a case for a Scottish Data Strategy, we also advocate that Scottish data resources continue to collocate within the wider UK data infrastructure. This approach will allow the successful continuation of many existing data series. Inevitably, from time to time some tough decisions might have to be taken regarding which datasets continue to be supported. It would be more beneficial if such decisions were taken as part of a wider and longer term, data strategy.

We are aware that the multiple-purpose survey datasets cannot be operationalised to investigate every research question. Therefore there should always be capacity (and funding) available for more ad hoc social surveys that are directed towards collecting data appropriate for more specific research hypotheses. These survey data resources are valuable because they often provide additional opportunities for secondary analysis. Through the work of the UK Data Archive and ESDS substantial progress has been made in relation to the longer term storage of ad hoc survey datasets. We expect that in the emerging era of e-Social Science further progress will be made, especially in terms of data management, data delivery, and documentation. A clear data strategy could help curate these ad hoc survey data resources within a more unified survey data framework.

Scotland is sometimes referred to as the 'British policy laboratory', although this idea is contentious. Criticisms notwithstanding, there is scientific merit in comparing Scotland with the rest of the UK for some policy evaluations. Because comparisons with England and Wales often provide suitable counterfactual examples, we emphasise the importance of Scottish survey data being collected in a wider UK context.

Scotland is a small northern European country and we are aware of a growing political desire to compare progress in Scotland particularly with Eire and the Nordic countries. In order to facilitate valid cross-

[43] http://www.esds.ac.uk/ .
[44] http://www.data-archive.ac.uk/ .

national comparisons of this kind it is imperative that Scotland has appropriate large-scale social science data. Again, it can be argued that a Scottish Data Strategy can make a clear contribution to raising standards of comparability.

A Scottish Data Strategy could be used to promote best practice in survey data collection. A contribution could also be made by endorsing best practice in areas such as data curation and the development of equivalent, comparative and harmonised measures. In our experience there has not been strong guidance in these areas and overall management of the data portfolio has been largely unplanned. This often results in ad hoc decision making, which increases the risk of the sub-optimal provision of data resources.

At the current time, we are aware that a number of Scottish and UK data analysts have highlighted gaps in Scottish data portfolio. A notable example is in the area of ageing. At the current time Scotland does not have an individual (i.e. person level) dataset, similar to the English Longitudinal Study of Ageing[45]. The ageing population is seen as a priority, and often an unfolding social problem, and social care policies and arrangements differ in Scotland. Therefore it is easy to jump to the conclusion that developing such data resources is an imperative. Longitudinal data are expensive to collect and take time to mature analytically. Is the Scottish ageing population sufficiently different that a unique, and potentially expensive, data source is required? What would the additional analytical advantage be? We don't know the answers to these questions and we suspect that others don't either. A strategic approach to the development of new data resources should however assist in making such decisions.

As well as gaps within Scottish data resources we are aware that a number of Scottish data analysts have highlighted overlaps between datasets. The most notable example is the commonality between the data provided by the Growing up in Scotland and the Millennium Cohort studies. The timing, design and content of the two studies is obviously not identical, but they data coverage is similar. Therefore some commentators have questioned the extra analytical value that Growing up in Scotland will provide weighed against its costs. We argue that a more considered approach to developing new data resources in Scotland is essential and that an explicit data strategy would necessitate this.

[45] http://www.ifs.org.uk/elsa/ .

We believe that appropriate data resources are a vital part of Scotland's research infrastructure and after a decade of devolution Scotland needs a full, and scientifically informed, review of existing data resources. The expansion of research capacity in Scotland is essential if survey data are to return their full analytical value, and training and knowledge transfer are integral to improving research capacity. Our closing remark is that wide consultation within the Scottish social research community is fundamental to developing an appropriate data strategy. The development of a full-blown Scottish Data Strategy that dovetails with the UK National Strategy is imperative.

References

Bell, D.N.F., Jack, G. and Wright, R.E. (2004) 'Micro-Level Data Sources for Scottish Policy Studies, *Scottish Journal of Political Economy*, 51(1), 143-147.

Economic and Social Research Council (2007) *National Strategy for Data Resources for Research in the Social Sciences*, Swindon, ESRC.

Ermisch, J.F. and Wright, R.E. (Eds.) (2005) *Changing Scotland – Evidence for the British Household Panel Survey*, Bristol, Policy Press.

Harkness, J.A., van de Vijver, F. J. R., & Mohler, P. P. (Eds.). (2003) C*ross-Cultural Survey Methods*, New York: Wiley.

Hattersley, L. and Boyle, P. (2007) 'The Scottish Longitudinal Study – An Introduction', *LSCS Working Paper 1.0*, [http://www.lscs.ac.uk/sls/LSCS%20WP%201.0.pdf].

Hinds, K., Sproston, K. and Taylor, R. (2001) *Data Sources for Social Research in Scotland: Results from a Scoping Study on Longitudinal Research,* Scottish Executive Central Research Unit, [http://www.scotland.gov.uk/Resource/Doc/156505/0042021.pdf].

Hoffmeyer-Zlotnik, J. H. P. and Wolf, C. (Eds.) (2003) *Advances in Cross-national Comparison: A European Working Book for Demographic and Socio-economic Variables*. Berlin: Kluwer Academic / Plenum Publishers.

Iannelli, C. and Paterson, L. (2006) 'Social mobility in Scotland since the middle of the twentieth century', *Sociological Review, 54*(3), 520-545.

Laurie, H. and Wright, R.E. (2000) 'The Scottish Household Panel Survey', *Scottish Journal of Political Economy*, 47(3), 337-339.

Martin, J., Bynner, J., Kalton, G., Boyle, P., Goldstein, H., Gayle, V**.,** Parsons, S. and Piesse, A. (2006) *Strategic Review of Panel and Cohort Studies - Report to the ESRC*, Swindon, ESRC [http://www.longviewuk.com/pages/documents/ESRCstrategicreview website2bedit27.03.06a.pdf].

Rayner Report (1980) *Sir Derek Rayner's report to the Prime Minister*, London, HMSO.

Williamson, L. (2008) *Guide to data sources for Scotland*, Economic and Social Data Service, [http://www.esds.ac.uk/government/docs/ScotlandGuide.pdf].

Vernon Gayle: University of Stirling & Institute for Social and Economic Research, University of Essex,
Christopher Playford: University of Stirling)
Paul Lambert: University of Stirling
Contact: c.j.playford@stir.ac.uk