Social Science Data Services During the Last Five Years of the Millennium: Developments in the Delivery and Support of Data Services for Academic Research in Europe and North America.

by Adam Lubanski[1],
Information Systems Manager
ESRC Centre for Economic Performance

Introduction, Aims and Background

In general, data librarians are supported by researchers and computer staff in their view that demands for data services are likely to increase. Even where there have been technology related savings[2], other technology based tasks have arisen to add to the number of tasks performed by data support services - for example, the management and/or construction and maintenance of Web interfaces to data and associated documentation and literature.

The case for investing in data support services may seem clear to members of organisations such as IASSIST, CSS and Cause. And in the more recent reports produced by research/teaching support funding bodies such as the UK's ESRC and JISC there has been a marked increase in references to the data support environment. The main aim of this paper is to see if there is some empirical basis for the claim that investment in the continued development of data support services is worthwhile. The establishment of this claim will provide a sound basis from which to present the likely development scenarios of academic data services up to 2000.

The background to this study is associated with observations of a number of trends in institutional policy in the broad areas of social science support and administration in European and North American universities and associated research centres. These trends include:

- increasing funding pressures on researchers and research supervisors to speed-up submission rates - for example, from 1998, the UK's Economic and Social Research Council, (ESRC) will only fund PhDs at institutions where 60% or more students submit their PhDs within 4 years (currently, this is set at 50%)[3];
- the development of institutional measures to ensure researchers (and teachers) have necessary data resources and appropriate information systems (IS) infrastructure - about 70% of US academic institutions claim to have IS strategies[4]; and
- changes in IS infrastructure which have affected the resource demarcation between hitherto autonomous entities - for example, the integration of some or all of audio-visual, computer, data, library, network and telecommunications services[5].

To this can be added societal changes, such as the rise of so-called *meritocratic* practices such that "good policy" requires that position/status and associated resourcing have some empirical basis, as in *evidence-based planning* requirements[6].

The first set of facts gathered in this study relates to the current and projected growth rates in empirical research. Whichever way these are indicated, this growth is dramatic. The results of three methods of assessing empirical trends are summarised as follows:

- article-content analysis shows a consistent growth in the proportion of empirically based journal articles (Oswald, 1992; Figlio, 1994; Stigler, 1995; and Platt, 1996)[7];
- data access enhancements, particularly those associated with networking and interfacing, continue to speed-up the process of acquiring data and associated bibliographic references - for example, BIRON, BLS, IBSS, ICPSR and many local developments such as the data subsetting services at the NBER, SSDC and CEPIS[8]; and
- IT enhancements (storage and processing) have enabled major increases in productivity - recorded in studies of empirical research outputs (CEP/LSE) and business productivity measurement (MIT's CCS)[9].

The Fulbright Study is the basis for the second part of this investigation. It captures data support experiences from three perspectives: research, data services and IT/computer support. On the issue of efficiency of research and in-house data support, views are summarised as follows:

- the researcher-teacher view (28/30) is that data support (local and central) is an essential component of an efficient research environment - but, according to some (10/30), this may not be for ever;

• the data support service/person view (25/27) is that data and information services are experiencing a major upturn in demand - from both research and teaching activities (4 respondents also cited an increase in administration demands for data advice); and
• the majority IT/computer support service/person view (17/26) is that the acquisition of information and data support skills has become vital to their career prospects - others felt that networking and teaching support together with some integration of audio-visual support appeared a more fruitful path.

What seems obvious to the stakeholders, however, may not be fully recognised by the funders and planners. Ultimately, *in the long run*, data support funding will be determined by economic criteria.

The bad news at the present time, is that many data services are not well placed within the *order of things* to ensure that their strong economic arguments are well represented.

The good news is the data.

Background Data on Data Support Services
The selection of ninety or so interviewees, split about equally by the above types, was based on publication-citation methods (Gutman Library, February 1995). Briefly, this method adopted the following research sequence:

Data was captured mainly during 30-40 minute interviews (during some thirty visits to North American research institutions, March to May 1995); additional data was gathered from preliminary Internet searches and follow-up email to interviewees - typically, clarification of interview notes.

Supplementary *environmental evidence* was gleaned from institutional policy documents - as they related to data services/ support, and collected during the study. These included institutional responses from LSE, ESRC, national and state archives, data suppliers (e.g. the BLS) and a sample of North American universities.

The Sample Population:

Interviewees	Researcher	Data Support	IT Support
Total interviewed	30	27	26
Female[1]	6	18	5
Male	24	9	21
Job	9 SRAs (Senior Research Assistant/Officer	8 Data Librarians 3 Data Archivists	7 IT Assistants
	13 Professors	3 Data Consultants	14 IT Managers
	8 Directors (i.e. Directors of Research Centers	2 Info. Managers 4 Res. Managers 7 Data Managers	5 IT Directors
Research Center = 8 (Data Center = 7)	8 (of 400 fte) 5 (of 100 fte)	8 (of 20 fte) 7 (of 22 fte)	8 (of 14 fte + IT) 7 (of 15 fte)
University = 20 (17 Research)	17 (of ? fte)	12 (of 25 fte)	11 (of ? fte)
Published	27 (IBSS)	15 (Cause/effect, IASSIST, IBSS)	8+ (Cause/effect IASSIST, IBSS)
Cited	23 (ISI) 90? citations	12 (Cause/effect, IASSIST - approx.	Not counted
(Size/scale - average) Research Center Data Center University - research University - teaching	 80 fte 50 fte 5,000s+?f 6,000s+?f	 2.5 fte 2 fte 2 fte 1 fte	 2 fte 3 fte 30 fte (LSE) 30 fte (LSE)

1. An empirical basis for the prominence of women in computer-based data support is reported in Anderson, R.E., 1987.
? Denotes that figures were not noted at time of interview.
(All figures are for economic/social sciences. IT Support includes networking and systems staff.)

Of seventeen research universities visited, virtually all (16) had a level of local data support far higher than that found (informally) in the UK. The one university which did not claim to have any formal arrangements for data support did, however, provide a very competent IT and computing advisory service together with a catalogued tape library facility. Basic advice about dataset management tasks was given by a Program Advisory team which referred detailed queries to an "analyst programmer with experience of databases".

Of the sixteen research universities claiming to provide a "resourced data service", four were classified as having basic data support[10] ; nine were classified as having intermediate data support[11]; and three fitted the classification of full data support[12].

Level of Data Support	Basic Data Service	Intermediate Data Service	Full Data Service
Research Universities (17)	4	9	3
Research Centers (8)	1	3	4

- of all forty-eight respondents interviewed at the 17 research universities, 45 expressed unprovoked favourable opinions of data support services - although nearly all said this was an under resourced area (and were actively lobbying for better funding); only two researchers said that their level of data support was adequate for local needs; and
- about half of the data support services/libraries were managed by the library service - a trend which was generally welcomed, but opposed by 4 respondents (3c and 1d - i.e. three computer staff and one data support person) who favoured independence.

Of seven research universities with large independently funded economic/social research centers - i.e. similarly configured to LSE and CEP:

- all seven had data support facilities (often named data libraries) both centrally based - typically, managed by the library (1) or IT services divisions (2), sometimes independent (4) - and devolved in the research centers themselves (data from interviews held at Harvard/NBER, Princeton/OPR, Cornell/CISER, Syracuse/CPR, Wisconsin/SSC, Ohio/NLSY and Michigan/ICPSR);
- both central and devolved models of data support appeared to function and coordinate well (according to interviewees), and were associated with high levels of researcher (and support staff) satisfaction; and
- all seven researchers (all experienced professors) interviewed had recently visited research universities in the UK (typically, the LSE and one or two others), and expressed some dismay at the poor level of data and IT support facilities for researchers - although conventional UK academic library facilities were rated highly.

Teaching universities provided data services through the library and IT/computer services. The VP of one university described an "innovative plan" for creating information (and data) support teams attached to academic departments and managed by the Library Service. Each team would comprise a "subject librarian", a "computer/network adviser" and an "AV-graphics-teaching resources manager".

Overall, although a few data staff had major reservations, this sort of reorganisation - the CLIO Model - was expected to be a feature of the IS future in teaching institutions. Of nine such support staff interviewed, all looked forward to re-defined jobs, some with enthusiasm (5) others with apprehension (4).

Resources and costs associated with data support

Facilities (IT infrastructure)

All respondents seemed aware of the major time-savings enabled by technological advances. In particular, researchers were keen to cite benchmarks for various modelling and statistical tasks. The following examples are typical of a dozen or so proffered. These were provided by an Industrial Relations researcher (LSE and DTI) and a trade/productivity research economist at ESRC's CEP:

Year - Stata v3	Machine	Cost (new)	Time (approx.)
1985	PC-XT	$=£1,400	6,000 secs
1990	386DX20	$=£2,000	500 secs
1993	486DX33	$=£2,000	90 secs
1995	Pentium90	$=£2,000	23 secs
1996	PentiumPro200	$=£3,000	7 secs
1996	Sun Model20-71	$=£7,000	30 secs*

** Sun will run two identical jobs in less than twice this time (actually, 51 secs).*

The following times correspond to the running time of the same Gauss program solving a non-linear equation system for a grid of points.

Year - Gauss	Machine	Cost (new)	Time (approx.)
1993	486DX66	$=£2,200	68 secs
1995	Pentium90	$=£2,000	19 secs
1996	Dell Latitude (laptop P120)	$=£2,600	17 secs
1996	1996 PentiumPro	$=£3,000	7 secs
1996	Sun Model20-60	$=£7,000	20 secs*

** Unix times cannot be guaranteed if multi-user.*

Typical hardware platforms found in the research centers included several Unix boxes (HP, IBM and Sun) and about one 486/Pentium per fte researcher (excluding part-time postgraduates who typically shared pooled 386/486 facilities).

- Novell (stable), NT (expanding) and Unix (stable) network servers were typical - 486/Pentium servers (1 Gb to 5Gb) and Unix cluster (5 to 50Gb) were typical storage capacities
- A typical mid-sized research center had 52 dos/windows PCS, 10 Macs (Classic), 3 Unix, 1 VMS and 2 Novell servers - supporting about 200 postgraduate students and 40 fte research staff
- use of campus-wide email (with approx. allocation of 1Mb space per user) was typical in research centers - as opposed to own email installation
- most common/popular packages were WordPerfect/Word, Netscape/Mosaic, ELM/Pine/cc:Mail, Gauss, Excel/123/ QuattroPro, SAS/SPSS/Stata - little evidence of the use of programming languages such as FORTRAN and C.

Most interviewees reported major changes in the pattern of IT/computer support. For example, the central Program Advisory Service, still prevalent in many UK universities, had all but disappeared in the US institutions in the Fulbright Study. Typically, programmers had been relocated and redesignated as departmental or cluster IT support staff. In the department, experienced programmers were often expected to provide a wide range of skills, covering software and hardware installation as well as teaching support duties. Some common responses to this major structural change were as follows:

The majority of “ex analyst-programmers” experienced what they saw as a “deskilling process” - a minority were optimistic about the challenge/value of learning new skills. The majority of this group expressed disquiet over “cost recovery” policies, and some expected this to lead to their extinction.

Some data staff felt that lone researchers in particular had lost a valuable resource, the program advisor. Nearly all data staff said they now found it necessary to provide some programming support for basic data management tasks - typically, SAS, SPSS and Stata. Just over half of all data staff interviewed (i.e. 14 of 27) appeared familiar with one or other of these programs - most of these said they had always seen data management programming as part of their remit, although they also reported less demand for detailed program advice.

In turn, nearly all IT support staff (23 of 26) reported concern that their skills needed to be upgraded (19), or had already been upgraded (4), to cope with new information and data management tasks. Many computer staff reported a major decline in the demand for their programming skills, and some said they had stopped all programming activity "many years ago".

The following IT/computer issues were cited frequently:

- around one third of data staff stressed that "computer skills were a basic requirement for data support staff" (10d=10 citations)
- some IT support staff (below managers) were concerned that IT managers were not offering appropriate/relevant training for IT staff, particularly data skills (5c)
- the use of public PC facilities by students was often 100% with queuing at peak times, indicating that demand exceeded supply - although some universities experienced a decrease in use of public facilities, as students stayed off-site (long journeys, bad weather, good support for modem links or local networking, etc. encouraged purchase of laptops)

The 1997-2000 IT Outlook

> As predicted by Richard Rockwell (IASSIST, 1993), more powerful personal desktop PC-Workstations (running Unix and Windows NT/95), have continued to enable researchers to process large-scale datasets extracted from local and wide area networks. All respondents in the Fulbright Study expected continued performance improvements in desktop processing and data management, enabling further gains in research output. There continues to be general optimism about the contribution of IT hardware and software advances and their contribution to greater research productivity. The demand for IT support of remote laptop and home computing (distance learning) is expected to increase, stimulated by the growth in quality teaching software. In the light of so much concern expressed about reorganisation of support services, we may expect a professional review of all research support services. IT, library and data support staff will take the initiative (data piloted) and produce a more user-oriented information service.

Data Consultancy/Services (local data support)

According to researchers and data staff, the delivery of data support has become far more proactive. All data staff provided examples of "going out there" to find datasets, to advise on the best use of the data service (and other larger data facilities) and to help construct enhanced services through interlinked Web pages.

Library based data staff were most enthusiastic about the contribution of CD-ROM based datasets; some others, particularly experienced data support staff, seemed sceptical about the ultimate value of this form of data dissemination.

Researchers, closely followed by everyone else, were perplexed by the management and demarcation CD-ROM data (typically supported by the library service) and data on other media (typically found in data centers). This was generally put down to some form of *historical determinism,* and there was little evidence of plans for change in this respect!

Researchers and data staff reported that Internet type enhancements to data services had become expensive to maintain. Expectations were high, following the early lead taken (voluntarily) by data staff in constructing useable interfaces to datasets.

Web weavers reported time costs between 2 hours and two days per week for basic to comprehensive coverage of data services. Much of this work had been undertaken without additional funding. Data and research staff had become *de facto* Web Advisors.

Invariably, data support staff expressed "grave concerns" about data security and quality - particularly, in environments of decreased IT support services.

The following data issues were cited frequently:

- researchers in research groups/centers appeared less interested in programming support - although lone researchers still

needed help (4d+1r=5 citations)
• researchers seemed more concerned about quality of data accessibility, particularly with respect to speed (21r+10d+6c=31 citations)
• data staff and some IT staff were troubled by the ease of passing on large-scale undocumented (or poorly documented) datasets (22d+12c=34 citations)
• a small number of experienced researchers were concerned about a possible decline in the quality of data analysis - due to the trend to increase in accessibility (2r)
• European data could be difficult to locate, and often impossible to acquire (8d+9r= 17 citations)
• the majority of researchers prefer to download data directly to their personal machines, using their own data checking skills (18r)
• some data bureau seem reluctant to develop user services - so ICPSR (good at data checking) were playing an essential role (5r)
• the majority of researchers preferred to download entire file - rather make "front-end decisions" - even in the case of very large datasets (21r)
• researchers were keen to support the central university data repository - saying good local availability was important to research productivity (23r)
• about half of the experienced researchers interviewed said they liked to send their research assistants to the IT/data center - the other half tended to seek assistance directly from IT and data staff as appropriate

The 1997-2000 Data Outlook

General expectation of increased researcher self-sufficiency (with network infrastructure and data support) in programming/computing tasks. Alongside this, more effort in enabling access to datasets through the Internet. Later rather than sooner (evidence suggests) someone brave will pull CD-ROM data together with other data media. Expectation, *in the long run*, of large investment in distributed data services via Web/Internet - economists/accountants will work with data staff, network/communications staff and higher education planners to produce properly resourced infrastructure. In the mean time, data staff will continue to produce prototype Web Data Servers without proper funding, and to experiment with the linking of datasets, documentation and bibliographic information. Many data staff will change from being *de facto* Web Advisors to *de jure* Information Managers.

Data Archives and Data Services

• a significant number of data support staff (and others) were concerned about small-scale institutions - in particular, their inability to manage and afford big datasets (7d+5r+4c)
• even in larger institutions, data staff said that if funding problems persist, central archives such as ICPSR would become still more important (3d)
• a few IT support staff and researchers stated their preference for getting data directly from central large-scale/national archiving (Essex, ICPSR, Roper, etc.) which might assure the quality and security of important datasets (4c+2r)
• some experienced researchers appeared keen to get data direct from source, and to bypass both local and national data services and archives (6r)
• researchers and data staff based in specialist research centers expected to play a major role as data resource centers, claiming the "full set of research, data and computing expertise" at the necessary level of expertise to advise specialist research projects (6 of 7r + 7 of 7d + 5 of 7c)

The 1997-2000 Data Archive Outlook

There was some expectation of devolution of large-scale data archives - major research universities and research centers will negotiate to get data associated with their specialism direct from source. Specialist research centers will work with major archives to distribute datasets, associated materials and expert advice to high level research projects. Data archives will continue to distribute datasets to the majority of non-specialist institutions, and to provide some further "one-stop-shop" support for institutions unable to resource a local data service. Data archives will combine with national data services and social science information gateways to lead the management and coordination of specialised data services and

associated expertise based at universities and research centers. At an international level, they will plan and manage the network of specialist data servers, and they will jointly work towards making national datasets statistically comparable.

Structural/organisational trends

Whilst the majority of researchers and data staff supported developments in the integration of data services and libraries, some had major reservations - citing loss of autonomy, deskilling and reduced service as likely outcomes.

IT support devolution and cost recovery continued to alarm support staff and over-exercise IS managers. It appears that every institution has completed or is considering a major reshaping of teaching and research support services.

Most researchers (17) supported the development of "one-stop-shopping" - i.e. the integration of IT and data services - although some (4), typically experienced, researchers questioned whether extending the ranges of skills might dilute the expertise. One experienced researcher and a few (4) data staff viewed integration plans as "cosmetic", and counter-productive in that experienced data staff were likely to be lost or become disaffected in the transition.

Integration of data and library catalogues was fully supported. About half of data staff reported that datasets and library records "have been or are being fully integrated"; a third said it was "being planned"; and the others said they "expected integration of catalogues to happen soon".

Most researcher-teachers (18 of 21 interviewed) supported "the trend" to deliver research (project) based courses to undergraduates using "real datasets".

Some "research-led teachers" discussed the need for a more effective Information Systems structure to enable appropriate support for courses which required a range of data and information inputs together with more advanced information management and processing techniques - cf. courses which employ artificial intelligence methods[13]. In this scenario the popularity of the one-stop-shop was very evident.

Network and Communications remain central services - albeit with evidence of growth in the number of local Servers. The move towards full integration of voice and network services continues - reaching over 50% in research universities[14]. About half of the institutions visited charged for ethernet (or token ring) connections, and some added a rental charge - $130 for network installation (+ $5 additional annual rental in some) was typical.

There were reports of an increase in (hitherto flatish) demand for remote computing which IT support staff expected to further stretch their reduced (in real terms) resources. Teachers, students and researchers are already expecting computer advice from remote locations - i.e. from home, conference locations, etc).

The 1997-2000 Structural Outlook

The continued devolution of large-scale central IT services seems likely, although a few of the very successful central systems should be able to construct a professional/economic case. Professional groups such as IASSIST and CSS will cooperate to produce documentation of "models of successful research support systems". Joint work on teaching support will enable teachers to deliver remote/distant learning courses using real datasets extracted from central archives (for general/introductory courses) and specialist data servers (for advanced courses).

Problem areas

Data staff were seriously concerned over a number of data security issues. All experienced staff (23 of 27) said they had initiated (8 of 23) or were initiating (10 of 23) or would/should initiate (5 of 23) procedures for data checking in light of bad dataset transfers.

Large scale data transfers using FTP were commonly cited as error prone[15], and bad Windows transfers (via File Manager, particularly from CD-ROM) and tape backups were also reported.

The following problems were cited frequently:

• Proliferation of forms (8d)
• a few staff mentioned the importance of getting away from the "format statement", which was seen as problematic for researchers and time consuming for support staff (3d+1c+2r))
• variable extraction via Web interfaces was generally supported - but there were some fears that speedy extraction would mean misuse of data, particularly if "data alerts" were not built into the system (5d)
• researchers complained about the work overload at ICPSR which meant they had to plan for up to eight weeks delay from data order (via ICPSR and local Data Library) - one researcher said she advised colleagues to "order data on the expectation that you may need it!" (6r)

The 1997-2000 Problems Outlook

> Expect "contents to check contents", i.e. auto check for data consistency. Data users to feedback errors though speedy "feedback system". Overdue replacement of paper forms by electronic forms. Data staff will make a professional case for greater investment in the integration of metadata with datasets, and experienced researchers will advise Web Data Server designers on the attachment of appropriate data documentation to subsets.

Success factors and performance indicators associated with data support
All researchers interviewed showed a keen awareness of research technologies and their contribution to research productivity. About half said they were sceptical of Windows style GUIs, but all research respondents said that productivity gains from advances in operating systems (for example, multi-tasking and large memory management) had made a major contribution to their (and others) empirical research. About a third (9 of 30) volunteered detailed benchmark figures consistent with those cited earlier in this paper.

Most respondents said that the quality of output was higher due to both IT and data support (roughly equally, when prompted). Two respondents (senior/experienced researchers) said their own research benefitted very little from local data services and a great deal from IT support -system programmers advice. In one case, local data services did not feature at his institution - he tended to use highly skilled systems programmers to assist with data management tasks. Six researchers argued that their research could not be undertaken properly without assistance from local "highly skilled data support staff".

As might be expected, productivity issues cited by data staff invariably reflected the content of the recent IASSIST Newsletter/Journal coverage. The importance of documentation featured in all interviews. Content analysis of respondence to an open-ended question on "what matters most?" shows the following recent articles to be representative of the range of issues cited: for example, general issues covered in Rasmussen, 1995, on-line codebooks by Sheih 1995, quality and accessibility by Beedham 1995, production by Winstanley and standards by Greene 1995.

IT/computer support staff were much more likely than researchers and data staff to mention shortfall in training, both in terms of their own needs and the requirements of end-users.

Data staff were most concerned about getting additional resources for new developments such as the delivery of subsetting services and related documentation.

The Fulbright Study showed a strong association between high levels of local data support and good performance[16]. Nearly all researchers were keen to empathise with the data services view that local data services are a key factor in the production of highly cited research publications.

Unprompted, over half of all respondents expected data support to be a major component of developing teaching methods, particularly new and redesigned undergraduate courses.

There are also strong *a priori* grounds for associating data support with good research performance. The evidence for growth in quantity and quality of empirical research is very strong, and it is also clear that academics are rewarded for research performance measured by publications and citations.

The variables that most distinguished the academics in the sample who had been promoted from those who had not included rate of publication in refereed journals, level of citation, research grants applied for and obtained and the number of PhD students under a person's supervision. Likelihood of promotion was correlated negatively with self-reported commitment to teaching[17].

The 1997-2000 Resource Outlook

> Expect the organisation of local research support to be investigated more rigorously with a view to expansion in light of its proven contribution to research and teaching productivity.

Some economic conclusions

One thing is for certain in this study: Researchers, data services managers and IT staff all feel the funding future to be uncertain. While they may have clear visions of the what the direction of research support services ought to be, they are nervous about policy-making.

The *bad news* is that this *concern is well-founded*. Most researchers and virtually all research support staff (outside the library) are badly placed *(in the order of things)* to make a big impact on resource policy. And, as virtually every interviewee in the study has mentioned, failure to compete professionally for the appropriate level of resources will not enable their data utopia to become reality - or even *virtual reality*.

The *quite good news* is that they do have lots of real data on the productivity benefits of data support to enable the construction of a strong case for further investment. They also have the skills to disseminate this evidence. What is required is a framework for evaluating the contribution, and for this they may need to find time to review the small but growing literature in *information economics*. Recent work on the *economics of the Internet* and on the *contribution of IT to business* may provide some clues as to what to look for and what to measure in the context of research inputs.

As in other areas, the returns to investment in research support services can be measured in terms of productivity changes, performance and consumer benefits.

A number of recent research papers claim that investment in IT is associated with increased productivity, increased consumer benefits but unchanged business performance[18]. According to Brynjolfsson (1993), these results are compatible with conventional economic theory, i.e. "... firms are making the IT investments necessary to maintain competitive parity but are not able to gain competitive advantage".

Productivity gains to business and benefits to consumers due to investment in IT have been found to be strong. However, the impact of IT on Business Performance seems to be slight, sometimes negative. It appears from a stream of IT literature on business performance (1989-1993, cited by Hitt) that firms are unable to increase their profits through IT investment; indeed, while IT may be creating enormous value, it may simultaneously be intensifying competition and enabling entry, and thus lowering prices.

The *really good news* for data support services is that their contribution to empirical research is *truly, widely and deeply* recognised.

It is time to invest some of the energy and enthusiasm of research and teaching support services into the production of an empirically based case for expansion.

References

URL References:

BIRON: ESRC's Data Archive system for data searching on-line at Essex University:
http://dawww.essex.ac.uk/biron.html

CCS: MIT's Center for Coordination Science at MIT:
http://ccs.mit.edu/CCSWP190.html

CEP: ESRC's Centre for Economic Performance based at the LSE,:
http://cep.lse.ac.uk/

IBSS: ESRC/JISC's LSE based International Bibliography of the Social Sciences at BIDS, via
http://www.niss.ac.uk/ or http://www.lse.ac.uk/ - restricted access

ICPSR: Inter-university Consortium for Political and Social Research:
http://www.icpsr.umich.edu/ICPSR_homepage.html

JISC (1996): Policy for JISC Dataset Services Provision at NISS, via:
http://www.niss.ac.uk/it/JISCdatapol.html

NBER: National Bureau for Economic Research PWT data subsetting service:
http://nber.harvard.edu/pwt56.html

SSDC: University of California at SD's Social Science Data Center,:
http://ssdc.ucsd.edu/ - restricted access to some dataset services

Bibliographic Notes:

Anderson, R.E. and Coover, E.R., 1976, "Academic social research organisations and computerization", *Social science information* 15 (4/5): 741-754.

Anderson, R.E., 1987, "Females surpass males in computer problem solving", *Journal of Educational Computing Research*, vol.3(1).

Brynjolfsson, E., 1993, "The Productivity Paradox of Information Technology: Review and Assessment.", Centre for Coordination Science, MIT.

Cave, M., Hanney, S., Kogan, M. and Trevett, G., 1989, *The Use of Performance Indicators in Higher Education*, Jessica Kingsley, London.

Centre for Economic Performance, 1993, unpublished *Report to ESRC Datasets Policy Committee*, prepared by Layard, R., Lubanski, A. and Wadsworth, J. on behalf of the Centre for Economic Performance, LSE.

Figlio, D., "Trends in the publication of empirical economics", *Journal of Economic Perspectives*, 8 (Summer 1994): 179-87.

Gaspar, J. and Galeser, E., 1996, Preliminary Draft, Communications technology and the future of cities, Stanford, Harvard University and NBER.

Hitt, L. and Brynjolfsson, E., 1995, "Productivity without profit? Three measures of information technology's value.", *MIS Quarterly*. (See CCS under URL References above).

Munson, J.R., Richter, R.L. and Zastrocky, M.R., 1994, Cause ID 1994 Profile.

Oswald, A. (1991) "Progress and microeconomic data", *The Economic Journal*, 101 (January 1991), 75-80.

Over, R., 1993, "Correlates of career advancement in Australian universities.", *Higher Education*, vol.26, no.3, pp.313-329

Platt, J., 1996, "Has funding made a difference to research methods?", *Sociological Research Online*, vol.1, no.1,<http://www.socresonline/1/1/5.html>.

Ramsden, P. (1994) "Describing and explaining research productivity", *Higher Education*, V.28, N2, p 207-226.

Ruus, L.G.M., 1990, "Planning a data service facility", *Data Library Service*, University of Toronto, 30/11/90 (internal document).

1.Paper presented at the annual meetings of IASSIST, May 15, 1996, Minneapolis

2 For example, in some areas the requirement for help with acquisition of data documentation has been reduced due to investment in on-line services.

3 ESRC Annual Report, 1994-95, December 1995.

4 From Cause ID Survey 1994. In the UK, the ESRC and JISC have recently produced their respective policies for dataset services - see URL References.

5 From Cause ID Survey 1994, supplemented by evidence from my own study - see Fulbright Report in URL References.

6 Verified from NCDS studies at University of Sussex, 1994.

7 Refer to Bibliographic Notes

8 Refer to URL References

9 Refer to bibliographic notes - Hitt and Brynjolfsson, 1995.

10 As defined by Laine Ruus' 'Planning a data service facility', in Ruus (1990).

11 op. cit.

12 op. cit.

13 For example, Richard Freeman's new economics course at Harvard University.

14 Cause ID 1994 Profile.

15 A CEP researcher reported 5% failure rate (identified in data checking) in FTP transfers of some 60 data files of between 5 and 25 megabytes.

16 As indicated by publication and citation rates. See Kogan, M., et al, 1991, in Bibliographic Notes below.

17 See Over, R., 1993, in Bib.

18 Hitt and Brynjolfsson, 1995.