Scoping study for social scientific research materials based on public records and allied archival materials held at The National Archives

(Summary version)

Janet Foster

and

Matthew Woollard

November 2007

CONTENTS

This report was authored by Janet Foster (The Archive-Skills Consultancy) and Matthew Woollard (UK Data Archive)

1 Executive summary

The social science community in Britain is highly heterogeneous, and its current and potential needs for public records available from The National Archives (TNA) could be perceived as moderate, especially when compared with those of historians. Of TNA's resources, the National Digital Archive of Datasets (NDAD) is seen as most relevant, while access to public records is frequently sought by social science researchers directly from government departments. Many administrative statistics are presently available only through government department websites and because of their dynamic nature are not being preserved for future use. Most, but not all survey data is being collected by relevant data archives.

We recommend that:

- TNA in collaboration with the Economic and Social Research Council (ESRC) and UK Data Archive (UKDA) investigate the feasibility of a (historical) statistics portal to complement the proposals from the Statistics Commission.

- TNA, possibly in collaboration with ESRC and UKDA, create an on-going inventory of government sponsored surveys and statistics (see p21).

- TNA disseminate OSP 30[1] more widely amongst social scientists to inform researchers on current policy.

Social science researchers should not work in an historical vacuum; however resources must be targeted towards the enhancement of high use and expected high use materials. We recommend:

- ESRC to include historical resources within its data policy more explicitly.

- ESRC to consider a call for academic researchers to carry out three research projects in a non-response mode making heavy use of materials at TNA. The research projects should cover: the cost of living in the twentieth century; land use in a long-term perspective; and criminal records. This may be supplemented by a call for additional research projects based on materials situated at TNA. These projects should all have a public face.

- ESRC to consider a call for academic researchers or digitisation bureaux to implement two digitisation projects in a non-response mode relating to the published reports of the Social Survey and historical statistics. This might be supplemented by a call for other digitisation projects based on high TNA content usage.

- TNA/ESRC to investigate widening participation in use of census microdata with existing Licensed Internet Associates (LIAs).

- ESRC to invest in enhancing census microdata for the academic/social science community.

TNA can also widen participation from the social science community in other ways. We recommend:

---

[1] The National Archives, Operational Selection Policy - *Government and People: The Interaction of the State with the Citizen as Documented in Electronic Case Files and Datasets*. See http://www.nationalarchives.gov.uk/documents/osp30.pdf

- TNA to produce a list of records which might be used by social science researchers.

- TNA to examine the costs of further enhancement and dissemination of existing datasets by NDAD/AHDS History or other services.

- TNA to initiate procedure of informing NDAD/UKDA when a new series is included on DocumentsOnline for their assessment for retention.

- TNA to consider enlarging scope of OSP 30 to include datasets from independent non-governmental organisations and non-departmental public bodies.

- TNA to produce focused and directed promotional information for social scientists, via wiki and print media.

- NDAD to produce focused and directed promotional information for academic researchers (via Intute and other resource services).

- TNA to improve visitor statistics by having more meaningful user categories and to include a greater scope of their collection.

Each of these recommendations should be backed up with the relevant standards. We recommend that:

- the ESRC and TNA ensure that any proposals carried out must adhere to JISC and other standards relating to management, web usability and functionality, dissemination and preservation.

Furthermore we recommend that:

- to prevent duplication of effort all relevant organisations should investigate the benefits of shared services for the creation, enhancement and dissemination of any resources.

- any and all enhancements should be free at the point of use for academic users.

2 Introduction

The National Archives (TNA) in collaboration with the Economic and Social Research Council (ESRC) published a tender for a scoping study "to explore further the scope that exists for the development of research resources for social scientific research based upon archived records and allied material." These archived records were explicitly "public", and hence either presently within the collections of TNA or such that might fall within TNA's selection policy.

The specific terms of reference of the study were to:

a) assess current and potential need and demand within the social science community for research resources based on the public records and allied archival material;

b) advise on the types of research resources for which there is likely to be a demand, both qualitative and quantitative, and to comment on the potential new costs of creating, disseminating and preserving such resources consistent with TNA's vision and strategy, and research strategy;

c) evaluate and develop a priority list of research resources (candidate series or clusters of records) which are likely to be of greatest value and interest to social science research;

d) advise on Higher Education centres of expertise which might be interested in submitting proposals for projects to develop research resources, if possible matching these to the candidate series and in line with TNA's research strategy;

e) advise on the feasibility of setting up a resource development programme, based on TNA's collections; and

f) propose up to five priority projects to be scoped out in more detail as the next stage of this work.

Each of these terms of reference has been addressed implicitly in detail in the main body of the report which follows, and explicitly within Section 5 (Recommendations) below.

The original TNA tender used the term development to mean "improve"; for this study we have chosen to use the word enhancement as it is probably more commonly used amongst the academic community. We noted in the questionnaire (see below) that in this context enhancement may range from more detailed cataloguing to digitising and making available via the internet.

This study has been informed by a questionnaire as well as related documentation, in particular ESRC's key strategic documents *Data Policy* and the *National Strategy for Data Resources for Research in the Social Sciences*.[2] The development of key data resources in the social sciences enabling "increasingly more sophisticated social science research to be undertaken" is an important responsibility of the ESRC, and while its major commitment must be to data creation and support of existing surveys it has a responsibility to social and economic history; an area where some sophisticated social science research can only be carried out where stringent confidentiality requirements are not imposed.

---

[2] Economic and Social Research Council, *Data Policy* (April 2000). Available at: http://tinyurl.com/2gh72e and *The National Strategy for Data Resources for Research in the Social Sciences*. Available at: http://tinyurl.com/3c6lcs

This scoping study complements the *National Strategy for Data Resources for Research in the Social Sciences*, where no reference is made to historical resources which can be used to inform social studies: as most social research benefits from long-term trend analysis we feel that historical material should explicitly form part of this strategy.

Where possible we have followed the four main criteria mentioned in that strategy for data collections: inclusion of very detailed data; promotion of data which allows the monitoring of long-term trends; inclusion of data which allows links to be made between other data and thus increases the value of any previously collected data; and attempting to include the most diverse populations wherever possible. We are also of the opinion that administrative data (ie statistics) as well as microdata are important to the remit of this study. Consequently these forms of data are highlighted in our recommendations.

This study is also informed by the JISC-funded study on *Digitisation in the UK*[3] which provides evidence of materials which have already been digitised, but largely ignores 'born digital' data which are highly used by social scientists.

Wherever possible we have attempted to focus on user needs. Our analysis has been hampered by:

- a very diverse and heterogeneous user community;

- a low ratio of responses to our questionnaire from the social science community, possibly linked to survey fatigue;

- problems associated with the usage data from TNA;

- the weighting of responses in our questionnaire by social and economic historians as opposed to all social scientists.

These flaws do not invalidate the main conclusions of this study, but lead to the assumption that the social science research community has highly specific needs in relation to public records which require a diverse range of approaches to encourage their use of public records and historical sources.

---

[3] JISC, *Digitisation in the UK: The case for a UK framework*. Available at: http://tinyurl.com/2pfl7k

3 Methodology

The proposed methodology for this study was based on four interlocking workpackages. For a number of reasons we were not able to carry out our proposed workpackages to the letter. The workpackages were:

WP1: Data review

This package was to review existing literature, in particular the Research Libraries Group's *Information needs in the social sciences* published in 1989 and the more recent British Academy Policy Review for impacts on this study.[4] We also examined the ESRC's *Data Policy*.[5] These have informed and directed both our methodology and our conclusions. We had intended to produce a short document summarising earlier research into the provision of social science research materials, in particular those housed in TNA. However, this proved impractical as most other studies were not relevant to this particular study.

WP2: Data collection I

Workpackage 2 was to be undertaken in collaboration with TNA. For benchmarking purposes TNA was to provide use-figures for commonly used collections which include social science research resources. We had originally hoped to look specifically at resources which could only be properly analysed through ICT based-methods. The proposal stated that "It should be noted that current demand (ie use) does not necessarily provide accurate evidence for potential demand. Demand for a valuable resource may be low because it cannot easily be used in its current form, or the costs of digitisation make this impossible for use for a single researcher."

The stated outcomes for this workpackage were a) a list of highly used resources in TNA's collection for social science research and b) a list of less used (or unused) resources in TNA's collection for social science research, which might benefit from digitisation and c) a list of digital resources in TNA's collection which could be repurposed for social science research. The figures which TNA was able to provide were not particularly helpful for these outcomes as TNA lacks experience in the identification of material suitable for social science research and TNA does not collect usage statistics which allow the identification of social science researchers.

WP3: Data collection II

This workpackage was to address current and potential need and demand within the social science community for research resources available from TNA, and was to be carried out through a dual strand user survey. First, a web-based questionnaire was directed towards social scientists generally, and second a series of "focus groups" was to be held. The first was reasonably successful despite a low ratio of responses to invitations, but the second was not, for a number of reasons including a reluctance of a critical mass of people to attend each of the planned groups, despite almost 80 personalised invitations being sent out to those from relevant disciplines who had indicated in the questionnaire that they would be willing to attend a focus group. We asked

---

[4] *E-resources for Research in the Humanities and Social Sciences: A British Academy Policy Review* (April 2005).
[5] Economic and Social Research Council, *Data Policy* (April 2000). Available at: http://tinyurl.com/2gh72e.

permission from the ESRC to cancel this element of the project and instead concentrate on a smaller number of selected telephone interviews.

The email notifying cancellation of the focus groups invited participants to volunteer for a telephone interview. This solicited five potential interviewees representing medieval history, time-use study, contemporary economic datasets, living standards and race relations, although the latter interviewee incidentally provided information about ships log books as well. This reluctance to self-select is probably an indication that academics wish to protect the resources that they use for their research. Of these interviewees four responded when asked for a convenient date and time for an interview. Interviews were conducted by telephone and, according to their research interests, interviewees were either asked about their use of TNA resources and which resources they would recommend for enhancement, or whether they were aware of or had used NDAD and which Government datasets they would like to access.

Other interested academics were contacted by email asking for comment on particular series in their field but, with one notable exception, this approach was not successful. Representatives at the Royal Statistical Society were also consulted and, at the suggestion of a TNA staff member, the project director of the Reconstructing Consumer Landscapes project at Surrey University was contacted and interviewed. This small sample cannot provide overall conclusions, but individual comments have been noted as appropriate to corroborate and support the questionnaire results and to inform the selection of series in Section 5 as well as contribute to the recommendations. The list of telephone interviewees and email contacts are given in the full report.

WP 4: Gap analysis and report

This workpackage was intended to carry out an informal gap analysis estimating the market potential for different collections within TNA's holdings based on the online questionnaire and subject-based focus groups. This was to be compared with an estimate of existing usage of these resources based on TNA use figures. This was also to estimate the value added potential created by digitisation of these resources. Again with the usage figures available from TNA, and with variable quality results from the questionnaire, this has not been as successful as we had originally anticipated.[6] However, the conclusions to this study, as outlined in Section 5 do cover the main objectives, and are based on a mixture of qualitative and quantitative evidence.

We feel confident that the methodologies employed for this study give acceptable results. However, both the response rates to our questionnaire and the highly heterogeneous nature of the potential user community for this material means that all statements and conclusions need to be tempered.

The following section discusses the usage of TNA by social scientists, and the demand for material at TNA. The information in this section has been provided by TNA.

---

[6] We do not intend to imply that TNA's production figures are inaccurate, simply that production figures only relate to actual physical productions at TNA and that the classification of user types is too broad for a detailed analysis of academic reader types.

Usage and demand

All usage/demand information received from TNA only refers to services provided at TNA at Kew; it neither includes visits to the Family Records Centre (FRC) or unmonitored self-service microfilm use either at Kew or the FRC, nor does it include documents made available electronically via the DocumentsOnline service.[7]

Actual reader visits to TNA at Kew, by self-registered but pre-defined user-type categories, for 2006 are shown in the table below. TNA do not use social science as a research type, but we have assumed that those describing themselves as social and economic historians form the majority of social science users at TNA. The table demonstrates how small a proportion of use is potentially related to social science usage at TNA.

Table 1: Actual reader visits to TNA at Kew in 2006, by user-type categories

Research Type	Number of reader visits
Family history or genealogy	23,889
Military history	15,962
Foreign or colonial history	4,516
Social or economic history	4,496
Political / diplomatic history	4,238
Local history	3,133
Other	1,619
Legal history	1,367
Transport history	1,318
Design, cartography etc.	764
Total	**61,302**

There are some difficulties in meaningfully interpreting these figures as the grouping "family history or genealogy" might be selected by some academic historians of the family, and some genealogists might describe themselves as social or economic historians. We can only conclude that TNA might improve visitor statistics by having more detailed categories, and by possibly including main place of work.

Usage of TNA by social scientists could also be examined by looking at productions of materials in the reading room. However there are problems here as no materials will be exclusively examined by one of these "research types". Furthermore all the information provided specifically related to usage at Kew rather than at the FRC which houses many records fundamental to historical social research and thus falls into the category of social science research.

There are other issues which are highlighted in this statement[8] regarding usage from TNA of material which was considered to be of value to social scientists:

> Statistics of document productions are also pretty basic. Arranged by record series, they show that Foreign Office general correspondence was the most popular in 2006 (276,050 productions), followed by WO 97 soldiers' pensions records, some way behind at 123,774. The other top 10

---

[7] DocumentsOnline allows access to digitised material from TNA's collection. It is predominantly, but by no means exclusively, for personal family history research. Many resources have a small charge for external viewing, but all use is free within TNA and the FRC.
[8] Email from TNA staff to Matthew Woollard 15 December 2006

... are primarily genealogical with one local/house history source (although with some potential for detailed analysis of housing conditions) in IR 58. Home Office general correspondence in HO 45 puts up a fairly strong show at 17,644 but many of these may relate to naturalisation papers ordered up by genealogists. Treasury Papers in T 1 (9,862 productions) could be used for a myriad different purposes but may include some social scientists. Heavy usage by a few dedicated individuals may distort the statistics eg 6,448 Common Pleas Plea Rolls are largely explained by one man and his daughter taking digital snaps. However not one Ministry of Labour (LAB) series, which one might expect social scientists to make use of, makes the listing…which goes down to the level of 3,164 productions.

Of these materials WO 97 are records from the Royal Hospital at Chelsea, and are primarily of interest to genealogists, but given that they also contain the service record of pensioned soldiers they could provide information for social science research. These records have recently been re-catalogued at an individual level which explains their heavy use. IR 58 contains the Field Books for the Valuation Office records of 1910. Records in HO 45 are predominantly used for genealogical purposes and also have enhanced catalogue records.

TNA also provided usage figures (again, for materials at Kew only) for "social and economic historians" in 2006, but only for material which could be dated 1900-1999. These did inform our analysis, but only to a limited extent. The table below shows only those materials used by more than 50 individuals.

Table 2: Usage figures for "social and economic historians" in 2006 (for material dated 1900-1999)

Code	Series	Series Title	Orderable Units	Requisitions	Readers
FO	371	Foreign Office: Political Departments: General Correspondence (1906-66)	3199	3657	333
HO	45	Home Office: Registered Papers	634	780	157
MEPO	2	Metropolitan Police: Office of the Commissioner: Correspondence and papers	321	382	90
WO	32	War Office and successors: Registered Files (General Series)	240	272	90
CAB	21	Cabinet Office and predecessors: Registered Files (1916 to 1965)	190	233	85
HO	144	Home Office: Registered Papers, Supplementary	231	302	76
DO	35	Dominions Office and Commonwealth Relations Office: Original	252	270	75
PREM	11	Prime Minister's Office: Correspondence and Papers, 1951-1964	214	258	72
CO	537	Colonial Office and predecessors: Confidential General and Confidential Original Correspondence	246	287	72
CO	323	Colonies, General: Original Correspondence	218	254	71
ADM	1	Admiralty, and Ministry of Defence,	213	245	55

Code	Series	Series Title	Orderable Units	Requisitions	Readers
		Navy Department: Correspondence...			
PREM	13	Prime Minister's Office: Correspondence and Papers, 1964-1970	153	187	53
PRO	30	William Akroyd: Papers	106	122	53
BT	31	Board of Trade: Companies Registration Office: Files of Dissolved Companies	676	843	52

TNA also provided some information about numbers of productions of particular series of documents, but after some consideration decided that productions (ie usage) were of limited use because, if not carefully analysed, they could show particularly distorting results. Single productions are not comparable across series especially in terms of volume of relevant information. In terms of enhancement a single document production may yield information which provides a single record in a database, or 100,000 records. For example, production of a single will from PROB 11 might yield a single entry in a research database, whilst production of a single register in HO 26 might yield many thousands of records in a database. Therefore the workload involved in digitisation/enhancement cannot be judged simply by a calculation based on number of pieces.

Furthermore the number of productions may vary according to the level of detail provided in the catalogue. LAB 17 was little used during 2006 according to this table, but LAB 17 is not catalogued to the same level of detail as LAB 2.

Table 3: How usage is influenced by level of cataloguing

Code	Series	Series Title	Orderable Units	Requisitions	Readers
LAB	17	Ministry of Labour and successors: Statistics Department and Division: Registered Files and other records	8	9	7
LAB	2	Ministry of Labour and Predecessors: Correspondence	179	204	72
RG	23	Government Social Survey Department: Social Survey: Reports	21	26	16
IR	19	Board of Stamps: Legacy Duty Office and successors: Specimen of death duty accounts	129	196	24

IR 19 (Specimens of Death Duty Account) is shown in this table because it is currently being used by Dr David Green of King's College London for an ESRC-funded project on women investors in England and Wales 1870-1930. A database is being created amalgamating sources from IR 19 on shareholding and the census enumerators' books. LAB 2 is mainly made up of Trade Board files and the relevant catalogue entry has recently been enhanced.

This discussion demonstrates that assessment of social science usage from existing production or barrier figures is relatively unsuccessful, and does not help to assess the overall current demand from social scientists. This is troublesome for this study because any development or enhancement should only be carried out with evidence supporting likely usage. We will return to this question in the discussion about the questionnaire in Section 5 below.

4 Analysis

The full version of this report contains two additional chapters. The first comprised a detailed investigation into 26 discrete series at TNA which were either identified by researchers or our own additional research as of value to a wider social science community. The second chapter comprised a detailed analysis of the returns from the questionnaire. This analysis informed the recommendations which are made in Section 5 below.

In the full version of this report some additional conclusions were made which might impact on the decision making-process. It is important to note that the questionnaire was circulated to around 45,000 email addresses, representing the UKDA's registered users and all recent grant holders of the ESRC. The imprecision of the population is based on the probability of a level of overlap between the two email lists. Only 785 responses were returned, and these were predominantly from the UK Higher Education (HE) sector, and were broadly representative of the ESRC core community, with over-representation amongst economic and social historians and psychologists. A little under a third of respondents had previously used TNA, while a little over three-quarters of those describing themselves as economic and social historians were TNA users.

Usage of TNA materials by all respondents was considerably diverse and we concluded that there are only a few sources being used by a number of social scientists.

Of those who had not used materials at TNA the majority were unaware of resources relevant to their research, and for many of them, it is likely that their unawareness was related to the lack of *directly* relevant materials. We concluded that awareness-raising amongst the social science community is not only needed for TNA but also for its holdings; and that enhancement of the catalogue might provide better value for money in raising awareness of TNA's holdings amongst the social science community than a small number of large scale digitisation projects. However, it is clear that for some researchers, awareness of resources has been increased by existing, or forthcoming digitisation projects.

We explicitly asked researchers about their awareness of resources which might assist researchers within the ESRC's strategic research areas. The responses were on the whole inconclusive, though they did provide some evidence for possibilities for further enhancement, and these recommendations are made in Section 5. We also explicitly asked about the value of enhancing some particular series held by TNA. These pre-selected materials short-listed by the project steering committee. The final list of 12 series related to a combination of current usage, a narrow historical spectrum predicated on social scientists' projected needs and material which could clearly benefit from enhancement.

The resource names and analysis of responses are shown in the table on the following page. These results suggest two things: first, social scientists are not particularly interested in historical material; and second that only the broadest interdisciplinary material was highly regarded. However, the relatively low percentages should not be taken as discouraging, if, for example, 27 per cent of those who answered the questionnaire believe that digitising and enhancing the Ministry of Labour's Family Budget Enquiry of 1937-8 would improve *their own* research, then we might assume that the impact of enhancing any of these resources would produce rewards to social science research. The responses to

this question suggest that the greater the interdisciplinary range the better the response.

Table 4: Analysis of responses for enhancement of 12 suggested resources

Q	Name	No answer	Yes - would improve my research	No - would not improve my research
		%	%	%
13.10	RG 23: Government Social Survey Department: Social Survey: Reports and Papers	22.9	43.2	33.9
13.3	CAB 108: Statistical Digests	24.5	35.5	40
13.1	BN 35: DHSS statistical reports (1959-1997)	23.2	35.3	41.5
13.7	LAB 24 and LAB 81: Ministry of Labour: Household and Family Expenditure Survey returns and sample	24.3	34	41.7
13.9	MH 55 and MH 66: Ministry of Health: Records relating to the Public Health Survey 1930-4	25.9	28.3	45.9
13.6	LAB 17: Ministry of Labour: Statistics Dept. Family Budget Enquiry, 1937-38	25.6	26.5	47.9
13.2	BT 70: Board of Trade: Census of Production and of Distribution	26	21.5	52.5
13.12	Moving Here website	36.3	15.4	48.3
13.8	MAF 32 and MAF 73: Ministry of Food: National Farm Survey, Individual Farm Records 1941-43 and associated maps	29.3	14.9	55.8
13.5	HO 140: Home Office: Calendar of Prisoners	29.2	14.6	56.2
13.11	T 267: Treasury: Historical Memoranda 1957-76	29.4	13.4	57.2
13.4	IR 58 etc.: Inland Revenue: Valuation Office: Field Books	29.9	12.6	57.5

A more detailed analysis of these materials would be necessary to carry out a full cost-benefit analysis, but by ranking the approximate costs of enhancement one could use these results to inform a digitisation programme. Interestingly, the materials which could lead to research projects (and thus cost more) were generally lower ranked than straight-forward digitisation projects.

Respondents were given ample opportunity to make comments on their responses. The homogenous nature of those who did respond made it very hard to make wide-ranging conclusions. We found two major problems. First, that some respondents found these suggestions were too modern and second, more commonly, that these suggestions were too historic.

The difficulty we have faced is that the remit of the study was to investigate social science resources in a repository which by its very nature mainly deals with material which is more than thirty years old and often more than fifty years old, and the majority of "social scientists" generally prefer to concentrate on more contemporary material. Hence, the sometimes seemingly biased conclusions to the questionnaire, where, if perhaps an economist grudgingly

suggests that enhancing one resource over another might be of benefit to them for "background material", this is weighted more than if a historian whose research would be vitally affected by the enhancement of the same resource.

We feel that the responses to this question and the previous question show that identification of material at TNA is challenging for social science researchers who are not used to working with historical material. However, what we did extract from the respondents was some suggestions relating to technical issues surrounding digitisation which have led to our particular recommendations surrounding usability testing.

Respondents also gave some support for other collections of materials being digitised. Amongst the social and economic historians these fell into four main categories: medieval records, census/demographic sources, cabinet minutes and memoranda and some other more miscellaneous sources. Among other users, census-type materials were most prominent, which led us to suggest that TNA might wish to consider prioritising the enhancement of material which is already electronic for the social science community rather than material which is currently in paper-only form. Some specific examples for enhancement were given by researchers in other disciplines, but two interesting additional possibilities also surfaced. First, the concept of a wiki for TNA – which has subsequently been launched, second, collaboration with citation agencies on improving catalogue access from these agencies portals.

We were particularly underwhelmed by the responses to enquiries about the enhancement of previously existing guides and we felt that short promotional guides to these published guides may better promote their use and sales.

We discovered that NDAD had rather a low visibility amongst the HE sector, and many of those who indicated that they had used the NDAD service had in fact used the UKDA. We conclude NDAD needs to promote itself better, especially within the social science community, and that its selection policy could be reviewed to become more coherent.

Researchers were more likely to use other electronic data sources than NDAD, and most, but not all, of these materials were also available elsewhere. We concluded that either most "important" government datasets are already available for use from the UKDA/ Economic and Social Data Service (ESDS) or from the government departments themselves or that researchers' use of contemporary data is driven by the source of dissemination, ie the UKDA. However, there are many data collections created by government departments which are not in either the NDAD or UKDA collections, are not available from government departments, and are not archived in TNA's web archive as they exist in separate databases. This study firmly supported the preservation and archiving of digital material located in databases linked to websites as standalone datasets, with supporting documentation. We understand that some of the original functionality may be lost, but it is better to lose some functionality than the whole dataset.

The full report's section on analysis ended with a quote from one researcher which we felt might, to a small degree, sum up the responses of the non-historical researchers.

> "The only thing I can think of likely to add value for me would be digitising the actual data contained in historical statistical reports, so that one could simply download it eg into an excel spreadsheet. But that would be a

mammoth task - better funded on a case-by-case basis for particular historical research projects. Not sure this is where I would invest scarce resources - let historians dig around the historical archives and instead spend scarce money improving datasets for current scientific research."

We do not feel that the results of this questionnaire bear these feelings out. While there was a considerable lack of unanimity amongst respondents about what, or even how, material within TNA might be enhanced, there was a clear message that more material should be made freely available, with better search facilities and as reliably as possible.

5 Recommendations

Objective One of this study was to assess current and potential need and demand within the social science community for research resources based on the public records and allied archival material.

This study shows that the social science community[9] is highly heterogeneous, and that the current and potential need of this community for research resources based on public records available from TNA could be perceived as moderate, especially when compared with those of historians. Of TNA's resources, the NDAD datasets are seen as most relevant, as well as those which are not present in NDAD, but fall within the remit of NDAD.

A similar general conclusion may be drawn at a discipline-specific level, but there is evidence of a range of public records, both within TNA and government departments which would significantly improve researchers' ability to conduct research. The need is not great, but neither is it slight. For example, many researchers, and across disciplines, would benefit from access to microlevel census data from the nineteenth and early twentieth centuries. The need for this material can be demonstrated by the fact that some researchers have already made local level samples of these data for their own use, but real need cannot be ascertained in this manner. Demand for some resources is present and could be measured in a number of ways, explicitly current use and in terms of response to an explicit list of materials provided in the questionnaire. We can demonstrate, for example, that HO 144 which includes naturalisation papers for the period 1879 to 1934, is in demand by respondents to our questionnaire, but not whether many social scientists would use it. Demand in general has been shown to increase when better cataloguing has occurred. We believe that even making some material available in a machine-readable manner would increase demand further. For example, if the "metadata" which has been added to the catalogue records for HO 144 was available in a database rather than simply through the catalogue it would increase demand. There are other series which would benefit from this form of enhancement, and there are series which would benefit from both improved cataloguing as well as making these records available in machine-readable form.

Amongst social scientists, demand for resources currently within TNA seems to be less than demand for government records which are within government departments (and thus technically not public records until accessioned by TNA). This may be because of a general disregard of what are perceived as historical sources, as one interviewee noted "most social scientists working on contemporary problems never set a foot in TNA, or think history is anecdotal and soft."

The potential needs of researchers must also take into account what currently exists but is not yet available through either TNA (especially NDAD) or UKDA. There are clearly some datasets, for example, the London Area Transport Survey, the Household Below Average Income (HBAI) and the Work and Pensions Longitudinal Study (WPLS) both created for the Department of Work

---

[9] Amongst others, anthropology, area studies, demography, economics, education, environmental sciences/planning, gender studies, government policy, hospitality and catering, human geography, socio-legal studies, law, linguistics, management and business studies, political science, psychology, social policy, social welfare, sociology, travel and tourism.

and Pensions (DWP) that have material which is not currently publicly available and only parts of the Offenders Index database are available through UKDA. There are doubtless other surveys and collections of statistics which are not available through the UKDA which may once have been available on government department websites. For example one interviewee cited the New Earnings Survey which used to be distributed by the Office of National Statistics (ONS) but is now on restricted access for a daily charge.

In general, the potential demand for government datasets is mostly met by the UKDA or other similar services.

The current use may be relatively small, but it is very wide. We also note that increased digitisation, and improved resource-discovery tools generally lead to wider access. Many social scientists, especially historians, will find uses for material which they would not have used previously once it is easily available.

Objective Two was to advise on the types of research resources for which there is likely to be a demand, both qualitative and quantitative, and to comment on the potential new costs of creating, disseminating and preserving such resources consistent with TNA's vision and strategy, and research strategy.

There are three general types of materials for which there may be demand. Each of these types will be of value to different segments of the ESRC/TNA stakeholders. We suggest evaluating material in three broad conceptual classifications: a) statistical; b) microdata/database and c) reproductions. We have used these classifications to inform later discussion.

The costs relating to the complete digital life-cycle for each of these groups is reasonably homogenous. For example, statistical material is relatively cheap to re-key, easy to maintain, disseminate and preserve. Microdata are much more expensive to create — even from paper originals and harder to maintain and preserve. Reproductions, ie digital facsimiles, have a high overhead in terms of appropriate metadata creation, and may be much more expensive to disseminate, if a fully functional website is needed.

In general terms however the life-cycle costs of digital materials relevant to this study are likely to be positively correlated with their potential demand. One of the purposes of this study is to identify materials which may prove anomalous to this rule of thumb. The cost, for example, of re-purposing previously existing digital indices (ie to the census enumerators' books) or TNA catalogue material for social science use will be much less than creating them from scratch, so can be considered to be cost-effective. Whether or not these provide best value for money is another question.

Evidence from the questionnaire and follow up calls suggest that any materials enhanced or developed for the social science community should be free at the point of use within academia, and if applicable, should be available on an Open Access basis for all-comers. Sustainability of any of these resources will also be an issue to contend with. Again, as a rule of thumb, the more complex a resource the greater the costs are to sustain it. Careful consideration should be given to ensure that enhancement minimises the challenge of long term accessibility and preservation issues.

Objective Three was to evaluate and develop a priority list of research resources (candidate series or clusters of records) which are likely to be of greatest value and interest to social science research.

17

Most of the evidence for this section comes from the questionnaires and telephone conversations. The materials within TNA which are most likely to need enhancement for social science research are detailed in Section 4 above. Though there is clear evidence for born digital materials within the scope of TNA which are not mentioned in that section.

Objective Four was to advise on Higher Education centres of expertise which might be interested in submitting proposals for projects to develop research resources, if possible matching these to the candidate series and in line with TNA's research strategy.

None of the researchers mentioned in this section has been approached relating to this particular question. However, there are existing researchers who might play a part in providing academic advice on the creation of datasets. For example, David Heald (Sheffield) on Public Expenditure; Virginia Berridge and/or Martin Gorsky (both London School of Hygiene and Tropical Medicine) on health matters; Bernard Harris (Southampton) on health relating to social policy; John Hills (LSE) on social exclusion, Jane Lewis (LSE) on gender, social policy and family policies; Frank Trentmann (Birkbeck) on consumption, Robert Page (development of welfare state and social policy in general) and Eilidh Garrett and others for the use of military records for demographic purposes.

We recommend that any project to develop resources should be set up as a partnership, between TNA, academic researcher(s) within the field, and a third-party (but probably academic) web-development team. Examples, within the historical field of the latter include the Centre for Computing in the Humanities (King's College), the Humanities Research Institute at the University of Sheffield, or even the UK Data Archive at the University of Essex. While we do not exclude the possibility of commercial partnerships in projects *with a significant academic research element*, we would recommend that these are considered cautiously, and should generally be used for dissemination and sustainability purposes.[10] It will also be worth exploring partnerships with other research councils, most specifically the AHRC to meet the needs of extended communities.

Objective Five was to provide advice on the feasibility of setting up a resource development programme, based on TNA's collections.

There is clear evidence that initiating a resource development programme would be feasible. However, the form of this programme would, to a very large extent, depend on the levels of external funding, and the necessary will within TNA. It is always difficult to assess the impact of digitisation/development projects. A well-promoted and well-organised academic project, like the Old Bailey Online Project,[11] has achieved considerable success within both the academic and more popular user communities. However, it is self-evident that disseminating any material via the internet increases usage. The vital question is whether or not there is a positive relationship between the cost and the benefits of such a project. We believe that there is enough relevant material at TNA to set up a resource development programme which will benefit both the core social science user community as well as the public at large.

---

[10] We do not mean to imply that commercial partnerships with a high personal family history element are not considered cautiously.
[11] Details at: http://www.oldbaileyonline.org/

This programme would need to scope the resources to be developed in more detail than this project could, but these recommendations and those below provide a starting point.

Without clear and unambiguous usage figures from the main target community, it is difficult to measure the potential benefits of any development project, and the programme would need to analyse in detail the costs of such projects. This will be addressed briefly in the following paragraphs.

Objective Six was to propose priority projects to be scoped out in more detail as the next stage of this work.

There is evidence from the questionnaires that a very wide range of material might benefit from digitisation, though the overwhelming feeling we find from this survey is that there is a positive relationship between the complexity of material (and thus its cost to digitise) and its utility to the widest range of stakeholders. (Essentially, these results show that while there is some relationship between the needs of the ESRC on one hand and of social scientists on the other, this relationship is at best limited.)

The ESRC might wish to provide resources for a limited response-driven "programme" for the enhancement of certain resources within TNA, where a user-based consultation is carried out in selecting projects to be funded, in a similar manner to the recent JISC digitisation projects.

Furthermore, we would recommend that all material is free at the point of use for academic users.

While we believe that such projects need to be supported by evidence of likely usage, it is also necessary to examine cost mechanisms for sustainability and level of ESRC/TNA funding. This can only be carried out with a more focused study, but the results from the questionnaire on specific materials show that there are some which would benefit from immediate development. In these examples there is a clear demonstration of demand from within the user community. These are outlined below.

We have divided our recommendations for project work into five categories. All relate to the development and promotion of use of TNA's materials, but not all are traditional resource enhancement projects.

5.1.1 Research projects

Under this heading are those projects which benefit the widest academic communities and will influence both current ESRC strategic aims and those of TNA. These projects combine research objectives as well as digitisation objectives. These projects could not be undertaken without specialised academic input at the design and conception stages.

5.1.1.1 Cost of living in the twentieth century

This project would bring together materials from disparate sources both within and without TNA. It would probably have an output of a popular website for public use allowing prices and wages to be analysed. Amongst other material, LAB 17, 24 and 81 could be integrated in such a project with published materials.

5.1.1.2　National Farm Survey

Again we recommend funding of a research project examining land use in the UK across the twentieth century which would be heavily influenced by materials housed at TNA, but with a popular (ie genealogical and pedagogical) element. The materials for this project are mainly within MAF 32 and 73. These particular sources are discussed in the full report for this study.

5.1.1.3　Long term criminal records

This project would use and combine information from HO 27 and HO 140, and potentially the Offenders Index data from the Home Office. The linkage of these datasets would allow very long term crime trends to be examined.

Each of these three research projects would be likely to have costs (if carried out in the academic arena) of around the half million pound mark, although the cost of living project could be developed into a major resource costing, in the long run a lot more. The National Farm Survey with its emphasis on creating digital versions of coloured maps, and possible use of GIS (Geographical Information Systems), would be more discrete, but the opportunity of raising some revenue for a "value-added" public service would be greater. The final proposal here would also have some opportunities for a fee-paying service, but also long-term academic value. A project like this could easily be integrated into a site such as the Old Bailey Online.

The sustainability of digital outputs of each of these projects would need to be examined most carefully, as without dedicated project staff these outputs might rapidly (ie within five years) become redundant. Placing these projects in the academic environment might attract additional institutional support which would not be forthcoming elsewhere. However, two of these proposals might attract a limited income stream from the personal family research community.

Recommendations

- consider call for academic researchers to carry out these three research projects (non-response mode) and/or call for research projects with high TNA usage. [ESRC]

- any proposals carried out must adhere to JISC and other standards relating to management, web usability and functionality, dissemination and preservation. [ESRC/TNA]

- any proposals carried out under this heading should be free at the point of use for the academic community. [ESRC/TNA]

5.1.2　Digitisation (image and text capture)

Any digitisation project carried out under this study needs to be informed by current standards. The various JISC standards are preferred as these are internally consistent and allow a level of project-specific latitude.[12] Digitisation projects currently being funded tend to use the basic building blocks of Dublin

---

[12] Minerva, *Technical Guidelines for Digital Cultural Content Creation Programmes.* Version 1.0: revised 8 April 2004. Available at:
http://www.minervaeurope.org/publications/technicalguidelines_en.htm and JISC Information Environment Technical Standards. Available at: http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/standards/ (last updated May 2006).

Core (DC) for resource discovery metadata, MODS[13] (for more detailed bibliographic material and based on MARC 21)[14], MIX[15] (for digital images) and wrapping within METS[16]. The use of EAD[17] and ISAD(G)[18] for archival based digitisation projects probably does not help to promote interoperability.

5.1.2.1   Social survey reports

To digitise and make available via a bespoke website the complete collection of 526 printed social surveys in TNA's collection. This is probably the most beneficial non-genealogical and non-manuscript digitisation project of TNA materials which relate to current social science research. The materials are considerably diverse, though they are held as a discrete collection. Sample titles from each decade are: *Food Problems: Studies of consumption and distribution of various foodstuffs* (1941); *Enquiry into the feeding habits and diets of older people* (1950); *Enquiry into the Clydebank anti-smoking campaign* (1964); *Enquiry into the effectiveness of the present family planning services in terms of demand* (1970).

We would estimate a cost of between £250,000 and £400,000 depending on the precise number of pages being digitised and the level of metadata created, to support resource discovery. Evidence from the questionnaire suggests that this project would be of use to the greatest audience of social scientists.

5.1.2.2   Other statistical materials

We have identified a number of printed and "grey" statistical material in TNA (eg AN 12 (British Transport Commission: Transport Statistics)), some of which, (eg BN 35 (Department of Health and Social Security: Statistical Reports)) are no more than photocopies of computer print outs. A "statistical" archive could be developed on the back of these materials, along with materials currently available through ONS and other government departments. A "one-stop shop" for statistical information on the UK across all time would be the ultimate aim here, but could be developed in a modular fashion. This suggestion to "portalise" contemporary UK statistics is highlighted in a recent report from the Statistics Commission and should also be the preferred tactic for historical materials.[19]

A resource which contained both historical and contemporary statistics would have considerable start-up costs, would probably have to be provided as a continuous service, and would involve the close cooperation of researchers, TNA, some government departments and possibly the British Library.

Some 'historical' datasets with contemporary analogues are already not easily discoverable. For example the HSSA data (Housing Strategy Statistical Appendix) is available online for the period 2001/2-2006/7, but the same website asks users to "please contact us directly to check whether the data [for

---

[13] (Metadata Object Description Schema) http://www.loc.gov/standards/mods/
[14] (Machine-Readable Cataloging)
[15] (Metadata for Images in XML Standard)
[16] (Metadata Encoding and Transmission Standard)
[17] (Encoded Archival Description)
[18] (General International Standard Archival Description)
[19] Statistics Commission, *Report No. 34*: *Data on Demand – Access to Official Statistics* (June 2007).

1976/77-2001/02] are available"[20]. Most government departments publish some statistics on their websites, but these are not always easily discoverable. A portal for all statistics would be a great boon for social scientists, and TNA could play a valuable role in the provision of such a service.

Recommendations

- consider call for academic researchers or digitisation bureau to implement these digitisation projects (non-response mode) and/or call for digitisation projects with high TNA content usage. [ESRC]

- investigate feasibility of (historical) statistics portal to complement the proposals from the Statistics Commission [ESRC/TNA/UKDA]

- any proposals carried out must adhere to JISC and other standards relating to management, web usability and functionality, dissemination and preservation. [ESRC/TNA]

- any proposals carried out under this heading should be free at the point of use for the academic community. [ESRC/TNA]

5.1.3 Enhancement of existing catalogue records and indices for "academic" research

Wherever possible materials previously created either directly by TNA or in conjunction with its LIAs should be made available to researchers at no direct cost.

For example, the census enumerators' books provide national coverage microdata which could be used across the social science disciplines both as historical contextual material, but also to push back long term trend data.

The costs of developing the census enumerators' books for any given year for academic purposes will vary on the level of checking performed. We would estimate it would cost at the very least £250,000 to transform a single census index for academic use and this would not include keying-in the data deemed insignificant for genealogists.

Recommendations

- investigate widening participation in use of census microdata with LIAs [TNA/ESRC]

- invest in enhancing census microdata for academic/social science community [ESRC]

- produce a list of enhanced catalogue records which might be used by social science researchers [TNA]; examine costs of further enhancement and creation and dissemination of datasets by NDAD. [TNA]

- any proposals carried out under this heading should be free at the point of use for the academic community. [ESRC/TNA]

20

http://www.communities.gov.uk/housing/housingresearch/housingstatistics/housingstatisticsby/localauthorityhousing/dataforms/

5.1.4 Cataloguing and resource discovery aids

There are clear directions in which TNA could move to improve cataloguing functionality and detail. For social and economic historians of the medieval period the Inquisitions Post Mortem and E 159 would benefit from catalogue enhancement (though it should be noted that a considerable body of printed calendars exists for the Inquisitions Post Mortem). CAB 108 would clearly benefit from this form of development.

There is definite evidence that improved (ie more detailed) cataloguing promotes use. This form of enhancement probably provides increased use at the lowest cost. Some could be carried out by volunteers; others may need a greater level of professional skills.

Furthermore, making some of the catalogue records available in a database format would make some research considerably easier, without much cost to TNA. Much of the material in DocumentsOnline could be repurposed at a reasonable cost, and where appropriate augmented by supplementary materials within the original records. For example, the WAAC database has been suggested as a candidate for this form of enhancement, and other "family history" resources including the Death Duty registers might also benefit academic researchers.

TNA might investigate collaboration with citation agencies on improving catalogue access from citation records. For example, the ISI Web of Knowledge might provide access to relevant metadata for enhancement. We also understand that a card index to citations used to be kept and still exists. TNA could consider examining this (and any other supplementary finding aids) to assess the potential for digitisation.

Furthermore, the creation and maintenance of a bibliography of references of citations of TNA's sources is likely to benefit all academic researchers. This could be carried out on the Your Archives wiki, or as a stand-alone project.

Recommendations

- publish strategy for enhancement of current catalogue [TNA]
- open strategy up for public/academic discussion and revision [TNA]
- initiate procedure of informing NDAD/UKDA when a new series is included on DocumentsOnline for their assessment for retention. [TNA]
- investigate collaboration with online citation agencies. [TNA]
- create bibliography of references of TNA's materials in scholarly journals, etc. [TNA/ESRC]

5.1.5 Further steps: policy and promotion

This study has shown that there is a lack of understanding of the role and function of TNA amongst a broad range of social science researchers. The relationships and distinctions between TNA, NDAD and UKDA need to be clarified.

The acquisitions strategy of TNA in regard to datasets notes[21] that the UKDA will be the repository of choice for datasets created in government departments which concern the individual in society while NDAD will be the repository for all other datasets. Detailed criteria are laid down in OSP 30.[22] These policies need publicity among the social science community to demonstrate where researchers are likely to find government related materials. We also suggest that an inventory of government department sponsored surveys should be commissioned in order to inform the selection or collections policies of both UKDA and NDAD.

This study has shown conclusively that the selection or collections policies surrounding born digital material needs to be improved and made explicit. There is a reasonable awareness amongst social scientists of materials within governmental departments which are not in the public domain, or only available for restricted use. More detailed research on digital datasets including those on departmental websites needs to be carried out. We would recommend that an up-to-date inventory is created with departmental advice and cooperation.

If further investigation is to be carried out on the value of its holdings to social scientists TNA will need to collect more detailed and appropriate user information to inform such study.

TNA needs to promote itself much more pro-actively towards social scientists. The guides published in the 1970s are out of date, but contain considerable ancillary information. Shorter, more "popular", themed guides could be produced. A cost-effective solution could be to use TNA's Your Archives wiki. The ESRC could be used as a dissemination vehicle for this form of promotion. We would suggest that these short guides should in the first instance be discipline specific, and could probably draw on some of the material collected from this report.

NDAD, TNA's dataset archive, also needs better promotion within the social science community. It is clear that there is a certain amount of confusion amongst respondents to our questionnaire about the roles of the UK Data Archive and NDAD. Users of TNA catalogue should be able to access datasets directly. It might be worth considering interoperability with the UKDA collections (and in particular those of the ESDS). Furthermore, better use of the JISC/ESRC-funded Intute portal for promoting these resources may improve visibility.

The online catalogue creates frustration by not taking researchers directly to the full description — the short description is experienced as a barrier. Also the screen layout requires a considerable scroll down before finding the actual description of what a series contains. Some re-configuration of the online catalogue presentation could be considered as a means of immediately improving researcher's engagement with TNA. One researcher commented on some rather curious search effects found, most notably searching for IR 127 through the catalogue. Typing this into the "Go to Reference" box on the web takes one to the page for ADM 259/365, as this is the new reference for documents previously referenced as IR 127.

---

[21] The National Archives, *Acquisition and Disposition Strategy* (March 2007). Available at: http://www.nationalarchives.gov.uk/documents/acquisition_strategy.pdf
[22] http://www.nationalarchives.gov.uk/documents/osp30.pdf (and not as cited in the *Acquisition and Disposition Strategy*).

Recommendations

- disseminate OSP 30 more widely amongst social scientists [TNA]

- create on-going inventory of government sponsored surveys and statistics [TNA/ESRC]

- enlarge selection or collections policies to include datasets from independent non-governmental organisations. [TNA]

- produce focused and directed promotion towards social scientists, via wiki and print media [TNA]

- produce focused and directed promotion towards academic researchers (via Intute) [NDAD]

- improve visitor statistics by having more meaningful user categories and to include greater scope of collection [TNA]

5.1.6   Prioritisation

This section is additional to the full report. The evidence provided from the survey makes it difficult to prioritise the recommendations made in this report. However, we can make some tentative suggestions which are based both on the evidence of the survey and the authors' knowledge. The first priority must clearly be to ensure the preservation and access to government survey data which is not presently within the collections of either TNA or the UKDA. Since the risk of loss of digital born administrative data is greater than the loss of hard copy administrative data, we suggest that this is prioritised over the enhancement or promotion of existing paper materials. However, it seems vital that TNA (and NDAD) should promote itself in a more focused manner towards social scientists.

We have suggested a series of research and digitisation projects which could be funded by the ESRC. A separate call for proposals could be made by the ESRC which relate specifically to the enhancement of TNA's resources and a panel of researchers could assess the relative value of the proposals. We suggest that these proposals would be limited to the topics/collections/series discussed above. We note however, that the digitisation of the social survey reports would probably provide the widest usage amongst the widest audience base.

TNA's priorities should also be reasonably clear. Attempting to widen use of already machine-readable materials, especially the census enumerators' books, would be likely to have the greatest impact on the social science community. Less contentious and more time sensitive would be the creation of enhanced cataloguing and guides for social scientists and for the better promotion of TNA towards the social science community, via Intute and other sources. This could also include wider circulation of OSP 30 and information surrounding the role of TNA.

Longer term, collaborative efforts should be focused on the creation of an on-going inventory of government-sponsored surveys and statistics. This need not be restricted to the ESRC and TNA, as it will have much wider application both for the present and the future.

6 Bibliography

Note all web references were available on 30 January 2008. For the purposes of clarity we have made some "tiny URLs" for additionally long URLs.

Printed publications

The British Academy, *E-resources for research in the humanities and social sciences. A British Academy Policy Review* (April 2005)

Economic and Social Research Council, *Data Policy* (April 2000).
Available at: http://tinyurl.com/2gh72e

Economic and Social Research Council, *The National Strategy for Data Resources for Research in the Social Sciences.*
Available at: http://tinyurl.com/3c6lcs

Gould CC and Handler M, *Information needs in the social sciences: an assessment* (Mountain View, CA: Research Libraries Group, 1989)

JISC, *Digitisation in the UK: The case for a UK framework*.
Available at: http://tinyurl.com/2pfl7k

JISC Information Environment Technical Standards
Available at: http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/standards/ (last updated May 2006)

Minerva, *Technical Guidelines for Digital Cultural Content Creation Programmes*. (Version 1.0: revised 8 April 2004)
Available at:
http://www.minervaeurope.org/publications/technicalguidelines_en.htm

Statistics Commission, *Report No. 34: Data on Demand – Access to Official Statistics*. (June 2007)
Available at: http://www.statscom.org.uk/uploads/files/reports/Report 34 Data on Demand.pdf

The National Archives, *Acquisition and Disposition Strategy* (March 2007).
Available at:
http://www.nationalarchives.gov.uk/documents/acquisition_strategy.pdf

The National Archives, *Operational Selection Policy - Government and People: The Interaction of the State with the Citizen as Documented in Electronic Case Files and Datasets*.
Available at: http://www.nationalarchives.gov.uk/documents/osp30.pdf

Websites consulted

AHDS History: http://ahds.ac.uk/history

British Academy: http://www.britac.ac.uk/

Communities and Local Government: http://www.communities.gov.uk/

Economic and Social Research Council (ESRC):
   http://www.esrcsocietytoday.ac.uk

Metadata Object Description Schema (MODS):
   http://www.loc.gov/standards/mods/

National Digital Archive of Datasets (NDAD):
   http://www.ndad.nationalarchives.gov.uk/

Office of National Statistics (ONS): http://www.statistics.gov.uk/

Old Bailey Online: http://www.oldbaileyonline.org/

The National Archives (TNA): http://www.nationalarchives.gov.uk/

UK Data Archive (UKDA): http://www.data-archive.ac.uk/

7 Acknowledgements

8 Glossary

AHDS History	Arts and Humanities Data Service Centre for History
AHRC	Arts and Humanities Research Council
DHSS	Department of Health and Social Security
DWP	Department of Work and Pensions
ESDS	Economic and Social Data Service
ESRC	Economic and Social Research Council
FRC	Family Records Centre
GIS	Geographical Information Systems
HBAI	Households Below Average Income
ICT	Information Communication Technology
JISC	Joint Information Systems Committee
LIA	Licensed Internet Associates
NDAD	National Digital Archive of Datasets
ONS	Office of National Statistics
PRO	Public Record Office (now TNA)
Series	A discrete group of records on one subject, from the same administrative source. It has a two part archival reference (for example MAT 32, HO 27, IR 58) and a title.
TNA	The National Archives (formerly PRO)
WAAC	Women's Army Auxiliary Corps
WPLS	Work and Pensions Longitudinal Survey