



Economic and Social Data Service

Annual Report 2006-2007

Contents

Introduction	2
Background	3
Highlights	4
ESDS Management	5
ESDS Access and Preservation	7
ESDS Government	9
ESDS International	11
ESDS Longitudinal	13
ESDS Qualidata	15
Case Studies	17
Diary of the year 2006-2007	23
Statistics	25
Feedback	32
Advisory Committee	ibc



Introduction

This Annual Report marks the end of the first phase of ESDS. Established in January 2003, the initial contract for the service was granted until the end of September 2007. Following the successful Mid-term Review of the service, funding from the Economic and Social Research Council (ESRC) and the Joint Information Systems Committee (JISC) was agreed for a further five-year period.

In looking forward to what might be called ESDS – Phase II, I would like to emphasise that this provides an excellent opportunity to develop the service still further. Yet whilst change is important, for a service like ESDS, it must be tempered with continuity. Over the past four to five years ESDS has made building up a user community a key priority. In this sense ESDS is, and has become, far more than a facility whereby university-based social scientists can gain access to important data resources for research and teaching. ESDS prides itself on being user focused and aims to be support driven. Thus, Phase I of ESDS witnessed an emphasis on both pre and post data delivery user support. In the case of the former, this has included guidance for data producers on ethical and consent issues, data management, and preparing data for long-term preservation and access. In terms of post-dissemination, examples of support include training sessions and workshops on data use, and guides on methodological issues such as weighting. In this context it is important to realise that ESDS is clearly not just about data use, but rather actively engages with the whole process or lifecycle of data creation, management, use, analysis and preservation. In consequence, whilst we all too readily think of ‘users’ as being those who obtain data from ESDS, this is both narrow and misleading. The ESDS user community embraces many, who are producing or analysing data and are accessing a wide-range of web-based support materials, who have neither formally registered with the service nor obtained data from ESDS directly.

This emphasis on users, broadly defined, will continue as an important driver in ESDS Phase II, and the programme of producing high-quality guides and holding innovative workshops, conferences and bespoke training sessions will remain. Yet equally, Phase II will also see a number of new developments. Operating within a fast-moving technical environment, it is critically important that ESDS does not stand still. Thus over the course of Phase II ESDS will face many exciting challenges. In particular, two areas which will be addressed include building greater interoperability and embedding more Web 2.0 functionality in the web-based services that ESDS offers. A central task must be to make the discovery of, and access to, data resources held beyond ESDS – both nationally and internationally – more transparent to our user community.

The term ‘user community’ has been used several times, and I would like to end this introduction by commenting on how this is itself a changing entity. Conventionally, the core ESDS users are thought to be composed of UK-based academic researchers drawn from the central disciplines of social science – economics and sociology especially. Yet this can no longer be said to be the case. The number of international users has expanded and will undoubtedly continue to do so. Equally, those working in the government and public sectors account for an increasing proportion of use, whilst within the UK-university sector, increasing numbers of teachers and students, both postgraduate and undergraduate are ESDS users. Lastly, fewer ESDS users would claim to be ‘social scientists’, even in the broadest sense. Other discipline areas, especially the humanities and medical and health sciences, are making increased use of ESDS materials and data. This trend toward an increasingly diverse user community is already underway, and is set to continue in ESDS Phase II.

Professor K. Schürer, AcSS
Director, ESDS

Background

The Economic and Social Data Service (ESDS) is a national data service that came into operation in January 2003. ESDS provides access and support for an extensive range of key economic and social data, both quantitative and qualitative, spanning many disciplines and themes.

ESDS is a distributed service, based on a collaboration among four key centres of expertise: UK Data Archive (UKDA), Institute for Social and Economic Research (ISER), both based at the University of Essex; Mimas, and Cathie Marsh Centre for Census and Survey Research (CCSR), both located at the University of Manchester.

ESDS comprises six distinct yet inter-related functions:

ESDS Management – The overall direction and management for ESDS is the responsibility of the UKDA. Through this function, the UKDA provides consistency and standards across the service. It performs a broad strategic role, relating to a variety of stakeholders concerned with the supply, funding and use of social science data, and seeks to create a coherent publicity, promotion and outreach strategy for the whole service.

ESDS Access and Preservation – Also hosted by the UKDA, this function represents the core of the ESDS, focusing on the central activities of data acquisition, processing, preservation and dissemination, providing access to most of the data held by ESDS through a variety of web-based systems. The function also handles registration of users centrally across ESDS.

ESDS Government – The service, led by CCSR with support from the UKDA, promotes and facilitates increased and more effective use of large-scale government surveys in research, learning and teaching. This includes methodological guides, short courses and support for key statistical packages. The service also works closely with data producers and users to maximise the synergy between these groups.

ESDS International – This service disseminates and supports the major statistical publications of intergovernmental organisations and a range of international survey datasets. The central aims of the function are to promote and facilitate increased and more effective use of international datasets in research, learning and teaching across a range of disciplines. The function is run jointly by Mimas and the UKDA.

ESDS Longitudinal – The work of ESDS Longitudinal is undertaken jointly by the UKDA and the UK Longitudinal Studies Centre (ULSC) based at ISER. The function supports a range of longitudinal data collections by: creating data enhancements; providing user support and training; and promoting the increased use of longitudinal data.

ESDS Qualidata – This specialist function provides access to, and support for, a range of qualitative datasets, and is hosted by the UKDA. The function is responsible for generating a number of data enhancements, and for providing information and training resources that focus on strategies for re-analysing qualitative data.

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



Highlights

As part of the maturing of ESDS, the past year has witnessed the service extend and consolidate its involvement in a wide network of research- and data-orientated networks, both nationally and internationally. Examples of this activity are given in the ESDS Management section of this report, but taken together they reflect the position ESDS now holds as a key element of the developing social science research infrastructure. This pivotal role is equally reflected in the increasing emphasis placed by all sections of the ESDS on promotion and training events – ESDS staff contributed to a total of 176 workshops and conferences during the year, a daunting and exciting level of engagement with the user community by any measure. This outreach and user support activity has been accompanied by the creation of a wide-ranging collection of online guides, now totalling more than 100 in number, and a major achievement this year has been the generation of a suite of web pages providing easier navigation and access to this major resource.

The past year has also witnessed increasing concern, especially from government departments, over issues of confidentiality and statistical disclosure. ESDS has risen to this challenge through the further development of the Special Licence system. This has been applied to a growing body of data and has allowed for the successful dissemination of important micro data previously unavailable. With new statistics legislation due to come into force, this development, built upon a proven and trusted relation with government data producers, will become an increasingly important cornerstone of ESDS, ensuring access to key data for social science research.

ESDS Government has played an important part in this process as part of its overall vision of increasing the capacity of UK social scientists to conduct high-quality research using quantitative methods. The service has continued to provide both high-profile conferences and introductory workshops organised around a topic-based approach to show the relevance of the government surveys to a wide range of social issues. This year, workshops have focused on ethnicity and in particular on the research potential of the Office for National Statistics (ONS) Omnibus Survey. The ESDS Government annual research conference – specifically aimed at disseminating research based on the government surveys to government policy-makers and analysts, and this year focusing on ethnic differences – yet again proved to be highly productive and is fast becoming a key date in the social science researcher's calendar. Likewise the annual User Meetings which provide opportunities for collaboration and interaction with data producers have proved to be of tremendous benefit. Annual meetings are now being held for the Labour Force Survey and the Annual Population Survey, the General Household Survey, the Family Resources Survey, the British Crime Survey, and the health surveys.

Over the current year **ESDS International** has been busy negotiating further UK-wide academic data redistribution agreements for four additional datasets to enhance its macro data portfolio – the Organisation for Economic Co-operation and Development (OECD) Structural Analysis (STAN), the National Accounts databases, the OECD Education Statistics and the International Monetary Fund (IMF) World Economic Outlook all being acquired and processed. The micro data portfolio has also expanded, especially via the deposit of data by the European Foundation for the Improvement of Living and Working Conditions and the release of the seventh New Europe Barometer. A further significant development over the past year has been the updating of the CommonGIS geographical interface for a subset of the United Nations (UN) Millennium Development Indicators. ESDS International has also continued to develop a number of major e-learning resources, especially through the Linking International Macro and Micro Data (LIMMD) project. The web pages for international survey micro data have also been revamped and new pages added linking to important online collections held elsewhere, such as ZACAT, the German social science data archive's catalogue, which provides online access to cross-national data, including the Eurobarometers, and the International Social Survey Programme (ISSP).

The current reporting year has been an important and busy time for new data acquisitions for **ESDS Longitudinal**. In particular the service has seen the processing and release of a major new study – the Longitudinal Study of Young People in England (LSYPE), also known as Next Steps. The addition of this collection to its portfolio means that ESDS Longitudinal now supports, through web pages, user guides, value-added user resources and an outreach and promotion programme, twice as many studies as originally defined in 2003 when ESDS was established. This new acquisition was also accompanied by: the re-deposit of an extended version of the National Child Development Study (NCDS); the 2004-2005 follow-up of the 1970 British Cohort Study (BCS70); an additional survey from the Millennium Cohort Study (MCS) on mothers who had received fertility treatment; and several new English Longitudinal Study of Ageing (ELSA) files.

A key achievement for **ESDS Qualidata** this year has been the production of an online qualitative teaching resource for exploring diverse interview types. This teaching pack provides instructors and students with materials designed to assist in teaching qualitative interviewing. In addition, the ESDS Qualidata collection has been enhanced by a number of important new data acquisitions including Brian Jackson and Dennis Marsden's Education and the Working Class and the online version of Marsden's Mothers Alone study, together with a new suite of web pages for the Pioneers project, being a series of interviews with pioneer researchers in the fields of sociology, anthropology, and social psychology, including Peter Townsend, Michael Young, Frank Bechhofer, George Brown, Stan Cohen and Mildred Blaxter. Another major achievement has been the provision of access via the ESDS catalogue to some 140 virtual data collections held outside of the local ESDS Qualidata collection in a number of locations throughout the UK.



Management

ESDS Management

supporting high quality research, teaching, and learning in the social sciences by acquiring, developing and managing social and economic data and related digital resources; and by promoting, disseminating, and supporting the use of these resources as effectively as possible

Co-ordination, reporting and strategic planning

A major part of ESDS Management's co-ordinating role is to manage the ESDS web presence, providing a seamless interface between the various specialist services, and ensuring that they meet the most up-to-date web standards. Similarly, ESDS Management is responsible for ensuring quality control of catalogue records across the different services. Co-ordination and strategic planning of these were achieved through a variety of communication channels, both formal and informal, between staff within and across the different sections and institutions, and crucially by means of the Away Day in February, where key issues were discussed, new initiatives presented, and a list of priority tasks agreed for the year ahead. Another important role of ESDS Management is to keep abreast of new developments in metadata standards and tools. ESDS staff have taken the lead on a number of initiatives, such as the Data Documentation Initiative (DDI), and made significant contributions to innovative research and development projects, enabling the ESDS to continue to deliver an excellent service to its users.

This year has seen ESDS Management strengthen its external links, both nationally and internationally. In particular, it has continued to develop its links with its stakeholders, other ESRC/JISC investments and other related service providers. Angela Dale, head of ESDS Government, is director of the ESRC Methods Programme, while Nick Buck, head of ESDS Longitudinal, is director of the UK Longitudinal Studies Centre and Keith Cole, head of ESDS International, is Deputy Executive Director of the National Centre for e-Social Science (NCeSS) and Director of the Census Dissemination Unit. Other senior ESDS staff are members of the Advisory Committee of the Question Bank, NCeSS, the National Centre for Research Methods (NCRM), and the Timescapes programme, and have provided input for the new Northern Ireland Social and Political Archive Advisory Committee. Synergy between the ESDS and the ESRC was further strengthened with the launch in June 2007 of the ESRC's UK National Data Strategy for Data Resources for the Social Sciences, which informs the ESDS's data acquisition strategy.

ESDS staff have also participated in various JISC- and/or ESRC-funded projects, including the JISC-funded Data Exchange Tools and Conversion Utilities (DeXT) project. Kevin Schürer is a member of the expert panel of the Research Information Network Publication and Quality Assurance of Research Data Outputs project funded by JISC and the Natural Environment Research Council.

International presence

In the past year, ESDS's international standing has continued to grow. Kevin Schürer is the current president of the Council of European Social Science Data Archives (CESSDA), which was recently identified by the European Strategy Forum on Research Infrastructures (ESFRI) Roadmap process as a major European research infrastructure. He is also a council member of the Interuniversity Consortium for Political and Social Research (ICPSR), the nearest equivalent to the ESDS in the USA, while senior members of ESDS staff are leading officials of the International Association of Social Science Information Services and Technology (IASSIST).

The international reputation of the ESDS as leader in its field is reflected in the large and increasing number of visits it receives from overseas delegations wishing to explore first hand what ESDS does, and how it does it, in order to build on this experience in a local context. This year has seen visitors from, amongst other places, the Human Sciences Research Council (HSRC), Pretoria, South Africa, and from the Chinese Academy of Social Sciences (CASS).



www.esds.ac.uk





Promotion and user support

ESDS participated in 176 educational and promotional events last year. It was the principal sponsor for 34 workshops and conferences. In total, over 300 people were trained at ESDS workshops. ESDS organised a conference, 'Cooking numbers and eating words: using data to investigate food, lifestyle and health', as part of the ESRC Festival of Social Science. ESDS staff also contributed papers and panels to many other workshops and conferences, both nationally and internationally. Examples are listed in the Diary section of this report.

An important part of ESDS's promotional activities is the production of online guides. These now exceed 100. In order to make it easier for users to find a specific guide, a suite of web pages has been created, where guides are grouped either by service or by topic and theme. The guides introduce the various ESDS services and specific datasets within the ESDS collection, as well as the online systems for finding and exploring the data. Two new important guides that were produced this year are the data processing and documentation processing guides, which provide an overview and discussion of ESDS's in-house processing standards and procedures, and which are aimed at data depositors, users and other interested parties. ESDS also collaborated with the Rural Economy and Land Use Programme Data Support Service (RELU-DSS) to produce a guide on data management, which provides information and examples of good practice which apply more broadly to all kinds of research data.

In its support role for ESRC grant applicants and grant holders, ESDS has produced a suite of web pages devoted to legal and ethical issues surrounding informed consent. The pages are innovative in providing exemplar consent agreements.

Acquisitions

This year saw the acquisition of 285 new datasets, and 200 new editions. Notable additions to the quantitative study collection were: the second sweep of the Millennium Cohort Study; the English Longitudinal Study of Ageing; the OECD Structural Analysis database and the OECD National Accounts database; and the Flexible Working and Leave Arrangements Module of the ONS Omnibus Survey. An important addition to the qualitative study collection was the Health and Social Consequences of the Foot and Mouth Disease Epidemic in North Cumbria, 2001-2003 with its inclusion of audio files. ESDS continued to collaborate with ONS on the release of Special Licence studies, bringing their total number to 23. These studies make previously unavailable data available under more restrictive access conditions. ESDS also liaised with the ESRC over revision of the datasets policy and the end-of-award reporting and depositing process.

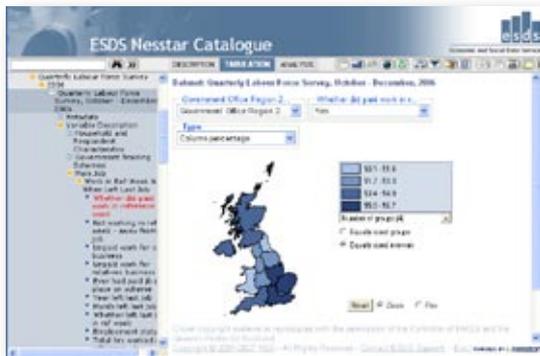
Access

ESDS Access and Preservation

ESDS Access and Preservation largely represents the unseen back office of ESDS, being responsible for data and metadata ingest, preservation, and dissemination of all data with the exception of international macro data. It is also responsible for primary user support for all data not covered by one of the specialist services

Service enhancements

The Special Licence system, implemented last year for more controlled access to previously unavailable sensitive or potentially disclosive data from ONS, has proved a tremendous success. This year has seen the expansion of the Special Licence remit to include 17 previously unavailable studies from both ONS and also the Centre for Longitudinal Studies. Other data producers, such as the Department for Work and Pensions, are in the process of reviewing the feasibility of depositing previously unavailable data. In the past year 73 datasets were disseminated to 25 different users under Special Licence. With the passing of new statistics legislation, many government data producers are coming to see the Special Licence as a way of maintaining data supply in the face of stronger controls on the dissemination of personal-level data. ESDS is also in discussion with the National Datasets Co-ordinator and the ONS about the integration of a potential 'secure data service' to handle data deemed too sensitive even for Special Licence.



In March 2007, the ESDS service upgraded to Nesstar v3.5, the latest version of the Nesstar online data browsing software and the attendant back office data publishing suite. This version has improved branding and layout, with the catalogue separated into datasets for research and datasets for teaching. In particular the new version provides users with the following additional functionality:

- computation and recode of variables;
- calculation on categories within tables;
- correlation analyses;
- the ability to include descriptive information within bookmarks.

The new version of the Nesstar Publisher greatly improves the efficiency of the workflow for publishing and republishing data, and allows for different data structures beyond simple flat rectangular files.

Also this year, as part of its commitment to good practice, ESDS has produced a service promise with a complaints procedure and explicit service standards.

New data

In addition to ingest, enhancement, publication, and preservation of data, ESDS Access and Preservation is also responsible for user support for all data not covered by one of the four specialist services. This year saw the acquisition of some important new datasets. In particular, three datasets from the British Election Study (BES) 2005 were released this year: Internet Rolling Campaign Panel Data and British Parliamentary Constituency Database; Comparative Study of Electoral Systems; and the Face-to-Face Survey which is also available for online browsing and analysis via Nesstar.

Other important ESDS Access and Preservation data released this year include enhanced data from the Workplace Employee Relations Survey (WERS). Data from the 1998 sweep, previously held under dissemination by explicit depositor permission only, were successfully reclassified as standard access. Access to data from the 2004 sweep was expanded and enhanced by its publication in Nesstar (see the ESDS Access and Preservation case study for an account of this successful collaboration between data producers and ESDS).



Lastly, another important acquisition, which will hopefully serve to fill a significant gap in the ESDS collection, was the Datastream Historical Company Accounts, 1965-2005. Datastream, which is part of Thomson Financial, provides an extensive database covering financial, economic and accounting data. This includes unique micro data on financial data for companies and market prices, such as the Financial Times Stock Exchange (FTSE) index. In 2002, Datastream accounting data began to be supplied by Worldscope, whereby much of the detail that was provided by the original accounting data for UK companies was lost. The new historical dataset acquired provides detailed data dating back to the 1960s that will enable continued access by all UK academic researchers, current and future, to a comprehensive electronic database of company accounts. The new licence has been negotiated for a two-year period, and was made possible by a supplementary award from the ESRC.

Information development

Web site improvements continued after the very successful launch of the new web site last year. One recent development has been to provide more dynamic information on datasets and data series via the catalogue and the major series web pages. Programming linking the catalogue to internal databases allows instant display of news and relevant links without requiring editing of individual web pages or catalogue records. This both increases the relevance and immediacy of information available to the user, and greatly streamlines the work required to maintain up-to-date web site and catalogue information.

Responding to the large number of requests it receives every year as a centre of expertise in data management, ESDS has made available its internal data processing procedures and data management guides. Data depositors, users, data organisations and other interested parties can access the guides via the web site. Data management is an area of increasing profile, and ESDS is recognised as an international leader in the field.

Preservation

Data preservation is a key (and often underestimated) responsibility of ESDS Access and Preservation. With 2007 marking the 40th anniversary of the original funding of the UKDA by the ESRC, the importance of ensuring that machine-readable studies are available and usable over time, despite rapid changes in technology, is made abundantly clear. Due to its proactive policy of digital preservation, the UKDA is now custodian of the only usable copies of several government surveys created in the 1970s and 1980s – surveys which in many cases are no longer held by the originating government department. The service received an important boost when UKDA appointed Matthew Woollard as its Associate Director, Head of Digital Preservation and Systems. Although Matthew is working more widely than just ESDS, the development of UKDA's leadership and activity in the area of digital preservation has significant benefits for ESDS Access and Preservation, ensuring that data assets created under research council and government funding remain accessible and usable for academic research and teaching over the long term.



Government

ESDS Government

facilitating more effective use of large-scale government surveys

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) and the Health Survey for England (HSE). The renewal of funding will allow ESDS Government to consolidate and extend existing work as well as to take a strategic look at where resources are best focused to produce the greatest gains.

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 are 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 are 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, thus ensuring that the value of the government surveys in research is widely recognised. 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 reporting year, workshops have focused on ethnicity and on the research potential of the ONS Omnibus Survey. Workshops also include a practical element to introduce software such as SPSS, Stata and Nesstar. Supporting web resources are designed to back up and extend the workshops. A web-based index of modules available within the Omnibus allows potential users to see at a glance the large number of topics covered. 'Starting analysis' guides provide continuing online support for new users and are updated regularly. Similarly, guides to using SPSS and Stata are available online with examples based on specific government surveys. These have been updated in the past year to reflect new software developments.

For more experienced users, workshops are held on topics such as the application of weighting and the use of government surveys to understand change over time. All guides and workshop materials are available from the ESDS Government web site.

Contact and collaboration with data producers

Collaboration and interaction with data producers is of enormous importance. Annual User Meetings provide a specific vehicle for this. They provide an opportunity for users to hear about proposals for survey changes and discuss with data producers where they feel improvements could be made to data quality. User presentations highlight the importance of the specific survey in addressing some of the key questions facing social science. Annual meetings are now being held for the Labour Force Survey and the Annual Population Survey, the General Household Survey, the Family Resources Survey, the British Crime Survey, and the health surveys. These meetings have been attended by researchers across a range of sectors, including government, the voluntary sector and also academia.



www.esds.ac.uk/government

ONS has expressed an interest in holding user meetings for the Expenditure and Food Survey. Consultation meetings have been held for the Integrated Household Survey and how this survey is included with other user meetings will be reviewed during the coming year.

ESDS Government works collaboratively with data producers when appropriate on both documentation and events such as the Omnibus workshop. For example, a new GHS Time Series Dataset which combines data from the 1972-2004 GHS into a single dataset has been released. This dataset was produced by ONS, with ESDS creating additional materials and providing dissemination and support.

Access

The standard End User Licence (EUL) underpins restricted user access to the ESDS data collections. However, the Special Licence developed jointly between ONS and ESDS provides protected access to survey micro data that are potentially disclosive and therefore allows researchers to access more detailed data than ever before. The Special Licence, which requires more tightly controlled conditions than the EUL and imposes heavier penalties for misuse, has become more widely used in the last year.

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 regularly emailed information sent to relevant JISCmail lists. A web-based database allows a search for publications that have been produced from government surveys. This is currently based on a search of the Web of Science and will be extended in the coming year 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.

Research conference

The ESDS Government annual research conference is specifically aimed at disseminating research based on the government surveys to government policy-makers and analysts. This year's conference, held at the British Academy in November 2006, on ethnic differences was very well attended and generated excellent discussion and feedback. Summaries of papers and presentations are all on the ESDS Government web site.

Web resources and help desk

As user numbers increase, the help desk becomes more heavily used and it is important that as much guidance as possible is placed on the web. These include Frequently Asked Questions, thematic guides that help the new users to find data sources relevant to their research topics, and guides that help users identify the key features of a specific dataset. Over the reporting year these have been continually expanded and updated, with new resources being added.



International

ESDS International

providing access and support for a range of international macro and micro data sources

ESDS International continues to be an exceptionally well-used service – now accessed by over 14,000 individual users from 185 institutions with the highest ever usage of the international databanks being recorded in March 2007. Over the past year, work has continued on promoting and facilitating increased and more effective use of international datasets in research, learning and teaching.

Data portfolio developments and liaison with data providers

ESDS International has developed excellent relationships with each of the international macro data providers and over the last year has been busy negotiating further UK-wide academic data redistribution agreements for four additional datasets to enhance its data portfolio; the OECD Structural Analysis (STAN) and National Accounts databases were released in October 2006, the OECD Education Statistics were released in January 2007 and the IMF World Economic Outlook was released late in 2007.

In addition to continuing to facilitate access to data from the Eurobarometers, the International Social Survey Programme (ISSP), the European and World Values Surveys and the European Social Survey (ESS), the reporting year has seen a number of important new survey data being acquired. These include data deposited by the European Foundation for the Improvement of Living and Working Conditions, such as the European Survey on Working Time and Work-Life Balance. Also deposited was the seventh New Europe Barometer, conducted by the Centre for the Study of Public Policy since spring 1991 to monitor trends within, and comparison across, post-Communist countries.

Additional ESRC funding has also enabled licensed access to the Latinobarómetro via Latinobarómetro Corporation, a not-for-profit non-governmental organisation based in Santiago, Chile, which produces and publishes the data. The Latinobarómetro is an annual public opinion survey of approximately 19,000 interviews in 18 countries in Latin America representing more than 400 million inhabitants – data are currently available from 1995 to 2004, surveying the development of democracies, economies and societies applying attitudinal, opinion and behavioural indicators.



Online resources and value-added services

The ESDS International web site remains a rich source of information for international data users, with each of the macro databases hosted by the service accompanied by a dataset user guide containing information on topics and countries covered, collection methods used, periodicity and time range and the terms and conditions associated with data use.

The international survey web pages have been revamped and new pages have been added about survey data in order to provide a new look and feel in the form of step-by-step guides for each of the major survey data series. Links have also been added to ZACAT, the German social science data archive's Nesstar catalogue, which provides online access to cross-national data, including the Eurobarometers and the ISSP. As well as the key tasks of data delivery and user support, ESDS International provides value-added services aimed at raising awareness of the potential of international data in research and teaching. These include visualisation interfaces, a publications database, dataset and software user guides and a guide to other sources of international data. In addition, the service has continued to develop its library of thematic guides with the creation of a new guide on CO₂ emissions intended to provide users with advice on finding both macro and micro data sources on this topic and information about the cross-country comparability of this data.

www.esds.ac.uk/international



Over the past year ESDS International's visualisation interfaces to freely available data have been updated and a new CommonGIS interface to a subset of the UN Millennium Development Indicators has been released (see Figure 1). User statistics obtained from the Intute World Guide web site show that the visualisation interface was accessed over 11,000 times during the reporting year.

Developing research capacities

During the past year the service has also been involved in the development of a major e-learning materials project – LIMMD. This project was funded as part of the ESRC's Researcher Development Initiative (RDI) to raise awareness, and provide training and support for researchers and postgraduate students in the use of the interrelated survey and aggregate data resources. Each unit has been written by a subject specialist and has been designed as a self-guided teaching resource containing extensive information on the topic, examples, teaching datasets and step-by-step hands-on activities. The resource is available on the ESDS International web site and also as a learning and teaching resource on the Jorum web site. Jorum is a JISC-funded collaborative venture in UK higher and further education helping to build a community for the sharing, re-use and re-purposing of learning and teaching materials.

ESDS International on the road

The service continues to place high priority and emphasis on user support. In particular the team have continued their outreach work attending relevant conferences to promote the service and providing an active programme of workshops and training. Over the last year training courses have been held at universities including Manchester, Bristol, London, Leicester and even the Fundación Juan March in Madrid. A highlight was the ESDS International annual conference, held in November 2006. With over 80 delegates, the conference attracted speakers from the World Bank, IMF and International Labour Organization, as well as researchers from the academic community presenting their work based on the international databanks. This annual, high-profile event creates a valuable opportunity for users and providers of international data to meet and share their ideas on issues relating to international data research. Of particular note was the lively and informative data providers' panel discussion on various key international topics such as scalable metadata and making publicly funded data freely available.

During May 2007, representatives from ESDS International were invited to attend OECD's Annual Forum on Innovation, Growth and Equity, which brought together business and labour leaders, civil society personalities, government ministers, leaders of international organisations and academics to discuss key international issues. Also in May, the service was represented at the IASSIST conference in Montreal where a paper explaining the complexities of converting raw data into a format deliverable via the data delivery software Beyond 20/20 was presented.

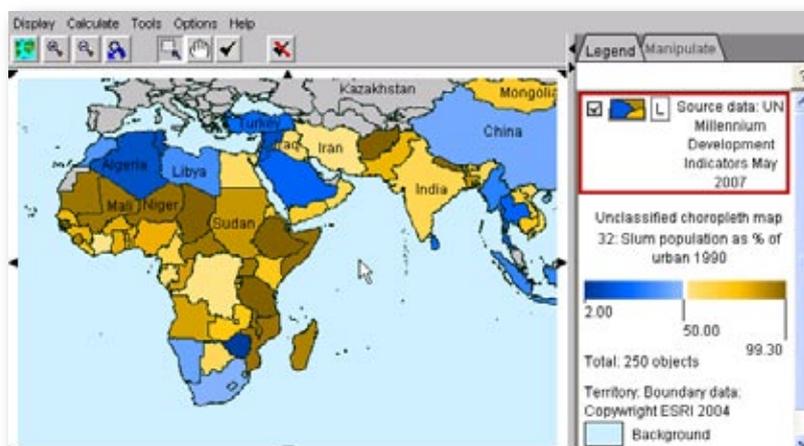


Figure 1 CommonGIS interface to UN Millennium Development Indicators

Longitudinal

ESDS Longitudinal

supporting the use of key UK longitudinal data collections

Additions to the data portfolio

December 2006 saw ESDS Longitudinal add a sixth major study to its portfolio of key longitudinal data resources – the Longitudinal Study of Young People in England (LSYPE). LSYPE, also known as Next Steps, was commissioned by the former Department for Education and Skills (DfES) and is now managed by the Department for Children, Schools and Families (DCSF). It is an innovative panel study which will follow a cohort of young people from secondary education through to their early labour market experiences. The acquisition of the LSYPE means that ESDS Longitudinal now supports, through web pages, user guides, value-added user resources and an outreach and promotion programme, twice as many studies as originally defined in 2003.



www.esds.ac.uk/longitudinal



The year 2006–2007 has also been a busy one in terms of acquiring new data for the established major longitudinal studies. The National Child Development Study (NCDS) data were re-deposited by the Centre for Longitudinal Studies (CLS) in their entirety, with the new data benefiting from a reorganisation of files (fewer in number and labelled more intuitively), further data cleaning, and the provision of a response dataset allowing users to track cohort members' participation through the study. In addition to this, sweep 7 of the study was made available to users in March 2007.

For the 1970 British Cohort Study (BCS70), the 2004–2005 follow-up data were also made available in March 2007 and a response dataset, equivalent to the NCDS one, was released in April 2007. County-level data for the BCS70 were also supplied by the CLS for use under a Special Licence. Users of the Millennium Cohort Study (MCS) benefited from the deposit of a follow-up study of the cohort mothers who had received fertility treatment and the linking of birth registration and maternity hospital episode data to the main collection.

Several new English Longitudinal Study of Ageing (ELSA) files were made available during the reporting year. Chief among these were the data from wave 0 (created using data from the 1998, 1999 and 2001 Health Surveys for England) and an index file for the study which contains details of all age-eligible individuals within households identified at wave 0 and includes information on response and mortality, as well as analytical identifiers. Finally, the fifteenth wave of the most heavily used of the ESDS Longitudinal studies – the British Household Panel Survey (BHPS) – was made available to users in April 2007.

Longitudinal research resources

A good deal of new material has been added to the ESDS Longitudinal web pages and many of the existing resources have been extended and updated.

Each of the major studies now has its own online step-by-step guide to using the data, which includes links to external resources. The Database of Comparable Indicators, created last year, has been updated to cover ELSA and the latest sweeps of the NCDS and BCS70. Additionally, two



new resources have been created: a Database of Geography Variables, which enables users to identify and locate variables that are available via the standard End User Licence and provides information on those for which access has special conditions attached; and a Version History document, which allows users to track data releases and identify which version of the data they hold. As in previous years, ESDS Longitudinal has also focused on making longitudinal data accessible to students and first-time users – an LSYPE teaching/sampler dataset is the latest bespoke file to be produced for this purpose by the service.

Access to sensitive data

As has been the case across all the ESDS services, access to sensitive data – be it occupational coding, detailed information on geography, links to administrative records etc. – has become an important issue for ESDS Longitudinal and its key stakeholders. The Boyle et al. report, *Adding geographical variables and identifiers to longitudinal datasets: a report to the National Longitudinal Strategy Committee and ESRC*, on the disclosure risk attached to geography variables included in the major longitudinal datasets has pushed the issue further up the agenda.

It is in this context that ESDS Longitudinal produced the Database of Geography Variables. Data depositors, too, have responded to the challenge of balancing the risk of disclosing the identities of respondents with the utility of the data for secondary analysis. The availability of county-level information for BCS70 cohort members under Special Licence conditions is likely to be a precursor for further detailed geography information being made available under similar conditions for the birth cohort studies. For the other major longitudinal studies, whilst restrictions on using sensitive data quite sensibly remain in place, avenues for making such data available to researchers via the ESDS are being actively explored. Where sensitive data are not available via the service, ESDS Longitudinal continues to work as a gatekeeper, advising users on how, and under what conditions, they might access data from the data producers themselves.

More of the same

While 2006-2007 saw the unveiling of a host of new data and online resources, it has also been a year of consolidation for ESDS Longitudinal. Users of longitudinal data continue to draw on the advice of the service's help desk, which is itself able to draw on the expertise of the data producers. The service also continues to work closely with the data producers to promote the major longitudinal studies through an expanding outreach programme. One- and two-day workshops for new and potential users of the birth cohort studies have again been held in London in collaboration with the CLS, with around 100 new users introduced to the mechanics of the surveys, access to the materials, user support and exploratory analysis. Presentations have also been made by ESDS Longitudinal staff at a variety of seminars and conferences, both in the UK and abroad. In July 2007, for example, a presentation was given at the University of Essex Summer School in Social Science Data Analysis and Collection. At the May 2007 IASSIST conference in Montreal, ESDS Longitudinal gave a presentation on a project in which it has played a key role, namely a Medical Research Council project to open up access to the National Survey of Health and Development (the 1946 British birth cohort study). Finally, also in May 2007, ESDS Longitudinal staff helped organise and chair a one-day workshop in London entitled 'Families, youth, community and justice: new survey datasets for social research'. The workshop introduced users to four key ESDS datasets, including the LSYPE and another panel study, the Families and Children Study (FACS).

Qualidata

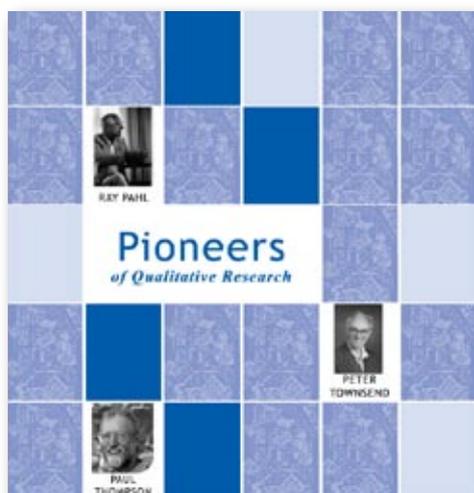
ESDS Qualidata

providing enhanced access and support for a range of multimedia qualitative data sources

Online resources and datasets

A highlight of the reporting year has been the production of a qualitative teaching resource for exploring diverse interview types. This teaching pack provides instructors and students with materials designed to assist in teaching qualitative interviewing, and is particularly aimed at instructors who have limited research materials of their own. The teaching resource focuses on distinguishing different types of interviewing, offering summaries of seven interview types: structured, unstructured, semi-structured, feminist, psycho-social, oral history and life story interviews. Each interview typology begins with a summary of what characterises that particular type of interview and is illustrated and exemplified by selected extracts from studies held by ESDS Qualidata. The entire pack is available online.

This year, a new suite of web pages were initiated to provide information on the Pioneers project. This project draws on a series of full life-story interviews conducted by Paul Thompson with pioneer researchers in the fields of sociology, anthropology, and social psychology. Those interviewed include: Peter Townsend, Michael Young, Frank Bechhofer, George Brown, Stan Cohen, Mildred Blaxter and Dennis Marsden.



Model transcript

The initial format in which data are received is critical in order to undertake systemic processing and enhancement of qualitative data. Therefore, ESDS Qualidata has defined a model transcript for use in dissemination as well as for prospective depositors.

A number of factors were considered in developing the template, including: the provision of adequate metadata; compatibility with Computer Assisted Qualitative Data Analysis (CAQDAS) packages; and time requirements placed on transcribers. The template comprises three sections: a header that contains collection and interview information and appears on every page; a respondent details section that appears only on the first page; and the body of the transcription itself.

The template provides a flexible and useful model that meets minimum preservation requirements.

Significant collections processed

A number of important new data acquisitions have been processed and enhanced during the year. These include data from Brian Jackson and Dennis Marsden's Education and the Working Class: Some General Themes Raised by a Study of 88 Working-Class Children in a Northern Industrial City, 1946-1960. The available data are their original interview notes from 1962, along with a user guide and data listing, and all 265 interviews from the study converted from paper to PDF files.

Another major study added to the ESDS Qualidata Online data browsing system was the Mothers Alone: Poverty and the Fatherless Family, 1955-1966 study interviews, also by Dennis Marsden. This brings the total number of interviews available for online interrogation to approximately 800.

Providing access to virtual collections

Pre-dating ESDS Qualidata, the Qualicat catalogue of qualitative resources listed collections held outside of the local Qualidata collection. An important task for ESDS Qualidata has been to research and rewrite these records so that they can be integrated into the Data Catalogue. This process is now complete with the addition of the final virtual collection records to the catalogue. Almost 140 records were created and added to the catalogue and include titles such as: Family and Social Life in Barrow, Lancaster, Preston, 1940-1970; National Sound Archive: Oral History Collection, 1890 – present; and Development of Television Literacy, 1989.

www.esds.ac.uk/qualidata





Collaboration in qualitative data sharing and archiving research

ESDS Qualidata staff are continuously involved in providing expertise on a range of external projects funded by the ESRC and others. As part of this support aspect of the service, ESDS led the ESRC Qualitative Archiving and Data Sharing Scheme (QUADS) programme and provided tailored advice to many of the projects on data formats, data preparation, data context, metadata and mark-up, and web presentation. Equally, a partnership with the ESRC Timescapes qualitative longitudinal study has been established, providing expert advice and training on many aspects of data management, archiving and access.

ESDS Qualidata staff have also collaborated closely with text miners in the UK, from the National Centre for Text Mining (NaCTeM) and the National Centre for e-Social Science (NCeSS) in Manchester to consider how information extraction tools may be applied to qualitative data.

Enabling the sharing of multimedia data

ESDS Qualidata has always welcomed opportunities for further technical development of its online data access systems. As part of this strategy, ESDS Qualidata have worked with the ESRC Smart Qualitative Data: Methods and Community Tools for Data Mark-Up (SQUAD) project to provide tools that enable the automated mark-up and anonymisation of data. These use text natural language processing with eXtensible Markup Language (XML) standards and open source tools. Likewise the UKDA was also successful in attracting a JISC award under the Repositories and Preservation Programme for the Data Exchange Tools and Conversion Utilities (DExT) project. This project aims to explore the feasibility of developing data exchange models and data conversion tools for primary research data collected in the course of empirical research. The qualitative component is developing, refining and testing models for data exchange for qualitative research data based on XML schema, from software packages such as Atlas-ti, MaxQDA and Nvivo to an open intermediate standard.

Publicity and outreach

ESDS Qualidata continues to promote the re-use of data with staff giving various presentations on how to approach secondary analysis from practical and methodological perspectives. A workshop was organised as part of the QUADS scheme, where the projects' tools were showcased. Presentations on qualitative data re-use were given to the ESDS-sponsored conference, 'Cooking numbers and eating words: using data to investigate food, lifestyle and health', as part of the ESRC Festival of Social Science. Staff also presented at the international CAQDAS event and held two summer school courses.

ESDS Qualidata staff also continue to publish on various aspects of data sharing, archiving, accessing and re-using qualitative data, a key example of which was the special issue of the journal *Methodological Innovations Online*, edited by Louise Corti.



Case Studies

Case Studies

ESDS Access and Preservation

Including the Workplace Employee Relations Survey in Nesstar



April 06

ESDS attends the first Workplace Employee Relations Survey (WERS) support service advisory committee meeting. It is agreed that WERS 2004 should ideally be included in Nesstar, in order to promote access to, and use of, WERS. To achieve this as quickly as possible the National Institute of Economic and Social Research (NIESR) agrees to resource some of the value-added data processing work necessary to prepare the file for publishing in Nesstar. ESDS undertakes a full evaluation of the work required, highlighting potential problems, time-scale, and staff resources needed. With a total of almost 3,500 variables, all requiring question text addition/grouping and quality control checking, the size of the task is not insignificant.

August 06

WERS steering committee agrees plans for including WERS 2004 within Nesstar.

September 06

ESDS begins working with NIESR. An initial problem to be tackled is the length of variable labels. Every label has to be edited in order to make it consistent, concise and of a reasonable length. This task is performed by NIESR with support and instruction from ESDS.

December 06 – February 07

The prepared Nesstar-specific files with edited variable labels are sent to ESDS, so that ESDS can add question text and group variables for easier navigation of the dataset within Nesstar.

March 07

WERS 2004 is ready for publishing within Nesstar. It is agreed that the study should be published once the new version of Nesstar (v3.5) is made available by ESDS. Due to a change of decision by the depositor over the withholding of financial performance data, region codes and detailed industry codes, a second edition of the study with these variables is deposited necessitating more, last-minute data processing. WERS 2004 is finally published to Nesstar on 4 April 2007 in time for the WERS Information and Advice Service Steering Committee Meeting on 11 April 2007.

Summary

Throughout this process close contact between ESDS and NIESR created a positive and efficient working relationship through which problems were solved and decisions were made, building on each others' experience and expertise.



Case Studies

Case Studies

ESDS Government

Ethnic differences in labour market outcomes, social capital and health



ESDS Government held its 2006 Research Conference at the British Academy on 'Ethnic differences in labour market outcomes, social capital and health: evidence from the UK government surveys'. The presentations highlighted some key research findings based on the analysis of UK government surveys.

Anthony Heath (University of Oxford) in his presentation 'Ethnic penalties: what role does discrimination play?' suggested that ethnic penalties can be thought of as the disadvantages that ethnic minorities experience in the labour market in comparison with people of white British ancestry of the same age and education. Discrimination is likely to be one cause of these penalties but it is not the only cause. Analysis of reports of job refusals from the Home Office Citizenship Survey 2003 showed that self-reported discrimination accounts for around one-third of the ethnic penalty in unemployment. First generation ethnic minorities report higher rates of job refusal than the white British; second generation ethnic minorities report even higher rates.

Sin Yi Cheung (Oxford Brookes University) studying 'Ethnic penalties in the labour market: the public-private sector divide' used the Labour Force Survey (LFS) 2001-2004 to show that, within the private sector, ethnic minorities tend to be under-represented in professional

and managerial occupations and over-represented in semi-routine and routine occupations. Ethnic minority men have lower earnings than white men. These patterns are less marked for women and are not found in the public sector, where ethnic penalties are lower than in the private sector. The British Social Attitudes Survey 2001-2003 shows that levels of self-reported prejudice amongst white employers and employees are significantly lower in the public administration, education and health sector than elsewhere and significantly higher in manufacturing, construction and transport.

Ken Clark (University of Manchester) and Stephen Drinkwater (University of Surrey) in their presentation on 'Earnings differences between ethnic groups: evidence from the LFS' showed that men from minority ethnic groups earn at least 10 per cent less than comparable white men (using standard controls for age and education), whilst for women the differentials are less. The earnings deficit is highest for black African women at 18 per cent, Pakistani men (21 per cent) and Bangladeshi men (27 per cent). Controlling for occupation reduces the earnings differences significantly and suggests that choice (or constraint) of occupation is an important factor in explaining earnings differences.



Case Studies

Case Studies

ESDS International macro data

Energy scenarios for Europe



Across Europe, policymakers face the challenges of reducing the impacts of air pollution on human health and ecosystems, mitigating climate change and its consequences, and securing long-term energy supplies adequate for providing services to people and the economy. Using data hosted by ESDS International, Mark Barrett of University College London, in a study funded by the Swedish Environmental Protection Agency, has constructed six low emission energy scenarios for twenty-five countries of the European Union that achieve these multiple environmental, social and economic goals.

The scenarios are based on the energy production and consumption statistics found in the International Energy Agency (IEA) databanks. In his report, Dr Barrett comments that the IEA databanks are the "single most useful" data source for international energy modelling. The scenarios are built using the IEA data combined with cross-country data on population, economic growth and behavioural changes, such as travelling less or improving household insulation.

The central scenario shows that, as compared to 1990 levels, carbon dioxide (CO₂) reductions of more than 30 per cent are feasible by 2030, coupled with similar reductions on pollutants such as sulphur dioxide and nitrogen oxides. Variant scenarios further explore the impact of a range of behavioural and technological changes. The findings have been presented at a series of intergovernmental events and conferences including the European Commission's 'National Emission Ceilings Policy Instruments Group', Vienna, March 2007, and the United Nations Economic Commission for Europe's 'Task Force on Integrated Assessment Modelling', Prague, May 2007.



Case Studies

Case Studies

ESDS International micro data Experiences and Expressions of Ageism



The age profile of most European countries is changing rapidly. In the UK the proportion of older people is growing at a rate that is likely to generate significant transitions in both employment and the need for care. The scope to adapt to these changes, economically and practically is likely to be affected not only by legislation relating to age discrimination, but also by social expectations and conventions.

One of the rotating modules in the 2008 European Social Survey (ESS), Experiences and Expressions of Ageism, will focus on beliefs, stereotypes and attitudes on ageing and about younger and older age groups. This module, led by Dominic Abrams (University of Kent), Luisa Lima (University of Lisbon), and Geneviève Coudin (Université de Paris V), is based on two surveys conducted in 2004 and 2006, directed by Dominic Abrams, and funded by Age Concern England. These were the first national surveys of ageism in the UK. The surveys used well-founded theory and research from social psychology to explore how different stereotypes and social experiences (in particular intergenerational contact) map on to prejudice and discrimination.

Several findings in the 2004 survey were notable. There were marked disparities in younger and older people's perceptions of ageing such that respondents aged 25 were likely to believe

old age starts before respondents aged 65 believed youth has ended. Although overt hostility toward older people was rare, people of all ages reported experiencing ageism more than other types of prejudice. Whereas prejudices against some ethnic groups tend to be rooted in hostility and mistrust, evidence of social stereotypes about ageing showed that ageism is likely to be reflected in perceptions of warmth (e.g. younger people as less friendly), incompetence and dependency (older people as less capable and likely to be pitied). Such perceptions were more likely among people who did not have close intergenerational relationships or friendships. It seems likely that aspects of both culture and the demographic age profile may have an important part to play in the stereotypes, attitudes and expectations people have about ageing. Given that different European countries differ markedly in these respects, the ESS module will give us a unique opportunity to evaluate their impact.

A guide to the ESS is available on the ESDS International web site. To access the data, users need to register on the European Social Survey Data Archive web site, hosted by the Norwegian Social Science Data Services. ESS registered users can browse, analyse and download the data online free of charge.



Case Studies

Case Studies

ESDS Longitudinal

World's largest household panel survey



In May 2007, ISER, together with colleagues from the University of Warwick and the Institute of Education, were awarded an initial grant of £15.5 million by the ESRC to establish the world's largest household panel survey available for research. The UK Household Longitudinal Study (UKHLS) will collect data from a panel of 90,000 individuals, from a sample of 40,000 households, at annual intervals. The new, multi-purpose survey will assist with understanding the long-term effects of social and economic change, as well as policy interventions designed to impact upon the general well-being of the UK population. Initial funding supports the collection of the first two waves of interview with each sample household. The study is planned to begin in 2009 and continue over several decades.

The UKHLS will supersede the BHPS – though an important feature of the UKHLS is that the BHPS sample will be incorporated and its members will continue to be tracked – and will extend and advance the work done on the latter. Significantly, the large sample size will allow small population sub-groups to be examined more reliably and an ethnic booster sample of over 3,000 households is a key feature of the study. The study will also benefit from the inclusion of an 'innovation panel' of around 1,500 households that will enable methodological experimentation. A further key feature of the study is that the survey will collect a wide range of biomarkers and health indicators, opening up the prospect of collaborative research between social scientists and epidemiologists.



Case Studies

ESDS Qualidata

Masculine preserves? Cooking, gender and family, c. 1945-1985



Peter Jackson and Sarah Olive (University of Sheffield) together with Graham Smith (University of London) are researching into cooking and gender and making extensive use of the data collection Families, Social Mobility and Ageing, an Intergenerational Approach, 1900-1988, spanning almost a century of recorded experience.

Rather than approaching gender in terms of a set of standard roles and looking for exceptions to these roles, their work emphasises the active construction of gender as a relational category within families. They focus, in particular, on men's relationships to the preparation of family meals. In so doing, they explore the making of masculinities (fatherhood and sonship) as a contested process and parenting as a series of dynamic social practices. They are particularly interested in the period from 1945 to the mid-1980s and the recorded recollections covering these 40 years include: descriptions of men learning to cook; men cooking to cope when women are absent from the family kitchen; as well as men 'dabbling' in cooking to show off professional skills, as weekend cooks, or to pursue a hobby. Throughout the research they explore the multiple and complex constructions of masculinity and motherhood that are evident in the oral histories, and conclude that the active nature of these constructions is especially present at points of change and crisis within their respondents' life-course. They also find, however, that responses to change, as illustrated by men's cooking, are shaped by broader family trajectories.



Peter Jackson

Diary of the year 2006-2007

August 06

ESDS Management hosts visits from N. Yakeda, Musaski University, Tokyo; I. Kuraishi, Tokyo University of Foreign Studies; and Warren Hilder and Mike Wadsworth, Medical Research Council (MRC)

ESDS International presents at Brooks World Poverty Institute Summer School, Manchester

September 06

ESDS staff organise a workshop, The Internet for Social Research, Birmingham

ESDS Qualidata presents at the Association for Survey Computing Conference, London

ESDS Longitudinal presents a workshop for potential users of the NCDS and BCS70, London

ESDS International staff present on international data resources for energy and the environment, Tyndall Centre, Manchester

ESDS Management presents at the Infrastructure for Spatial Information in Europe workshop on multilingual issues, Italy

October 06

ESDS Qualidata presents at the Fifth Annual Language and Computation Day, Colchester, and presents at the CESSDA Expert Seminar, Open Access to Data: Anonymisation, Data Protection and Confidentiality, Athens

ESDS International organises training courses at Leicester, Manchester, and London

ESDS Government hosts an introductory workshop on analysis over time; organises an introductory workshop, The Samples of Anonymised Records (SARs): Focus on the Small Area Microdata, Manchester; and delivers a seminar at ONS, Newport

ESDS staff contribute to the DDI 3.0 planning meeting, Germany

ESDS Management delivers an invited paper to a workshop for national statistical organisations on managing micro data access, Luxembourg

November 06

ESDS International runs the second Annual ESDS International Conference, Measuring the World Around Us, Institute of Materials, London

ESDS Government hosts an introductory workshop on government micro data, Bristol, and organises its annual research conference, Ethnic differences in labour market outcomes, social capital and health: evidence from the UK government surveys, London

ESDS Longitudinal organises an MCS workshop on data management and analysis, London

ESDS staff attend the Statistics User Forum annual conference, New Statistics from Administrative and Customer Files, London

ESDS Management leads a plenary discussion at New Opportunities for Research Funding Co-operation in Europe conference, Data and Research Infrastructure, Netherlands, and chairs a session at the Digital Curation Conference, Glasgow

ESDS staff attend a workshop: An Introduction for Geographical Referencing for Social Scientists, Southampton

December 06

ESDS Government organises a British Crime Survey user meeting, Royal Statistical Society, London, and delivers a workshop, Introduction to Stata, Manchester

ESDS International organises a workshop: Linking Macro and Micro International Data, Manchester

ESDS Longitudinal staff attend a NCDS/BCS70 user group meeting, London

January 07

ESDS staff organise a Statistics User Forum meeting, Improving access to government datasets, London

ESDS Government presents at a Labour Force Survey user meeting, London

ESDS International presents on the service at School of Oriental and African Studies, London

ESDS hosts visits from delegations from the Chinese Academy of Social Science; Human Sciences Research Council, South Africa; National Survey of Health and Development; Napier University, Edinburgh, and The National Archives

February 07

ESDS Longitudinal organises a workshop for potential users of NCDS and BCS70, London

ESDS staff present at a NCDS workshop, Imagining the future: researching children and the lifecourse, London

ESDS Management contributes to a National Data Strategy meeting, London

March 07

ESDS International organises a workshop: Introduction to ESDS International, Bristol

ESDS Qualidata organises a conference, Cooking numbers and eating words: using data to investigate food, lifestyle and health, Leeds

ESDS hosts visits from Pedro Diaz Munoz, Eurostat; a team from the Information Centre and Social Care, Leeds; Deb Mitchell, Australian National University; John Rae and Patrick Tate, CACI; and Mark Thorley, Natural Environment Research Council

ESDS Management gives an invited presentation to a European Science Foundation meeting, Paris

ESDS staff provide a stand at the JISC conference, Birmingham

ESDS Management staff are invited participants at the joint National e-Science Centre and NCeSS workshop, 'Realising and Coordinating e-Research Endeavours', Edinburgh

April 07

ESDS Government organises a workshop, Starting SPSS, and presents at the British Sociological Association conference, London

ESDS International organises the second workshop on Linking International Macro and Micro Data, Manchester

ESDS Qualidata presents at CAQDAS 07 conference, London, and Methods conference, Plymouth

ESDS staff attend a Producers and Users Group meeting between ONS and Statistics User Forum representatives, London

ESDS Management provides input into the selection committee for a new national data archive in Sweden

May 07

ESDS Government organises a workshop focusing on the Omnibus Survey, Cardiff

ESDS Qualidata presents a course on secondary analysis of qualitative data, Colchester

ESDS staff organise a workshop: Families, youth, community and justice: new survey datasets for social research, London

ESDS staff present and chair sessions at IASSIST 2007 conference, Montreal, and attend the British Urban and Regional Information Systems Association conference, London

ESDS International staff are invited participants at the OECD Forum 2007, Paris

June 07

ESDS Longitudinal organise a NCDS and BCS70 workshop, Institute of Education, London and a two-day workshop on MCS data management and analysis, London

ESDS International organises an introductory workshop, Manchester

ESDS Management contributes to the 4th European Conference on Research Infrastructures, Germany

ESDS staff contribute to meetings with ONS on new plans for access arrangements to restricted micro data

ESDS hosts a visit from the Centre for Survey Research, Academia Sinica, Taiwan

July 07

ESDS Longitudinal presents at the Essex Summer School in Social Science Data Analysis and Collection, followed by hands-on laboratory exercise, Colchester, and contributes to the BHPS 2007 conference, Colchester

ESDS Qualidata staff give invited presentations to the Changing Families, Changing Food ESRC Programme meeting, Sheffield, and present at the Timescapes Data Management workshop, Leeds

ESDS staff present at the Mimas Open Forum, Manchester

ESDS staff contribute to a reception at the House of Commons and organise a workshop to celebrate the 40th Anniversary of the UKDA, University of Essex

ESDS hosts a visit from Nathalie Lefever and Claire Lobet-Maris, University of Namur, Belgium, investigating the creation of a new national data archive in Belgium

Statistics

Statistics

As usual for the ESDS Annual Report, a section is included aimed at providing a statistical summary of activity across all sections of ESDS. This is presented in order to give an indication of the levels of activity and engagement between ESDS and the social science community, and equally to demonstrate changing trends over time. For this latter reason figures are also produced where appropriate for the previous reporting year and in certain cases a longer time period. ESDS routinely collects a huge body of statistical information concerning just about every aspect of its operation. Thus the information presented in this report represents only a brief digest of what might be considered the headline indicators. 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 all available from www.esds.ac.uk/news/jiscmu.asp.

Data acquired and processed

All data for ESDS, with the exception of the international macro data series, are acquired, processed and mounted on the download service for online access by ESDS Access and Preservation at the UKDA. Due to the special licensing arrangements relating to international macro data series these are acquired directly by ESDS International at Mimas, then processed and loaded onto the Beyond 20/20 data delivery interface. The total number of data collections

acquired, processed and made available to users (Table 1) increased slightly over last year, reaching 485 compared with 460. Importantly the acquisitions of entirely new collections, as opposed to new editions and updates to existing collections, which are more resource-intensive to process increased at a higher rate (18 per cent). Within these totals the levels of activity increased for all the specialist services, while the number of data acquisitions unassigned to a specialist service declined. In part this is a reflection of a strategic development resulting from the fact that despite some productivity gains, ESDS data acquisition and processing is working at maximum capacity, therefore any additional new important specialist data collections have to be acquired, pro rata to a trade off in non-specialist data collections. Thus the decline in number of core data collections acquired and processed can be seen as a quid pro quo for the marked increases in new additions to the ESDS collection in relation, especially, to ESDS Government, ESDS Longitudinal and ESDS Qualidata.

Value-added data preparation (Table 2) was another growth area this year, particularly in the case of ESDS Government data where the number of data collections processed doubled in comparison to the previous year. This was largely due to significant improvements in the data enhancement tools provided by a new version of the Nesstar Publisher software suite.

Table 1 Data collections acquired and processed for online delivery*

Service	2005-2006			2006-2007		
	New datasets	Updates and new editions	Total	New datasets	Updates and new editions	Total
ESDS Core	125	39	164	82	26	108
ESDS Government	63	24	87	90	32	122
ESDS International	19	126	145	24	131	155
ESDS Longitudinal	10	3	13	26	11	37
ESDS Qualidata	24	27	51	63	0	63
Total	241	219	460	285	200	485

*includes data acquired and prepared for ESDS International macro Beyond 20/20 service

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

Service	2005-2006			2006-2007		
	New datasets	Updates and new editions	Total	New datasets	Updates and new editions	Total
ESDS Core	17	1	18	4	0	4
ESDS Government	19	11	30	32	28	60
ESDS International*	3	49	52	2	41	43
ESDS Longitudinal	2	1	3	2	1	3
Total	41	62	103	40	70	110

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

Table 3 Datasets delivered to users by access method

2005-2006	Download	FTP	Via/access	Portable	Nesstar	Total
ESDS Core	10,538	281	1,372	34	65	12,290
ESDS Government	22,203	124	4,437	18	525	27,307
ESDS International*	309	569	507	30	13	1,428
ESDS Longitudinal	6,630	27	410	11	51	7,129
ESDS Qualidata**	927	13	73	2		1,015
Total	40,607	1,014	6,799	95	654	49,169

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		1,033
Total	39,479	986	7,090	53	326	47,934

* international micro datasets only, macro data series counted separately (see Table 4)

** Nesstar does not include qualitative data

Datasets delivered

Usage of ESDS in terms of the delivery of datasets to end users is summarised in Tables 3, 4 and 5 showing the number of data collections delivered to users by the various methods of data access offered by ESDS for each of the separate services.

Following an historic trend of year-on-year growth in the number of data collections disseminated, the figures for this reporting year show a plateauing of the pattern of ever-increasing usage, with most ESDS services recording levels of data dissemination similar to those of the previous year. Indeed, slight declines are recorded in some areas, especially Core and ESDS Longitudinal. However, it would be inappropriate to interpret this flattening out of the amount of data disseminated in isolation since it is only one element in the overall measure of ESDS resource usage. Whilst delivering datasets is part of the ESDS bread-and-butter, as the service-specific sections of this report clearly show, ESDS provides a user-focused family of support functions which go beyond the simple delivery of data. As the web statistics indicate, usage of the value-added support materials which are available through the ESDS web site has increased significantly. Equally, the number of new registrations and active registered users increased even more rapidly this year than last. These trends combined might lead to the interpretation that usage of the service is changing its character, with more users using fewer (but more complex) datasets in a more intensive way, supported by a growing body of sophisticated value-added guides and support web pages.

In contrast to the data access and dissemination discussed so far, value-added data dissemination – online data browsing and analysis systems via the Beyond 20/20 for

ESDS International macro data and Nesstar for all other ESDS micro data – continued to increase (Tables 4 and 5). Overall ESDS International macro data sessions increased by 22 per cent from an already high base, with the number of unique users rising slightly from 5,282 to 5,753. On average over the reporting year international macro data were accessed 8,571 times per month by 959 individual users from 77 different institutions. In March 2007 the service recorded its highest ever usage with 1,601 individual users from 82 different institutions accessing the data 14,967 times (Figure 2). Equally, the past year saw the number of Nesstar data analysis sessions rise by 78 per cent, with the number of individual users (as measured by unique IP addresses) also rising slightly from 5,980 to 6,289. As can be seen in comparing Tables 4 and 5, the number of sessions pro rata per individual user are significantly different between the 20/20 interface and Nesstar. This could be accounted by the fact that the two systems function rather differently. In Nesstar univariate statistics can be viewed and generated prior to authentication. Another factor is that the international macro databanks in Beyond 20/20 are time series databanks which are regularly updated with new data releases. This observed difference in pattern of use warrants further investigation and will be examined over the coming year.

When the usage reports for ESDS International macro data held in Beyond 20/20 and other ESDS micro data accessible via Nesstar are viewed by month (Figures 2 and 3) it is clear that the pattern of use is variable over the year in both cases. The period February to April appears to be when usage peaks in both instances, a result most probably of the use of these online resources in classroom teaching.

Table 4 ESDS International macro datasets accessed via Beyond 20/20

	Total number of data analysis sessions	Individual users
2005-2006	84,324	5,282
2006-2007	102,855	5,753

Table 5 Use of Nesstar online data browsing and analysis service

	Total number of data analysis sessions	Individual users
2005-2006	6,888	5,980
2006-2007	12,313	6,289

Figure 2 ESDS International macro data usage

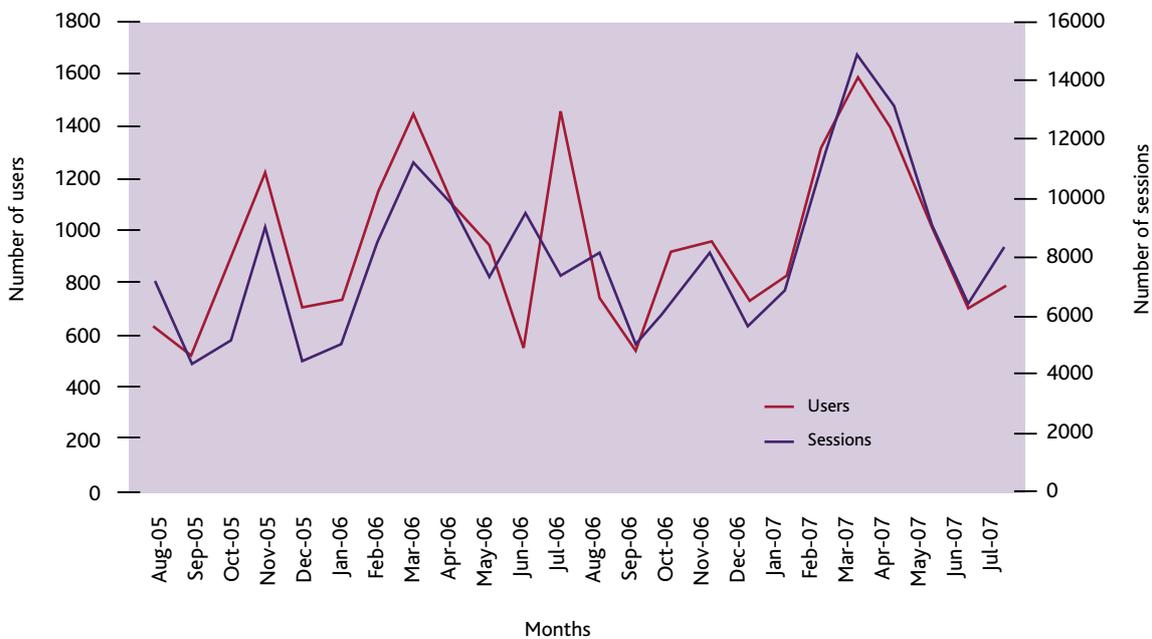
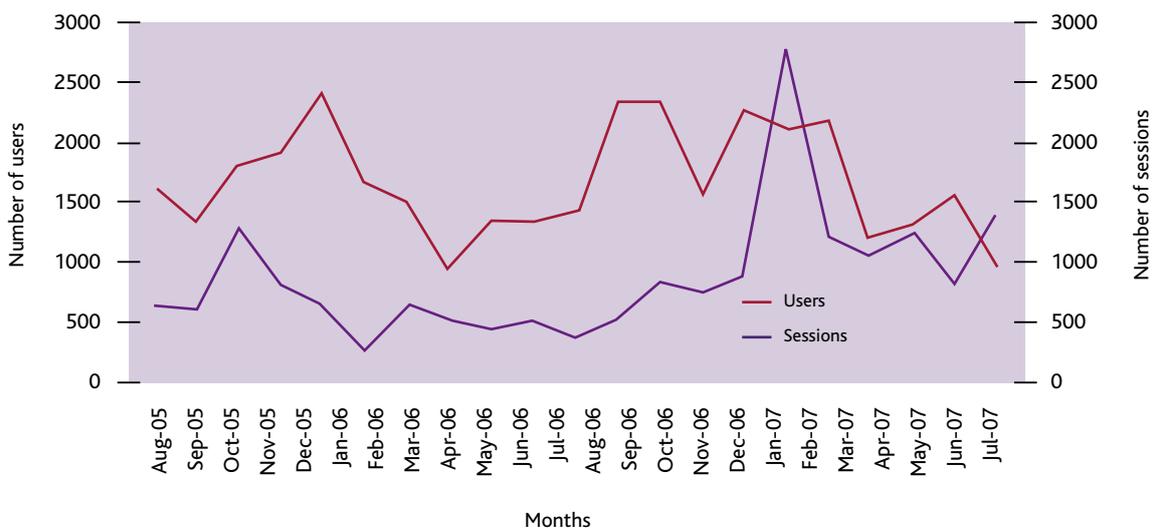


Figure 3 Nesstar micro data usage



User queries

Interestingly, for the first time in the life of ESDS the overall number of user queries fell in comparison to the previous year's figure, from 3,448 to 2,947. Curiously the fall was uneven across the composite services, with ESDS International seeing a dramatic reduction in queries with improved user guides and online FAQs. ESDS Longitudinal and ESDS Qualidata experienced an increased level of user queries as a result of an enhanced portfolio with the release of important new major collections.

Registrations

The reporting year witnessed 22,485 new user registrations, an increase of 17 per cent over last year, whilst the number of active users remained roughly the same at 47,635, a small increase on last year (up by 2 per cent).

Web site usage

The internet is the main vehicle through which ESDS delivers its various data dissemination and support services. The web site not only provides a means through which users can discover and access data, it is also a gateway for a wide range of information about data management, data preparation and the general use of data. The web site presents detailed information for Research Council award holders on preparing data for long-term preservation and archiving, together with advice on issues such as informed consent and confidentiality and pro forma consent forms. Additionally the web site contains a vast quantity of user guides and questionnaires, together with instructional guides on the use

of software and use of data, all of which can be downloaded without the requirement to register. In total, 942,755 PDF files were downloaded in the year, an increase of 17 per cent from 2005-2006.

The web statistics for the reporting period, once standardised and adjusted to take account of internal ESDS use, web-crawlers and page calls to image files, indicate another year of dramatic growth. The total number of page hits in 2006-2007 was 17,120,437, up 81 per cent on the figure for the previous year, and an increase of more than two-fold from the figure for the year 2004-2005. Likewise, the number of visits to the site also rose sharply, reaching a total for the year of 4,349,052, an increase of some 120 per cent on the previous year. The overall trend in web use over the twenty-four month period August 2005 to July 2007 is given in Figure 4. Interestingly, this shows that the number of page hits rose quite dramatically following the launch of the new web site in June. Whilst increased web use is to be welcomed, the fact that pro rata the pace of increase in the number of visitors has been out-stripped by page hits, the average number of pages viewed per visit was 3.28, down from 5 the previous year.

Examining this pattern of increased web usage in terms of the country of origin of the web hit, the general picture is one of continuity and change. Table 8 shows that seven of the current year's 'top-ten' most popular countries were also on the list for the previous year. The list continues to be headed by the United States and UK, with China moving up one place to third. Newcomers to the list include the Russian Federation, which records the fourth highest number of visits but pro rata

Table 6 User queries received by ESDS by service responsible for answering

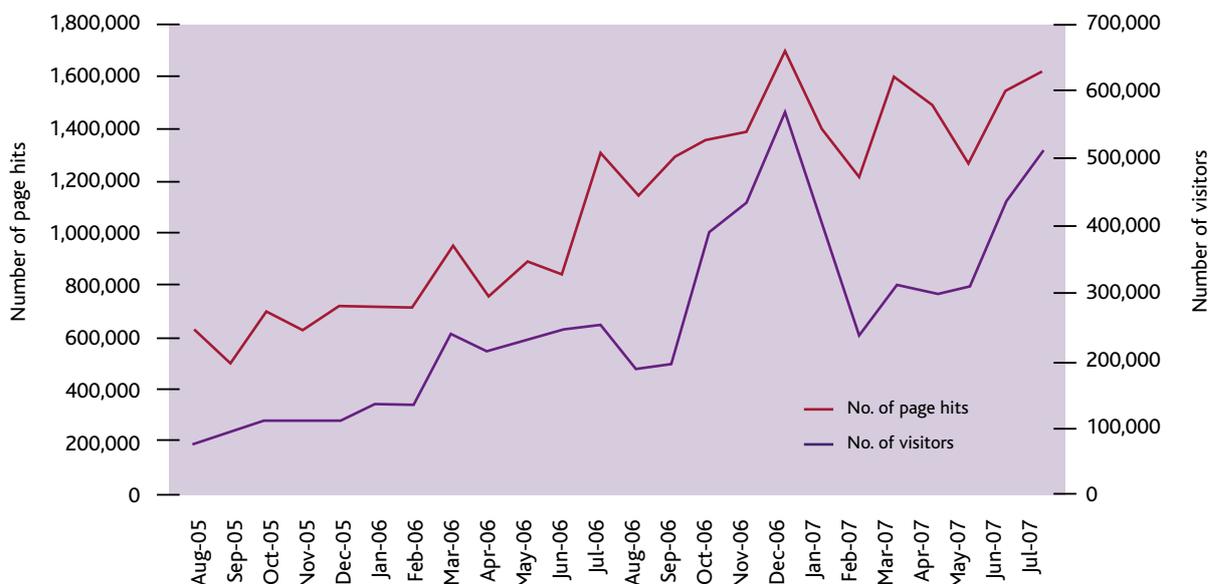
Services	2005-2006			2006-2007		
	Queries	Referrals	Totals	Queries	Referrals	Totals
ESDS Management	2,198	104	2,302	2,029	4	2,033
ESDS Government	382	22	404	285	51	336
ESDS International	497	19	516	284	16	300
ESDS Longitudinal	62	0	62	77	0	77
ESDS Qualidata	163	1	164	200	1	201
Total	3,302	146	3,448	2,875	72	2,947

Table 7 New registrations and active registered users*

	2005-2006	2006-2007
New registrations	19,258	22,485
Total active registered users	46,682	47,635

*includes users registered for ESDS via the Census Registration Service

Figure 4 Web site use: August 2005 to July 2007



relatively few page hits. The top-ten is also, encouragingly, joined by Brazil and India, both specific targets of the ESRC's National Data Strategy. Interestingly, calculating the average number of page hits per visit indicates that the pattern of use of the web site from those within the UK is significantly different to those for users from other countries, reaching 18.5 pages per visit, over three times higher than any other country. In total 42 countries recorded 5,000 visits or more to the web site, with 19 countries recording 20,000 or more visits.

Usage by user type and discipline

This year additional information is provided on patterns of usage across the various services of ESDS according to the status or type of user, as well as the discipline groups from which users are drawn. Turning first to user types, Table 9, shows that across most of ESDS the majority of users are either academic staff, postgraduate or undergraduate students at institutions of higher or further education. In contrast, public sector workers (central and local government, charitable organisations, non-governmental organisations, etc.) account for between 7.3 and 10.5 per cent of service usage. The exceptions to this general picture lie with ESDS Qualidata – who, as one might expect given the nature of the data collections have proportionally fewer users from the public sector – and ESDS International macro data. In the

case of the latter, a higher proportion of student users (post and undergraduate) is recorded. However, it is important to note that licensing arrangements restrict access to these macro data collections to those working or studying within UK higher or further education only, unlike other areas of ESDS.

Moving lastly to usage by subject or discipline groups, although those working in economics (broadly defined) and sociology are two of the largest groups across all the ESDS services, the general picture that is painted is one of variety. Clearly the ESDS user community is a very heterogeneous body. Indeed, it is important to note not just the spread of disciplines but also the fact that a number of users are drawn to ESDS from outside the core social science domain – the sometimes sizable 'Other' category including, in particular, those working in health and medical research, as well as sciences, mathematics and engineering, and library and information science. Again the main exception to this general pattern is with the ESDS International macro data service. Here usage is heavily clustered, with over three-quarters of all use falling with the economics, financial and business categories. ESDS Qualidata also shows a slightly different pattern than the quantitative services, with fewer users from economics and higher proportions coming from sociology, history and humanities.

Table 8 Web usage by country

Country	2005-2006			2006-2007			
	Number of page hits	Number of visits	Rank	Number of page hits	Number of visits	Rank	Mean pages per visit
United States	6,816,540	1,592,545	1	9,473,817	2,972,471	1	3.2
United Kingdom	1,464,418	160,840	2	3,889,283	210,424	2	18.5
China	128,876	17,786	4	235,516	133,668	3	1.8
Russia				77,647	68,686	4	1.1
Canada	96,175	14,282	5	282,306	64,501	5	4.4
Poland	19,674	12,442	6	70,267	59,949	6	1.2
Germany	136,590	17,871	3	219,058	44,430	7	4.9
Brazil				58,324	43,821	8	1.3
France	58,040	10,109	7	213,844	43,766	9	4.9
India				84,251	38,115	10	2.2
Australia	45,005	9,315	8				
Japan	50,611	9,243	9				
Netherlands	32,598	7,860	10				

Table 9 Usage by user status

	Percentages					
	Core	Government	International (macro)	International (micro)	Longitudinal	Qualidata
Academic	32.6	32.0	14.0	37.0	41.3	30.2
Undergraduate	19.2	20.9	39.2	13.5	10.6	22.1
Postgraduate	36.9	32.8	43.6	37.7	37.0	40.4
Public Sector	7.3	10.5	0.0	7.3	8.5	2.1
Commercial	0.8	1.1	0.0	1.0	0.4	0.7
Other	3.1	2.7	3.2	3.5	2.2	4.6
Number of users	1,814	2,372	5,174	289	1,198	285

Table 10 Usage by discipline of user

	Percentages					
	Core	Government	International (macro)	International (micro)	Longitudinal	Qualidata
Business/Accounting/Finance	9.8	5.8	16.6	10.4	3.5	16.1
Economics/Econometrics/Labour/ Employment	22.8	27	59.2	24.2	42.4	7.7
Geography/Area Studies	4.5	5.5	3.4	4.5	3.3	3.5
Health/Medicine	8.5	7.8	0.3	2.8	7.0	5.3
History/Humanities	2.5	1.4	1.3	3.1	0.6	7.0
Politics/International Studies	8.6	4.5	6.9	15.2	1.8	7.7
Psychology	3.5	2.6	0.2	3.5	4.4	7.4
Social Policy/Administration	5.4	5.7	1.0	4.8	6.7	2.8
Sociology	12.6	15.5	1.3	19.7	15.8	21.8
Statistics/Methodology/Computing	5.2	7.6	1.1	4.1	5.9	2.8
Other	16.6	16.7	8.7	7.6	8.5	17.9
Number of users	1,814	2,372	5,174	289	1,198	285

Feedback

"Thank you very much for your two kind messages. They are very useful. Sorry for not replying to you in time because of being outside for a research project. I was moved by your great kindness and warm-heartedness."

"Many thanks for this – just what I needed. Grateful for the rapid response as well!!"

"This is great. The Sponsors are very happy about the level of usage."

"I've finally managed to get a look at the web pages that the Archive has put up for LSYPE and I'd like to congratulate you on a really excellent job. I'm sure the combination of this and the data use workshop in the spring is going to encourage public use of the data in a big way."

"This is perfect. Thanks for your fast response!"

"Mystery solved, and thank you very much for putting me onto those other variables!"

"Good grief, this gets better and better. Thanks you so much for the speedy, helpful service."

"Many thanks for turning the DA into such a professional service!"

"Thank you for the information. I have downloaded the latest versions and have had no problems so far. Thanks once again."

"I did manage to download the data. Thanks a lot for your support. Much appreciated."

"Got it! You guys are fab."

"I have downloaded the data successfully. Thank you once again for your persistence in obtaining this data. It is very much appreciated."

"Wow, that was quick. Thanks!"

"Thanks for all your help with processing this request, it's great that we can now download the material."

"Thanks so much for your help. I knew that it is usually easy to download the data from the data archive but this afternoon I obviously made a mistake during the download process. Now I have the data. Thank you."

"Thanks very much for your help with this. I have been able to download and access the files without a problem."

"Gained a lot from the talks."

"Structure of conference was excellent, by mixing applications from researchers and presentations by international organisations."

"The conference talks were very informative and raised the advantages of the ESDS data services as well as the challenges that still remain to improve access and make more data widely available."

"I feel it has opened up a whole new world of research opportunities!"

"I had no idea of the existence of this data – it could be very useful for my research."

"This will be very helpful for future research and will help speed up the research process."

Advisory Committee

Chair

Jil Matheson

Office for National Statistics

Vice-Chair

John Curtice

University of Strathclyde

Members

Sheila Anderson

Arts and Humanities Data Service

Sara Arber

University of Surrey

Richard Blundell

Institute for Fiscal Studies

Keith Dugmore

Demographic Decisions Ltd.

Peter Elias

University of Warwick

Nigel Fielding

University of Surrey

Peter Halls

University of York

Corrine Moy

National Opinion Polls

Stephen Penneck

Office for National Statistics

Patten Smith (from November 2006)

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

UK Data Archive

Kevin Schürer

Louise Corti

Melanie Wright

In attendance

Martin Bulmer

Question Bank

Siân Bourne (from April 2007)

ESRC

Abi Obafemi

JISC

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

