The Value of Libraries for research and researchers

Annexes A, B and C

www.rin.ac.uk/value-of-libraries

A Project methodology

A.1 Methodology and approach

A.1.1 Our overall approach in this project was to:

- conduct initial qualitative research with researchers and library staff at two institutions to guide the construction of the datasets for quantitative analysis;
- design the parameters for the quantitative analysis, including:
  - the sample;
  - the input measures (*eg* library characteristics) and output measures (*eg* research performance);
  - constructing the datasets for analysis;
- conduct quantitative analysis on the sample to test and analyse any correlations between input measures and output measures over the last ten years; analysis was conducted to describe characteristics at the overall HE institution level, but also undertaken on cross-sections of data (*eg* subject domains);
- investigate the factors underlying any correlations identified in the datasets, using focus groups and interviews with stakeholders from selected institutions;
- pull together the information gathering and analysis activities to provide an informative, clear and evidence-based report of the key findings.

A.2 Quantitative analysis

A.2.1 This sub-section provides a summary of the quantitative analysis that was conducted as part of this project. See Annex B for the results of the quantitative analysis.

The sample

A.2.2 A sample of 67 higher education institutions that engage in academic research was selected, including all Russell Group and 1994 Group institutions along with other institutions selected to ensure that the sample covered a good mix of institutions with:

- good performance in the 2001 and 2008 RAEs (in terms of Grade Point Averages (GPAs));
- improved performance between RAE 2001 and RAE 2008 (in terms of ranking order of GPAs).
- worsened performance between RAE 2001 and RAE 2008 (in terms of ranking order of GPAs);
- isolated high ranking units of assessment in the RAE 2008;
- a high proportion of postgraduate research students (as a proportion of all students);
- high expenditure on serials per FTE (SCONUL statistics 2007/08 7b/7m);
- high income per FTE of users (SCONUL statistics 2007/08 8x/3h);
- high levels of research grants & contracts income measured against the QR income they receive from the Funding Councils (Performance Indicators 2007/08).

A.2.3 Checks were also made to ensure that the sample included were performed to ensure that it included a good mix of institutions covering:

- specialist and non-specialist institutions;
- across the various geographic locations;

- single site and multiple site institutions;
- single library and multiple library institutions.

Input and output factors

A.2.4 Input variables were defined as the library characteristics that support research activity, obtained from the latest set of SCONUL statistics (2007/08). These cover library information such as stock, expenditure and staffing. The output variables were defined as the research characteristics of institutions such as research student and staff numbers, income from research grants and contracts, academic staff costs and results of the Research Assessment Exercise (RAE) 2008. In addition, results of Web of Science searches were used to identify numbers of publications and numbers of publication citations per institution.

A.2.5 Input factors:

- Total catalogued book stock (2a).**Error! Bookmark not defined.**
- Number of serials, electronic databases (2d+2k).
- Number of serials, electronic databases & electronic books (2d+2k+2l).
- Number of library staff hours spent delivering orientation sessions, post-orientation information, skills etc (4k).
- Number of successful requests for full text articles (journals only) (4s).
- Staff expenditure on Professional posts (6a).
- Information provision spend on serial subscriptions (7b).
- Information provision spend on electronic resources, excluding serials (7g).
- Operational expenditure of library (7r).
- Total library income (8x).

A.2.6 Output factors:

- Grade point averages of RAE 2008 results.
- Higher Education Management Statistics (HEMS) Research statistic A - Market share of research grants & contracts income 2007/08 (HESA).
- HEMS Research statistic B – Market share of teaching & research/ research only staff 2007/08 (HESA).
- HEMS Research statistic C – Market share of research council research studentships 2007/08 (HESA).
- PhDs awarded per academic staff costs 2007/08 (Performance Indicators – published by HESA on behalf of the funding councils).
- Publications 2009 per FTE of library users.
- Number of citations per publication in 2009.

Analyses

A.2.7 Correlation techniques were carried out on the input and output variables in an attempt to identify relationships between these sets of variables.

A.2.8 Two types of correlation analysis were undertaken: canonical correlation which identifies relationships between isolated input variables and the set of output variables (and similarly an output variable and the set of input variables); and Pearson's method of correlation which looks for relationships between pairs of input and output variables.

A.2.9 Although a time series of the SCONUL statistics are available, it was judged that gaps in the data affecting around a third of the sampled institutions would undermine robust auto-regressive modelling, so this technique was not used. Unless specified, throughout this report, all variables relate to the academic year 2007/08.

A.2.10 Two phases of analyses were conducted, the second following consultation with the project board. In Phase 2, the correlations found in Phase 1 were explored in more detail and adjustments were made to some of the variables used. To remove the bias caused by size of institutions, ratios were used (*eg* per FTE of users; per FTE of researchers) in some analyses.

A.2.11 Analyses were also conducted at the sub-group level:

- By Mission Group: the correlations found in the initial work were analysed at the Russell Group and 1994 Group level to assess any relationships acting at this lower level.
- By Cost Centre: disaggregation of data to cost centre level was possible for the output variables, although not possible for any of the input variables as the SCONUL statistics are not broken down at this level. For the output variables, the cost centres were grouped into 'Medicine, Science, Engineering & Technology' and 'Arts, Humanities & Social Sciences' for analysis against the institution level SCONUL statistics.

A.3 Qualitative analysis

The sample

A.3.1 A sub-set of ten institutions from the quantitative sample were identified to approach for qualitative information gathering. Institutions satisfying one or more of the following criteria were chosen:

- high RAE scores;
- specialist institutions (social sciences, arts and STEM);
- increase in RAE rating;
- decrease in RAE rating.
- representation across mission groups;
- postgraduate only institution.

A.3.2 The sample of ten institutions was approved by the Project Board. Because of the small sample size and the use of information gathered in case studies within this report, we have not named the institutions. However, the project team are extremely grateful to these institutions for being involved in this work and organising the focus groups and interviews.

A.3.3 One institution (University Alliance institution #2) had poor staff availability over the summer period and so the number of institutions from which we gathered information was nine.

Institution	Rationale
Russell Group institution #1	– High RAE scores – Specialist institution
Russell Group institution #2	– Scored highly in the research activity criteria – High RAE scores – Postgraduate focused
Russell Group institution #3	– Scored highly in the research activity criteria – Offers a postgraduate-only space

University Alliance institution #1	– Increase in RAE rating – Offers a postgraduate-only space
1994 Group institution #1	– Good score in research activity criteria
1994 Group institution #2	– Good all round research institution
Russell Group institution #4	– Specialist institution – Revamped library
Russell Group institution #5	– High RAE ratings
University Alliance institution #2	– Decrease in RAE rating

Information gathering

A.3.4 Focus groups were conducted with researchers and librarians from ten institutions. One-to-one interviews by telephone or face-to-face were also conducted other relevant stakeholders from each institution, including:

- the Library Director;
- the PVC for Research;
- a representative from the Research Support Office (or equivalent);
- the finance director (or representative).

Areas of interest

A.3.5 All focus groups and interviews were semi-structured to allow conversation to flow. The starting points for discussions with each stakeholder group are listed below.

A.3.6 Researchers:

- Do you use the library?
- What services does the library provide to support your research?
- Are they valuable?
- What do you value most?

A.3.7 Librarians:

- What services does the library offer to researchers, and in support of research?
- What is your view on what is most valuable to researchers, and to the institution?
- Are there are any areas where the library could better support researchers?

A.3.8 Library Director:

- What are the key characteristics of the library and the services it delivers to support research?
- What are the main benefits that your library offers to researchers?
- What benefits does the library offer to the wider institution in supporting research, and outside of your institution (other academic institutions, government and the public sector, the public, the economy, *etc*)?
- What evidence do you provide of the impact of your library on research when in dialogue with funders? How could this be improved?

A.3.9 Finance Director:

- How do you see the library supporting research?
- What informs your decision-making with regards to library budget and the services it offers to support research'?
- What evidence do you require about the library's performance and impact?
- How do you view the role of the library within your institution with respect to researchers?

A.3.10 PVC Research:

- How do you see the library supporting research?
- What role does the library play in research?
- What do you see as the major benefits the library offers
  - directly to researchers?
  - to the institution as a whole in supporting research?
  - outside of the institution (other academic institutions, government and the public sector, the public, the economy, *etc*)?
- How could the library change to support better research and researchers?

A.3.11 Research Support Office

- How do you see the library supporting research?
- What role did the library play in the RAE 2008 submission? Was it valuable?
- How do you value the library's role as a support service for researchers? Could they be doing more?
- What do you see as the major benefits the library offers to the institution and directly to researchers?

B Quantitative analysis results

B.1 Introduction

B.1.1 This annex presents the results of the quantitative analysis, summarising where correlations were found and also where there was lack of correlation between the selected variables. Results are presented for Phase 1 and Phase 2 of the analysis, as outlined in the methodology at Annex 0. Unless specified, all variables in this annex relate to the academic year 2007/08.

B.1.2 Interpretation of the correlation results is not presented in this annex, but rather in the relevant sections of the main document. All of the correlation results reported here are relevant to Section **Error! Reference source not found.** of the main report

B.1.3 It should be noted that the correlations do not in themselves establish causation but provided avenues to explore in the qualitative phase of the project.

B.2 Correlation techniques

B.2.1 Correlation techniques were carried out on the input and output variables in an attempt to identify relationships between the sets of variables. Although Principal Component Analysis (PCA) was cited in the project proposal as a possible technique, the project group defined an initial set of variables which was sufficiently small to render PCA unnecessary. Two types of correlation analysis were undertaken: canonical correlation which identifies relationships between isolated input variables and the set of output variables (and similarly an output variable and the set of input variables); and Pearson's method of correlation which looks for relationships between pairs of input and output variables.

B.2.2 Canonical correlation analysis highlighted some significant correlations between input and output measures. Given the small number of variables, Pearson's method was easily implemented using the full set of input and output variables and the granular nature of the results provided rich material which is presented in this annex.

B.3 Summary of results

B.3.1 Some of the moderately strong significant correlations identified in the analysis are highlighted in the table below. Those that are positive are indicated via a 'black dot'; those that are negative via a 'red dot'.

	GPA RAE 2008 results	Research Grants & Contracts Income per FTE researchers	Number of PhDs awarded per academic staff costs	Number of publications per FTE of researchers
Proportion of information provision spend on serial subscriptions (7b/7m)	●		●	●
Total library income per FTE of users (8x/3h)		●		
Total catalogued book stock per total FTE of users (2a/3h)	●		●	
Number of successful requests for full text articles (journals only) per FTE of users (4s/3h)[1]	●	●	●	
Number of serials, electronic databases (2d+2k) per FTE of researchers	●	●	●	
Number of serials, electronic databases & electronic books (2d+2k+2l) per FTE of researchers	●		●	
Spend on serial subscriptions (7b) per FTE of researchers	●		●	
Spend on electronic resources, excl. serials (7g) per FTE of researchers	●	●	●	

1 Note that whilst a positive correlation was observed for 'Number of of successful request for full text articles **per FTE of users**' no correlation was observed per **FTE of researchers**.

B.4 Phase 1 results

Raw input and output variable analysis

B.4.1 The table below presents the results of the initial analyses on the raw input and output variables.[2] A blue dot indicates a strong significant correlation (correlation coefficient greater than 0.7 and significant at the 1% level), a black dot indicates a moderate significant (correlation coefficient >0.4 and ≤.0.7 and significant at the 1% level).

	GPA RAE 2008 results	Research Grants and Contracts Income	Total FTE of student and staff researchers	Total number of library users	Number of PhD awarded	Academic Staff Costs
Total catalogued book stock (2a)	●	●	●	●	●	●
Staff expenditure on Professional posts (6a)	●	●	●	●	●	●
Information provision spend on serial subscriptions (7b)		●	●	●	●	●
Information provision spend on electronic resources (7g)		●	●	●	●	●
Total library income (8x)	●	●	●	●	●	●
Number of library staff hours (4k)				●		●
Number of successful requests for full text articles (4s)	●	●	●	●	●	●
Operational expenditure of library (7r)				●		●

B.4.2 The full results of the correlation analysis are summarised in the table below.[3]

---

2 This report and the RIN report "E-journals: their use, value and impact" (April 2009, pp39-40) found significant correlations between *article downloads* / *successful requests for full text articles* and both *PhDs awarded* and *Research grants & contracts income*. This report uses a smaller sample of institutions than the RIN report giving slightly different results.

3 The figures show the correlation coefficient; significance level and number of observations. Yellow indicates moderately strong positive correlations (correlation coefficient >0.4 and ≤.0.7); green indicates strong positive correlations (correlation coefficient >0.7 and ≤.0.9); orange indicates moderately strong negative correlations (correlation coefficient ≥-0.7 and <-0.4) and grey indicates that the significance level is above the 1% level (figures 1-4) or above the 5% level (figure 3 - results for the specialist and non-specialist institutions only).

	Output variables					
Input variables	**Grade point average RAE 2008 results**	**Research Grants and contracts income**	**Total research student & staff FTE**	**Total number of library users**	**PhDs awarded**	**Academic staff costs**
Total catalogued book stock (2a)	0.4278	0.70921	0.78828	0.54907	0.77147	0.59116
	0.0003	<.0001	<.0001	<.0001	<.0001	<.0001
	66	65	66	61	66	66
Staff expenditure on Professional posts (6a)	0.44116	0.82688	0.82436	0.82372	0.77829	0.76708
	0.0002	<.0001	<.0001	<.0001	<.0001	<.0001
	65	64	65	60	65	65
Spend on serial subscriptions (7b)	0.55721	0.77785	0.85619	0.85917	0.84672	0.92222
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	66	65	66	61	66	66
Spend on electronic resources (7g)	0.38921	0.60654	0.65013	0.63964	0.61943	0.63411
	0.0019	<.0001	<.0001	<.0001	<.0001	<.0001
	61	60	61	57	61	61
Total library income (8x)	0.47726	0.8314	0.8704	0.85668	0.83212	0.81599
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	64	63	64	59	64	64
Number of library staff hours spent delivering orientation sessions, post-orientation information, skills etc (4k)	-0.01005	0.28973	0.34283	0.42445	0.28444	0.44175
	0.9387	0.0247	0.0068	0.001	0.0263	0.0004
	61	60	61	57	61	61
Successful requests for full text articles (4s)	0.50753	0.77233	0.83155	0.63976	0.82469	0.80105
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	56	55	56	51	56	56
Operational expenditure of library (7r)	0.26723	0.57543	0.56591	0.53801	0.52394	0.39268
	0.0314	<.0001	<.0001	<.0001	<.0001	0.0012
	65	64	65	60	65	65
Number of serials, electronic databases (2d+2k)	0.14881	0.40689	0.45651	0.41283	0.43035	0.44333
	0.2406	0.0009	0.0001	0.0012	0.0004	0.0002
	64	63	64	59	64	64
Number of serials, electronic databases & electronic books (2d+2k+2l)	0.08615	0.20667	0.28708	0.36267	0.25403	0.36248
	0.4985	0.1041	0.0214	0.0048	0.0428	0.0032
	64	63	64	59	64	64

B.4.3 Unsurprisingly, strong and significant correlations (correlation coefficient greater than 0.7 and significant at the 1% level) were observed between some of the raw input and output variables, most of which can be accounted for by the size of the institution. Larger institutions are likely to have more researchers, award more PhDs, spend more on staff and have bigger libraries with more stock and higher costs. The strongest correlations are seen against the numbers of research students and staff and amount of research income received. Although less strong, there appears to be correlation between the various library characteristics and grade point average RAE 2008 results.

Input ratio analysis

B.4.4 To remove the bias caused by size of institutions, the following input ratios were generated:

- Total catalogued book stock per total FTE of users (2a/3h)
- Staff expenditure on Professional posts per FTE of Professional posts (6a/5a)
- Proportion of information provision spend on serial subscriptions (7b/7m)
- Proportion of information provision spend on electronic resources, excluding serials (7g/7m)
- Total library income per FTE of users (8x/3h)
- Number of library staff hours spent delivering orientation sessions, post-orientation information, skills etc per FTE of users (4k/3h)
- Number of successful requests for full text articles (journals only) per FTE of users (4s/3h)
- Proportion of total gross library expenditure on operational expenditure of library (7r/7x)

B.4.5 Although some moderately strong correlations were observed and seen to be significant at the 1% level, there were no strong correlations, the table overleaf highlights moderately strong significant correlations (black) and a negative moderately strong significant correlation (red).

	GPA RAE 2008 results	HEMS Research statistic A	HEMS Research statistic B	HEMS Research statistic C	PhDs awarded per academic staff costs	Publications 2009 per FTE library users	Citations per publication 2009
Total catalogued book stock per total FTE of users (2a/3h)	●			●	●		
Staff expenditure on Professional posts per FTE of Professional posts (6a/5a)							
Proportion of information provision spend on serial subscriptions (7b/7m)	●				●	●	
Proportion of information provision spend on electronic resources, excl. serials (7g/7m)	●						
Total library income per FTE of users (8x/3h)							
Number of library staff hours spent delivering orientation sessions etc per FTE of users (4k/3h)							
Number of successful requests for full text articles (journals only) per FTE of users (4s/3h)	●	●			●		
Proportion of total gross library expenditure on operational expenditure (7r/7x)							

B.4.6 Moderately strong correlations were seen for (correlation coefficient and significance levels are quoted in brackets):

B.4.7 Total catalogued book stock per FTE of users against

- *Grade point average RAE 2008 results* (0.40, 0.13%)
- *Market share of research council research studentships 2007/08* (0.50, <0.01%)
- *PhDs awarded per academic staff costs 2007/08* (0.41, 0.10%)

B.4.8 Proportion of information provision spend on serial subscriptions

- *Grade point average RAE 2008 results* (0.49, <0.01%)
- *PhDs awarded per academic staff costs 2007/08* (0.44, 0.02%)
- *Publications 2009 per FTE of library users* (0.53, <0.01%)

B.4.9 Number of successful requests for full text articles (journals only) per FTE of users

- *Grade point average RAE 2008 results* (0.42, 0.20%)
- *Market share of research grants & contracts income 2007/08* (0.43, 0.19%)
- *PhDs awarded per academic staff costs 2007/08* (0.53, <0.01%)

B.4.10 A negative significant correlation was also seen between the *proportion of information provision spend on electronic resources, excluding serials and Grade point average RAE 2008 results* (-0.42, 0.07%).

B.4.11 No significant correlations were seen between the number of library staff hours spent delivering orientation sessions, post-orientation information, skills *etc* per FTE of users or the proportion of total gross library expenditure on operational expenditure of library and any of the output variables. This was also true of the number of citations per publication 2009 which was only significant against proportion of information provision spend on serial subscriptions but weak. These input and output variables have therefore been excluded from the onward analysis.

B.4.12 Although a time series of the SCONUL statistics is available, it was judged that gaps in the data affecting around a third of the sampled institutions would undermine robust auto-regressive modelling, so this technique was not used. Correlation analysis using time lags between input ratios and output variables was carried out successively to see if any of the correlations strengthen in a particular year implying a delay between changes in an input variable and any related change in the output variable (*eg* does an increase in a spend figure have an immediate influence on research performance or is the influence seen after one year, two years etc?). The output year was fixed at 2007/08 and tested against input ratios for the last four years; this did not however uncover any appreciable differences in the correlation results.

B.4.13 The full results of the correlation are summarised in table on the following page.[3]

	Output variables						
Input ratios	**Grade point average RAE 2008 results**	**Market share of research grants & contracts income 2007/08**	**Market share of teaching & research / research only staff 2007/08**	**Market share of research council research studentships 2007/08**	**PhDs awarded per academic staff costs 2007/08**	**Publications 2009 per FTE of library users**	**Number of citations per publication 2009**
Total catalogued book stock per total FTE of users (2a/3h)	0.40	0.34	0.28	0.50	0.41	0.33	-0.10
	0.0013	0.0073	0.0290	<.0001	0.0010	0.0094	0.4517
	61	61	61	61	61	60	59
Staff expenditure on Professional posts per FTE of Professional posts (6a/5a)	0.26	0.21	0.18	0.15	0.24	0.36	0.13
	0.0320	0.0944	0.1552	0.2402	0.0544	0.0036	0.3052
	67	67	67	67	67	65	64
Proportion of information provision spend on serial subscriptions (7b/7m)	0.49	0.33	0.39	0.21	0.44	0.53	0.33
	<.0001	0.0071	0.0013	0.0863	0.0002	<.0001	0.0081
	66	66	66	66	66	64	63
Proportion of information provision spend on electronic resources, excluding serials (7g/7m)	-0.42	-0.27	-0.30	-0.22	-0.23	-0.34	-0.23
	0.0007	0.0367	0.0183	0.0939	0.0716	0.0078	0.0781
	61	61	61	61	61	59	58
Total library income per FTE of users (8x/3h)	0.32	0.28	0.17	0.23	0.31	0.34	0.05
	0.0125	0.0322	0.1948	0.0804	0.0166	0.0077	0.7010
	59	59	59	59	59	59	58
Number of library staff hours spent delivering orientation sessions, post-orientation information, skills etc per FTE of users (4k/3h)	-0.18	-0.17	-0.18	-0.24	-0.06	-0.06	-0.04
	0.1742	0.2153	0.1713	0.0780	0.6570	0.6664	0.7606
	57	57	57	57	57	56	55
Number of successful requests for full text articles (journals only) per FTE of users (4s/3h)	0.42	0.43	0.37	0.39	0.53	0.38	0.25
	0.0020	0.0019	0.0074	0.0043	<.0001	0.0057	0.0840
	51	51	51	51	51	51	50
Proportion of total gross library expenditure on operational expenditure of library (7r/7x)	0.07	0.15	0.12	0.19	0.13	0.02	-0.03
	0.5549	0.2376	0.3583	0.1302	0.3072	0.9047	0.8145
	65	65	65	65	65	64	63

Mission group analysis

B.4.14 The correlations found in the initial work were analysed at the Russell Group and 1994 Group level to assess any relationships acting at this lower level. As might be expected, given the smaller number of institutions in each group, fewer correlations were seen to be significant at the 1% level. However, moving to the 5% level (still considered as an acceptable significance

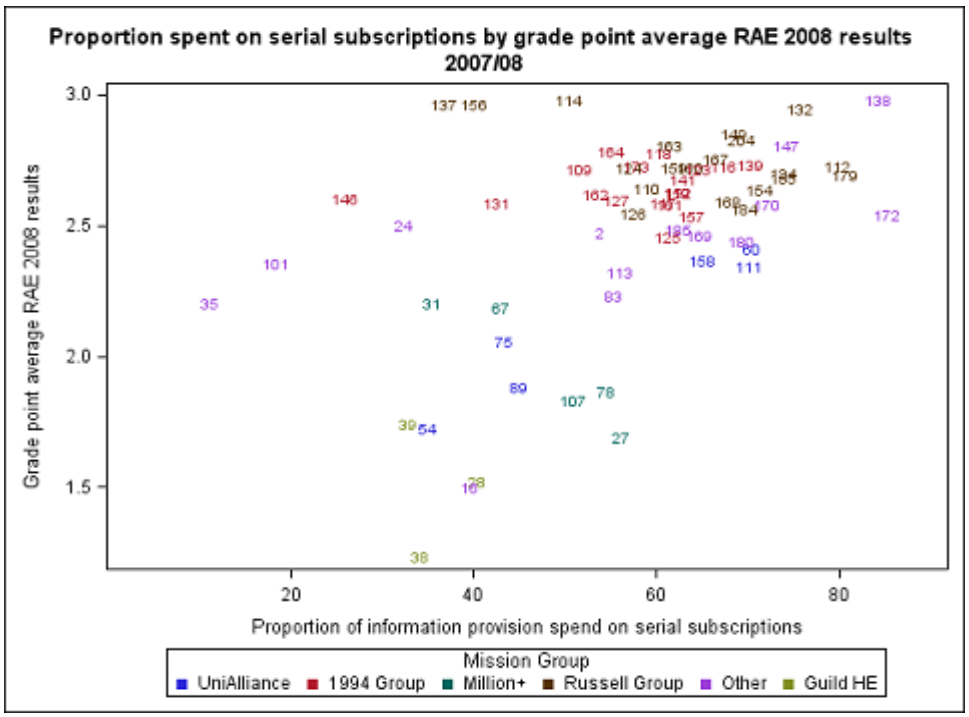
level), correlations were seen between similar pairs of input and output variables as with the whole-sample analysis, but with some notable differences in the nature of the relationship.

B.4.15 For the Russell Group, the *percentage expenditure on serial subscriptions* was negatively correlated with *GPA RAE 2008 results* (-0.47, 3.54%) and *PhDs awarded per academic staff costs* (-0.53, 1.55%). Both of these correlations were positive for the whole sample.

B.4.16 For the 1994 Group, stronger correlations were seen between the *percentage expenditure on serial subscriptions* and the *market share of research grants and contracts income* (0.60,0.86%) and *market share of teaching and research/research only staff* (0.71, 0.09%).

B.4.17 Plots of the correlations show some obvious outliers, especially from the Russell group which, when analysed at group level give a negative correlation. The outliers are some of the most well-known research intensive institutions, which perhaps attract high-quality researchers through reputation. Removing the institutions which fall in the top 10% (6 institutions) of the RAE 2008 results has the effect of strengthening some of the correlations, particularly for percentage expenditure on serial subscriptions across all output variables.

B.4.18 The numbers on the plot indicate institution identifiers which can be found in Annex C.



Cost centre group analysis

B.4.19 Disaggregation of data to cost centre level was possible for the output variables, but not for any of the input variables. For the output variables, the cost centres were grouped into 'Medicine, Science, Engineering & Technology' and 'Arts, Humanities & Social Sciences' for analysis against the institution level SCONUL statistics. All correlations in this section are quoted at the 1% significance level.

B.4.20 Within the Medical cost centres, *Catalogued book stock per total FTE of users* (0.45, 0.08%) and *percentage of expenditure on serials* (0.57, <0.01%) were significantly correlated against *GPA RAE 2008 results*, all of which were stronger and more significant than for the sample as a whole. *Total income per FTE of users* was positively correlated against *market share of research council research studentships* (0.50, 0.08%) and *GPA RAE 2008 results* (0.40, 0.40%), correlations which were not significant for the sample as a whole. A negative

correlation was seen between *expenditure on electronic resources* and *GPA RAE 2008 results* (-0.5, 0.01%).

B.4.21 Within Science, Engineering & Technology cost centres stronger positive correlations were seen between *catalogued book stock per FTE user* and *market share of research grants and contracts income* (0.49, 0.02%) and *market share of teaching & research/research only staff* (0.43, 0.11%) than the sample as a whole. *Proportion expenditure on serial subscriptions* was also correlated with *GPA RAE 2008 results* (0.48, 0.02%) and *proportion of expenditure on electronic resources* with *GPA RAE 2008 results* (-0.46, 0.08%).

B.4.22 Within Arts, Humanities and Social Sciences cost centres stronger positive correlations were seen between *catalogued book stock per FTE user* and *market share of research grants and contracts income* (0.63, <0.01%) and *market share of research council research studentships* (0.67, <0.01%) than the sample as a whole.

B.4.23 Also within Arts, Humanities and Social Sciences, *proportion expenditure on serial subscriptions* was correlated with *GPA RAE 2008 results* (0.45, 0.02%), as was *full text article requests per FTE user* with *GPA RAE 2008 results* (0.49, 0.03%).

B.5 Phase 2 results

Revised input and output ratios

B.5.1 After further consultation with the project board, the correlations found in the initial work were explored in more detail and the following adjustments were made.

- For historical reasons, the costs of journals and electronic media varies by institution and therefore using numbers of stock rather than spend on stock may be more appropriate. It was also noted that some electronic material is difficult to categorise and therefore data on electronic media should be analysed as a group, rather than disaggregated. In response, consider also number of serials, electronic databases (2d+2k) and number of serials, electronic databases & electronic books (2d+2k+2l).
- Material such as serials and electronic resources tend to be aimed more at researchers than the general library user, therefore dividing by the FTE of researchers (student and staff) may be more appropriate than total library users. In response, Total FTE of users (3h) was replaced with Total FTE of research students and staff (HESA student and staff data).
- Consider the ratio of expenditure on information provision (7m) to expenditure on professional posts (6a), rather than Staff Expenditure on Professional posts per FTE of Professional posts (6a/5a).
- The HEMS ratios look at the market share of research based aspects which will generally be larger for larger institutions. In response, replace the HEMS statistics with Research Grants and Contracts income per FTE of research students and staff (HESA Finance, Student and Staff data).
- Consider subgroup analysis for RLUK members, rather than for the mission groups and carry out some analysis on the data, excluding Oxford and Cambridge which are structurally quite different from most other institutions. Also consider the more specialist institutions (less than 24 cost centres) versus the less specialist (24 of more cost centres).

B.5.2 The following table illustrates the moderately significant correlations between the revised input and output ratios. As above, a positive correlation is marked with a black dot, a negative one with a red dot.

	GPA RAE 2008 results	Research Grants & Contracts Income per FTE researchers	Number of PhDs awarded per academic staff costs	Number of publications per FTE of researchers
Number of successful requests for full text articles (journals only) (4s) per FTE of researchers				
Spend on serial subscriptions (7b) per FTE of researchers Number of serials, electronic databases & electronic books (2d+2k+2l)	●		●	
Spend on electronic resources, excl. serials (7g) per FTE of researchers	●	●	●	
Number of serials, electronic databases (2d+2k) per FTE of researchers	●	●	●	
Number of serials, electronic databases & electronic books (2d+2k+2l) per FTE of researchers	●		●	
Total library income per FTE of users (8x/3h)		●		
Total catalogued book stock per total FTE of users (2a/3h)	●		●	
Ratio of expenditure on information provision (7m) to expenditure professional posts (6a)				

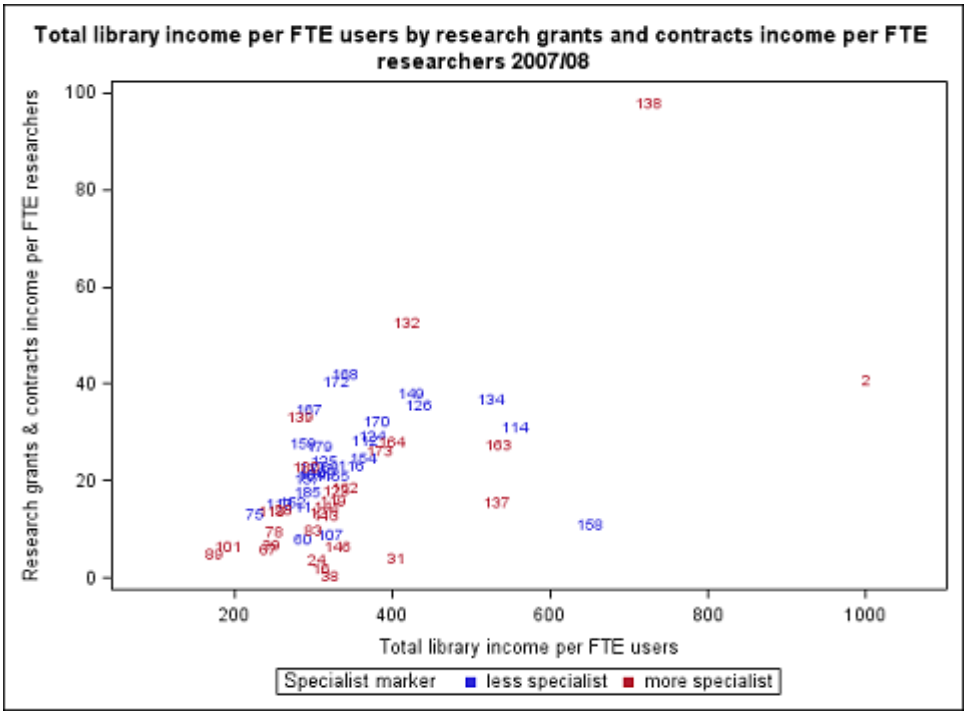
B.5.3 Using the new set of input and output ratios, a much higher proportion of the correlations were negative, most of which were significant at the 1% level. The full results are shown in the table on the following page.[3]

	Full sample			More specialist institutions (with <24 cost centres)[4]			Less specialist institutions (with ≥24 cost centres)[5]		
Input variables	Grade point average RAE 2008 results	Research grants and contracts per research students and staff	PhDs per academic staff costs (PIs)	Grade point average RAE 2008 results	Research grants and contracts per research students and staff	PhDs per academic staff costs (PIs)	Grade point average RAE 2008 results	Research grants and contracts per research students and staff	PhDs per academic staff costs (PIs)
Successful requests for full text articles per research FTE (4s/researchers)	-0.39	-0.29	-0.29	-0.31	-0.27	-0.18	-0.63	-0.27	-0.61
	0.0034	0.0302	0.0329	0.0974	0.1589	0.3388	0.0004	0.1747	0.0007
	56	55	56	29	28	29	27	27	27
Spend on serial subscriptions per research FTE (7b/researchers)	-0.51	-0.18	-0.45	-0.48	-0.07	-0.37	-0.63	-0.55	-0.69
	<.0001	0.1574	0.0002	0.0055	0.7225	0.0378	<.0001	0.0009	<.0001
	66	65	66	32	31	32	34	34	34
Spend on electronic resources per FTE of researchers (7g/researchers)	-0.66	-0.45	-0.64	-0.64	-0.42	-0.64	-0.69	-0.42	-0.51
	<.0001	0.0003	<.0001	0.0001	0.02	0.0001	<.0001	0.0203	0.0036
	61	60	61	31	30	31	30	30	30
Number of serials, electronic databases per FTE researchers ((2d+2k)/researcher s)	-0.69	-0.41	-0.54	-0.67	-0.33	-0.49	-0.68	-0.62	-0.61
	<.0001	0.0009	<.0001	<.0001	0.0672	0.0041	<.0001	0.0002	0.0002
	64	63	64	32	31	32	32	32	32
Number of serials, electronic databases & electronic books per FTE researchers ((2d+2k+2l)/resear chers)	-0.56	-0.33	-0.46	-0.61	-0.30	-0.44	-0.31	-0.36	-0.46
	<.0001	0.0094	0.0001	0.0002	0.1017	0.0124	0.0838	0.0412	0.0074
	64	62	64	32	31	32	32	32	32
Total library income per FTE users (8x/3h)	0.32	0.55	0.31	0.37	0.65	0.40	0.27	0.28	0.16
	0.0124	<.0001	0.0166	0.046	0.0002	0.0335	0.154	0.1375	0.3929
	59	58	59	29	28	29	30	30	30
Total catalogued book stock per FTE users (2a/3h)	0.40	0.16	0.41	0.41	0.07	0.35	0.51	0.34	0.58
	0.0013	0.2113	0.0010	0.0229	0.7181	0.055	0.0038	0.064	0.0007
	61	60	61	31	30	31	30	30	30
Ratio of expenditure on information provision (7m) to expenditure professional posts (6a)	0.00	0.04	0.11	0.42	0.34	0.48	-0.52	-0.37	-0.32
	0.9729	0.7389	0.4051	0.0187	0.0664	0.0057	0.0015	0.0291	0.0641
	65	64	65	31	30	31	34	34	34

[4] Significance levels are highlighted at the 5% level.

B.5.4 Positive significant correlations were seen for the full sample for:

- Total catalogued book stock per FTE users (2a/3h) against Grade point average RAE 2008 results (0.40, 0.13%) and PhDs awarded per academic staff costs (0.41, 0.10%).
- Total library income per FTE users (8x/3h) against research grants and contracts per research students and staff (0.55, <0.01%). This correlation is illustrated in the following plot.

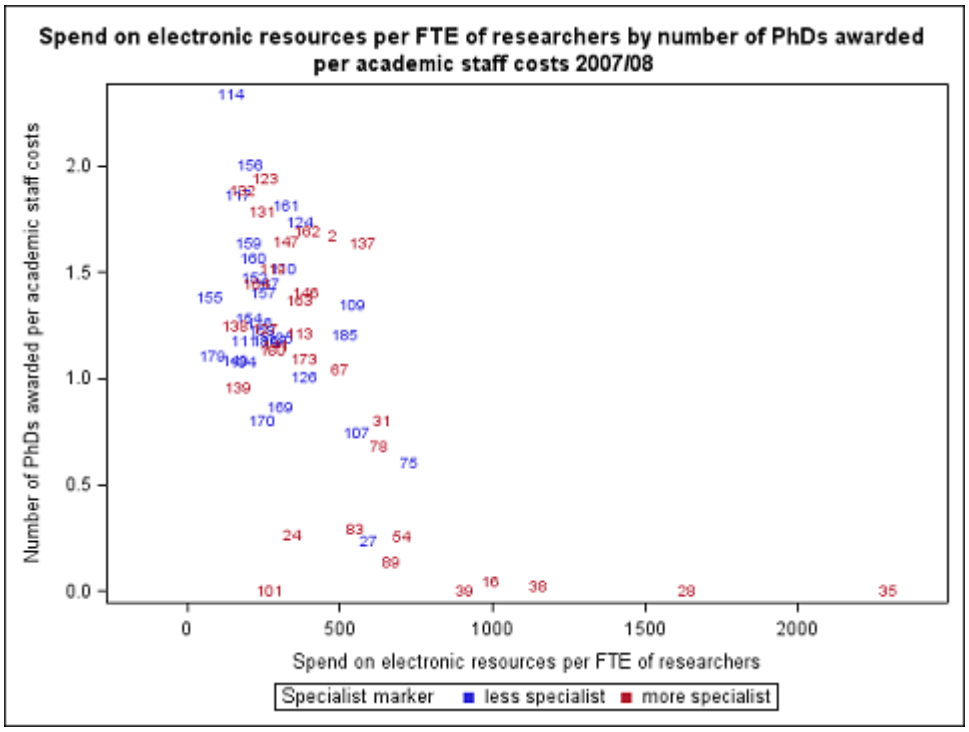


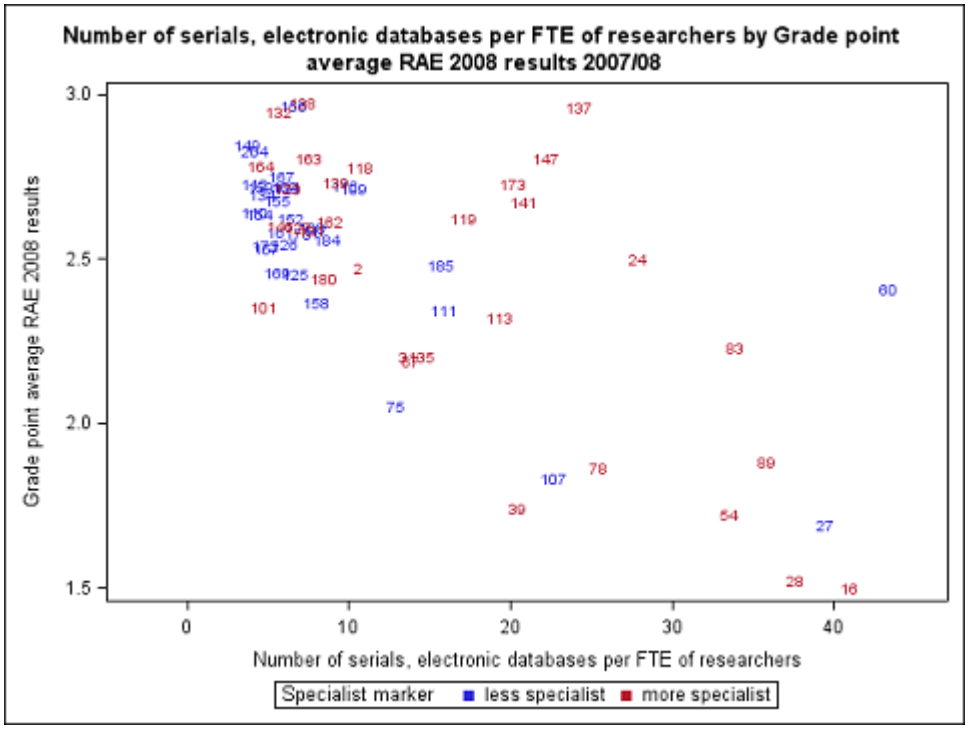
B.5.5 Negative significant correlations were seen for:

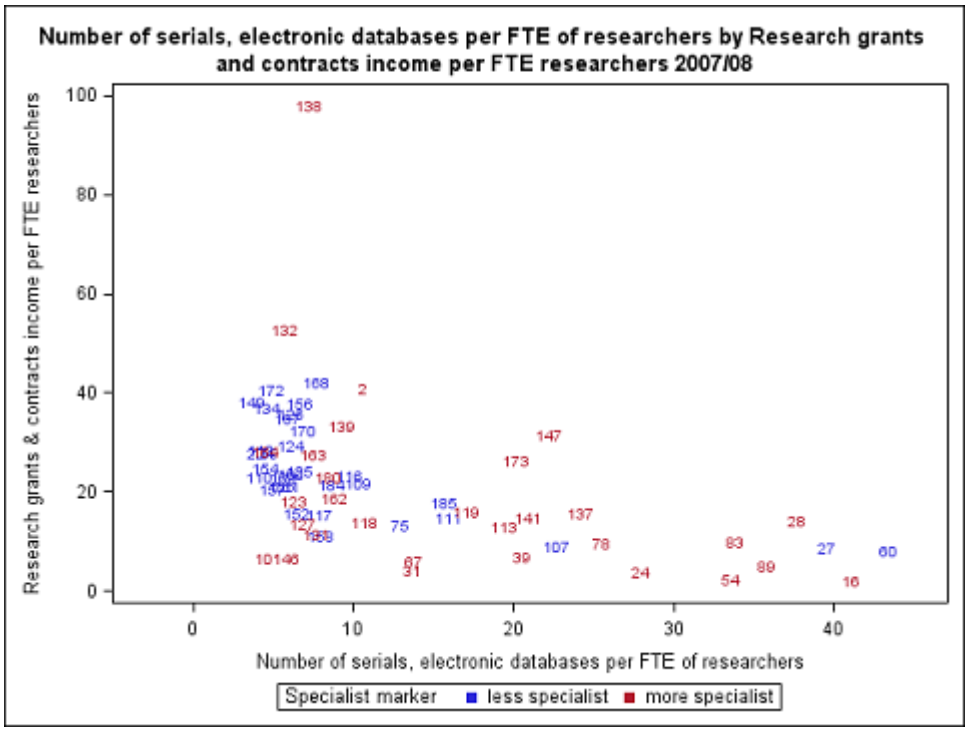
- Number of serials, electronic databases & electronic books (2d+2k+2l)[3] per FTE of researchers against *Grade point average RAE 2008 results* (-0.56, <0.01%) and *PhDs awarded per academic staff costs* (-0.46, 0.01%).
- Spend on serial subscriptions (7b) per FTE of researchers against *Grade point average RAE 2008 results* (-0.51, < 0.01%) and *PhDs awarded per academic staff costs* (-0.45, 0.02%).
- Spend on electronic resources, excl. serials (7g) per FTE of researchers against *Grade point average RAE 2008 results* (-0.66, <0.01%); Research Grants & Contracts Income per FTE researchers (-0.45, 0.03%) and *PhDs awarded per academic staff costs* (-0.64, <0.01%).
- Number of serials, electronic databases (2d+2k)[5] per FTE of researchers against *Grade point average RAE 2008 results* (-0.69, <0.01%); Research Grants & Contracts Income per FTE researchers (-0.41, 0.09%) and *PhDs awarded per academic staff costs* (-0.54, <0.01%). These correlations are illustrated in the following plots.[6]

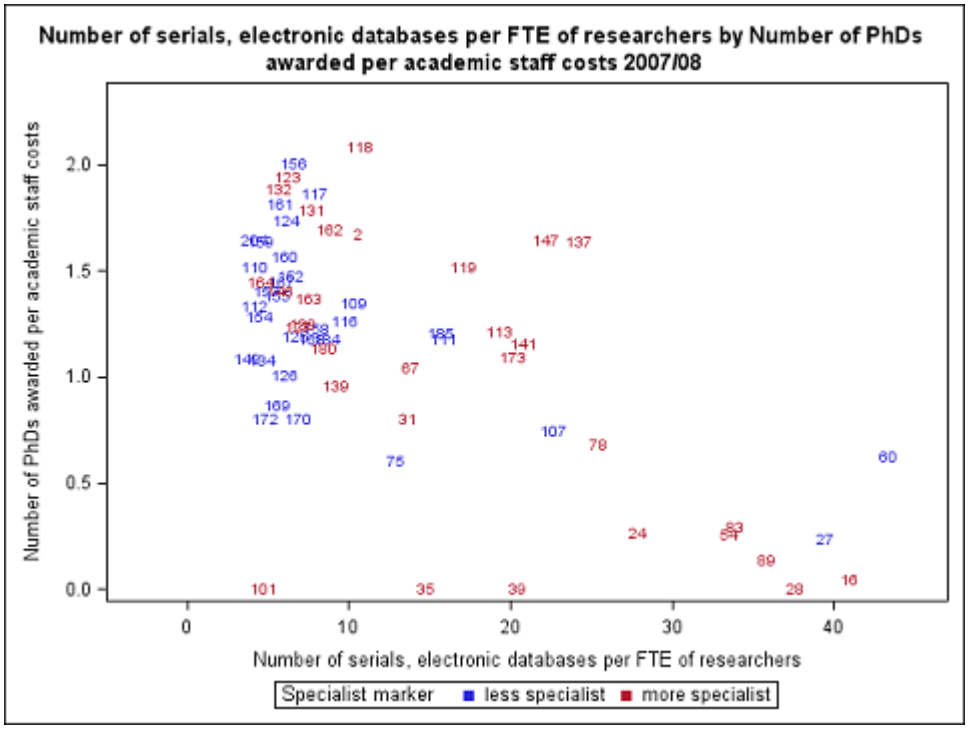
5 The correlations were slightly less strong with the inclusion of electronic books which could imply that electronic books are aimed at a wider audience than researchers alone.

6 These plots exclude one institution which is a clear outlier in the data, the correlation coefficients and significance levels have been altered accordingly.


Spend on electronic resources per FTE of researchers by number of PhDs awarded per academic staff costs 2007/08
Number of PhDs awarded per academic staff costs
2.0
1.5
1.0
0.5
0.0
0
500
1000
1500
2000
Spend on electronic resources per FTE of researchers
Specialist marker
less specialist
more specialist


Number of serials, electronic databases per FTE of researchers by Grade point average RAE 2008 results 2007/08
Grade point average RAE 2008 results
3.0
2.5
2.0
1.5
0
10
20
30
40
Number of serials, electronic databases per FTE of researchers
Specialist marker
less specialist
more specialist


Number of serials, electronic databases per FTE of researchers by Research grants and contracts income per FTE researchers 2007/08
Research grants & contracts income per FTE researchers
Number of serials, electronic databases per FTE of researchers
Specialist marker
less specialist
more specialist


Number of serials, electronic databases per FTE of researchers by Number of PhDs awarded per academic staff costs 2007/08
Number of PhDs awarded per academic staff costs
Number of serials, electronic databases per FTE of researchers
Specialist marker
less specialist
more specialist

Subgroup analysis

B.5.6 Restricting the data to RLUK members weakened all the negative correlations except for:

- *Spend on serial subscriptions (7b) per FTE of researchers* against *PhDs awarded per academic staff costs* (-0.66, 0.09%).

B.5.7 Compared to the sample as a whole, an additional, moderately strong positive correlation, significant at the 5% level, was seen between total library income per FTE users (8x/3h) and grade point average RAE 2008 results (0.57, 1.15%).

B.5.8 Additional analysis was carried out comparing the more specialist institutions with the less specialist institutions (see table on previous page). The sample was split rather crudely into two similar sized groups, one group of institutions who had staff employed in less than 24 cost centres (more specialist institutions) and the other with staff employed in 24 or more cost centres (less specialist institutions).

B.5.9 For the more specialist institutions, most of the negative correlations weakened but remained significant at the 1% level, except for number of serials, electronic databases per FTE of researchers against research grants and contracts per FTE of researchers and spend on serial subscriptions per FTE of researchers against PhDs per academic staff costs.

B.5.10 Some additional moderately strong positive correlations were seen for the ratio *of information provision spend to professional posts against Grade point average RAE 2008 results* (0.42, 1.87%) and *against PhDs per academic staff costs* (0.48, 0.57%).

B.5.11 For the less specialist institutions, additional moderately strong significant negative correlations were seen for:

- *Spend on serial subscriptions (7b) per FTE of researchers against Research grants and contracts per FTE of researchers* (-0.55, 0.09%);

- *Successful requests for full text articles per research FTE against Grade point average RAE 2008 results* (-0.63, 0.04%);

- *Successful requests for full text articles per research FTE against PhDs awarded per academic staff costs* (-0.61, 0.07%).

Departmental level analysis

B.5.12 Some final analysis was carried out at grouped department level, looking at medicine compared with science, engineering & technology and arts, humanities & social sciences.

B.5.13 The correlation seen for *total catalogued book stock per FTE users* and *grade point average RAE 2008 results* was strongest for the medical cost centres (0.45, 0.08%) as was the *total library income per FTE users* and *research grants and contracts per research students and staff* (0.57, <0.01%).

B.5.14 Amongst arts, humanities & social sciences cost centres, the strongest correlation was seen between *total catalogued book stock per FTE users* and *PhDs per academic staff costs* (0.61, <0.01%). All of the positive correlations weakened for the science, engineering & technology cost centres.

B.5.15 *Spend on electronic resources per FTE of researchers against Grade point average RAE 2008 results* was strongest for Science, engineering & technology (-0.8, <0.01%) and for Medical cost centres (-0.75, <0.01%).

B.5.16 Full results of this analysis can be seen in the table below.[3]

	Medical			Science, engineering & technology			Arts, humanities & social sciences		
Input variables	**Grade point average RAE 2008 results**	**Research grants and contracts per research students and staff**	**PhDs per academic staff costs (PIs)**	**Grade point average RAE 2008 results**	**Research grants and contracts per research students and staff**	**PhDs per academic staff costs (PIs)**	**Grade point average RAE 2008 results**	**Research grants and contracts per research students and staff**	**PhDs per academic staff costs (PIs)**
Successful requests for full text articles per research FTE (4s/researchers)	-0.23	-0.28	0.22	-0.28	-0.25	-0.24	-0.36	0.02	-0.07
	0.1101	0.0508	0.1363	0.0558	0.0789	0.1046	0.0066	0.878	0.5999
	49	50	46	47	49	47	55	54	52
Spend on serial subscriptions per research FTE (7b/researchers)	-0.32	-0.23	-0.25	-0.37	-0.39	-0.27	-0.55	0.04	-0.55
	0.0132	0.0803	0.0669	0.0056	0.003	0.0441	<.0001	0.7683	<.0001
	58	59	55	55	57	55	64	63	60
Spend on electronic resources per FTE of researchers (7g/researchers)	-0.75	-0.38	-0.17	-0.80	-0.65	-0.46	-0.65	-0.15	-0.64
	<.0001	0.0041	0.2274	<.0001	<.0001	0.0009	<.0001	0.247	<.0001
	53	54	50	50	52	50	59	58	55
Number of serials, electronic databases per FTE researchers ((2d+2k)/researchers)	-0.67	-0.41	-0.17	-0.67	-0.58	-0.36	-0.65	-0.36	-0.52
	<.0001	0.0016	0.2159	<.0001	<.0001	0.0083	<.0001	0.0046	<.0001
	56	57	53	53	55	53	62	61	58
Number of serials, electronic databases & electronic books per FTE researchers ((2d+2k+2l)/researchers)	-0.41	-0.23	-0.19	-0.46	-0.47	-0.25	-0.58	-0.39	-0.46
	0.0018	0.0805	0.18	0.0006	0.0003	0.0698	<.0001	0.0017	0.0002
	56	57	53	53	55	53	62	61	58
Total library income per FTE users (8x/3h)	0.40	0.57	0.11	0.09	0.28	0.42	0.35	0.42	0.21
	0.004	<.0001	0.4417	0.5433	0.0408	0.0024	0.0072	0.0012	0.1335
	51	52	50	50	52	51	58	57	54
Total catalogued book stock per FTE users (2a/3h)	0.45	0.10	0.53	0.29	0.29	0.14	0.39	0.19	0.61
	0.0008	0.4912	<.0001	0.0393	0.0401	0.3315	0.0023	0.16	<.0001
	53	54	51	50	52	51	59	58	55
Ratio of expenditure on information provision (7m) to expenditure professional posts (6a)	-0.11	-0.04	0.09	-0.11	-0.03	0.19	-0.02	-0.03	0.03
	0.4329	0.7858	0.5279	0.4319	0.8409	0.1638	0.8644	0.8232	0.807
	57	58	54	55	57	55	64	63	60

C Sample institutions for quantitative analysis

C.1.1 The following table lists the 67 institutions that were included in the final sample for the quantitative analysis.

Mission Group	Reference number	Institution
1994 Group	0109	The University of Bath
	0116	University of Durham
	0117	The University of East Anglia
	0118	The University of Essex
	0119	The University of Exeter
	0123	The University of Lancaster
	0125	The University of Leicester
	0127	Birkbeck College
	0131	Goldsmiths College
	0139	Queen Mary and Westfield College
	0141	Royal Holloway and Bedford New College
	0146	The School of Oriental and African Studies
	0152	Loughborough University
	0157	The University of Reading
	0161	The University of Surrey
	0162	The University of Sussex
	0164	The University of York
	0173	The University of St Andrews
Russell Group	0110	The University of Birmingham
	0112	The University of Bristol
	0114	The University of Cambridge
	0124	The University of Leeds
	0126	The University of Liverpool
	0132	Imperial College of Science, Technology and Medicine
	0134	King's College London
	0137	London School of Economics and Political Science
	0149	University College London
	0154	The University of Newcastle-upon-Tyne
	0155	The University of Nottingham
	0156	The University of Oxford

Mission Group	Reference number	Institution
	0159	The University of Sheffield
	0160	The University of Southampton
	0163	The University of Warwick
	0167	The University of Edinburgh
	0168	The University of Glasgow
	0179	Cardiff University
	0184	The Queen's University of Belfast
	0204	The University of Manchester
GuildHE	0028	Newman University College
	0038	University of Cumbria
	0039	St Mary's University College, Twickenham
Million+	0027	The University of Northampton
	0031	Roehampton University
	0053	The University of Central Lancashire
	0067	Middlesex University
	0078	The University of Sunderland
	0107	Edinburgh Napier University
University Alliance	0054	University of Gloucestershire
	0060	University of Hertfordshire
	0075	Sheffield Hallam University
	0089	University of Wales Institute, Cardiff
	0111	The University of Bradford
	0158	The University of Salford
Other	0002	Cranfield University
	0016	Edge Hill University
	0024	University of the Arts, London
	0035	Royal Northern College of Music
	0083	The University of Westminster
	0101	The Royal Scottish Academy of Music and Drama
	0113	Brunel University
	0138	London School of Hygiene and Tropical Medicine
	0147	The School of Pharmacy
	0169	The University of Strathclyde
	0170	The University of Aberdeen
	0172	The University of Dundee
	0180	Swansea University

Mission Group	Reference number	Institution
	0185	University of Ulster