



2007 - 2008

Welcome to the ESDS Annual Report for 2007-2008 from the Chair and Director



J N Matheson

Jil Matheson
Chair, ESDS Advisory
Committee 2007-2008

I am delighted to be able to present the Annual Report for the ESDS for the period August 2007 to July 2008. As Chair of the ESDS Advisory Committee until February 2008 I had the privilege of working with the Committee and with Kevin Schürer in delivering this excellent service for researchers and educators during a critical stage in its history. What last year Kevin called 'ESDS Phase I', the first five years of the Economic and Social Research Council (ESRC)- and Joint Information Systems Committee (JISC)-funded service, came to an end in September 2007, but was continued seamlessly into 'Phase II' in October 2007.

ESDS is one of the jewels in the ESRC crown. It is now established as the most important system for the acquisition, processing, enhancement, preservation and delivery of a huge variety of quantitative and qualitative datasets in the UK. The datasets range from government surveys, to international economic studies, long-term studies tracking social groups over decades, and oral histories covering an amazing spread of individuals and organisations. ESDS now employs a comprehensive team of highly skilled professional specialists to deliver the service, mainly using the ever-expanding internet, from its distributed base across the Universities of Essex and Manchester.

As this Annual Report shows, 2007-2008 continued the development of ESDS with a restless commitment to user satisfaction, which is what ESDS is all about. I have enjoyed my chairing the Advisory Committee enormously and look forward to next year's report which I am sure will take this commitment even further. I wish continuing success to my successor, Audrey Brown, and to all those associated with ESDS.



Kevin Schürer
Director, ESDS

It is my pleasure, as Director of ESDS since its inception in 2003, to support the Chair's foreword to the 2007-2008 report. Working closely with ESRC and JISC, ESDS works to deliver against its aims (as stated on page 2 of this report) by pursuing a range of objectives for 2007-2012. This report seeks to satisfy at least in part the last but arguably most important of these objectives: engaging with our stakeholders.

As the report shows clearly, during 2007-2008 ESDS built steadily on its already firm foundations and delivered a first class service. We must now rise to the challenge of an ever faster-moving technical environment to develop the service still further, in what I believe will be an increasingly collaborative and shrinking world.

Economic and Social Data Service

Annual Report August 2007 – July 2008

Contents

Welcome to the ESDS Annual Report for 2007-2008 from the Chair and Director	ifc
Aims and Objectives	2
This is ESDS	3
Executive summary	4
Demonstrating the value of ESDS	5
ESDS Government	11
ESDS International	13
ESDS Longitudinal	15
ESDS Qualidata	17
ESDS Access and Preservation	19
ESDS Management	21
ESDS communicating and engaging during 2007-2008	23
Measuring ESDS	26
Advising ESDS	32



Aims and Objectives

The overall aim of the Economic and Social Data Service (ESDS) is to promote wider and more informed use of data for research and teaching in the social sciences and to ensure that these data continue to be available over time.

To achieve this aim ESDS will:

- > develop a clear, co-ordinated acquisitions strategy in conjunction with the ESRC and JISC;
- > develop and promote common standards and agreed strategies for data preparation, processing, documentation and preservation;
- > integrate data services through the development of a universal single point of access and simple registration procedures and common conditions of use;
- > enhance underlying data sources through the development of value-added services around key datasets and series;
- > promote increased use of datasets through both general and specialist user support activities;
- > develop interface and analysis tools appropriate to differing levels of expertise amongst end users;
- > respond to key technical developments in the broader information environment;
- > engage on a regular basis with a wide range of stakeholders, including data suppliers, data funders and end users.

This is ESDS

ESDS is a distributed national data service established in 2003 by the ESRC and JISC to provide enhanced access and support for an extensive range of key economic and social data, both quantitative and qualitative, spanning many disciplines and themes. ESDS is based on a successful partnership between the University of Essex and the University of Manchester, and comprises six distinct yet interrelated services:

ESDS Government – Led by **Angela Dale** of the Cathie Marsh Centre for Census and Survey Research (CCSR), this service promotes and facilitates the use of large-scale government surveys in research, learning and teaching – providing methodological guides, short courses and support for key statistical packages. It is supported with input from the UKDA in the areas of data acquisition, processing and dissemination.

ESDS International – This service is led by **Jackie Carter** at Mimas which is responsible for macrodata, with support from UKDA, which focuses on microdata. It supports and disseminates data from major inter-governmental organisations, plus a range of international survey datasets. The service promotes and facilitates use of international datasets in research, learning and teaching across a range of disciplines.

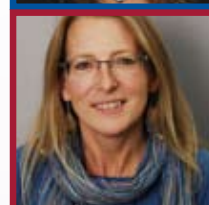
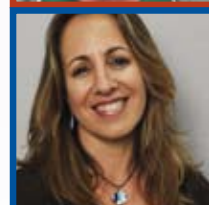
ESDS Longitudinal – Managed jointly by **Jack Kneeshaw** at the UKDA and by the Institute of Social and Economic Research (ISER), this service promotes the use of longitudinal data and supports a range of longitudinal datasets by creating data enhancements and providing user support and training.

ESDS Qualidata – Led by **Louise Corti** at the UKDA, ESDS Qualidata provides access to and support for a range of qualitative datasets. It generates a number of data enhancements and provides information and training resources for the re-analysis of qualitative data.

ESDS Access and Preservation – Also hosted by the UKDA and led by **Melanie Wright**, this service represents the core of the ESDS, focusing on data acquisition, processing, preservation and dissemination. It provides access to most of the data held by ESDS through a variety of web-based systems and also handles registration of users across ESDS.

ESDS Management – The overall management of ESDS is the responsibility of the UKDA, led by **Kevin Schürer**. ESDS Management provides consistency and standards across the other services and performs a broad strategic role in relation to the supply, funding and use of social science data. It also provides a coherent marketing and outreach strategy for the whole of the ESDS.

Further details of the activities of these individual services can be found on the ESDS web site at www.esds.ac.uk. This report covers all six services, and for brevity includes a combined set of performance statistics.



Executive summary

Each data delivery service presents a case study and a general report for 2007-2008. These are followed by a statistical report, and a diary of key events during the year.

In the case study section **ESDS Government** describes its 2007 conference on the theme of consumption. **ESDS International (macro)** highlights a study which asked whether macro-economic patterns might have predicted the US sub-prime crisis. **ESDS International (micro)** case study concludes that a sense of European identity depends in part on one's view of the European Union (EU). **ESDS Longitudinal** highlights 50 years of the *National Child Development Study* (NCDS). **ESDS Qualidata** focuses on re-use of the *Inventing Adulthood* study, looking in particular at young men's relationships with their mothers. **ESDS Access and Preservation** is illustrated by the use of the *Road Accident Data* (RAD) for a wide range of research purposes.

ESDS Government made great efforts during 2007-2008 to encourage greater use of secondary analysis amongst early researchers. Introductory workshops and support materials were produced and ESDS members played a major role in a very successful ESRC-funded project to increase the use of quantitative methods in undergraduate dissertations. Greater use of the repeated measures found in many of the government surveys was promoted for more experienced researchers.

ESDS International implemented a new data access system and produced additional user support and resource materials, as well as working hard to reach a new generation of data users. Enhancements to the portfolio this year included the *Latinobarómetro* survey and the International Monetary Fund survey, *IMF World Economic Outlook*. The data redistribution licence to replace the United Nations Common Database with the Comtrade database was also signed. Access to international macrodata was opened up to the House of Commons and House of Lords libraries, supporting evidence-based policy making by Government.

ESDS Longitudinal added the *Families and Children Study* (FACS) to the service's portfolio of major studies. Giving FACS major study status involved creating a special suite of web pages for this increasingly policy-relevant survey, including a guide for novice and potential users. A brand new acquisition was the first sweep of data for the *Growing*

Up in Scotland (GUS) birth cohort, which aims to provide important new information on young children and their families in Scotland.

ESDS Qualidata contributed to user support materials for the ESDS themes pages on health. Re-use case studies and literature on ethics, informed consent and secondary analysis of qualitative data were also updated. The service continued to work with other research resources and formed a partnership with Timescapes, the £4.5 million ESRC Resources qualitative longitudinal study. In March 2008, the JISC-funded Data Exchange Tools (DExT) project released its multi-media QuDEx metadata schema that will provide an intermediate format for longer-term preservation and sharing.

ESDS Access and Preservation during 2007-2008 dealt with changes to ways researchers can access ESDS data, including Special Licence data, or data containing sensitive or detailed variables. The UK Data Archive (UKDA) published version 3 of its Preservation Policy, which underpins all the activities of the ESDS and explains how and why the UKDA ensures the authenticity, reliability and logical integrity of all resources in its care, while also providing usable versions for research, learning or teaching.

ESDS Management promoted ESDS to the widest possible range of stakeholders around the world, including researchers from Canada, New Zealand, Australia, Russia, Japan, Taiwan and South Africa, as well as numerous visitors from around the UK and Europe. An important new service was UKDA-store, a self-archiving service for all ESRC-funded projects.

The statistical section summarises the key measures of the effectiveness of the ESDS over the reporting period. Difficulties inherent in presenting and interpreting the metrics are detailed, while any obvious trends in comparison with the previous years of the service are highlighted. An apparently large increase in processing statistics is attributed partly to new versions of key international macrodata becoming available. Following year-on-year growth during the first five years of ESDS, there now appears to be a plateauing of use. Various factors which may account for this are discussed, as are some other statistical trends, including the broadening range of user disciplines.

Demonstrating the value of ESDS

ESDS Government

All consuming



An enthusiastic audience from government, voluntary organisations and academia attended the ESDS Government 2007 Research Conference at the British Academy. The topic was consumption, and papers covered consumption by older consumers and children as well as the consumption of food, culture and private transport.

Joyce Dargay from the University of Leeds used the *National Travel Survey* to show that, while there has been an overall increase in car ownership and travel between 1989-1991 and 1999-2001, gender differences have declined and women are now catching up with men with respect to car ownership and usage. The use of buses is influenced more by bus frequency than by proximity to a bus stop. People who live nearest to local amenities are more likely to walk and less likely to own a car.

Lisa Farrell from University College Dublin presented results from an analysis of children's expenditure diaries in the *Family Expenditure Survey 1997-1998*. Children from higher income households are significantly more likely to buy alcohol and cigarettes than those from lower income households. For most types of purchasing, children's spending patterns mimic those of their parents; therefore policies aimed at influencing parental consumption of, for example, cigarettes and alcohol, may have an indirect beneficial effect on their children.



ESDS International (macro)

Forecasting the global financial crisis



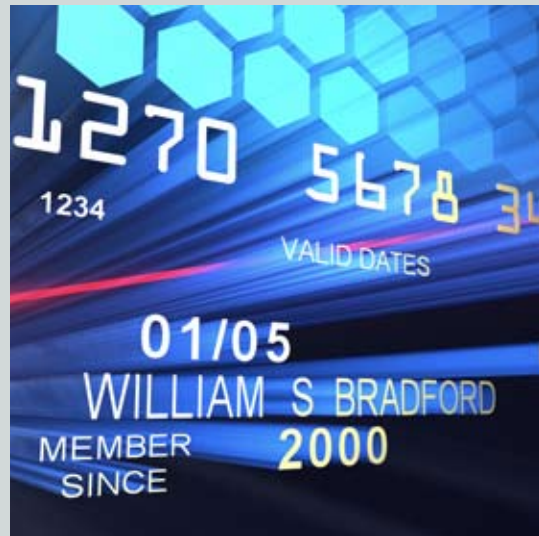
One of the features of the sub-prime crisis was its unexpected nature. It came as a surprise not only to most financial market participants but also in some degree to the policy community. That it should surprise the policy community is a paradox, since there has been extensive research, both in universities and in policy institutions, into the causes and predictors of financial crises. In this context, Diruba Karim and Philip Davis of Brunel University sought to assess whether some of the early warning systems that have been developed in the literature (but rarely applied to banking crises) could have helped in the prediction of the crisis.

To examine the extent to which the US sub-prime crisis was predicted by ex ante macro-economic and financial data, the researchers utilised two different early warning system approaches: the traditional multivariate logit method; and a binary recursive tree technique. By using ESDS International data from the World Bank and IMF for over 100 countries covering a 30-year period, the authors found that for

the majority of countries, there was no appreciable increase in crisis probabilities in the two years preceding the sub-prime episode, although the logit model performed best for the UK and USA. Iceland was an interesting exception, with a rise in crisis probabilities.

Their work shows that the sub-prime episode had been less to do with macro-economic fundamentals (with the exception of Iceland) and more to do with micro-economic banking risk. It was further demonstrated that these models would be of greater help to those undertaking the difficult task of macro-prudential surveillance if they could be better adapted for the specific features of advanced countries and include securities market instability.

Published as: Davis, E.P. and Karim, D. (2008) 'Could early warning systems have helped to predict the sub-prime crisis?', *National Institute Economic Review*, 206(35).



ESDS International (micro)

When do people feel European? European identity, EU attitudes, and questionnaire design



The legitimacy and popularity of the European Union (EU) are increasingly seen as hinging on a sense of European identity among its citizens. Evidence from the *Eurobarometer* survey series - which includes a variety of established measures of European identity - indicates that reported levels of European identity vary markedly over time. The University of Strathclyde's Rob Johns has employed the survey series to demonstrate that at least some of this variation can be attributed to differences in the context in which a question is asked.

Johns's key finding is that when respondents are asked about European identity immediately after questions about the EU, levels of European identity appear lower overall, and especially weak among anti-EU respondents. However, if the European

identity measures are located prior to, or a long way after, the EU questions, then European identity is relatively strong, even among respondents who are EU-sceptic, anti-immigration and supporters of extreme right parties. The results suggest that European identity is more fragile and volatile than attitudes to the EU. That in turn implies that the former is unlikely to drive the latter: in other words, cultivating European identity is not the route to boosting the popularity of the EU.

The results carry implications for survey designers, reinforcing the importance of question order, and warning *Eurobarometer* methodologists that they cannot expect to measure pure European identity if - as has become the norm - those questions immediately follow a long battery of questions about the EU.



ESDS Longitudinal

50 up – many happy returns!



2008 saw the members of the *National Child Development Study* (NCDS) - and, of course, the study itself - celebrate their fiftieth birthdays.

The ESDS makes all seven sweeps of data collected (eight including the initial perinatal survey) available for re-use and, as the NCDS and its members reached their impressive milestone, the study continues to sit among the top 20 most popular data series downloaded via the ESDS with over 730 accesses in 2007-2008.

To mark the happy occasion, the Centre for Longitudinal Studies (CLS) – which houses three of the four national birth cohort studies – produced a report summarising the main findings from the study and reported those findings against the background of half a century of social change in Britain.

The *Now we are 50* report, available to download via the CLS web site, not only looks back on the lives of the cohort members over five decades but also points to the future. Indeed, as the report itself suggests, now that the members of the 1958 birth cohort have entered their sixth decade, new opportunities for comparative research will open up as the cohort enters the territory of another longitudinal study whose focus is on retirement and ageing, the *English Longitudinal Study of Ageing* (ELSA).



ESDS Qualidata

Using secondary analysis to explore young men's relationships with their mothers



Martin Robb, senior lecturer in the Faculty of Health and Social Care at the Open University, has undertaken secondary analysis of data in his pilot study of young men's relationships with their mothers. Martin has previously carried out research on fathering and on men working in childcare, and has an interest in the formation of masculine identities and men's capacity to take on caring roles. For his current research Martin has analysed data from the ESRC funded qualitative longitudinal study *Inventing Adulthoods* based at South Bank University and archived in 2008 with ESDS under SN 5777.

Having examined qualitative interview transcripts from *Inventing Adulthoods*, specifically repeat interviews with five young men conducted over a period of 10 years, Martin is attempting to trace connections between maternal relationships and the

development of masculine identities. He has found that, although there are some drawbacks to doing secondary analysis of data from an existing study (secondary users have no control over the questions asked, and their own focus of interest will inevitably be only one among many topics covered), these have been outweighed by the advantages. In this case, the wide-ranging, longitudinal nature of the original research means that it has been possible to track the development of identities and relationships over time, and to set them in the broader context of participants' whole life experience.

Martin plans to write up the pilot as a case study and to use it as the basis for a broader study of young masculinities and family relationships, which will analyse both existing and newly-generated data.



ESDS Access and Preservation

Use of Road Accident Data



UKDA holds annual *Road Accident Data* (RAD) from 1985 onwards. The data are deposited by the Department for Transport and include information recorded on a report form by police officers in Great Britain. The details recorded include accident time, date and location; vehicle type, location and manoeuvre; driver details; casualty details and severity.

During the reporting period there were 76 individual users of the RAD data. These included 20 members of staff, 29 postgraduates and 10 undergraduates at institutes of higher education, and a small number of users from local and central government and the commercial sector.

Examples of research being conducted by staff at institutes of higher education included work commissioned and/or funded by the Department of Communities and Local Government, the Department for Transport, the Engineering and Physical Sciences Research Council and the Commission for Integrated Transport. Research areas, many using Geographic Information Systems (GIS), included the impact of road safety initiatives, accident modelling, and the

development of a spatial measure of risk to children from road accidents. The data were also used for teaching purposes to introduce students to working with real datasets.

Examples of postgraduate use of data included research for Ph.D and D.Phil. theses and for M.Sc. dissertations, for example for the module 'Traffic management' as part of the M.Sc. in Transport Planning and Engineering at Napier University, Edinburgh.

Undergraduate students used the data for dissertations and projects - one example was a study into how street layouts and topology affect both the likelihood and severity of road traffic accidents.

Users within the commercial sector were contracted to conduct research, for example for the Eastern Regions Fire Service and the Highways Agency.

Local government use of data included work being conducted by the Eastern Region Public Health Observatory on road traffic injuries in the East of England.



ESDS Government

Government surveys, especially repeated series, provide a key resource for UK social science. ESDS Government promotes and facilitates effective use of surveys such as the *General Household Survey* (GHS), the *Labour Force Survey* (LFS), the *Omnibus Survey* and the *Health Survey for England* (HSE).

ESDS Government teams at Essex and Manchester

The work of ESDS Government is divided between the UKDA at Essex and CCSR at Manchester. The Essex team is responsible for negotiating the deposit of data with the data producers, for checking the data quality and documentation, and ensuring the safe preservation of each file. They also disseminate the data online and, for most of the Government surveys, provide the data in Nesstar for the online exploration of metadata, frequency distributions and analyses.

The Manchester team is responsible for promoting and supporting the use of the government surveys in research and teaching, including answering user queries and producing enhanced ancillary materials, made available via the web site.

Promoting use of data in high quality research

ESDS Government plays an important role in helping the ESRC achieve its goal of increasing the capacity of UK social scientists to conduct high quality research using quantitative methods. Important elements in this strategy include presentations at high profile conferences and introductory workshops held at different venues around the country. Workshops take a topic-based approach to show the relevance of the government surveys to a wide range of social issues.

In the last year ESDS at Manchester has played a major role in an ESRC-funded project to increase the use of quantitative methods in undergraduate dissertations. The project involved running five half-day workshops that introduced students to the range of government surveys available and explained how to access data and how to conduct exploratory analyses. The project has been very successful in giving undergraduates the knowledge and the confidence to conduct secondary analysis. Three workbooks developed by this project are now available on the ESDS web site and are being promoted more widely to teachers and students across the UK.

The screenshot shows the ESDS Government website interface. The main heading is 'ESDS Government' with the tagline 'Economic and Social Data Service'. The navigation menu includes 'About', 'Data', 'Themes', 'Resources', 'News', 'Events', and 'Which service?'. The current page is 'Secondary Analysis for Social Scientists', featuring a section for 'SPSS workbooks for dissertation students'. This section includes introductory text and a table of available workbooks.

Topic	Dataset used	More...	Download
 Exploring the fear of crime	British Crime Survey Teaching Dataset 2000	Fear_of_crime (pdf)	Crime.pdf
		Fear_of_crime (word)	Crime.doc

On the right side of the page, there is a 'Latest News' section with links to various updates and a 'Upcoming Events' section.

ESDS Government also promotes its work overseas through conferences and invitations to participate in workshops and other meetings. The role played by ESDS in supporting the UK research and teaching community is invariably seen by overseas colleagues as setting an enviable standard of support and facilitation.

Getting results into the policy arena

ESDS Government is concerned to ensure that government departments who fund the major surveys are aware of the range and quality of research that is conducted via ESDS. Our research conferences, held in the British Academy, are designed to achieve this and are planned to be of particular interest to policy analysts. In 2007 our topic was consumption and papers covered consumption by older consumers and children as well as the consumption of food, culture and private transport. The meeting was very successful and papers are on the ESDS Government web site.

Contact and collaboration with data producers

Collaboration and interaction with data producers is of enormous importance. In the reporting year we have held six user meetings which have brought together data producers and data users to discuss survey developments and to hear about research findings, many of which have important policy implications.

A workshop on the *Expenditure and Food Survey* (EFS) – one of the most complex of the government surveys – brought together members of the Office for National Statistics (ONS) and the Department for Environment, Food and Rural Affairs (Defra) who are responsible for the survey, ESDS Government staff who support the survey, some expert users and some potential users who wanted more information about EFS. The afternoon included two hands-on sessions and materials from the workshop are being developed as a web-based resource to help others navigate their way through this complex but important survey.

Access

In the last year ESDS Government has contributed to ONS working parties on data access, necessitated by the Statistics and Registration Service Act 2007 which came into force on 1 April 2008. New procedures have defined the basis upon which distinctions should be made between identifiable (personal) data and anonymised or non-identifiable data. The Act now

gives researchers a clear basis for access to identifiable individual data through accreditation as an Approved Researcher. Approved Researchers enter a legally binding agreement with the National Statistician and are required to abide by tightly controlled conditions of data use with penalties of imprisonment for misuse. However, the standard End User Licence (EUL) remains the primary route through which users access the ESDS data collections.

Keeping in touch with users and producers

It is important to ensure that users are made aware of survey developments, forthcoming events and the availability of new resources. This is achieved through a twice-yearly ESDS Government newsletter and by regular emailed information sent to relevant JISCmail lists. A web-based database allows a search for publications that have been produced from government surveys. This has recently been extended to include a wider range of journals from more sources. An additional web-based resource holds information on the purposes for which the government surveys are being used, for those users who agree to this when they register for data.

Web resources and help desk

Web-based resources increase annually and provide an important source of information. In the reporting year a new search facility has been added that allows users to search across all ESDS government data sources to locate which surveys include specific topics of interest and, equally important, whether the same topic has been asked in successive year of the same survey.



Jo Wathan running an ESDS Government workshop

ESDS International

The last year has seen a great deal of activity within ESDS International. As well as the service implementing a new data access system, a greater range of user support and resource materials have also been produced to further promote the service.

Extending the user community

The development of a new community of users of international macrodata continues to be a high priority. In this regard 2007-2008 witnessed a highly significant and important landmark. In July 2008, access to socio-economic macrodata, previously restricted to UK academics and researchers, was opened up to the House of Commons and House of Lords libraries. Professor Ian Diamond, Chief Executive of the ESRC said: "As part of the joint ESRC and JISC investment in the ESDS Service, ESDS International was established to meet the growing need from researchers for access to key international databanks to make cross country comparisons and interpret their findings in a broader perspective. I am delighted that this excellent service will be opened up to support policy related research at the heart of government in the UK."

Online resources and value-added services

A key theme for the year has been the joining up of the macro and micro components of the service, providing users of both aggregate and survey data with a new entry point for online resources and raising awareness of the benefits of using both aggregate and survey data in analyses.

The last year has seen a number of resources added to the ESDS International web site under the new Resources section and an improved Support section.

The service launched its 'Case studies of international data use' at the beginning of 2008, highlighting the diverse use of international data in research and teaching. Researchers submit an outline of their research question, data sources, methods and key results for which they receive a £50 Amazon voucher if the study is published.

In keeping with its aim to provide value-added support materials, a new migration thematic guide has been added to the range of guides. For survey data users, construction of a new resource to promote the analysis of cross-national data has begun.



To keep the service's home page current and eye-catching, ESDS International has introduced a regularly updated 'On the international agenda' feature, commenting on key socio-economic events occurring world-wide as well as linking the feature back to data contained within the portfolio. Archived features are also available.

Developing research capacities

Countries and Citizens: Linking international macro and micro data (previously known as LIMMD), a major e-learning resource developed under the ESRC's Researcher Development Initiative, was updated and re-launched in June 2008. The open access materials comprise five units, each written by a subject specialist, and are designed as a self-guided learning resource combining international survey and aggregate data. In the three months since the re-launch, Countries and Citizens has been viewed an impressive 12,775 times by 5,238 users.

The service has continued its training programme with courses and workshops delivered around the UK including two introductory courses for parliamentary statisticians at the House of Commons Library. Hands-on workshops were also held at the School of Oriental and African Studies, the London School of Economics, Essex and Manchester.

Data portfolio enhancements

Over the year ESDS International has enhanced its portfolio with several exciting new additions. In November 2007, data from the *Latinobarómetro* survey – the licence for which was purchased with ESRC funding and was announced in last year's report – was made available to users within UK higher and further education.

The service has also enhanced its macrodata portfolio with the negotiation of access to the *IMF World Economic Outlook* database. This global time series database contains country-level projections for the next five years on topics such as national accounts, inflation, unemployment rates, balance of payments, trade and commodity prices.

"Every policy department looks beyond national boundaries... ESDS International is essential for the development of a cohort of researchers with the skills to undertake comparative analyses." Siobhan Carey, Head of Statistics, UK Department for International Development



Delivering a training course at the House of Commons library

The United Nations data re-distribution licence to replace the Common Database with the Comtrade database has been signed.

ESDS International on the road

ESDS International held their third annual conference on 3 December 2007 at the Institute of Materials in London. The event was well received by delegates who came from a wide range of educational, research and inter-governmental organisations to meet and share their ideas on issues relating to international data research. A full report and conference proceedings from all ESDS International conferences are available on the web site.

The team has continued with its outreach and awareness-raising activities including organising a session on international data at the IASSIST conference at the University of Stanford in California. A delegation of ESDS International was also invited to the Organization for Economic Co-operation and Development (OECD) Forum 2008: Climate Change, Growth and Stability in Paris. Members of ESDS International organised and presented a session at the Research Methods 2008 festival in Oxford in June 2008.

ESDS Longitudinal

From September 2007, Jack Kneeshaw took up the new position of Service Manager within ESDS Longitudinal. The creation of this post, responsible for both strategic vision and day-to-day management of the service, has enabled ESDS Longitudinal to develop its data portfolio further, to add and extend online resources for longitudinal data users and to work more closely with the depositors of the major longitudinal studies and also with potential depositors.

Developing the data portfolio

During 2007-2008 acquisitions activity was brisk for the newer longitudinal studies. Several new files were made available for the *English Longitudinal Study of Ageing* (ELSA), including the main data and documentation for wave 3 of the study. In addition, pensions data for waves 2 and 3 and mortgage grid data were deposited. The reporting period also saw the *Longitudinal Study of Young People in England* (LSYPE) 'go longitudinal' with the addition of waves 2 and 3 to the ESDS holdings. For the *Millennium Cohort Study* (MCS), data for wave 3 were acquired, alongside a number of Special Licence datasets allowing vetted users to match the main data to fine geographical identifiers and to each cohort member's hospital of birth.

For the longer-standing studies, activity concentrated on the deposit of updates to data already held, although one important acquisition was the National Child Development Study (NCDS) age 11 cohort members' essays. With Nuffield Foundation support, the Centre for Longitudinal Studies transcribed and coded a sample of around 500 essays. Cohort members were asked in 1969 to '*Imagine that you are now 25 years old. Write about the life you are leading, your interests, your home life and your work at the age of 25. (You have 30 minutes to do this)*'.

Although no new data were acquired for the *Families and Children Study* (FACS) during the year, this ongoing study was added to the longitudinal major studies pages in September 2007. Improving the visibility of the study has encouraged greater re-use of the data with more downloads in the 10 months following the study's promotion (i.e. to the end of the reporting period) than in the whole 12 months prior. This is



impressive as the release of a new wave or sweep of data generally creates a spike in usage and, as was noted above, the rise in the use of FACS has occurred despite no new data being made available.

It is also worth reporting on two new studies that have been made available in 2007-2008. Firstly, December 2007 saw the release of data from a new birth cohort study – *Growing Up in Scotland* (GUS). Secondly, adding to the expanding series of sampler datasets based on the major longitudinal studies, the Centre for Longitudinal Studies (CLS), under the aegis of an ESRC Researcher Development Initiative award, deposited a package of teaching datasets and associated pedagogical material based on the NCDS and *1970 British Cohort Study* (BCS70) for use in teaching quantitative methods. Each of these studies is expected to attract significant use in 2008-2009.

Longitudinal research resources

The existing user resources have been expanded or updated. Over the course of the year, two new resources have also been added. Using a software tool developed by ESDS Government, the project titles and abstracts of those users who have access to the seven major longitudinal studies – and who have agreed to share this information with others – may now be browsed via the 'Usage' link provided for each study. This provision of information on users' projects complements another new resource - the cited publications search.

Access to finer geographies

At the outset of the ESDS in January 2003, ESDS Longitudinal was charged with exploring methods by which researchers could make use of sensitive geographical data, without prejudicing the confidentiality of survey respondents. Five and a half years on, ESDS Longitudinal now provides Special Licence access to six low-level geography datasets for the three birth cohort studies, with plans in place to release six more datasets containing geographical markers for the *British Household Panel Survey* (BHPS). Over 40 users have so far taken advantage of these previously restricted data becoming available.

In addition, where access to sensitive geographical data remains restricted, ESDS Longitudinal continues to update the database of geography variables, enabling researchers to remain fully informed as to the procedures in place for access to controlled data outside the ESDS system.

Involving key stakeholders

Close relationships with the depositors of the major longitudinal studies are central to the operation of ESDS Longitudinal. The service continues to liaise directly with the key data depositors over issues such as promotion and publicity, monitoring study usage, access to sensitive data and training novice and established users. Equally importantly, the service continues to build relationships with a number of units managing longitudinal data resources not currently held by the ESDS. Most notably, the Service Manager continues to be a board member on the *National Survey of Health and Development* (NSHD, also known as the 1946 birth cohort study) Data Access Project, a Medical Research Council (MRC) funded initiative which aims to make data from the 1946 birth cohort study more widely available for secondary analysis.



ESDS Qualidata

In this reporting period ESDS Qualidata contributed to a number of ESDS user support materials. Notably, it provided content for the joint ESDS themes pages on health and health-related behaviour. A case study introducing the *Severalls Hospital: Interviews for 'Madness in its Place'*, 1913-1997 dataset and a web-based video demonstrating searching ESDS Qualidata for health topics were created. Further, re-use case studies and literature on secondary analysis of qualitative data were updated to profile new exemplars.

Significant collections processed

A number of important new data acquisitions were processed and enhanced. Most prominent is Dennis Marsden's study, *Mothers Alone: Poverty and the Fatherless Family, 1955-1966* (SN 5072). Based on 116 in-depth, semi-structured interviews with respondents from Essex and West Yorkshire, it explores the lives and experiences of mothers living alone through investigating what is poverty and who are the fatherless. It is available via ESDS Qualidata Online following extensive enhancement – scanning, full optical character recognition (OCR) and Extensible Markup Language (XML) mark-up – of the interview transcripts.

Another important acquisition was the NCDS age 11 cohort members' essays (see ESDS Longitudinal). Now available as downloadable anonymised files, these essays represent a rich and unique source for both qualitative and mixed methods researchers.

Collaboration in qualitative data sharing and archiving

Every year ESDS Qualidata works collaboratively with other UK institutions and projects. One important partnership has been formed with the £4.5 million ESRC Resources Board project, Timescapes. This qualitative longitudinal study tracks the lives of 400 people over 5 years. As well as sitting on their Advisory Board, Libby Bishop, the former manager of ESDS Qualidata and ongoing ESDS Qualidata researcher, works part-time as the Timescapes archiving manager. Recognised as a 'satellite' data archive, Timescapes examines methods and procedures for handling and sharing complex qualitative longitudinal multimedia data.



ESDS Qualidata staff continue to collaborate with both the National Centre for Research Methods (NCRM) (consent projects), the National Centre for e-Social Science (NCeSS) (multimedia display projects and ethics) and various JISC projects including the National Centre for Text Mining (NaCTeM) (text mining of qualitative data) and the Repositories Programme. In addition ESDS Qualidata staff have endeavoured to work with university research ethics committees to highlight and remedy the apparent tension between data protection (destruction or retention of personal data by researchers) with the need for data sharing, as increasingly required by research award contracts.

Access to multimedia data

ESDS Qualidata has always welcomed opportunities for further technical development of its online data access systems.

In October 2007, the JISC-funded DExT project, led by Louise Corti, released the initial version of its QuDEx schema followed by an updated version 3 in March 2008. This schema aims to enable the transfer of coded and annotated data from one Computer-Assisted Qualitative Data Analysis Software (CAQDAS) package to another. It also provides an intermediate format for longer-term preservation and sharing. QuDEx defines and describes several core CAQDAS concepts required for the interchange of multimedia data such as defining extracts, codes, classifications (variables), memos and relationships. The work has included liaising with CAQDAS software package vendors to



discuss interoperability and a willingness to export to a common format, as well as collaboration with the Open Data Foundation on the development of robust XML schemas. Further development and collaboration will continue with the schema being more widely promoted in the next reporting period.

Efficient audio dissemination has also been prioritised in this reporting year to bring in more interview audio collections. Providing access to social science audio-visual materials is challenging due to issues such as ethics, consent, recording quality and storage capacity. ESDS Qualidata aims to minimise data storage requirements while simultaneously maximising audio quality through sustainable open preservation formats and optimal and flexible delivery formats. Staff have investigated new compressed and uncompressed open formats such as Free Lossless Audio Codec (FLAC) and Ogg Vorbis, in addition to testing some optimal streaming and MP3 download solutions. The launch of collections using these procedures is planned for the next reporting period.

Publicity and outreach

As always, ESDS Qualidata places high priority and emphasis on supporting users. Users do not only download data, they are also data creators, methodologists, teachers and other groups wishing to learn more about the art of archiving qualitative data.

The team delivered workshops to Ph.D. classes in psychoanalysis, economics and sociology, as well as more focused ones on re-using data and consent and confidentiality. Four workshops, two as part of the ESRC Festival of Science, have been held in this reporting period on the complex issues of informed consent and anonymisation in qualitative research. A successful

workshop was also held with teachers introducing the qualitative interviewing methods resource created in the year 2006-2007.

The team attends key meetings and dedicated conferences where possible to promote the diverse elements of the service such as preservation. Staff were invited to speak at two conferences on qualitative data in Europe at Data Archiving and Networked Services (DANS), the Netherlands, and at the National University of Ireland-Maynooth, as well as at a major workshop on research ethics in Sheffield. In addition, ESDS Qualidata hosted numerous visitors interested in the challenges of archiving qualitative collections.

Staff continue to publish on various aspects of data sharing, archiving, accessing and re-use. A prominent example is Louise Corti's contribution, with other ESDS staff, to a chapter in Nigel Fielding et al. (eds.) (2008) *The Sage Handbook of Online Research Methods*.

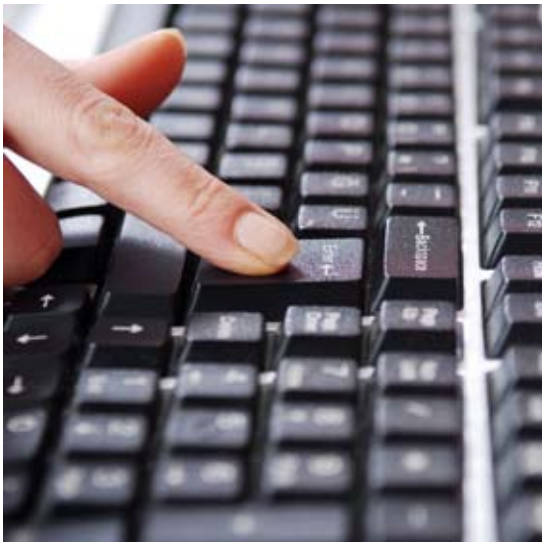


ESDS Access and Preservation

On 9 July 2008, ESDS launched a trial switch-over from Athens authentication to Federated Access (Shibboleth). Following the successful trial, Shibboleth access was launched on 1 August 2008, just outside the reporting period.

Shibboleth

Most institutions use Shibboleth in a simple one-to-one communication between the Identity Provider (IdP) and the Service Provider (SP). The IdP is the user's home organisation and is responsible for authenticating the user and passing information requested to the SP in the form of attributes. Based on the information sent in the attributes, the SP then decides whether to allow or deny access to that user. Shibboleth is set up to protect the identity of the user so that their personal details remain anonymous to the SP. The ESDS model is more complicated, however, in that it provides a one-stop registration service for multiple distributed SPs and ESDS needs to know the identity of a user in case there is a breach of its terms and conditions. Also some data collections held at certain ESDS SPs require an agreement to special conditions.



Existing users were contacted prior to 1 August and asked to log in to their account and provide a memorable word. To retain continuity of usage and avoid re-registration, existing users switching from Athens to the UK federation are required to provide some limited personal information, including the memorable word.

Approved Researcher

During 2007-2008 changes also occurred in the methods used to access ONS Special Licence data - data containing sensitive or detailed variables judged to require additional access controls. When the Statistics and Registration Service Act 2007 came into force on 1 April 2008 it brought an end to existing Special Licence arrangements. Access to these data continued after 1 April but under a new legal framework set out in the 2007 Act.

Any registered users wanting to access ONS data previously available through a Special Licence now will have to be accredited by the UK Statistics Authority as an Approved Researcher. Applicants for Approved Researcher status are expected to provide evidence of previous research projects and publications and new users may need to provide the contact details of a senior member of staff who can vouch for them. A research purpose form asks for information about the intended use of the data, a justification for access and a summary of planned outputs. The new system is being managed by UKDA on behalf of ONS.

LFS reweighting project

During 2007, the ONS undertook a reweighting project, whereby *Labour Force Survey* (LFS) and *Annual Population Survey* (APS) data were reweighted using the latest population estimates for 2007-2008. Reweighting work is undertaken by ONS approximately once every five years, so that data are kept as current as possible to enable accurate analysis of the labour market. The last reweighting project took place in 2002-2003.

As a result of the 2007 project, reweighted editions of the LFS series back to 1992 were deposited at the UKDA during early summer 2008. Some further amendments were also made to these data following collaborative negotiations between ESDS and ONS on disclosure risk, meaning that some previously unavailable variables were reinstated within the data, enabling researchers to undertake more detailed analysis. The initial deposit of reweighted LFS and APS data comprised around 300 new studies and new editions, causing a significant increase in processing workload and subsequently a large peak in dataset releases during July and August 2008. Further deposits of new and updated data generated by the reweighting project are set to continue throughout the remainder of 2008.

Back office developments

This reporting period saw a focus on data workflow, from acquisitions through to ingest, enhancement, preservation and dissemination. As part of the process, existing procedural documents have been enhanced, and new ones produced. These include a new document consolidating the robust data security procedures already in use at ESDS, which will also inform the upcoming ESRC Secure Data Service. The new workflow arrangements have also introduced a tight programme of periodic document review, enabling consideration of the effects of relevant external developments such as the 2008 Cabinet Office report on Data Handling Procedures in Government (also known as the Hannigan Review) on ESDS procedures.

Preservation policy

In March 2008 the UKDA published version three of its Preservation Policy. It represents a complete restatement of the policy, and the principles which inform it. It articulates the actions underpinning the activities of the ESDS and provides a detailed explanation of how and why the UKDA ensures the authenticity, reliability and logical integrity of all resources entrusted to its care while also providing usable versions for research, learning or teaching. The policy is partially based on the international standard (ISO 14721) Open Archival Information System (OAIS) reference model, which provides a conceptual framework for entities and relationships between entities in digital preservation environments.



The complete implementation of this policy over the next few years will ensure that the UKDA and the services which it provides will function at an international gold standard, while continuing to meet the precise needs of ESDS user communities. A review of all existing internal processes commenced in the summer of 2008 to provide a high-level reformulation of preservation metadata needs and to see how JISC application profiles can be incorporated into future workflows. The JISC-funded DExT project has also influenced strategic direction in understanding the significant properties of statistical data files, which will promote digital preservation activities, and ensure the provision of continuous access to the ESDS's collections of data.

ESDS Management

ESDS Management is hosted by UKDA and exercises overall direction and management of ESDS. The management and co-ordination function provides a broad strategic role, relating to a variety of stakeholders concerned with the supply, funding and use of social science data.

One of the most important aspects of this function is liaison with stakeholders and this is to a large degree carried out through biannual meetings of the ESDS Advisory Committee (membership listed on the inside back cover). During the year the Committee met on 2 November 2007 and 1 April 2008. The Chair until February 2008 was Jil Matheson from ONS.

ESDS Management is active in promoting ESDS to the widest possible range of stakeholders around the world, and the individual ESDS sites host many visitors. During the reporting period this included researchers from as far afield as Canada, New Zealand, Australia, Russia, Japan, Taiwan and South Africa, as well as numerous visitors from around the UK and Europe.

Close co-operation between the Universities of Essex and Manchester is of course essential for effective delivery of ESDS, and this is ensured by regular meetings of staff. Throughout 2007-2008 there was a regular series of such meetings, culminating in an awayday in London on 28 January 2008 involving all key teams in ESDS.

UKDA-store

An important additional service launched at the end of the reporting year was the UKDA-store, a self-archiving service for all ESRC-funded projects. Research data and output sharing is an important part of publicly funded research, and research funders are increasingly implementing formal data sharing policies in line with high-level recommendations and policies made by the OECD and others. The ESRC was one of the first UK funding bodies to initiate a data sharing policy and fund an archive to house research data generated as a result of its funding. The fact that a rich and diverse stock of social and economic data is now available to the wider community is a testament to this forward thinking.

UKDA-store is complementary to the formal preservation and dissemination system for data that is offered by the ESDS via the ESRC Data Policy. While the UKDA-store system can hold all kinds of digital objects from numeric and textual datasets to technical and research reports, it can also link to outputs held in other repositories. UKDA-store will enable a greater number of research data outputs to be shared by investigators, in cases where ESDS may not have the resources to acquire and store these data, or where the data simply do not fit the ESDS collections development policy. UKDA-store, developed with funding support from JISC, uses a state-of-the-art open-source repository system based on Fedora to allow contributors to manage their own data and research outputs. The benefit to users searching for a range of research outputs is that the system allows linking between research funding information, research outputs, publications and archived data sources.

Ian Diamond, Chief Executive of the ESRC, speaking at the launch of UKDA-store at the Research Methods Festival, welcomed the system as "a tremendous advance for UK social science research".

Themed web pages

During 2007-2008, ESDS embarked on the development of a suite of web pages to aid researchers looking for data on a particular theme. Towards the end of the reporting year the first of these pages was launched, on the theme of health and health behaviour. Data have been collected and are made available via the ESDS on topics as varied as the experience of illness, child development, access to care, lifestyle behaviour, subjective physical and



mental well-being, diet and nutrition, immunisation programmes and attitudes towards health service provision. Data on health can cover not only a person's status, behaviour, attitudes and expectations but also the provision of health care, including the mechanics of policy making, government expenditure and service coverage.

The ESDS thematic health pages contain case studies, pointers to key data sources, ready-made searches demonstrating how to find health-related data, and web-based videos to showcase ESDS's online data exploration tools. The ESDS resources section contains overviews and work-through guides and presentations. The presentations are from workshops organised by ESDS where expert speakers talk about data they have created or re-use of data. The web-based videos demonstrate how researchers can view frequencies, conduct simple online tabulations, and produce graphs and subsets for a selection of health-related data, all using ESDS's online data exploration tools.

Some of the visitors welcomed to ESDS during 2007-2008

Anna Beletskaya and Anton Mityushin from the Sociological Data Archive based in Moscow, Russia were typical of many visitors coming from young, developing data archives. They spent three days at Essex, during which time they shadowed the work of most of the data processing and user support sections of UKDA, talking to staff about what they do and how they do it, as well as receiving presentations from staff on aspects of the UKDA work flow and processes. Equally, they also gave a lunchtime seminar to UKDA staff on the Russian Sociological Data Archive.

Data archiving and dissemination services such as ESDS are not as well developed in Africa as they are in Europe or North America. In part, this is due to an immature culture of data sharing, which is true of most of the academic disciplines. In order to gain insight of how this has been changed elsewhere Faye Reagon of the Human Sciences Research Council in South Africa visited the UKDA spending time with the Director and other senior staff. She was particularly interesting in the role funders can play to facilitate and widen the culture of research data sharing, and afterwards visited staff at the ESRC to continue her discussions.

David Ingram and Humphrey Gyde of the Clinical Decision Support Oral History Project, London – a oral history project which has undertaken extensive in-depth interviews in both the UK and USA with a number of pioneering medical practitioners – visited members of ESDS Qualidata to gain a better understanding of how interview texts are marked up and prepared for online browsing and searching. The importance of capturing adequate and appropriate contextual materials, which could also be used in a teaching and learning context was also stressed.

Dr Ly-Yun Chang, Executive Director of Survey Research at Academia Sinica in Taipei, Taiwan and her colleague Dr Lin spent time in Manchester with staff of both ESDS Government and ESDS International, and also met with the Director of ESDS at a specially arranging meeting in London. Thus they received a comprehensive overview of how ESDS operates and what the key drivers are. One particular concern, which was discussed at length, was the treatment of sensitive and potentially disclosive data, a subject of great concern and topical relevance.



l-r Kevin Schürer, Anna Beletskaya, Anton Mityushin and Margaret Ward

ESDS communicating and engaging during 2007-2008

August 07

ESDS Management meets with David Ryan, Buckingham Palace

ESDS Management meets with Stephen Pennick, Cynthia Clark, and Mike Hughes of ONS

ESDS Management contributes to the Information Communication Technology Meeting, Bristol

September 07

ESDS International contributes to a UK Beyond 20/20 User group meeting, hosted by Forvus, Bath

ESDS Management meets with John Pullinger, House of Commons Library

ESDS staff present at the Association for Survey Computing international conference - the challenges of a changing world: developments in the survey process, Southampton

ESDS Government presents at the European BSD conference, Copenhagen

ESDS contributes to the Producers and Users Group meeting with ONS and the Statistics User Forum, ONS, London

ESDS International presents at the SAGE Cultural Assets Fair, University of Manchester

October 07

ESDS staff present at NatCen, London

ESDS staff contribute to the NCeSS nodes commissioning panel, London

ESDS Management contributes to the Belgium Social Data Archive workshop, Brussels

ESDS Longitudinal attends a NDCS/BCS70 workshop for data users planning studies on psychological health, subjective well-being and life satisfaction

ESDS organises an ESDS training day for Maidstone and Tunbridge Wells National Health Service (NHS) Trust

ESDS Government organises a workshop: Working with survey files: linking, matching and using hierarchy, Manchester

November 07

ESDS Government runs its annual research conference on the theme 'All consuming: consumption research findings from government surveys', British Academy, London

ESDS International runs an introductory workshop, Manchester

ESDS International runs a workshop at the London School of Economics and Political Science, London

ESDS Management contributes to the Timescapes/UKDA collaborative meeting, Essex

ESDS Management contributes to the Data Centres Forum, London

ESDS International presents at the ESRC meeting 'Enhancing the UK Social Science Skills Base in Quantitative Methods: Developing Undergraduate Learning', Royal Institute of British Architects, London

ESDS Qualidata presents at the East of England Regional Archive Council 'Sounds Good – Improving Sound Archives in the East of England' workshop, Imperial War Museum, Duxford

ESDS Longitudinal runs an introductory workshop to the Millennium Cohort Study, London

December 07

ESDS International runs its international annual conference on the theme 'Data without borders', Institute of Materials, London

ESDS Management participates in the UK Vice-Chancellors and Funders workshop, Research Information Network (RIN), London

ESDS Government runs an introductory workshop to Stata, Manchester

ESDS Management presents at a Council of European Social Science Data Archives (CESSDA) meeting, Eurostat, Luxembourg

ESDS Qualidata contributes to Mixed Media Grid, London

January 08

ESDS staff meet with London Organising Committee for the Olympic Games, and ESRC concerning data outputs from the impact study, London

ESDS Management chairs and contributes to the European Strategy Forum on Research Infrastructures (ESFRI) meeting, Amsterdam

ESDS Government runs introductory workshop to Stata, Manchester

ESDS International runs a seminar on data resources for Africa and East Asia at School of Oriental and African Studies, London

ESDS staff contribute to ONS meeting of Statistical Disclosure Control microdata standards working group

ESDS Qualidata gives an invited talk on automated indexing of survey questionnaires and interviews, NaCTeM, Manchester

ESDS awayday, London

ESDS Management and ESDS Qualidata contribute to the Timescapes launch and Advisory Board meeting, Leeds

February 08

ESDS Management contributes to the RIN awayday, York

ESDS Management contributes to the ONS I-dissemination strategy meeting, Statistics User Forum, London

ESDS Management presents at the 1st African Digital Curation Conference, Pretoria, South Africa

ESDS Management contributes to the *UK Household Longitudinal Study* (UKHLS) Communications Advisory Group, London

ESDS Management is invited speaker at the JISC/RIN Data Preservation and Curation seminar, London

March 08

ESDS Qualidata runs a workshop 'Establishing and sustaining qualitative data archiving', National University of Ireland, Maynooth, Ireland

ESDS staff run a workshop on managing and sharing research data, University of Essex, Colchester, and University of Leeds

ESDS International presents at the Brooks World Poverty Institute meeting, University of Manchester

ESDS Qualidata gives a keynote speech at the Key Ethical Issues for Social Science Research workshop, Humanities Research Institute, University of Sheffield

ESDS Longitudinal runs a workshop on the NCDS/BCS70 for potential users, Institute of Education, London

ESDS Management contributes to the launch of the UK Statistics Authority, London

ESDS Government runs a workshop on starting Stata, Manchester

April 08

ESDS Qualidata presents at the Open Repositories 2008 conference, University of Southampton

ESDS International runs a workshop: Using the ISSP and the ESS: welfare provision and social inequality in a cross-national perspective, Essex

ESDS Qualidata runs a workshop on data resources for teaching qualitative methods, Essex

ESDS staff attend the Family Resources Survey user meeting, Royal Statistical Society (RSS), London

May 08

ESDS Government contributes to a meeting on Vital Statistics, ONS, Titchfield

ESDS Longitudinal runs an introductory workshop to the MCS3, Institute of Education, London

ESDS Government and ESDS International meet with Ly-Yun Chang, Center for Survey Research, RCHSS, Taiwan

ESDS staff give multiple presentations and workshops at International Association for Social Science Service and Technology (IASSIST) 2008, Palo Alto, USA

June 08

ESDS International contributes to the OECD Forum 2008: Climate change, growth and stability, Paris

ESDS Longitudinal runs a workshop on the NCDS/BCS70 for potential users, University of Stirling

July 08

ESDS Government presents at the 4th International Conference on e-Social Science, Manchester

ESDS Management presents at the 13th Government Statistical Service (GSS) Methodology Conference, London

ESDS Management meets with Sir Michael Scholar, UK Statistics Authority, London

ESDS Government presents at the International Conference on Survey Methods in Multinational, Multiregional, and Multicultural Contexts (3MC), Berlin

ESDS presents at the Research Methods Festival, Oxford

ESDS Management presents at the JISC/JANET dinner and debate: "UK Research is Failing", Oxford.

ESDS staff present at the Mimas Open Forum, Manchester

ESDS Qualidata presents at the British Sociological Association (BSA) Food Study Group Conference: Food, Society, and Public Health, London.

ESDS International delivers training courses to researchers and statisticians from the House of Commons and the House of Lords, London.

ESDS Government presents at the New Health Information Specialists Conference, RSS, London

ESDS staff present at the Digital Preservation Planning: Principles, Examples and the Future with Planets workshop, British Library, London

Measuring ESDS

This section produces a statistical summary of activity across all sections of ESDS. Figures are also produced in most cases for the previous reporting year in order to enable temporal comparisons to be drawn. However, it is equally important, if not more so, to place the statistical evidence in a long-term context in order to provide a broader perspective. This represents a headline digest of the most interesting statistics collected about the service. For those requiring more detailed statistical information on the working of ESDS, the service provides quarterly statistical returns reporting against a Service Level Definition (SLD) to the JISC Monitoring Unit (MU). These are available at www.esds.ac.uk/news/jiscmu.asp.

Data acquired and processed

With the exception of the important international macrodata series, the data provided by the various sections of ESDS are acquired, processed and mounted on the download service for online access via ESDS Access and Preservation at the UKDA. Due to the special licensing arrangements relating to the international macrodata series these are acquired directly by ESDS International at Mimas, then processed and loaded onto the Beyond 20/20 data delivery server in Manchester.

As shown in **Table 1**, overall throughput of data in terms of acquisitions and processing rose sharply, compared to the figure for the previous reporting year – 716 new data collections or updates, compared with 485, a rise of 48 per cent. Yet this figure requires further explanation. It is somewhat artificially high due to the fact that there was an influx of some 200 *Labour Force Survey* studies which were newly acquired, processed and made available to users due to reweighting of the underlying data. In turn this had an impact on the capacity to ingest additional materials, which is why, in part, the number of qualitative data processed in the year declined (from what was an unusually high figure the previous year).

Enhanced data processing, for release and analysis via either the 20/20 or Nesstar systems, also increased over the past year (**Table 2**). The increase of 40 per cent was largely due to a number of new versions of the key international macrodata series becoming available and being processed for use.

Data access

Increasingly users wish to receive data for either research, teaching or learning in a variety of ways and forms, and one of the challenges that continuously faces ESDS is meeting these changing needs and requirements. In order to accommodate the wide range of data needs ESDS provides a combination of data access tools. Because of these multiple routes to, and forms of, data access it is not a straightforward task to measure data usage and certainly not easy, or in some cases appropriate, to attempt to compare the various usage figures across the component parts of ESDS.

In outline, users of ESDS can access data in the following ways. For survey-type microdata, across all services, registered users have the option of downloading a data bundle directly from the data catalogue. These bundles relate to whole data collections or studies which in many cases will contain multiple datasets, and will also include files containing documentation, user guides and notes. Users can also request such data bundles to be made accessible to them by machine-to-machine file transfer protocol (FTP), or written to a fixed portable device (e.g. CD/DVD/USB) and sent to them, although the latter is declining significantly in demand. Data bundles can also be made available to users indirectly via third party access arrangements, most commonly university-based data libraries and sister data archives. All these forms of data download are reported in **Table 3a**. A subset of the micro survey data, primarily those supported by the four specialist units of ESDS, is also made available to users via the Nesstar online data browsing and visualisation tool. This allows users to download whole data collections as mentioned above, but also permits the user to define bespoke subsets, filtering on one or more variables within a given dataset. Additionally, the system can be used to generate aggregate tabular data in the form of two- and three-way tables, as well as graphical representations of these in the form of histograms, piecharts and in some cases maps. All of these can be saved to a local device (downloaded) or hyperlinked for sharing or future reference. Only the download of whole data collections is recorded in **Table 3a**. A further complication arises from the fact that any of the microdata collections enumerated in **Table 3a** can also be redistributed locally to students using the data for instructional purposes.

Table 1 Data collections acquired and processed for online delivery

Service	2006-2007			2007-2008		
	New data collections	Updates and new editions	Total	New data collections	Updates and new editions	Total
ESDS Core	82	26	108	87	22	109
ESDS Government	90	32	122	277	150	427
ESDS International	24	131	155	14	96	110
ESDS Longitudinal	26	11	37	26	20	46
ESDS Qualidata	63	0	63	22	2	24
Total	285	200	485	426	290	716

Table 2 Datasets further processed for value-added delivery via Nesstar and Beyond 20/20

Service	2006-2007			2007-2008		
	New datasets	Updates and new editions	Total	New datasets	Updates and new editions	Total
ESDS Core	4	0	4	5	2	7
ESDS Government	32	28	60	67	7	74
ESDS International*	2	41	43	1	68	69
ESDS Longitudinal	2	1	3	3	1	4
Total	40	70	110	76	78	154

*International macrodata are delivered via Beyond 20/20; all other services via Nesstar

Table 3a Datasets delivered to users by access method

2006-2007	Download	FTP	Via/access	Portable	Nesstar	Total
ESDS Core	9,658	283	1,192	10	15	11,158
ESDS Government	22,143	163	3,898	7	279	26,490
ESDS International*	551	499	1,485	24	6	2,565
ESDS Longitudinal	6,195	21	435	11	26	6,688
ESDS Qualidata**	932	20	84	1	0	1,037
Total	39,479	986	7,094	53	326	47,938

2007-2008	Download	FTP	Via/access	Portable	Nesstar	Total
ESDS Core	9,639	500	1,121	11	34	11,305
ESDS Government	22,052	689	4,639	2	353	27,735
ESDS International*	1,118	387	598	0	10	2,113
ESDS Longitudinal	6,073	71	357	3	125	6,629
ESDS Qualidata**	728	7	63	0	0	798
Total	39,610	1,654	6,778	16	522	48,580

* international micro datasets only, macrodata series counted separately (see Tables 3b and 4)

** Nesstar does not include qualitative data

Table 3b Data Reports generated for international macrodata via Beyond 20/20

2007-2008	Total
Total number of Data Report downloads	91,205*

* this figure includes Data Reports generated only for the period 1 September 2007 - 31 July 2008, to reflect when recording of this information began

Table 4 ESDS International macrodata accessed via Beyond 20/20

	Total number of data analysis sessions	Individual users
2006-2007	102,855	5,753
2007-2008	105,587	6,340

Table 5 Use of Nesstar online microdata browsing and analysis service

	Total number of data analysis sessions	Individual users
2006-2007	12,313	6,289
2007-2008	16,995	8,719

International macrodata are supplied to users via the Beyond 20/20 software interface. This allows the user to create tabular statistics in the form of Data Reports, as well as graphical outputs, similar to those of Nesstar described above. These can also be downloaded to a local device. **Table 3b** reports the number of Data Report downloads.

It is important also not to overlook the access given to registered users to online qualitative data collections. These consist mostly of open-ended interviews and responses and are not really downloaded or saved in the same ways as the structured data collections mentioned thus far. Users can download related audio and image files from the collection but more usually might cut and paste snippets of text, thus making measurement problematic.

With this outline of data access methods established, **Table 3a** shows that overall the number of microdata collections delivered to users rose slightly on last year's figure, but remained just under the 50,000 figure. Within this figure, the basic distribution of data access between the various services supporting microdata also remained roughly the same. The

available statistical information would therefore suggest that after a period of year-on-year increase in usage during the first five years of ESDS, data access appears to be plateauing at a new high level. However, this trend must potentially be evaluated in a longer-term perspective, since data usage and demand can be cyclical, influenced by such factors as the UK Research Assessment Exercise (RAE) reporting period. It is important to note that these figures do not include counts of data redistributed for us in classrooms. Unfortunately, these data are notoriously difficult to capture since lecturers do not always return class registrations in a timely fashion. As a result figures are always at best approximate, but these add roughly a further 50,000 users accesses, in broadly the same proportions by service as shown in **Table 3a**. It should also be noted that in addition to data provided to users from the collection curated and hosted by ESDS, 460 international microdata collections were acquired from sister archives and disseminated on behalf of UK-based researchers.

Turning to the experience of use for international macrodata, for the first time, this year's report produces a figure for the number of Data Report

Table 6 User queries received by ESDS service responsible for answering

Service	2006-2007			2007-2008		
	Queries	Referrals	Totals	Queries	Referrals	Totals
ESDS Management	2,029	4	2,033	2,335	15	2,350
ESDS Government	285	51	336	259	70	329
ESDS International	284	16	300	288	19	307
ESDS Longitudinal	77	0	77	74	0	74
ESDS Qualidata	200	1	201	199	0	199
Total	2,875	72	2,947	3,155	104	3,259

Table 7 New registrations and active registered users

	2006-2007	2007-2008
New registrations	22,485	20,938
Total active registered users	47,635	46,304

downloads. **Table 3b** records this as an impressive 91,205, clearly indicating the high demand and usage of these collections of data. Likewise access of the 20/20 system witnessed a further increase in the number of users, increasing by 10 per cent from 5,753 to 6,340 with the number of registered data analysis sessions rising slightly compared with those of last year, up 3 per cent to 105,587 (**Table 4**).

Those using the Nesstar system for the online analysis and tabulation of microdata also saw a notable increase, with both the numbers of users and the number of registered data analysis sessions increasing by a factor of some 38 per cent (**Table 5**).

Lastly, in addition to data access by users it is important to note a significant increase in the number of downloads of data documentation, user guides, reports and manuals. One measure of this is that PDF downloads increased by 109 per cent compared to last year, reaching a figure of 1,973,215. Such guides and reports are also available in a number of other formats including HTML, RTF and Microsoft Word files.

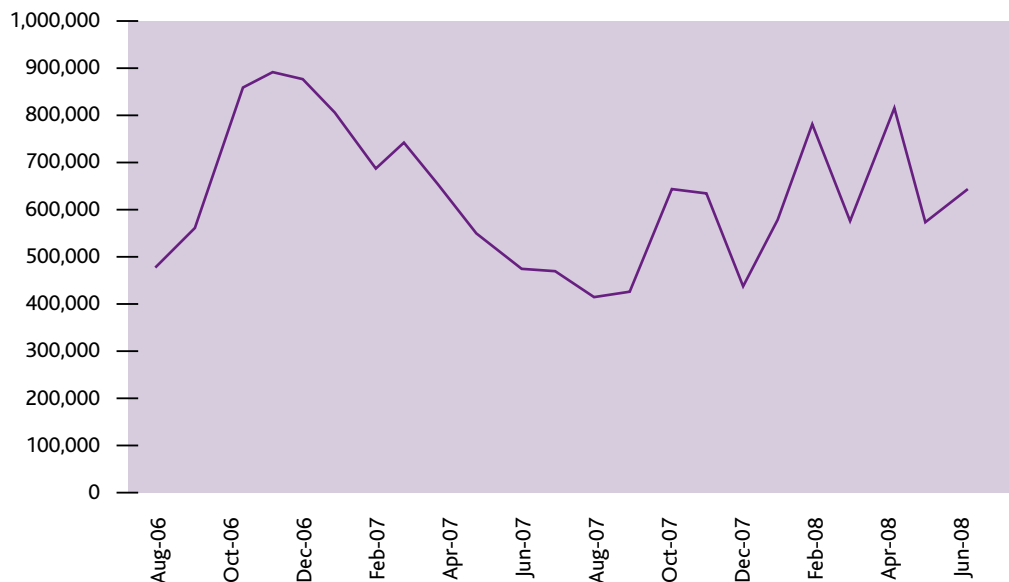
User queries

The number of user queries increased over last year's figures by 312, with all of the increase (11 per cent) being witnessed by the central ESDS help desk, which handles queries of a non-specialist nature (**Table 6**). In part this was due to the first phase of the switch over from the Athens to the Shibboleth-based UK Access Management Federation during this year. Queries to the specialist services remained at roughly the same levels as the previous year.

Registrations

The past year saw a decline in the number of new user registrations for the first time since ESDS was established in 2003, a drop of 1,547 or 7 per cent on the previous year (**Table 7**). Overall, however, the number of active registered users within the year remained broadly the same, at 46,304, a slight decrease. In part this small decline is a feature of technological change. In past years UK-based researchers wishing to gain access to important cross-national microdata surveys such as the Eurobarometers and ISSP, needed to register first with ESDS, whereas in the past year these data have been available directly from a sister data archive – the Zentralarchiv in Germany – and users are redirected there where they can obtain the data online directly.

Figure 1 Web site use August 2006 to July 2008 (page hits)



Web site usage

As the previous section on data access demonstrated, the internet is clearly the major tool for the delivery of ESDS products and services. However, many web statistics can be misleading since uncorrected counts will give an artificially high figure due to the site being trawled by so-called web-crawlers as well as page calls to image files (pictures, logos and the like). Once these corrections have been implemented, the total number of page hits in the reporting year was 7,068,221 compared to a figure of 8,076,208 for the previous year.

However, interpretation of these figures is not easy, as recourse to the monthly data for the two-year period August 2006 to July 2008 (**Figure 1**) shows. Web hits peaked at a figure of around 900,000 in December 2006, dropping to a monthly low of 400,000 in September 2007, then rising again through peaks and troughs to around 650,000 web hits per month in July 2008.

Usage by user type

Breaking usage of the various services down by type or section of user, as shown in **Table 8**, illustrates that the principal user groups are drawn from academic staff and students at institutions of higher education, ranging from 85 per cent (ESDS Government) to 96 per cent (ESDS Qualidata). Public sector workers (central and local government, charitable organisations, non-governmental organisations (NGOs), etc.) tend

to account for between 8 and 11 per cent of usage for most services. The exception to the general trend is with international macrodata, where licensing arrangements imposed by the data owners restrict use primarily to the academic sector. Here use is more heavily concentrated towards educational users, who account for some 81 per cent of use, supporting the building of future research capacities. However, the fact that student use is not uniformly recorded across all services must also be taken into consideration.

Usage by discipline

Table 9 shows the relative proportions of active users in the most popular disciplinary areas for each of the specialist services. With some variations, the largest disciplinary users across all services are economics and sociology, with international macrodata being particularly heavily used by the former, and less so by the latter. Whilst for qualitative data, the reverse is true. International macrodata, more so than others, are also widely used by the business/accounting/finance sector, which is not surprising given the nature of much of these data. What is perhaps most striking is the fact that all types of data have a broad and wide-ranging disciplinary appeal, suggesting that ESDS reaches much further than the core social sciences, including health, medical and humanities researchers in its varied user community.

Table 8 Usage by user status

	Percentages					
	Core	Government	International (macro)	International (micro)	Longitudinal	Qualidata
Academic staff	31.2	30.2	13.8	30.3	37.7	25.0
Undergraduate	19.5	21.0	34.6	8.9	13.1	26.6
Postgraduate	36.8	33.5	46.0	45.9	38.1	44.1
Public Sector	8.0	11.3	0.0	9.8	8.2	0.9
Commercial	1.2	1.6	0.0	0.9	1.0	0.3
Other	3.3	2.4	5.5	4.3	2.0	3.1
Number of users	2,025	2,799	6,340	327	1,467	320

Table 9 Usage by discipline of user

	Percentages					
	Core	Government	International (macro)	International (micro)	Longitudinal	Qualidata
Business/Accounting/Finance	11.8	6.1	16.0	10.7	4.4	13.4
Economics/Econometrics/Labour/ Employment	19.4	29.3	58.9	23.9	40.7	5.9
Geography/Area Studies	5.3	5.4	3.4	6.1	3.0	8.4
Health/Medicine	8.1	8.1	0.3	2.1	6.8	3.1
History/Humanities	2.4	1.2	0.9	0.9	0.7	5.3
Politics and International Studies	7.5	4.8	10.9	11.3	1.7	5.9
Psychology	4.0	2.4	6.3	3.1	3.9	5.3
Social Policy/Administration	5.4	6.6	0.2	6.4	7.0	5.0
Sociology	12.8	14.1	1.2	18.0	15.8	27.2
Statistics/Methodology/Computing	5.3	6.9	1.3	6.7	6.1	3.4
Other	17.9	15.1	0.8	10.7	9.8	16.9
Number of users	2,025	2,799	6,340	327	1,467	320

Advising ESDS

ESDS has an Advisory Committee to advise on the general direction and strategy of the service, and to provide the ESDS Director with comments on Annual Reports and reports for periodic reviews.

Chair

Jil Matheson (until February 2008)
Office for National Statistics

Vice-chair

John Curtice
University of Strathclyde

Members

Sara Arber
University of Surrey

Tanvi Desai
London School of Economics and Political Science

Richard Blundell
Institute for Fiscal Studies

Keith Dugmore
Demographic Decisions Ltd.

Peter Elias
University of Warwick

Jude England (from April 2008)
British Library

Nigel Fielding
University of Surrey

Myron Gutmann (from April 2008)
Interuniversity Consortium for Political
and Social Research

Peter Halls
University of York

Michael Hughes (from April 2008)
Office for National Statistics

Julia Lane (from April 2008)
National Science Foundation

Corrine Moy
National Opinion Polls

Stephen Penneck (until April 2008)
Office for National Statistics

Gillian Robinson (from April 2008)
Queens' University Belfast

Patten Smith
Ipsos MORI

Martin Weale
National Institute of Economic and Social Research

Economic and Social Data Service staff

Cathie Marsh Centre for Census and Survey Research
Angela Dale

Institute for Social and Economic Research
Nick Buck

Mimas
Keith Cole (until April 2008)
Jackie Carter (from April 2008)

UK Data Archive
Kevin Schürer
Louise Corti
Melanie Wright

In attendance

Martin Bulmer
Question Bank

Siân Bourne
ESRC

Abi Obafemi
JISC

Michael Hughes (from April 2008)
Audit Commission

Economic and Social Data Service
UK Data Archive
University of Essex
Wivenhoe Park
Colchester
Essex
CO4 3SQ

Email: help@esds.ac.uk
Tel: +44 (0)1206 872143
Fax: +44 (0)1206 872003
www.esds.ac.uk

