

# Data-Driven System Level Mapping for Strategic Decision Support: Economic Health of UK Universities and Other Higher Education Institutes

Fionn Murtagh (1, 2)

(1) Science Foundation Ireland, Wilton Place, Dublin 2, Ireland

(2) Department of Computer Science, Royal Holloway, University of London  
Egham TW20 0EX, England

Email: fmurtagh@acm.org

May 20, 2011

## Abstract

In exploring the HEI (higher education institute) financial health and safety data published in the Times Higher of 18 March 2010, we find some very interesting underlying patterns in the data. These patterns point to an interesting contrast involving diametrically opposite orientations of HEIs in the UK. This polarity goes considerably beyond the usual one of research-led elite versus more teaching-oriented new universities. Instead we point to the role of medical/bioscience research income in the former, and economic sectoral niche player roles in the latter. In addition to the 2010 data we also analyze data from the Times Higher on 7 April 2011, finding a similar set of outcomes.

## 1 Introduction

In Murtagh (2010) resulting from the Sixth Annual Boole Lecture (organized by the Boole Centre for Research in Informatics, <http://www.bcri.ucc.ie>) in 2008, we show how information focusing is carried out in data analysis, i.e. determining where the data is put under the analytic microscope. One issue addressed is coverage and completeness of research funding in technological sector domains. Another issue addressed is evolution of funding decisions over time. We show how the narrative of science and engineering policy – the story that policy decisions have to tell – can be mapped out from the raw data. The orientation of such narrative is crucial.

In this present work, we use the same data analysis approach, Correspondence Analysis. Based on the data on UK HEI (higher education institute)

financials provided by Newman (2010), we look for underlying patterns of particular interest. The data is due to accountancy firm Grant Thornton and is based on institutions' financial statements for 2008-2009. In section 5 we look at data from 2009-2010.

In Murtagh (2010) we provide background on the analysis approach which takes cross-tabulations as inputs – in this case of HEIs crossed with financials on a set of incomes or expenditures. *Profiles* of the (positively-valued) data, on either rows (i.e., HEIs) or columns (i.e. financial incomes or expenditures) are mapped into the same visualizable (hence Euclidean distance-based) space. *Profiles* are values in the row or column that are divided by the row/column total. Hence HEIs, or financial attributes, are normalized in this way – by dividing by their respective row/column totals.

A range of analysis options are opened up by the Correspondence Analysis: simultaneous display of HEIs and incomes/expenditures; optimal planar display; accounting for most of the information content (in a precise mathematical sense) of the data; among others.

## 2 Attributes and Interpretation of the Planar Visualization

Attributes used in the main analysis were as follows. These attributes constituted the primary data used on the 155 institutions.

- Attribute 3, Funding council grants (all grants of: HEFCE, Higher Education Funding Council for England; or HEFCW, Higher Education Funding Council for Wales; or SFC, Scottish Funding Council).
- Attribute 4, Research grants and contracts (from all sources other than HEFCE/HEFCW/SFC).
- Attribute 5, Tuition fees and education contracts (excluding overseas) (UK and European including short courses or other ancillary teaching).
- Attribute 6, Overseas fees.
- Attribute 7, Other income (from catering, residential, possibly from companies spun out).
- Attribute 8, Endowment and investment income.
- Attribute 11, Total staff costs (including social security and pension contributions).
- Attribute 13, Total borrowing.

We omitted net surplus (attribute 1) because of the remarks in the Times Higher, noting how Cambridge had the largest deficit but it was a very small

percentage of its total income; Bucks New University recorded a large deficit but then sold a campus to reverse this deficit; and Thames Valley University had a surplus but this disappeared when HEFCE was reimbursed for this university's over-reporting of its fundable student numbers. Our interest lies in financial health. Arising from this, we were interested not in the financial position as such but rather in determining underlying indications of where the sector is headed as it seeks to address the current economic climate. So we used the more basic financial data.

Attributes projected into the analysis subsequently were as follows. These were attributes derived from the more basic data.

- Attribute 2, Net surplus as % of income.
- Attribute 10, Funding council grants as % of income.
- Attribute 12, Total staff costs as % of income.
- Attribute 14, Total borrowings as % of income.

Fig. 1 summarizes the data. Shown in the figure is a principal plane projection, accounting for  $42 + 30 = 72$  % of the information content – most, therefore.

Factor 1 is dominated in influence by attribute 4, “Research grants and contracts”. Such domination is determined not just by its relatively extreme (positive or negative) projection on this first (newly determined) coordinate axis, but also by its *contribution* to, and its *correlation* with, the first axis. (Contribution, correlation, inertia expressing information, factor, and so on are all mathematically defined terms in the Correspondence Analysis data analysis and display context.)

For Factor 2, the dominant attribute is 13, “Total borrowing”, and attribute 11, “Staff costs”, is not far behind in terms of influence.

On Factor 2, it can be seen that attribute 14, “Total borrowing as % of income”, is in the same general region as 13, “Total borrowing”. We can note too that, counterposed to 13 and 14, there are attributes 2 (surplus-related), 10 (funding council grants-related), 12 (staff costs-related), and also 5 (tuition fees) and 3 (funding council grants) – all possible countervailing means relative to borrowing.

Interestingly, attributes 2 and 10 – “Net surplus as % of income”; and “Funding council grants as % of total income” – are closely located, indicating that the information conveyed is very similar. Attribute 6, “Overseas fees”, is close to the origin of the display, indicating where it is a not very discriminating attribute here.

Fig. 2 is the same as Fig. 1, just showing the areas where we will now mostly focus our attention.

**Times Higher 2008-2009 financial data. 155 institutes at dots.**

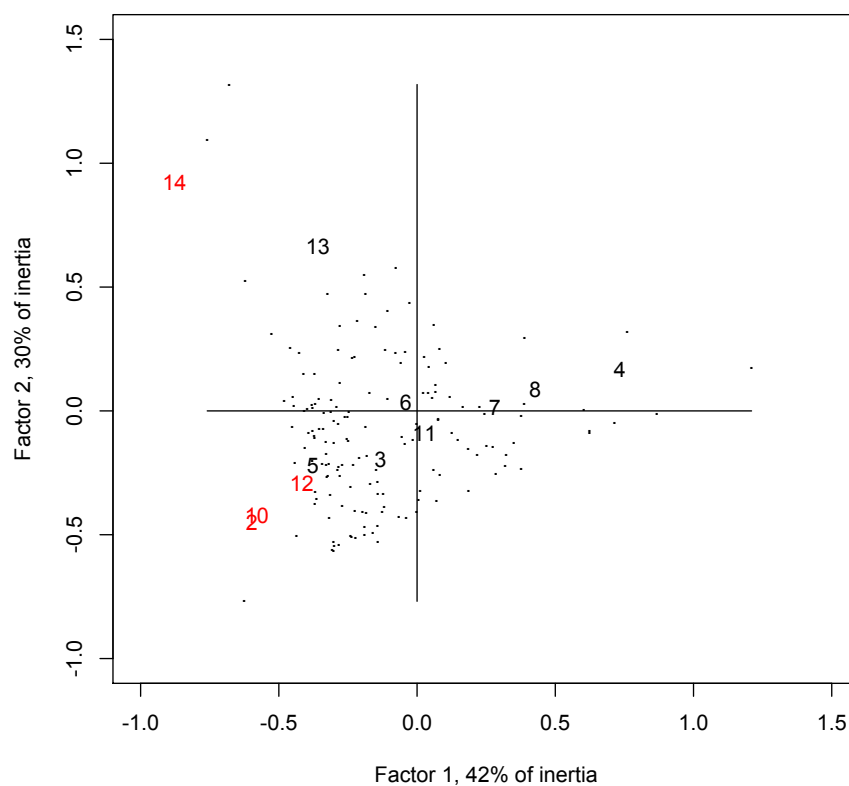


Figure 1: A first visualization of the data: the main analysis is based on the data table for 155 UK higher education institutions using the Times Higher attributes (see text for these) 3, 4, 5, 6, 7, 8, 11, and 13. Their locations can be seen in the (numeric values in the) planar projection. In addition, based on the analysis of the main data, the locations were found for the more “illustrative” attributes, 2, 10, 12, and 14. These latter are shown in red. The higher education institutes, in order not to crowd this initial display, are each shown as a dot.

**Times Higher 2008-2009 financial data. 155 institutes at dots.**

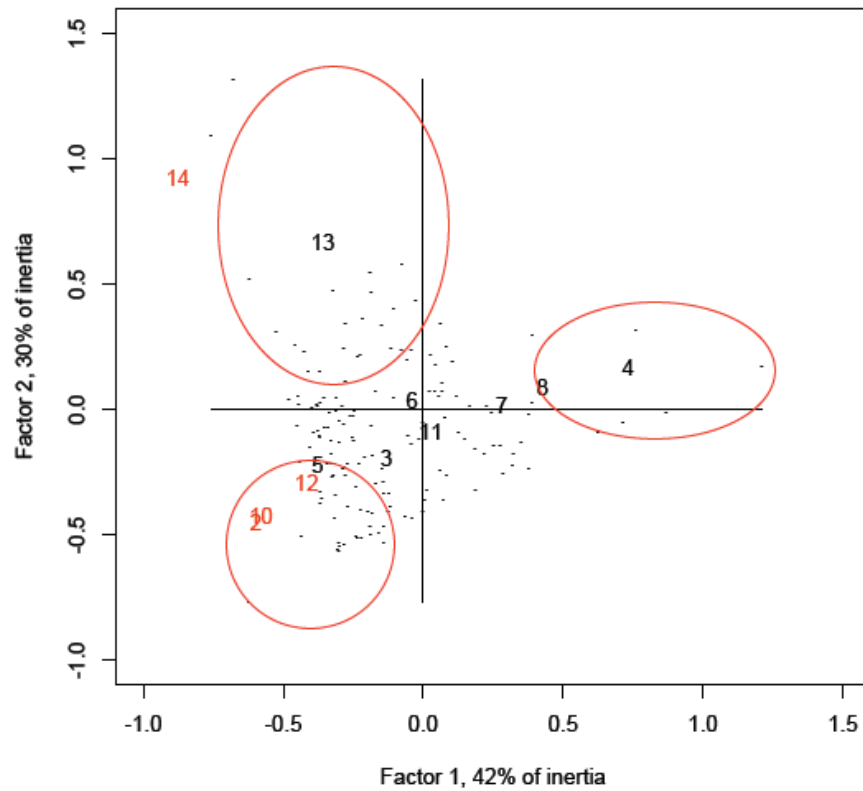


Figure 2: We will focus attention on the rightmost HEIs here; on the upper left ones; and finally on the lower rightmost. Meanwhile both HEIs and attributes that are close to the origin (coordinate 0, 0) are average, relating either to average HEI profile, or to average attribute profile.

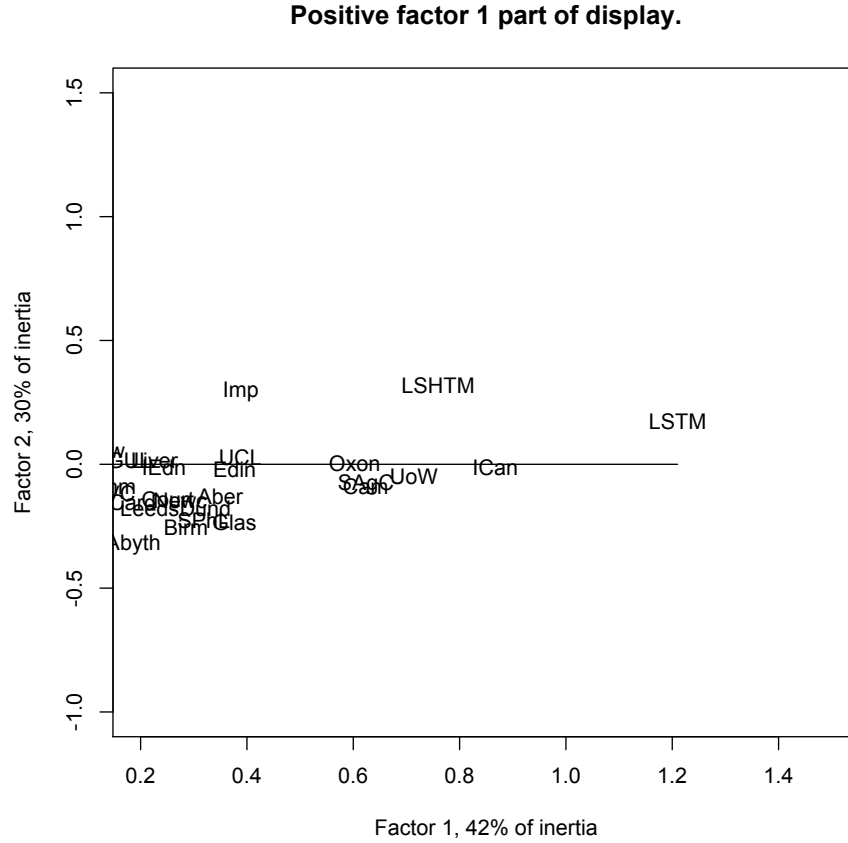


Figure 3: The higher education institutes: first, the rightmost part of Fig. 1, relating to the positive end of Factor 1.

### 3 Factor 1: Role of Medical Disciplines in HEIs that are Strong in Research Funding

The most positively linked institutes relative to Factor 1 are to be seen in Fig. 3. These are:

- Cambridge (“Cam”, overlapping “SAgC”)
- Institute of Cancer Research (“ICan”)
- Liverpool Sch Tropical Medicine (“LTSM”)
- Tropical Medicine – London Sch Hygiene & (“LSHTM”)

- Oxford (“Oxon”)
- Scottish Agricultural College (“SAgC”, overlaid on “Cam”)
- University of Wales (an administration only institute) (“UoW”)

Somewhat less pronounced in terms of this factor are: Imperial, UCL (University College London) and University of Edinburgh.

Apart from the traditionally strong Oxbridge research presence, what is also noteworthy is the medical and biosciences presence, albeit specialist, in this cluster.

Adams and Gurney (2010) point to how citation ratings from Thomson Reuters attribute the lion’s share of UK research outcomes to five HEIs: Oxford, Cambridge, Imperial, UCL, and LSE (London School of Economics). In our concluding section below we will return to this view of performance and achievement evaluation.

## 4 Factor 2: Borrowing

As noted Factor 2 is firstly and foremostly related to borrowing. Fig. 4 shows the positive end of this factor. We see a number of institutions that are flagged in the Times Higher article in terms of high gearing, i.e. “Total borrowings as & of income”: Queen Margaret University, 220.5% of income; Ravensbourne College, 171%; University of Worcester, 82.5%; University of Surrey, 63%; and Brunel University and the University of St Andrews, both 62%.

We will next look at the non-g geared end of Factor 2. We look at what is most opposite the research, Oxbridge, medical and biosciences, end of Factor 1. What we find in Fig. 5 is that the following institutes are to be found there:

- Conservatoire Dance & Drama (“CDD”)
- Bishop Grosseteste (“BiGr”)
- Bath Spa (“BSpa”)
- Swansea Metropolitan (“SwanM”)
- Newman College (“Newm”)
- Liverpool Inst Performing Arts (“LPerf”)
- UHI Millennium Institute (“UHIMI”)
- Leeds Trinity (“LTrin”)
- Manchester Metropolitan (“ManM”)
- Open University (“OU”)
- London Business School (“LBS”)

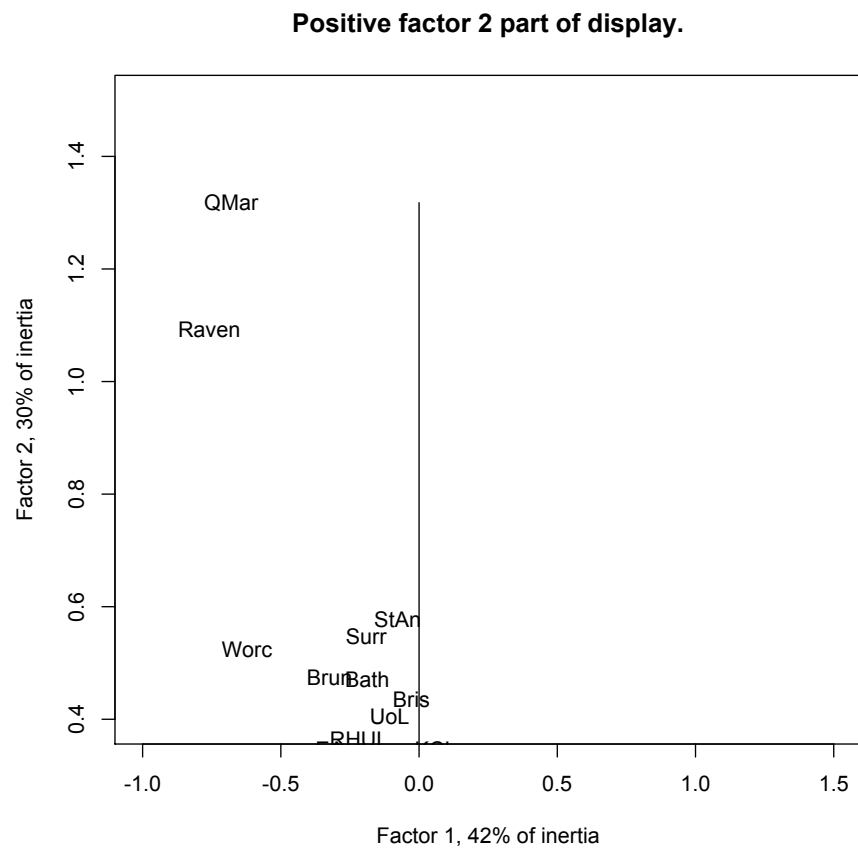


Figure 4: The higher education institutes: the positive end of Factor 2, cf. the complete view in Fig. 1.



- West of Scotland (“WoS”)
- Glasgow Caledonian (“GCal”)

We note specialist and/or business – or business sector – orientations that are well represented among these institutions. Note again that these institutions are not at all as highly geared as those institutions that are more towards the positive end of Factor 2.

## 5 From 2008-2009 to 2009-2010

In Baker (2011), data is presented for 2011. Some (small number of) universities differ in the list of 154 used in 2009-2010, compared to the list of 155 used in 2008-2009. It is seen though that the overall characteristics of the data are very similar: cf. Figs. 1 and 6.

In regard to the rightmost projections on Factor 1 of Fig. 6, we again find the following (in order of prominence, given by projections). (Fig. 4 had zoomed in on this part of display for the 2008-2009 data.)

- Liverpool Sch Tropical Medicine
- Institute of Cancer Research,
- Scottish Agricultural College
- London Sch Hygiene & Tropical Medicine
- Oxford
- Cambridge
- University of Wales
- University College London

With reference again to the 2009-2010 data, we find the most prominent on Factor 2 (cf. for 2008-2009, Fig. 5) to be:

- Queen Margaret
- Ravensbourne

These are then followed by: Surrey, St Andrews, Worcester, Reading, Bath, University of London, Bristol.

In regard, for 2009-2010, to the lower left quadrant of Fig. 6, and with reference to the year earlier of 2008-2009 shown in Fig. 7, on this occasion – 2009-2010 – we do not have data for the Conservatoire for Dance and Drama (labelled “CDD” in Fig. 5).

We do find others though, in order of prominence by projection on Factor 1:

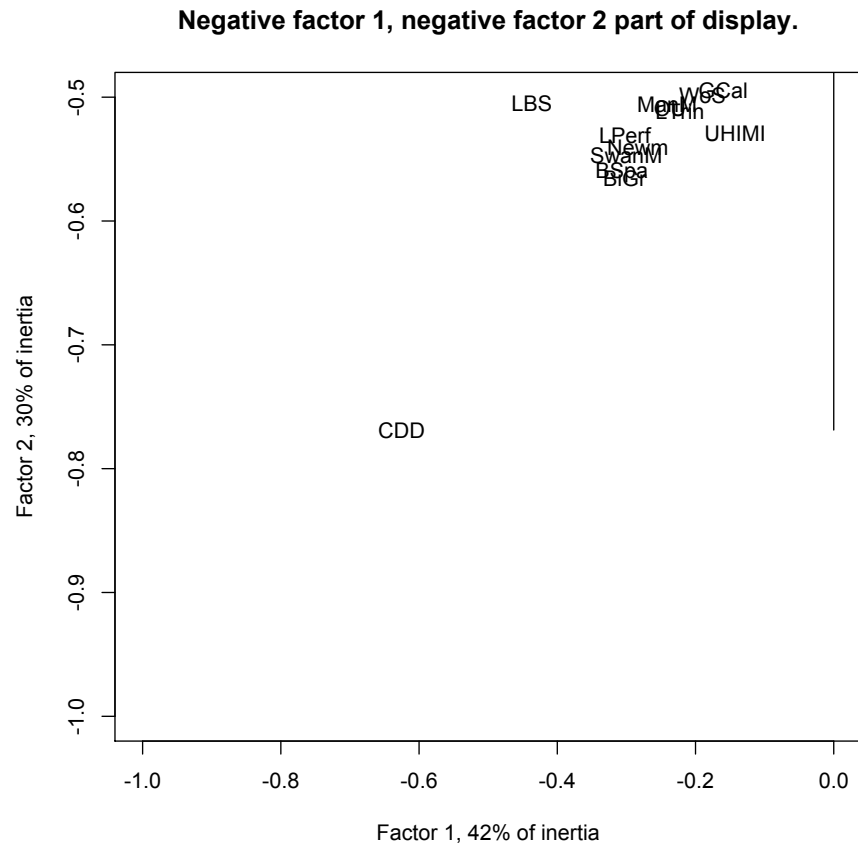


Figure 5: The higher education institutes: the negative ends of Factor 2 and of Factor 1, cf. the complete view in Fig. 1. These are less research funding-based, and also non-borrowings geared, institutions.

**Times Higher 2009-2010 financial data. 154 institutes at dots.**

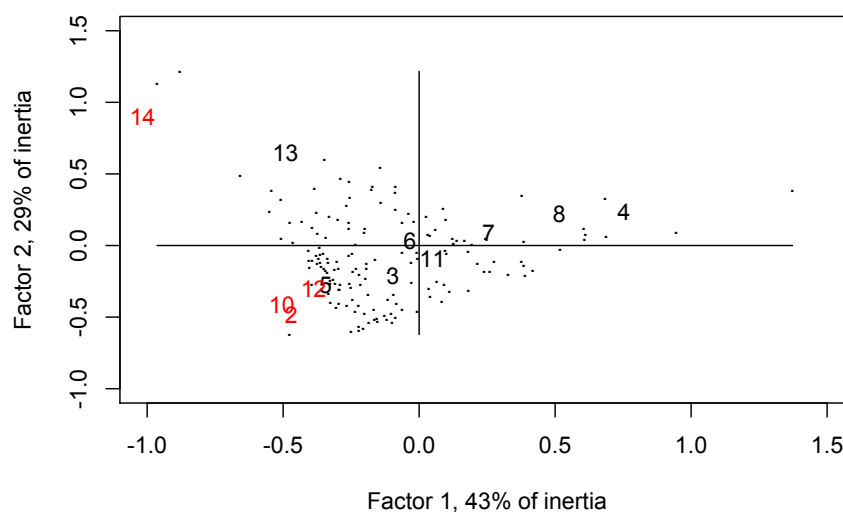


Figure 6: Fig. 1 was related to 2008-2009 and, here, we have 2009-2010 data. This is a visualization of the data: the main analysis is based on the data table for 154 UK higher education institutions using the Times Higher attributes (see text, section 2, for these) 3, 4, 5, 6, 7, 8, 11, and 13. Their locations can be seen in the (numeric labels in the) planar projection. In addition, based on the analysis of the main data, the locations were found for the more “illustrative” attributes, 2, 10, 12, and 14 (see also section 2). These latter are shown in red. The higher education institutes, in order not to crowd this initial display, are each shown as a dot.

- London Business School
- Bath Spa
- Newman University College
- Swansea Metropolitan
- Bishop Grosseteste
- Manchester Metropolitan
- Liverpool Inst Performing Arts

Overall we see that there is little relative difference between the two sets of data, for 2008-2009 and 2009-2010.

## 6 Implications and Conclusions

We conclude that:

- Factor 1 is primarily based on research funding, not from HEFCE and sister organizations outside England but rather from research councils, and also is indicative of the particular importance of medical and bio-science research funding which results in institutes that we have noted being strongly positioned on this underlying dimension in the data. As a part of this finding, we note this central role played by medical and closely related disciplines.
- Factor 2 is primarily borrowing, with the property of gearing (i.e., borrowing relative to income) being particularly useful to explain this. Newer institutes, with limited but focused course offerings, and with specialist business or industrial sector orientations, together with the London Business School, the Open University, and the UHI Millenium Institute – latter now the University of the Highlands and Islands – are all the most extreme in the low (or zero) borrowing sense. In section 4 we have noted the highly geared institutions.
- Our main finding therefore is the polarity between, on the one hand, traditional research, by now well swayed towards medical and closely related research; and, on the other hand, newer and more specialist, or business-oriented institutions.

To draw out implications of this polarity we can show – see Fig. 7 – the placements of any of the HEIs. Properties vis-à-vis the Factor 1 and Factor 2 oppositions can be appreciated. For example, LSE is seen to be in an average position.

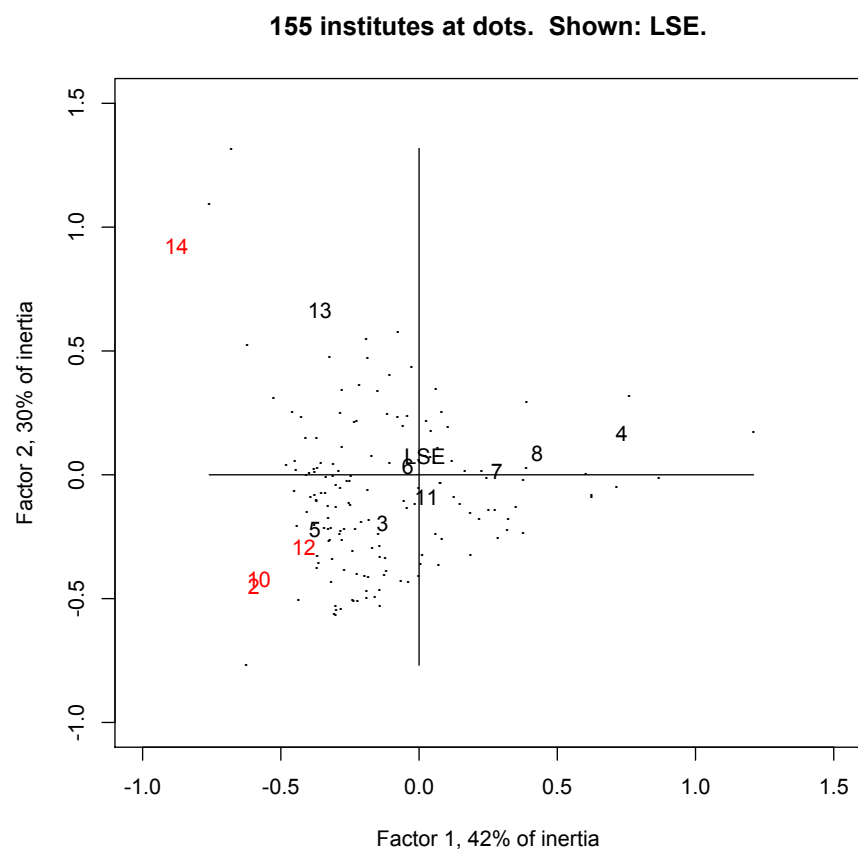


Figure 7: Again Fig. 1, with LSE (London School of Economics) displayed.

## 7 Discussion and Conclusions

In studying world leadership in research, Adams and Gurney (2010) find five institutes (Oxford, Cambridge, UCL, Imperial, and LSE) to be significantly separate from all others, including others in the Russel Group of universities. Adams and Gurney label the five universities the “Golden Triangle”. The criterion used by Adams and Gurney is citation impact, based on Thomson Reuters databases. Of course this is not necessarily a good basis for the measurement of impact in, for example, computer science (see e.g. Moed and Visser, 2007) due to more limited coverage of the literature in this area and also different citation practices and culture (involving conferences, for example).

When viewing the university system in its entirety, other forms of impact are clearly important also. These include human capital, sectoral and niche applications, and also engineering (as opposed to science) demonstrators and testbeds, and their depoloyment.

It is seen from our data analysis that the UK system is gravitating – or has largely already done so – towards two attractors: high impact research, as measured through citations, which as a performance measurement tool is very fit for purpose across a wide range of disciplines including the life sciences, biosciences, materials science and others; and what we have characterized as niche industrial/business sector application-oriented research, that also incorporates business and management, and human capital too.

It is our view that taking account of both “attractors” is advantageous. Among the benefits is overall system robustness, achieved both through coverage and breadth, on the one hand, and on the other hand complementarity of objectives and of achievements.

## References

1. Jonathan Adams and Karen Gurney, “Funding selectivity, concentration and excellence – how good is the UK’s research?”, HEPI report, Higher Education Policy Institute, 20 pp., 25 March 2010.
2. Simon Baker, “Underlying disorders: The financial health of UK higher education institutions”, Times Higher, 7 April 2011. Data tables in “Wealth and Health: Financial Data for UK Higher Education Institutions, 2009-2010”.
3. Henk F. Moed and Martijn S. Visser, “Developing bibliometric indicators of research performance in computer science: an exploratory study”, Research report to the Council for Physical Sciences of the Netherlands Organisation for Scientific Research (NWO), 117 pp., Feb. 2007.
4. Fionn Murtagh, “The Correspondence Analysis platform for uncovering deep structure in data and information”, Sixth Annual Boole Lecture in Informatics, Computer Journal, 53 (3), 304–315, 2010.

5. Melanie Newman, “Ready for the storm?”, Times Higher, 18 March 2010, pp. 36-44.