technopolis |group|

Review of the Research Excellence Framework

Evidence Report

technopolis|group|

Review of the Research Excellence Framework

Evidence Report

technopolis |group| October 2018

Erik Arnold
Paul Simmonds
Kristine Farla
Peter Kolarz
Bea Mahieu
Kalle Nielsen

Summary

This is the report of a review of evidence commissioned in 2016 by the Department of Business, Energy and Industrial Strategy (BEIS). It provided information to the review of the Research Excellence Framework (REF) undertaken under the leadership of Lord Stern.

What is the REF?

The UK like most countries has a 'dual support system' for university research. Under such a system, money for university research comes partly as 'institutional funding', which the universities can spend as they see fit, and partly from external, competitively won project-based funding, which in principle has to be spent on the subject matter of the projects. Government can use the combination of institutional and external funding to provide universities with incentives to do research that is of high quality and that addresses a range of societal and industrial needs in addition to responding to a solely scientific logic.

The REF is a performance-based research funding system (PRFS), which governs the allocation of much of the institutional funding for research provided to the UK universities. In most countries where a PRFS is in use, it governs only a small part of the institutional funding for research provided to the research-performing institutions. The UK is unusual in that a large proportion of institutional funding is determined by the PRFS.

The REF is based on peer review of outputs of research that are submitted to it by the UK universities. While a small number of other systems use peer review, most other countries' PRFS are indicator-based and rely on a mixture of scientometric indicators (typically relating to scientific publications and sometimes patents) and indicators of other outputs, such as the number of PhDs produced.

The UK research system context

The UK uses less of its gross domestic product (GDP) on R&D than many leading developed countries and this figure has tended to fall over the last 30 years or so. Expenditure on research in the higher education sector has risen over the same period to reach about the same proportion of GDP as France, Germany, Japan and the USA but more slowly than in the small countries that are the top performers in research, according to the available bibliometric indicators. By the yardstick of citations (how often other researchers refer to a scientific publication when they publish their own), UK science is more influential than that of the other medium and large sized countries while failing quite to reach the level of the very best of the small ones (Denmark, Netherlands, Switzerland). As in other countries with a strong research base, a large proportion of the UK's published scientific output is written in collaboration with authors abroad, indicating that the UK research community is well integrated into the global system. In the aggregate, UK researchers thus appear to generate a lot of quality on the basis of a level of funding input that is modest in international terms.

History of the REF[1]

In 1982, cuts to universities' institutional funding for research were implemented unevenly, with some universities subject to cuts of up to 30%. The universities saw this as lacking transparency and biased in favour of the traditional elite universities, especially Oxbridge. The next time these budgets were revisited, in 1986, the University Grants Committee (UGC) decided to abandon the traditional 'equity principle' that assumed all universities did research and therefore needed institutional funding for research. Instead, it introduced the idea of 'selectivity', which meant selectively allocating more

[1] The REF was called the Research Selectivity Exercise in the 1980s and the Research Assessment Exercise in the following two decades before being re-named as the REF in the 2014 exercise. Except where we describe specific earlier exercises, we refer to 'the REF' for simplicity.

resources to those which did best. This was implemented in the 1986 Research Selectivity Exercise, in which peer review panels evaluated research outputs. A proportion of the overall institutional funding available was redirected towards those judged to be the better performers. A second selectivity exercise was done in 1989.

Following the abolition of the UGC and the creation of the Higher Education Funding Councils (HEFCs), for England, Wales, Scotland, and Northern Ireland, the first Research Assessment exercise in 1992 allocated about 90% of institutional research funding based on judgements of quality, with those achieving higher grades getting many times as much as those which scored less well. While there was a constant process of evolutionary change over the succeeding two decades, the Research Assessment Exercises continued to focus on quality. The criteria were only broadened in the new REF in 2014, when quality governed 65% of funding, universities' descriptions of impacts of their past research governed 20% and data about the research 'environment' the remaining 15%.

The effectiveness of the REF

The purposes of the REF are to inform the selective allocation of institutional research funding to universities by rewarding excellence and international competitiveness, protecting research infrastructure and 'pockets' of excellence and encouraging dynamism in the research system as well as to provide benchmarking information and accountability.

In HEFCE' s terminology, the quality-related component of REF funding is referred to as 'quality=related- (QR). The bigger component of HEFCE's QR research funding, referred to as 'mainstream QR in 2015/16, is allocated to English HEIs based on the volume of research, subject cost weights and quality of research, as measured by the REF. Other components of QR that are driven by the REF include RDP (research degree programme) supervision funding and 'London weighting' which compensates for the higher costs of being located in London. The Charity Support Fund and the business research element are also a substantial part of total QR funding in 2015/16 and are used to compensate for the reluctance of both the charities and industry to pay the full overhead costs of research they commission from the universities. Altogether, total QR funding has increased in current money terms from 1994/95 (£590m) to 2010/11 (£1,603m) and thereafter dropped slightly to £1,559 in 2015/16. The component of QR funding that is more directly linked to the REF ('mainstream QR') rose slowly in money terms up to 2001/02. Thereafter, it remained close to constant in current money terms and fell gradually as a proportion of the total QR funding to about 65% in 2012/13. The difference was largely made up of the increase in the RDP fund and the Charity Support Fund (from 2006/07). Since there is in fact a quite close similarity between the way external funding from the research councils varies over the university population and the allocation of institutional funding based on assessment, the institutional funding system in practice provides the universities with at least some of the internal funding they need in order to be able to afford to do externally funded research (which rarely covers its full costs).

REF funding is concentrated, with over 50% of total QR funding going to the top 10 universities and 71% of total QR funding going to the larger Russell Group of research-intensive universities in 2015/16. However, under the RAE 2008, it was slightly less concentrated than research funding from other sources. The shape of the overall distribution is rather static over time, though within this there can be quite large 'wins' and 'losses' for individual universities. Adding up these individual gains and losses, they account for 18.6% in 1996/97 to 8.6% in 2015/16 of the total QR funding, although the net effect, accounting for the change in total funding is lower and ranges from 4.0% in 2002/03 to 8.6% in 2015/16 of the total QR funding. There are a few examples of universities emerging from comparative obscurity and consistently receiving growing shares of REF money over time, but mostly universities' shares go up and down from exercise to exercise. There is some evidence that money moves between disciplines over time.

The limited international evidence available suggests that PRFS are associated with increased production and quality of scientific outputs. However, this normally occurs in a context of rising performance so it is unclear how much of the increase can be attributed to the introduction of a PRFS and how much would have taken place even without it.

The first Research Selectivity Exercise took place in the UK at a time when the citation impact of UK research was declining. That trend was reversed after the 1989 exercise, since when UK citation impact has continued to grow. While it seems likely that the RAE played a role in this reversal, no solid statistical evidence is available to this effect and it is difficult to distinguish the effects of the RAE from other factors contributing to the performance improvement.

Effects on research

Concerns are raised in the literature on the REF and other PRFS that they may have negative effects on the amount of interdisciplinary and 'transformative research done, that they discourage non-conventional and applied research and reduce researcher autonomy and propensity to collaborate.

There is general agreement that discipline-based peer review assessment systems do not handle interdisciplinary research well. Like other assessment systems, the REF tries to tackle this by using special procedures: broad, sometimes interdisciplinary panels and the option to refer research outputs across panel boundaries in order to obtain a second opinion. We found no evidence to indicate that the REF handles interdisciplinary work less well than other systems or that it discriminates against it. However, the proportion of the outputs submitted to the REF that are interdisciplinary is lower than that in UK research as a whole, suggesting that university research managers are less likely to submit interdisciplinary outputs to the REF than other outputs. This could have a knock-on effect, discouraging interdisciplinary work in the universities, but we found no evidence about this and there are clearly also other forces operating that promote interdisciplinary work.

Transformative research, in the sense of high-risk research that challenges established theory or method, is understood to suffer in peer review of proposals and manuscripts, partly because of opposition to new ideas from the research establishment and partly because risky research is less attractive to funders than non-risky research. This concern is significant enough that some funders have set up schemes intended specifically to fund transformative research building on unconventional peer review approaches. Since the REF panels are populated by established scientific figures, there is a risk that they might rate transformative work poorly but this possibility does not seem to have been studied. The wider evidence that the presence of the REF in the funding system leads research managers to prefer low-risk research that produces outputs in the short term suggests that if there is a problem it is not in the REF process itself but in the way the REF encourages university research managers and researchers to behave.

There is evidence from the fields of economics and innovation studies that the role of mainstream researchers in the REF panels and research managers' use of 'approved' lists of mainstream journals in which they encourage publication tend to drive out non-mainstream, unconventional work.

Applied research can suffer in a metrics-based PRFS because a greater proportion of its results are published outside the indexed journals used for bibliometrics than is the case with more fundamental research. However, the REF panel system means that outputs are assessed by peers who understand how the relevant disciplines publish. There appears to be no evidence or *prima facie* reason to expect applied research to be disadvantaged in the REF. University career incentives to publish in indexed journals, however, may influence researchers' choice of topics.

There is evidence that PRFS can influence the extent to which researchers cooperate and co-publish. Whether this influence is positive or negative depends on specific aspects of the PRFS design. The effects of the REF in this respect are unclear.

Effects on researcher careers

The effects of the REF and other PRFS on individual researchers are mediated through the mechanism of careers. An immediate effect of the fact that PRFS analyse outputs is that research managers increase the importance of producing outputs in their assessment and management of research careers and de-emphasise other aspects. These outputs are expected to appear in high-status, indexed journals and this changes the character of the research done, especially in disciplines where journal publication has

traditionally not been the dominant mode of communication. The importance of outputs extends to the recruitment process.

Consultations with the academic community in the UK about research assessment and the REF tend to produce agreement that the REF functions as a performance management tool. Institutional comment (ie from research managers who use the REF as a tool) tends to be positive. Unsurprisingly, comment from individual academics (on whom the tool is used) tends to be negative in tone, as is much of the scientific literature on the REF.

When surveyed, few UK academics see the REF as contributing positively to the quality of UK academic research or as measuring that quality well. Many say that their university threatens them with career-related sanctions if their REF performance is poor. This negative perspective is slightly more prevalent among younger researchers and women.

The fact that the REF's focus on research has not been counterbalanced by a similar system for teaching has changed UK academics' understanding of their role. The sense that teaching has been downgraded in importance is reinforced by the use of transfers from academic to teaching-only contracts as an actual or potential sanction against academics whose research performance is seen as not 'REFable'.

National research assessment systems in general, and the REF in particular, foster transfer markets for academics, where those who produce highly rated outputs can find themselves in high demand. Others who do not rate well in the PRFS are likely to be disadvantaged in comparison.

International practice in performance-based research funding systems (PRFS)

The REF is probably the oldest PRFS but, especially since 2000, many countries now use such a system.

Performance-based research funding systems are intrinsically linked to the structure of the research system, the national approach to research governance and the policy context. As a consequence, PRFS are country-specific and respond to specific policy objectives. Four main categories of policy objectives can be observed, namely to

- Enhance the quality of research and the country's research competitiveness
- Steer behaviour in order to tackle specific failures in the research system
- Strengthen accountability
- Provide information for research strategy at institutional and/or national level

PRFS designers face a choice between peer review and metrics-based approaches. Peer review is generally seen as costly so more recently designed systems tend to make heavy use of metrics, sometimes combined with a small amount of peer review. On the other hand, peer review systems are better able to tackle inter-field differences than metrics-based ones.

Historically, PRFS have focused on research quality, but more recently there has been increasing interest also in understanding the effects of research on innovation and society more generally. Indicators of quality are relatively well developed and there is a broad consensus in the research community about what they mean as well as their weaknesses. While many impact indicators are available, they are not standardised to the same extent and the meaning of individual indicators in relation to societal effects of research is often unclear.

Both peer review and metrics approaches can be used to provide analysis down to the level of the individual researcher. However, this is in practice not done. Peer review is best adapted to working on fields or disciplines while metrics are better suited to providing assessments at the level of organisations. Metrics-based systems can operate frequently. Many are in fact run annually. The cost and complexity of a peer review approach means it is limited to longer periods.

The proportion of research income that universities obtain from external versus institutional sources varies considerably among countries and there is no clear link between this proportion and the

performance of the research system. In most cases, the PRFS drives only a small part of the institutional funding.

The UK system stands out from others in that a very high proportion of institutional funding is distributed based on quality criteria and that the performance-based funding itself comprises a high proportion of institutional funding for research overall. The UK is one of the few countries where institutional funding comprises less than 50% of universities' research income, so it is already much more strongly competition based than others. The UK funding formula selectively allocates almost all the institutional funding for research to the best performers. This leads to a degree of concentration of institutional funding not seen in other systems.

Research Information Systems (RIS) and performance-based funding

Research information systems (RIS) play an increasingly important role for the governance of research at the institutional, funding body and national levels.

At the national level, governments increasingly develop national research information systems that would constitute a consolidated basis for the collection and analysis of information related to research performance in the country as a whole and the identification of short, medium or long-term impacts. Research assessment and in particular the need for information informing PRFS is among the drivers for government research information systems.

Some small countries have gone a long way towards creating a national database of research outputs that can be used in a PRFS. Often this has been driven by the need to approve and assign quality ratings to journals and other outputs written in a minority language and not necessarily indexed in the commercial bibliographic databases. Other countries are making progress in developing and interconnecting RIS systems, aided by the availability of international standards for such systems and for communication between them.

In the UK, the research councils have a joint RIS called Researchfish, which collects and stores information about research they have funded. A number of other funders, notably charities, also use the system but it does not provide a comprehensive set of information about UK research outputs. Most universities maintain their own RIS. The Higher Education Statistics Agency maintains a comprehensive but separate set of data about the universities. It is recognised that these could form useful building blocks in constructing a comprehensive national RIS, which could in turn inform a future REF exercise. However, there is no project in progress to create a comprehensive national RIS.

Table of Contents

Tables

Figures

1 Introduction

1.1 The role of this report

This is the report of a project commissioned in 2016 by the Department of Business, Energy and Industrial Strategy (BEIS). It provided evidence to the review of the REF undertaken under the leadership of Lord Stern (Stern, 2016). It explores evidence about the following.

- The context of the REF in the UK research system
- The history of the REF and modifications made to the process over time
- The effectiveness of the REF, in so far as this can be discerned from quantitative evidence
- The effects of the REF on the research process and research careers
- International practice and experience with equivalent forms of research assessment and funding
- The extent to which Research Information Systems (RIS) could be used to support the REF in future

1.2 The REF

The REF is the UK's performance-based research funding system (PRFS), which assesses the 'goodness' of UK university research along a range of dimensions and uses the resulting judgements to inform the allocation of institutional funding for research to the universities. It has antecedents going back to the Research Selectivity Exercise in (RSE) 1986 and a series of Research Assessment Exercises (RAEs) in intervening years. For simplicity, we refer to all of these as the 'REF' except where we need to focus attention on a specific exercise.

The REF was introduced as a policy instrument in a time very different from the present, during the early stages of 'massification' of higher education when the university sector as a whole was much smaller than today and in the early stages of a shift from educating a small elite towards an aim to offer higher education to up to half of the people leaving school. Despite continuous and mostly small changes in approach, the REF has been in important respects rather stable[2], remaining within one part of the 'design space' available to designers of PRFS systems. As consultation exercises in recent reviews of the REF show (Wilsdon, 2015) (Stern, 2016), some of its central characteristics are strongly defended by the academic and university management communities and has become embedded in the wider culture of the UK university system.

The costs of the REF have been regarded both nationally and internationally as large. The most recent analysis shows that the main cost is the effort the universities make in order to feed into the assessment process (Farla & Simmonds, 2015). The structure of these costs suggests that achieving a major reduction in cost and burden would depend upon changing some of the central principles of the REF, especially the role of the universities in the process. That is why we not only look at the REF itself in the UK and try to discern its effects on the research system but also look abroad in order to understand how others achieve similar purposes by other means.

1.3 Some terminology and definitions

Like any institution, the REF has generated its own jargon. Since part of our task is to look at equivalent systems outside the UK, we need a more general language, which we introduce in the following discussion.

This report – and the REF itself – relates to institutional funding for research. In the UK we speak of a 'dual support system' for research (which is in fact the norm in Europe and much of the rest of the world). Under such a system, money for university research comes partly as 'institutional funding' provided directly to the universities and which they can spend as they see fit, and partly from external,

[2] Only in 2014 were substantially new assessment criteria introduced

competitively won project-based funding, which in principle has to be spent on the subject matter of the projects.

Internationally, the general way to provide institutional funding to state universities in the post-War period was through a single block grant. As university systems grew, funders tended to make an explicit distinction between funding for education and that for research. In each case, this is 'institutional funding' in the sense that the money is passed from the responsible ministry to the institution in one block. In most European countries, university autonomy means that no matter what formulae, processes of negotiation or incentives are built into institutional funding, once the money is in the hands of the universities they themselves decide how to spend it[3]. The next step in many countries has been to make some of the institutional funding dependent upon past performance. Making institutional funding conditional upon performance changes its role in that it becomes retrospective, while external funding is prospective in character.

The REF is a PRFS. Any such system has two components: an assessment process, which makes judgements or decisions about performance using specific criteria; and a funding process, which takes the results of the assessment and uses them to guide the allocation of institutional funding – normally to universities, although there are PRFS that also allocate funding to research institutes (for example in Norway, the Czech Republic, Latvia and Lithuania).

In most countries where a PRFS is in use, it governs only a small part of the institutional funding for research provided to the research-performing institutions. The UK is unusual in that a large proportion of that funding is determined by the PRFS.

The REF is based on peer review of outputs of research that are submitted to it by the UK universities. While some other countries use or involve peer review in various ways, most use indicator-based PRFS, which rely on a mixture of scientometric indicators (typically relating to scientific publications and sometimes patents) and indicators of other outputs, such as the number of PhDs produced. Indicator-based PRFS can be designed to assess the work of individual researchers irrespective of their field and then aggregate these to the level of groups, faculties or entire institutions. However, peer review systems must view research outputs through the lenses of disciplines or fields. They therefore in practice tackle groups of people within individual institutions who research in similar areas, typically disciplines. Like the UK REF, we refer to these groupings as 'Units of Assessment'.

A final clarification on institutional funding is necessary. While the distinction between institutional and project-based funding in a dual support system is conceptually neat, it has to tackle some messy realities. The traditional UK view was that institutional funding for research should finance the internal infrastructure of people and facilities needed for the university to play 'host' to externally funded projects and to some extent to do internally-funded research. The university was to provide a 'well-found laboratory'. The UK Research Councils would provide additional money to pay the variable costs of competitively-awarded research projects but expected the institutional funding to cover the fixed costs. Hence, they did not pay 'overheads' to the universities but only the marginal costs of research (Parliamentary Office of Science and Technology, 1997). Non-academic funders such as industry were expected to pay the marginal costs plus an 'overhead' contribution to fixed costs.

From 2005, a principle of 'full economic costing' of research was introduced across UK higher education. This required universities to calculate the total costs of any project they won competitively. The principle was retained that the Research Councils should get a discount (typically 20%) because the state had already provided institutional funding for research while others who had not made a contribution had to pay the full amount. It therefore became important that the distribution of institutional funding for research among universities is not too dissimilar from the distribution of Research Council income,

[3] This is a truth with modification. University systems are in different stages of progress towards full autonomy. UK universities' charters have always meant that they are supposed to be outside the control of the government. In other cases, some of the budget may be 'hypothecated', in the sense that the university has to spend it on particular things, such as academics' salaries. But the trend towards full autonomy is clear and strong

otherwise at least some universities will have to find ways to cross-subsidise Research Council work, for example from the teaching budget.

The bigger component of institutional finding for research that is governed by the REF is known in the UK as 'mainstream QR' (quality related) funding. In practice, other external funders than the Research Councils are also reluctant to pay overheads. The UK system provides a second category of 'institutional' funding in the form of 'Other QR' (non-mainstream QR). This encourages the universities to work with research charities (like the Wellcome Trust) and at least parts of industry by compensating them for the loss of overheads. Much of the growth in total QR funding for research since the full cost reform is in relation to the QR Charity support fund.

1.4 The role of PRFS in the wider research funding system

Historically, both in the UK and elsewhere, the trend has been for the ratio of external, competitive project funding to institutional funding to rise (Lepori, et al., 2007), suggesting increased competitive pressure for research resources and increased shaping of the research agenda and the research-performing institutions themselves by external forces. It also places more pressure on universities' institutional research funding, which needs to support a growing volume of externally funded research.

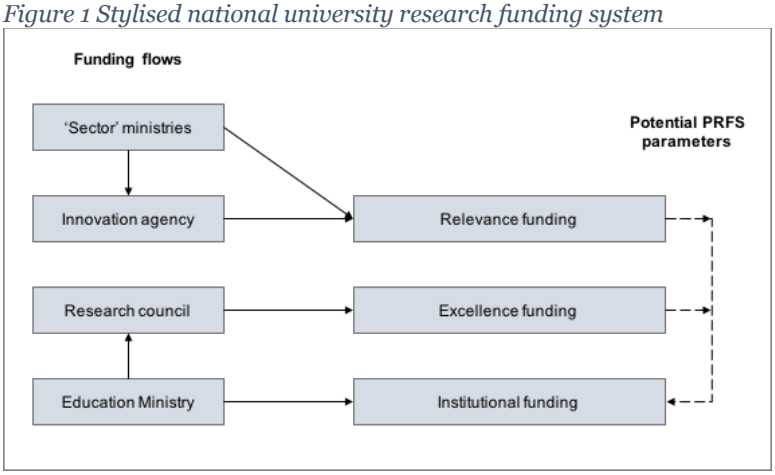
In most countries, research councils initially dominated the external research-funding stream, responding to investigator-initiated ('bottom up') proposals. This imposed quality control through peer review of project proposals but did not involve overtly directing research activities towards particular themes[4]. As a consequence of the OECD's work to promote 'science policy' in the form of a linkage between national (especially industrial) and scientific priorities, a new set of institutions ('innovation agencies') developed in many countries from the late-1960s that programmatically funded 'relevant' research. The innovation agencies thus generate 'focusing devices' (Rosenberg, 1976) (Arnold, Good, & Segerpalm, 2008) in the form of projects and programmes, implementing science policy through research-funding incentives.

Government university funding at the national level today tends to comprise the blocks shown in Figure 1. An education ministry typically provides both institutional funding and money for 'excellence' research, the latter through a research council. In principle, the institutional funding is infrastructural in nature: it provides a basis for strategy and planning and for capacity to do research in the universities. Ministries responsible for industry and various other 'sectors' of society (energy, environment, defence, etc) may fund research in the university system, directly or through an innovation agency. (Often they also use captive government labs or contracts with industry to do research.) Business provides a further funding stream for university research, offering resources for industrially relevant work. The external funders thus provide incentives for particular kinds of research. The balance among the various flows would be expected to influence the overall shape of the university research system.

Many PRFS now use the amount of external research funding from various sources as quality indicators, so they can be used to magnify the effects of external funding – in the direction of 'excellence' or 'relevance' or internationalisation via participation in the EU Framework Programme of research and technological development. However, while PRFS clearly interact with their context, there is no systematic research yet done that tries to understand this interaction more clearly. Nor have we been able to identify explicit policy discussions about this.

[4] The allocation of budget among thematic or disciplinary research councils, of course, did generate some degree of thematic steering but these research councils tended not to programme their resources, largely responding to 'bottom-up' proposals and prioritising among them on the sole criterion of scientific quality.

Figure 1 Stylised national university research funding system



2 The UK research system context

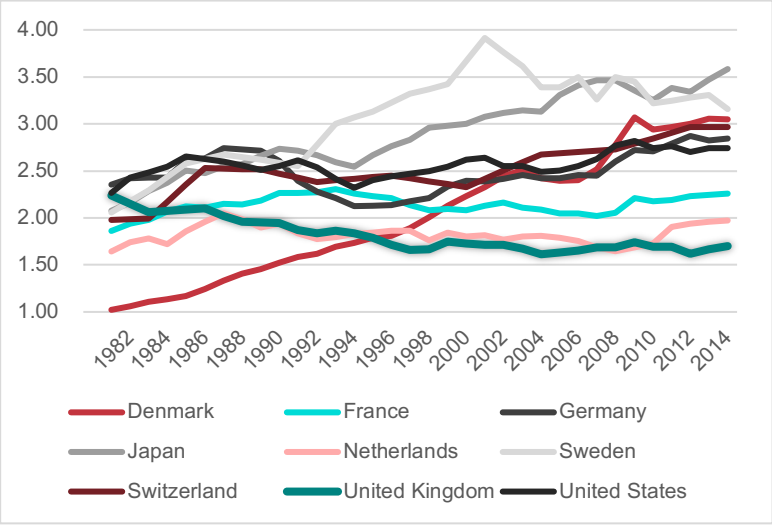
Research funding policy is made in specific national contexts, including the pattern of need implicit in the industrial structure and the extent to which the business community invests in research. This Chapter sets out a few key aspects of the UK context. Additional charts and data are provided at the Appendix. We first discuss the structure of production and of research in the UK. We observe the dramatic increase in the size of the higher education system during the life of the REF. We provide some bibliometric evidence to support the idea that UK research performance is strong and note the importance of internationalisation in supporting that strength. This provides needed context for our subsequent discussion of the effectiveness of the REF.

2.1 Spending on R&D in the UK

IMF data show that the structure of UK production is roughly 80% services and 20% industry. Agriculture accounts for less than 1% of Gross Domestic Product (GDP). In this respect, UK production is similar to that of France, Denmark and the USA, where industrial output is similarly one fifth of the total. Germany, Japan and Switzerland get more like 30% of their output from industry. Sweden and The Netherlands sit somewhere in between (see Table 27 at the Appendix). By implication, since industry is the main user of research[5], the UK research system is likely to experience less demand for scientific and technological research than some of the other countries considered. Conversely, a policy objective could be to use the science system to drive increased use of research in industry and to create an attractive environment for inward investment.

The amount of R&D funding input into the UK research system is rather low compared with a number of other mature and successful economies. Figure 2 shows that the UK's Gross Expenditure on R&D (GERD[6]) has declined as a share of GDP since the early 1980s, while in a number of comparator countries[7] it has tended to increase.

Figure 2 Gross Expenditure on R&D as a Percentage of GDP, 1981-2014



Source: OECD Main Science and Technology Indicators (MSTI), accessed 21/5/16. Some values are interpolated in this and subsequent Figures derived from the MSTI

[5] The distinction between industry and services can, however, be slippery. Some services (IT, design, North Sea supply, and so on) can be knowledge-intensive, relying heavily on research and the production of research-capable manpower

[6] GERD is literally the sum of all the money spent on research by government, industry and others in a country

[7] Denmark, France, Germany, Japan, The Netherlands, Sweden, Switzerland and the USA. These were chosen for their size, advanced stage of economic development and, in the case of the smaller countries, for their excellent research production, based on citation indicators

In contrast to overall GERD, the UK's Higher Education Expenditure on R&D[8] (HERD) has grown from 0.3% of GDP in 1981 to 0.44% in 2014, though it is still well below the levels seen in some of the comparator countries (Appendix, Figure 24). France, Germany, Japan and the USA spend roughly the same proportion of GDP on HERD as the UK does. HERD in Denmark, Switzerland and The Netherlands has grown particularly fast. That in Sweden has grown more slowly, but all these are well above the UK level.

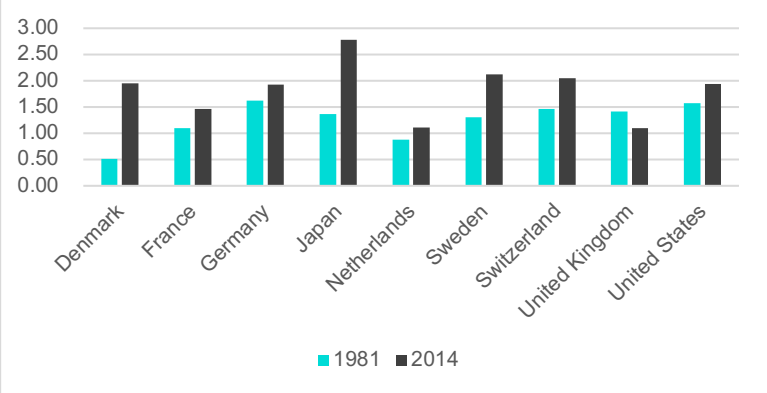
A number of small countries have therefore dramatically increased HERD over the period to very high levels. These are the small countries that perform very well in bibliometric terms. The increase in Denmark results primarily from sustained budget increases, but note that most of the government laboratories were absorbed by the universities in 2008, which has the statistical effect of moving them out of the Government Expenditure on R&D[9] (GOVERD) category and into HERD.

In the GOVERD category, the UK government lab sector has shrunk dramatically since the 1980s (Appendix, Figure 25). UK GOVERD has fallen from 0.46% to 0.13% of GDP between 1981 and 2014. Most of the comparator countries have reduced their GOVERD, but none by as much as the UK. Germany and Japan have increased their spending in this category. Sweden and Switzerland have in any case historically done little research in government labs.

The GOVERD data should be viewed also in the understanding that this is the category containing the bulk of government defence research and the fact that this has tended to decline since the end of the 1980s as countries took a 'peace dividend' after the end of the Cold War.

The main driver of GERD in developed countries, however, is not government's or universities' R&D expenditure but that of business (BERD[10]). Figure 3 shows that, uniquely among the countries considered, the UK's BERD has gone down as a share of GDP over the period, paralleling the decline of UK manufacturing industry. The pattern, then is that in the UK, both the government and business sectors have been reducing their R&D efforts, while the higher education sector research effort has gone up – but only by a comparatively small amount.

Figure 3 Business Expenditure on R&D as a percentage of GDP



Source: OECD, MSTI, accessed 21/5/16

[8] That is the sum of all the money spent on research in the university system, irrespective of source. It includes both institutional and external funding,

[9] That is the sum of all the money the government spends on research outside the university sector, for example through government labs or by buying contract reach from industry. In some countries, defence accounts for a big part of GOVERD.

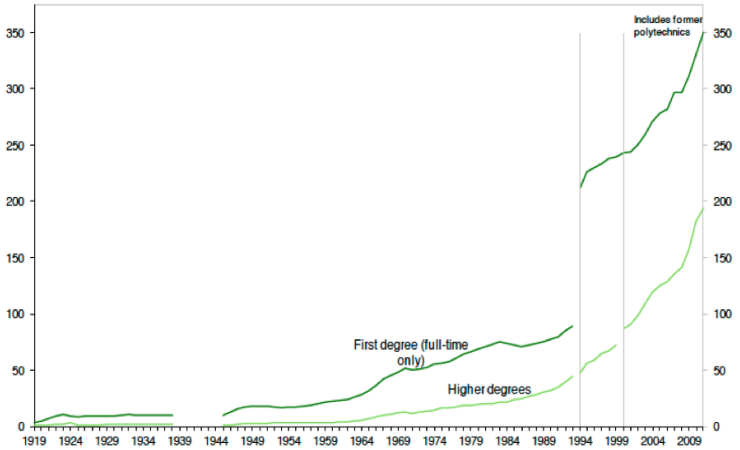
[10] This is literally the sum of all the money business spends on R&D, irrespective of whether companies are privately or publicly owned.

2.2 Expansion of the higher education system and growth in university research funding in the UK

At the same time, the university sector as a whole in the UK has expanded dramatically, driven by a desire to 'massify' higher education. Most developed countries, like the UK, have decided to invest in increasing the proportion of young people who complete higher education from a handful of percentage points to something in the range 40-50%. This puts pressure on national budgets (because it costs several percent of GDP to do it), redefines the meaning of degrees in the labour market and creates a need for a large amount of teaching capacity, raising the question whether in the massified universities the old assumption that all academics should both teach and do research can remain valid. The Research Selectivity Exercises in the 1980s addressed this problem, aiming to concentrate institutional research funding on areas of high quality (OECD, 2010).

The historical data on UK first degree production in Figure 4 show low levels before the Second World War, followed by a boost in the 1950s and a rapid increase with the wave of new university building in the 1960s and then further expansion in the 1970s. The award of university status to the polytechnics at the start of the 1990s caused the discontinuity shown at that point, as the number of institutions allowed to award degrees roughly doubled. So the number of first degrees awarded per year has roughly quintupled during the lifetime of the REF while the number of institutions that could in principle do higher education research has roughly doubled. Over the same period, the proportion of GDP devoted to HERD has risen by about 50% (Appendix, Figure 24).

Figure 4 Students Obtaining University Degrees in the UK (000s), 1919-2009

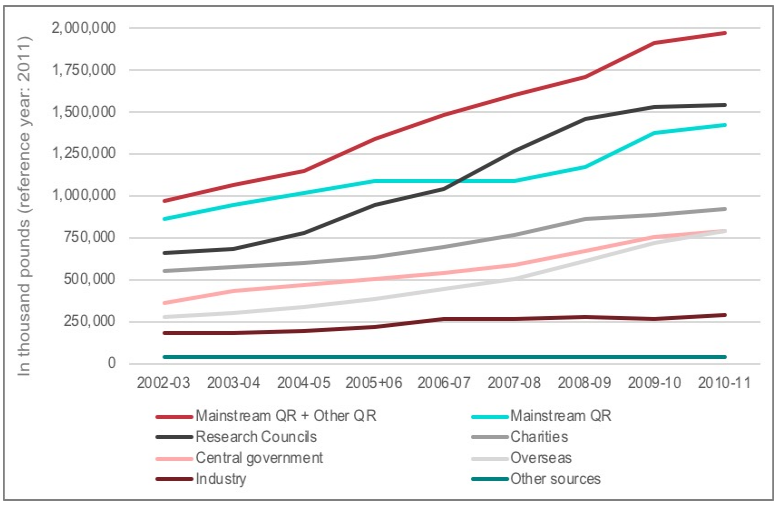


Source: (Bolton, 2012) based on statistics from HESA
Note: The break in the 'higher degrees' line is caused by the inclusion of graduations from 'dormant' status, which were previously not counted. Typically, this would involve a PhD student finally submitting a thesis after some years of apparent inactivity.

The 'dual support' funding system used in the UK (and in most other countries) is generally seen as an important factor in maintaining and developing research capacity and performance in the face of the wider economics of massification. Figure 5 shows the various sources of funding for UK university research since the Millennium. (In effect, it provides a breakdown of the HERD category discussed above.)

Figure 5 illustrates that total QR funding is larger than other university research funding (2002/03 – 2010/11). Income from research councils overtook that from 'mainstream QR' in 2006/07 after the creation of the QR Charity support fund and increase in the research degree programme (RDP) PhD supervision funds over time. The growth of industry funding is the slowest of all the sources shown on the chart at the same time as the share of GDP allocated by business to R&D declines.

Figure 5 Funding of UK university research: dual support and other sources, 2002-3 to 2010-11 (in 2011 prices)



Source: Based on (Hughes A. , Kitson, Bullock, & Milner, 2013), calculations based on HESA Financial Statistics

2.3 The performance of the UK research system

These trends suggest a sort of 'pressure cooker' effect on the UK universities, whose research has dramatically increased in both quantity and quality despite the (in international comparison) limited spending by both government and business. A recent study by Elsevier (BIS, 2013) suggests that the UK research system is highly efficient in terms of scholarly production and inventions and produces good quality research results. The comparator countries for that study were Brazil, Canada, China, France, Germany, India, Italy, Japan, Russia and the USA. (Note that these omit a number of small, high-performing countries.)

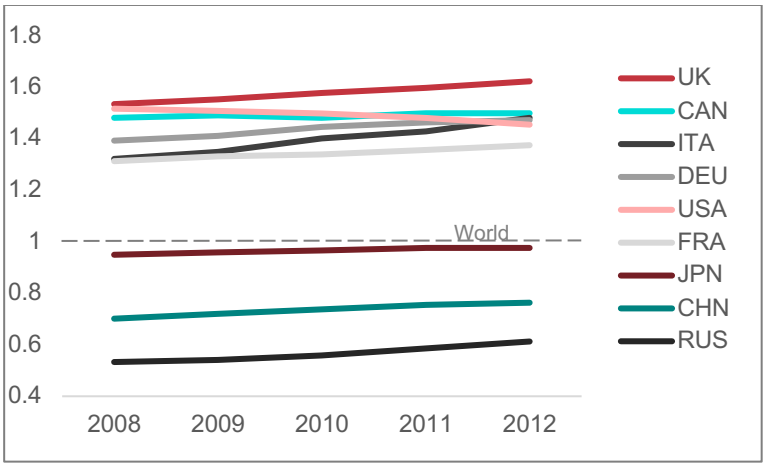
- The UK ranked first among Elsevier's comparator countries for the number of scientific articles (in Scopus) per $m of GERD (3.9 articles)
- The UK produced 0.5 articles per researcher in 2012, ranking third after Italy and Canada
- UK articles generated 5.87 citations per researcher in 2012, ranking second after Italy
- "While the UK represents just 0.9% of global population, 3.2% of R&D expenditure, and 4.1% of researchers, it accounts for 9.5% of downloads, 11.6% of citations and 15.9% of the world's most highly-cited articles" (pp. 2)
- "The UK is a focal point for global research collaboration and researcher mobility" (pp. 2)
- "The UK exhibits strong cross-sector knowledge exchange" (pp. 2) such that a high number of UK academics download UK corporate-authored articles and a high number of UK corporate users download UK academic-authored articles.

The study also shows that the field-weighted citation impact[11] of UK research is higher than the equivalents for the comparator countries and that it is on an upward trend (Figure 6). Altogether, the

[11] Bibliometric analysis conventionally treats the number of times a scientific publication is cited in other scientific publications as a proxy for its impact on science and therefore as an indicator of quality. However, researchers in different fields or disciplines have different publication patterns, both in terms of the publication channels they use and in the number of publications a researcher can produce per year. For example, humanities scholars publish a lot in books (which are not indexed in the databases used for counting citations) while natural scientists tend to prefer writing papers in scientific journals (which are). In some disciplines like chemistry it is possible to publish many papers per year while in others such as mathematics a good researcher may only produce one publication in a period of a few years. There are also more researchers in some fields than others, so a good paper in a large field will be cited many times more than an equally good paper in a small one. Comparison among fields or aggregates of fields (as when comparing countries) is therefore done by normalising citations by the average level of citation in each field. These field-weighted citation impact values can then be compared directly or they can be aggregated to the country level, as has been done to produce the curves in Figure 6

study suggests that the UK research base is internationally competitive and is increasing its performance on some dimensions.

Figure 6 Field-weighted citation impact for the UK and comparators, 2008-2012



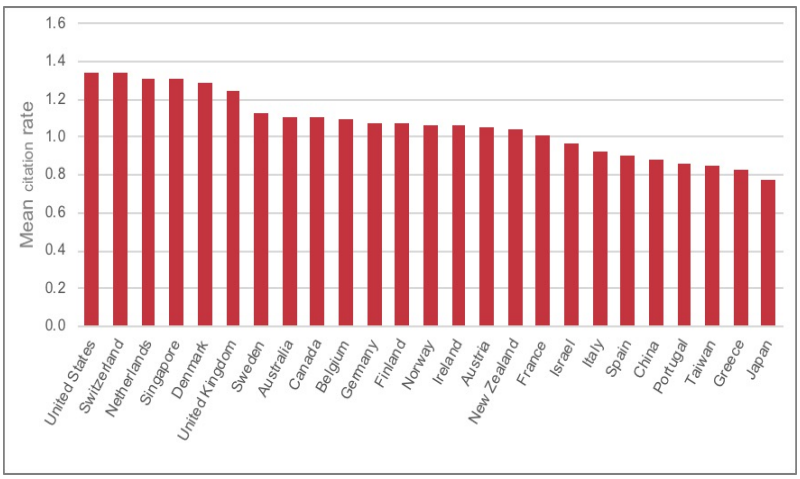
Source: Based on (BIS, 2013)

While UK bibliometric performance is strong compared with that of other large countries and Canada, a number of smaller countries appear to do as well or better – particularly those identified above as substantially increasing their investment in HERD. Notably, these high performing countries do not operate dual support systems that involve high degrees of competition. Those that have introduced PRFS have done so within the last ten years. (Chapter 7 provides more information about PRFS systems internationally.)

- In Denmark, some 72% of universities' dual support income comes in the form of institutional funding for research. 10% of the institutional income is allocated through a PRFS
- In The Netherlands, some 70-80% of universities' dual support income comes in the form of institutional funding for research. Research assessment is not connected to funding allocation
- In Sweden, some 57% of universities' dual support income comes in the form of institutional funding for research. 10% of the institutional income is allocated through a PRFS
- In Switzerland, some 82% of universities' dual support income comes in the form of institutional funding for research. There is no PRFS
- In contrast, in the UK, some 48% of universities' dual support income comes in the form of institutional funding for research. Currently, about 68% of that institutional income is allocated through the REF assessment exercise in the form of mainstream QR and most of the rest as 'other QR'

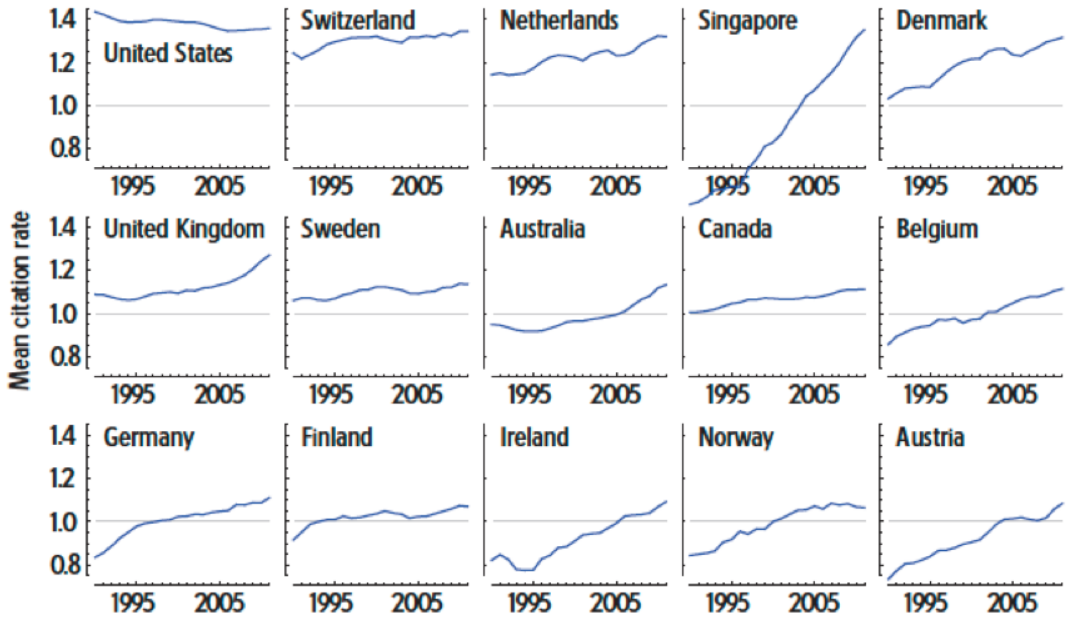
Figure 7 shows the countries with the highest field-normalised citation rates in 2009-10. (These are indexed so that the world average is 1.0.) Figure 8 shows how these rates have developed over two decades in a sample of these countries. With the exception of the UK (where the REF has been in place for the whole period) the improvers manage to do so in the presence of either a weak or no PRFS. Clearly, it is difficult to attribute performance in any clear way to the use of a PRFS in these cases; equally clearly, given the variations in the degree to which research funding is contested, investment and institutional factors must play a large role, though probably in different proportions in different countries.

Figure 7 Countries with the highest normalised mean citation rates, 2009-11 among the 39 countries with at least 4,000 publications per year



Source: Based on (Karlsson & Persson, 2012)

Figure 8 Trends in mean citation rates, 1990-2011 for fifteen of the currently most highly cited countries

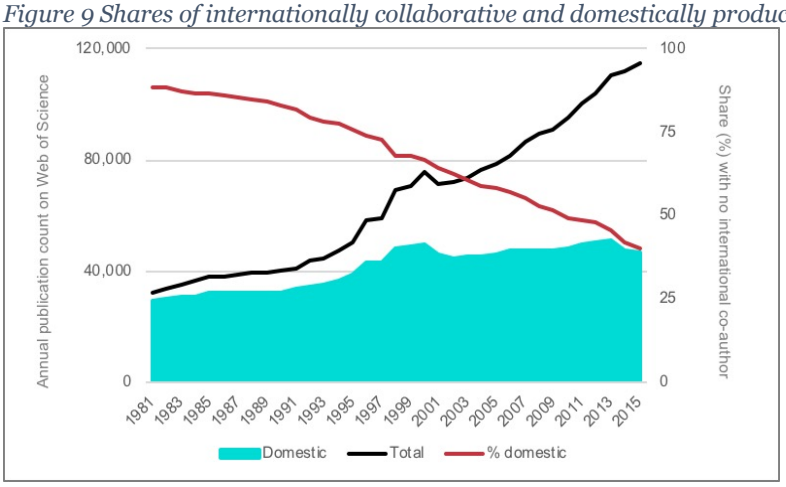


Source: (Karlsson & Persson, 2012)

2.4 Internationalisation

The improvement in citation rates observed in the UK and other countries is accompanied by an increasing amount of international collaboration. This is most easily visible in the proportion of publications in the Web of Science that are written in collaboration between people with UK and non-UK addresses, compared with the wholly domestic production (Figure 9).

Figure 9 Shares of internationally collaborative and domestically produced UK scientific publications over time



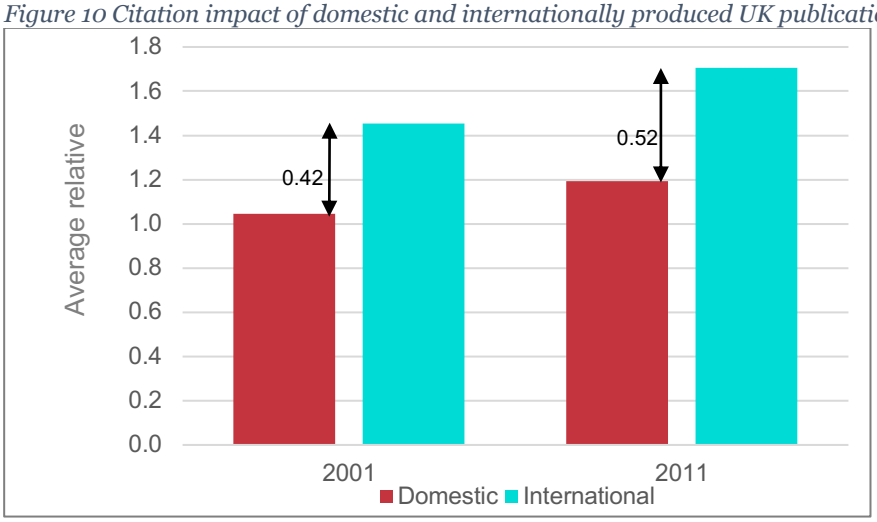
UK total and domestic research output of article and reviews in journals indexed on Thomson Reuters Web of Science™. The left hand axis is annual output where (i) the continuous black line refers to total and (ii) the blue shape refers to domestic output. The right hand axis refers to the percentage of total output that is domestic (has no international co-author) in each year (red line).
Source: (Adams & Gurney, 2016) based on Web of Science data

Collaboration has an important and positive effect on the citation impact of papers (Figure 10) and reinforces the widely-accepted idea that science is a global rather than a national activity. As Adams and Gurney (2016) observe

> *The most research-active universities are intensively engaged with partners in other countries … and the publications that emerge become the most highly cited part of the UK (and international) research base. This shifts the leading edge of research from a national base to an international network and the gap between that network and the domestic research base is growing. By implication, universities that are not partnered with institutions in other countries are less closely engaged with that leading network. If they are not actively involved, then they risk being left behind the advancing research front. Their reduced awareness impairs their ability to competitively acquire resources, and that will further compromise their research capacity. To retain access to the network of knowledge and innovation and to remain competitive in accessing resources, it is essential that each university develops a strategy to expand its international engagement.*

It seems that the networking and collaboration is most intense among the international research elites. Collaboration among the top 100 universities according to the CWTS (Centre for Science and Technology Studies, Leiden University) ranking is rising more quickly than that with other universities (Gazni & Thelwall, 2014). Presumably this intensifies the tendency of elite groups of researchers to form global 'invisible colleges', communicating about their work ahead of publication and building barriers to entry by those outside the colleges (Price, 1963). This reinforces the need to pursue excellence as well as international research cooperation and mobility at the national level, in order to ensure membership of these colleges and maintain a competitive edge.

Figure 10 Citation impact of domestic and internationally produced UK publications 2001 and 2011



Note: citation impact of UK articles and reviews indexed on Thomson Reuters Web of Science™ relative to world average. The citation impact is greater for papers with international co-authors than for papers that have only UK (domestic) authors and the disparity is increasing.
Source: Based on (Adams & Gurney, 2016) based on Web of Science data

3 History of the REF

This Chapter provides a straightforward account of the origins and development of the REF. It charts changes in the way the exercise was performed at each stage and provides and account of feedback to the system from a range of sources. We take account of the links to policy where it is possible to identify them.

Overall, the story is of the introduction of an assessment system that drove part of the allocation of institutional funding for research. By 1992, it determined the allocation of nearly all such funding, so that almost all university research funding has become contested, not only that received from external funders. There have been evolutionary changes in the way the assessments have been done. There has been little policy response affecting the 'REF' to the changed circumstances of massification. The REF2014 is arguably the first major discontinuity in the development of the REF, reducing the focus on scholarly excellence somewhat and taking account of the development of the research environments in the universities and the universities' links with societal needs and impacts.

However, the growing amount of research output overall and its improving quality over time poses a problem: 'quality' rises faster than the total budget that the REF can distribute. The result is a devaluation in money terms of the funding reward per publication – an effect also observed in other PRFS (Bloch & Schneider, 2016) (Good, Vermeulen, Tiefenthaler, & Arnold, 2015). The REF has responded to this problem increasingly by only rewarding performance in the highest grades and by skewing the rewards so that the highest grades generate rewards several times those produced by lower ones.

3.1 The origins of the REF

As indicated above, the UK's research assessment exercises have been introduced against a background of dramatic expansion in the size of the UK higher education system and a resulting transformation in the accessibility of higher education. Overall participation in higher education increased from 3.4% in 1950, to 8.4% in 1970, 19.3% in 1990 and 33% in 2000 (National Committee of Inquiry into Higher Education (Dearing Report), 1997).

At the start of the 1980s, the UK still had a relatively large non-university research institute or government lab sector but during the decade a number of the non-university publicly funded research institutes were privatised, merged or closed (Martin & Whitley, 2010), resulting in some of the changes in GOVERD described earlier and increasing the relative importance of the universities as research performers.

Before the introduction in 1986 of the first Research Selectivity Exercise (RSE), the predecessor to the Research Assessment Exercises (RAE) and the Research Excellence Framework (REF), universities' institutional funding for research was allocated largely on the basis on student numbers and it was assumed that all academics were engaged in research (Lee & Harley, 1998). Broadly, the fifty or so universities were treated as if they were of equal status (the 'equity principle'), although the Universities of Oxford and Cambridge were relatively generously funded compared with the rest (Martin & Whitley, 2010). Demands for transparency required a response, which justified an unequal allocation of funds. The idea of 'selectivity' – in the sense of directing money towards those who produced the highest quality research – was therefore adopted in the 1986 exercise.

In the context of the UK's poor economic performance in preceding years, the Conservative government elected in 1979 imposed significant budget cuts to many parts of government expenditure, including that on the universities. In order to protect the best universities and departments, an exercise done by the University Grants Committee (UGC) in 1982 led to a very uneven reallocation of institutional funding for research, to the extent that some universities were subject to reductions as large as 30 percent. This selective distribution of funding was based on an unclear set of criteria and reductions in funding were largely unexplained, leading to considerable protest (Martin & Whitley, 2010). Both the Jarrett review of efficiency in the universities (Jarrett, 1985) and the 1985 Green Paper (DES, 1985) on the future of

higher education pointed to a need for a more prudent and selective distribution of funding (Martin & Whitley, 2010). It was argued that this selectivity in funding would contribute to maintaining excellence in research (Lee & Harley, 1998).

The debate on selectivity, however, went beyond the RSE. Notwithstanding the equity principle, the University Grants Committee (UGC) had in fact selectively been allocating parts of university block funding to support areas that attracted a lot of research council money or that corresponded to government priorities since 1947. Both the ABRC and the Advisory Committee for Applied Research and Development (ACARD) promoted the idea that research funding should increasingly focus on excellence, with government agencies focusing on the best institutions and the universities themselves allocating internal resources towards the best performing researchers (Adams & Gurney, 2014). In 1987, the Advisory Board of the Research Councils (ABRC) argued that current funding policies would not lead quickly enough to the degree of concentration required to maintain the competitiveness of UK university research. It proposed to classify universities as research universities (R), teaching universities (T) and mixed-mode universities that had research capacity in some but not all areas (X). Only the 15 or so R universities and selected departments in the X universities would then be eligible for research council funding (ABRC, 1987). The idea was roundly rejected by the academic community but helped trigger a decision to be more selective in the 1989 RSE than in the 1986 exercise, in the sense of allocating more of the institutional funding based on quality (Martin & Whitley, 2010). This would have a similar effect to the ABRC proposal in concentrating resources for research on research-performing universities.

3.2 The Research Selectivity Exercise, 1986

Major change	Introduction of an exercise to distribute funding more prudently and selectively

In 1986, responding to the criticisms of lack of transparency in the 1982 exercise, the UGC led by Peter Swinnerton-Dyer launched the first Research Selectivity Exercise (RSE). This allocated about 40% of institutional research funding based on performance but the majority of the funding was still distributed on the basis of student numbers. Universities were asked to submit information on their research income and expenditure, research planning priorities, the five best publications per unit or department from the previous five years and up to four pages of information about their links with industry. Based on this material, 'subject committees' ranked the departments (Jones & Sizer, 1990) and generated a formula that was used to allocate funding in the next few years. HEIs varied in the amount of effort made in their submission and this had some effect on the results (Martin & Whitley, 2010).

It came as a surprise to some that the exercise became part of a more regular evaluation of the research landscape. A consultation on the exercise was run and this received close to 300 responses. The main point made was that research outputs should be evaluated on the basis of a peer review process (Bence & Oppenheim, 2005). The RSE was also criticised on the following grounds (see (Jones & Sizer, 1990).

- The criteria for assessing research quality had not been made clear to universities
- The identity of assessors whose advice had been sought was withheld
- Evaluation of research on the basis of UGC cost centres/university departments had not allowed proper assessment of the work of interdisciplinary research groups and of joint departments
- Information sought from universities had biased judgements in favour of larger departments
- The descriptive terminology in announcing the ratings was confusing; 'below average' had been understood to imply a low absolute standard
- Different assessment standards had been used for different subjects
- The exercise had taken little or no account of work in progress and research potential
- There was no appeals mechanism against individual ratings
- There had been insufficient consultation with subject and professional groups

The 'Committee of Vice-Chancellors and Principals' (the predecessor of Universities UK) argued for an increase in the number of outputs to allow better illustration of strength (Jump, 2013).

3.3 The Research Selectivity Exercise, 1989

Major change	Formalisation of a peer review system that evaluated two 'publications' for each member of staff

The successor to the UGC, the Universities Funding Council (UFC), carried out a second exercise in 1989 with the intention of developing a more selective and systematic approach (Jones & Sizer, 1990). The proportion of institutional funding for research governed by the exercise increased, so close to half of institutional research funding was distributed on the basis of it (Martin & Whitley, 2010).

The amount of information universities had to submit increased substantially and included (Jones & Sizer, 1990)

- Staff returns
- Publication and other output including two publications and other noteworthy publications or public output per member of staff
- Total numbers of publications in relation to the number of full-time staff
- Data on students and studentships
- Data on research grants and contracts
- Statement of research plans
- General observations on other matters including consultancy and external recognitions

No definitions of what qualifies as a publication were imposed. However, unlike under the 1986 RSE, this exercise, where only five publications were required per department/unit, the 1989 RSE demanded quality research from all academic staff rather than from the department as a whole. The exercise was expanded beyond the 37 cost centres that were analysed previously (Jones & Sizer, 1990) and introduced 152 Units of Assessment (UoAs), roughly corresponding to all disciplines and subject areas of departments. One the basis of the material submitted, UoAs were now assessed using a 5-point rating scale (see Table 1).

The increase in the volume of publications to be assessed resulted in a substantial increase in burden on the system. In response to this increase in workload and in order to increase accountability, 70 peer review panels were set up to review the applications. The names of panel members were made public after the completion of the exercise (Jones & Sizer, 1990). The panels were also allowed to seek confidential advice from external experts. According to Jones and Sizer (1990) the exercise was set out as an informed peer review but

- No guidance was given on how to identify and recognize quality
- Publications lists served as a guide but only the two publications per person actually submitted were expected to be read by the panels
- Additional quantified data served to inform peer judgements

In 1988, the UGC commissioned a study to look into the value of bibliometrics. As described in Jones and Sizer (1990, p. 321), this exercise concluded that "bibliometric analysis will never be a substitute for peer review, but, if the analysis is comprehensive and sound, it should inform peer review". As summarised in the analysis by Bence and Oppenheim (2005) the study also raised a number of points of criticism of RSE 1989.

- The 5-point rating scale was not published until very late in the process making it difficult to prepare in advance
- The full list of UoAs was not settled in advance

- Some of the forms were unnecessarily complex and there was need for a more precise definition of 'publications'
- No facility for systematic verification of the accuracy of the submissions was built into the exercise, and there was some evidence of deliberate misreporting in the publications lists, for example by altering publication dates and making exaggerated authorship claims
- The exercise favoured large departments, particularly in the physical sciences
- Science ratings favoured excellence in basic and strategic research to the detriment of applied research
- By assessing all staff in post for any part of the 5-year review period, the exercise was unduly retrospective

Moreover, according to Jones and Sizer (1990) the publication lists were not found to be informative because no distinction was made between refereed and non-refereed papers or books and hence the total number of publications was a poor indicator of quality.

On the topic of costs, a 'representative' sample of eight universities was asked to estimate the costs associated with submitting to the RSE 1989. Total costs to Universities were estimated at £3.8M and an additional £0.3M cost was attributed to the UFC for running the exercise. Grossing up, the total cost was therefore £4.1M and approximately £500M funding was distributed for each of the following 3 years (Jones & Sizer, 1990). Thus, the cost amounted to roughly 0.3% of the funding distributed (which is somewhat lower than was the case for the REF2014).

3.4 The Research Assessment Exercise, 1992

Major change	Increase in the funding distributed on the basis of the RAE results

The Higher Education Funding Councils (HEFCs), for England, Wales, Scotland, and Northern Ireland were established by Act of parliament in 1992[12]. The respective governments set the amount of funding to be distributed and the HEFCs were now were responsible for the distribution (Lee & Harley, 1998).

Following the 1992 Further and Higher Education Act, 35 of the former polytechnics became eligible for research funding. Most of the polytechnic universities had been formed during the 1960's in the expansion of higher education. This new group of universities was granted the power to award academic degrees. Because the former polytechnics had large teaching staffs the funding formula needed to be adjusted to avoid redistributing a substantial amount of funding to these new universities, at the expense of the pre-92 universities. The solution was to increase the distributive power of the next assessment exercise, which was launched in 1992 and was named the Research Assessment Exercise (RAE). Over 90% of the institutional research funds were distributed on the basis of the results of the exercise, so the large number of teaching-oriented faculty in the former polytechnics did not attract much institutional research funding.

RAE 1992 used the same 1-5 scale as the previous exercise but this time a non-linear relationship was established between the result of the research assessment and the funding awarded. No funding was distributed for research output that was judged a 'grade 1' (see Table 1). 'Grade 3' received twice as much funding as 'grade 2', 'grade 4' received three times as much and 'grade 5' four times as much (Otley, 2010).

The purpose of the RAE 1992 was described as to inform funding of basic research but the funding bodies also sought to assess the quality of research in general (HEFCE, 1997).

[12] The funding councils are the Higher Education Funding Council for England (HEFCE), the Scottish Funding Council (SFC), the Higher Education Funding Council for Wales (HEFCW), and the Department for Employment and Learning, Northern Ireland (DEL).

For this exercise, HEIs were allowed actively to select the staff that they would include in the submission, which meant that for the first time they could trade off the number of people they submitted and the perceived quality of the work they presented. The submission required the inclusion of two publications and two other forms of output for research-active staff that were formally employed after the 'census date' and whom the university chose to submit to the exercise. By only allowing the inclusion of staff employed after this census date, the exercise became slightly less retrospective in the sense that only researchers that were still employed at a department could help increase the departments' score. This change also sought to put an end to the 'transfer market' where researchers moved in the year of the census, though it is not evident that it did so. Moreover, UoAs were asked to submit information on the research environment, future research activity and quantitative data on all publications. However, these quantitative data were not extensively used in the evaluation and the focus of the exercise was that of informed peer-review of the submitted research output (HEFCE, 1997).

The RAE 1992 allowed submissions to distinguish between the assessment of pure and applied research in science and engineering in an attempt to balance funding. However, few institutions took up this option (HEFCE, 1997).

A major criticism of the RAE 1992 was that despite the formalisation of the system, uncertainty remained about the exact relationship between assessment results and the funding distributed. There was an increased recognition among the heads of university of the importance of a 'quality submission' but the optimum combination between research volume (partially measured using the number of active research staff) and research quality remained uncertain. There were calls for the system to become more transparent and robust. The following lists some of the concluding comments from the HEFCE study of the impact of the RAE 1992 (HEFCE, 1997).

- The RAE was said to have influenced positively the internal management and strategy of the institutions, generating a planning cycle as well as a broad programme review. HR seems to have become more formal and research focussed
- The process of designating staff as research active is undoubtedly intrusive
- The RAE was associated with stress on researchers
- The extent of staff movement promoted by the RAE had probably been overstated
- Researchers seemed to be developing strategies to get outputs into the public domain within relevant RAE periods, while also giving themselves the opportunity to produce a crafted publication.
- There seemed to be little evidence that the RAE has had a negative impact on interdisciplinary or applied research
- The RAE was probably not driving the research agenda as strongly as the activities of other funders
- There was undoubtedly a general concern that the RAE has had a detrimental effect on teaching activity.

Moreover, the evaluation referred to the notion that the drive towards increased transparency made it more "difficult to conduct assessment exercises without them becoming instruments of policy (or at least being perceived as such). There is then the danger that the process of assessment may change the behaviour being measured." (HEFCE, 1997).

3.5 The Research Assessment Exercise, 1996

Like the previous exercise, the 1996 RAE awarded funding based on the combination of the volume of submitted staff and the quality of output. By the time of the 1996 RAE, the submission was normally managed by a university manager familiar with the procedures for submission. This may have influenced the increase of high-quality quality research submitted.

As before, the main objective of the exercise was to produce ratings of research quality that the funding bodies could use to allocate money for research (Funding Councils, 1997). In the 1996 RAE, several major changes to the exercise were introduced.

- UoAs were invited to submit up to four publications per member of 'research active' staff. It was argued that increasing the number of publication from two to four made it somewhat easier to demonstrate higher levels of excellence achieved (Page, 1997). Some academics found it difficult to select among eligible publications because it was unclear how, for example, joint papers and frequently cited papers were reviewed by the panel and there likewise was uncertainty among panellists
- UoAs were invited to include 'indications of peer esteem' (such as journal editorships and conference presentations)
- The rating scale was expanded to seven grades (1, 2, 3b, 3a, 4, 5, 5*) but no funding was given to departments that were given the lowest two (see Table 1)
- Unlike in the previous exercise, there no longer was a requirement to submit a list of all publications
- Whereas the RAE 1992 distinguished between pure and applied research this practice was discontinued in the RAE 1996 because of the difficulty of distinguishing clearly between the two (Jump, 2013).

As in the previous exercise, the 1996 RAE awarded funding based on a combination of the number of submitted staff and the quality of output. As before, this resulted in universities making judgments about the trade-off between the number of people submitted and the average quality of their work.

As customary, following the exercise a consultation took place and some recommendations for change were fed into the following RAE. The consultation concluded, amongst other things, that

- Not enough time had been allocated for the development of software and this had caused some complications with the submission
- There had been challenges for some departments to run the software on their computer, including difficulties with the font chosen, etc.
- Some panellists made reference to the possible benefits of consulting a Citation Index; but not all disciplines had citation indices
- Concerns were voiced over the development of interdisciplinary research (Funding Councils, 1997)

A long list of recommendations was given but the majority were minor and overall the consensus was that only minimal changes should be introduced for the next exercise (see also the review by Page (1997)).

One of the key questions among panellists was how much of the output they should actually read and what approach they should take to the assessment (Funding Councils, 1997). It was noted that panellists in the arts and humanities tended to read more than those in engineering and medicine. There were problems in getting access to the publications even though HEIs should have had the copies at hand but sometimes there were concerns about whether works were published and, especially in the arts, having access to non-text works such as performances.

The Dearing report (1997) made 93 recommendations for the future of the UK higher education system, including the introduction of higher student fees. In reference to the latest RAE, the report noted that, albeit it had some imperfections, the system had become generally accepted within the community and provided a relatively cost-efficient way to allocate funding for research. Dearing recommended an inquiry into the funding of interdisciplinary research and the introduction of international panel members in the next RAE exercise.

3.6 The Research Assessment Exercise, 2001

Major change	Expectation from panellists to treat each publication or output on its merits

Ahead of the 2001 exercise, the funding bodies reaffirmed that the purpose of the RAE was to produce ratings of research quality, which would be used to determine funding for research. An additional set of principles was presented as follows (HEFCE, 1999).

- Clarity: A large and complex body of regulations and procedures governs the RAE. All written documents and statements about the RAE should therefore be clear and consistent, and misinformation about the RAE should be challenged wherever possible.
- Consistency: Assessments made through the RAE should be consistent, especially across cognate areas and in the calibration of quality ratings against international standards of excellence.
- Continuity: The RAE develops through an evolutionary process, building on and learning from previous exercises. With each successive exercise a balance has to be struck between continuity and development. In general, changes are only made where they can bring demonstrable improvements which outweigh the cost of implementing them.
- Credibility: As was demonstrated through responses to the funding bodies' consultation on the conduct of the RAE, the methodology, format and processes employed in the exercise are credible to those being assessed. It is essential that this credibility be maintained.
- Efficiency: Previous RAEs have been extremely cost-effective given the value of public funds distributed according to their ratings. The cost of the exercise, including the cost to Higher Education Institutions (HEIs), should continue to be the minimum consistent with a robust and defensible process.
- Neutrality: The RAE exists to assess the quality of research in HEIs. It should carry out that function without distorting what it is measuring. In other words, the RAE should not encourage or discourage any particular type of activity or behaviour, other than providing a general stimulus to the improvement of research quality overall.
- Parity: The RAE is concerned only with assessing the quality of research in participating HEIs, regardless of the type, form or place of output of that research.
- Transparency: The credibility of the RAE is reinforced by transparency about the process for making decisions. All decisions and decision-making processes will be explained openly, except where there is a need to preserve confidentiality (for example in panels' discussions, or when dealing with the names of nominees for panel membership or with the strategic research plans of institutions).

Under the 2001 RAE, the rules of the game were only moderately revised. Overall, more emphasis was placed on accountability and robustness.

- Staff members with special circumstances were permitted to submit a lower number of publications and outputs without this being weighted negatively in the score of the UOA
- Panels were expected to treat each publication or output on its merits (rather than eg on the basis of the journal in which the publication was published). This of course increased the burden on panellists. Because of the reading burden, Bence and Oppenheim (2005) found that panellists began to assume that work already subjected to rigorous peer review should be given more weight
- Five groups of panel chairs in related disciplines gathered with the objective to increase consistency in the assessment and rating and to provide a fairer assessment of interdisciplinary work (Jump, 2013)
- The assessment results were made public and some feedback was provided to the Pro Vice-Chancellors (ratings were already published under the previous exercise)

After the announcement of the RAE 2001 results, HEFCE realised that funding distributed on the basis of these ratings would cost £200M more than expected but, because the total volume of funding to be

distributed had already been determined in the government's Comprehensive Spending Review, it was decided that all funding for grades lower than 5* would be cut (Martin & Whitley, 2010)[13]. Moreover, HEFCE decided that departments that had received a 5* rating in the RAE 1996 and the RAE 2001 should receive a new 6* grade. This change resulted in steeper differentials than initially envisaged but prevented the overspend that would otherwise have happened (Martin & Whitley, 2010).

Following the RAE 2001, there was a growing dissatisfaction with the burden that it placed on academics and there was a question as to whether a metrics based system could help reduce that. Roberts (2003) conducted a review of the RAE. Key proposals for change included the following.

- Peer review should stand central to the evaluation of research excellence. Only such a system "will enjoy both the confidence and the consent of the academic community (…) and is sufficiently resistant to unintended behavioural consequences to prevent distorting the very nature of research activity". The recommendation also noted that experts may choose to employ performance indicators to inform their judgment. However, according to Roberts (2003), a metric system was not a viable option to help reduce the burden of the RAE.
- The assessment should be designed in line with the expected rewards
  - A separate approach for the least research intensive institutions, which in essence allows these institutions to opt out of the competition for Research funding in exchange for base level funding.
  - Assessment of the less competitive departments in the remainder of institutions based on indicators, to ensure that they exceed a minimum quality level
  - Expert review assessment similar to the old RAE for the most competitive departments

  This proposal was not adopted.
- The output of the Research Quality Assessment should be a 'quality profile' indicating the quantum of 'one star', 'two star' and 'three star' research in each submission and there should be a continued rating scale.

Roberts considered the possibility of including all staff in a department but concluded "*there is a real risk that institutions would respond with an even more damaging form of games-playing – removing reference to research from the contracts of large numbers of teacher-researchers*" (Roberts, 2003). He recommended that departments should include at least 80% of their staff but this was not taken up by HEFCE.

The report by Roberts builds on a study conducted by Wooding & Grant (2003) in which nine facilitated workshops with 142 academics and research managers examined how research could be assessed. The advantages and disadvantages of four systems were discussed (Wooding & Grant, 2003, p. 20):

1. Algorithms (based entirely on quantitative metrics): A system based purely on an algorithm for combining metrics. Such a system would be automatic – leaving no room for subjective assessment. Various metrics could be included: measures of reputation based on surveys; external research income; bibliometric measures (publications or citations); research student numbers (or completions); and measures of financial sustainability.
2. Expert Review (including Peer Review): A system in which experts make a professional judgement on the performance of individuals or groups, over a specified cycle, and/or their likely performance in the future. The groups could be research groups, departments or consortia. Assessment may be undertaken entirely by peers or may incorporate other experts such as representatives of user groups, lay people and financial experts.
3. Historical Ratings: A system in which ratings of groups/departments/universities are determined entirely by their performance in the past. Research would, in effect, be presumed to be strongest in those departments or institutions with the strongest track record. Various measures could be used

[13] We focus on the steps taken by HEFCE. The other funding councils may have taken different measures.

to determine past performance in addition to RAE rating, such as amount of research funds attracted.

4. Self-Assessment: In this assessment system, institutions, departments or individuals assess themselves. A proportion of the assessments are reviewed in detail. Although the assessment is made internally, external assessors could challenge the self-assessment.

With reference to these alternative systems (and using the expert review system as the starting point), the workshop participants were asked to design their ideal system in groups. Of the 29 systems proposed, 25 were based on expert review, three mixed peer review with other techniques and one was unclear about method, indicating overall that the community consulted overwhelmingly preferred peer review.

3.7 The Research Assessment Exercise, 2008

Major change	Introduction of quality profiles which made it possible to identify and reward 'pockets of high-quality research

Because of a delay in the publication of the assessment criteria, the next RAE was postponed by one year, (resulting in a seven year cycle) allowing for time between the publication of assessment criteria and the assessment (Otley D. , 2010).

The RAE 2008 changed the threshold of quality definition – see Table 1. Whereas in the RAE 2001, a single point rating was awarded for each Unit of Assessment, in the RAE 2008 the results were presented as quality profiles in order to give a sense of the spread of quality within the Unit of Assessment. Each quality profile showed the percentage of the publications submitted that was given 1,2, 3 or 4 stars. The introduction of the quality profile was an important change for it made it possible to identify and reward 'pockets of high-quality research wherever these may be located' (REF, 2008). The RAE 2001 had distributed substantial funding to larger institutions when they received 5* funding for the academics submitted, concentrating research funding in these institutions. Excellent researchers in lower rated departments were unable to secure substantial funding. By distributing funding on the basis of quality profiles, the RAE 2008 made the system slightly more dynamic. In fact, the outcome of the RAE 2008 suggested that high quality research was widely distributed across the HE sector (REF, 2008). Following this change in policy, an additional 25 universities and colleges were successful in gaining Mainstream QR funding (see also Adams and Gurney, (2014)).

There were also some additional changes to the submission process. The full list is shown in the RAE 2008 Guidance on Submissions (RAE, 2005). The RAE 2001 principles were for the most part maintained in the RAE2008 (although no specific mention was made of 'parity', see section 3.6). And, four additional principles were introduced (RAE, 2009).

- Equity. All types of research and all forms of research output shall be assessed on a fair and equal basis.
- Diversity. Submissions to RAE 2008 will reflect the diversity of excellent research conducted across the UK higher education sector, in line with the quality profiles introduced
- Equality. HEIs are encouraged to submit the work of all their excellent researchers, including those whose volume of research output has been limited for reasons covered by equal opportunities guidelines.
- Expert review is central to the RAE, reaffirming the results of the consultation by Sir Gareth Roberts (2003).

Although the RAE was judged as cost-effective, there were still questions about whether the peer review based system was an effective system for the allocation of resources. In fact, the Chancellor of the Exchequer (Gordon Brown) had refuelled this debate already in 2006. In spite of the report by Roberts (2003) and the analysis by Wooding & Grant (2003), the Chancellor of the Exchequer suggested that the next exercise should be replaced with a metrics approach to QR research funding – if the majority of HEIs were in agreement. The government report 'Next Steps' (HM Treasury, 2006) presented evidence

that the correlation between an HEI's QR and Research Council income streams, when measured at an institutional level, is as high as 0.98 on average, suggesting that *if* QR funding were allocated with reference to research council funding there would be little change to institutions' income. The RAE was seen as unjustified in light of this idea, but also because it imposed a burden on HEIs, influenced behaviour in publishing and recruitment and failed to support interdisciplinary research and excellent user-driven research (HM Treasury, 2006).

In response to the Chancellor of Exchequer, Bekhradnia (2006) published a short evaluation of alternatives to the Research Assessment Exercise and showed that because of differences in size, the percentage change in QR received would be substantial if the allocation of QR funding were based on income derived from research grants and contracts. Moreover, the report by HEPI concludes that a metric based system would cost 'very much more' and would likewise influence behaviour (see also the study by Sastry and Bekhradnia (2006)). In his report, Bekhradnia also lists additional downsides to a metrics system. A related study by HEPI also concluded, amongst other things, that a citation-based approach does not measure research quality (Bekhradnia, 2006).

Shortly after, the Department for Education (2006) published its consultation results on the design of a new metric system in which five models for allocating funding on the basis of income and research volume were presented. The consultation produced a range of arguments for and against a metrics based system (see also Barker (2007) for a summary). This consultation highlighted opposition to the reform, confusion around the design options, and concluded that the next exercise would (only) "use metrics as its basis where they are sufficiently robust and fit for purpose, and would help to stimulate development of metrics fit for the purpose of assessment in areas where they are currently less well developed" (pt. 86).

The University and College Union (UCU) likewise produced a post 2008 RAE-consultation (2008) which highlighted resistance to a metric system that would be based on success in obtaining research council grants as well as broad scepticism about the use of bibliometrics, to allocate funding. Moreover, the summary note of this consultation also argued that "the present RAE has already led to an unacceptable concentration of research funding which undermines the very principle of dual support" because the patterns of resource distribution among institutions via institutional and research council channels are very similar.

The Royal Academy of Engineering (2006) also published a report in reaction to the consultation led by the Department of Education. The key recommendations were

1. A shadow metrics exercise should not be carried out in parallel with RAE2008. The results from such an exercise based on any or all of the five metrics based models set out in the consultation would vary considerably from those of RAE2008. This could lead to QR allocations for RAE2008 being challenged.
2. RAE2008 should continue unaltered. Wholesale changes to the RAE system after 2008 would undermine rather than improve the QR funding allocation system. Instead, RAE2008 should be built on in an evolutionary fashion, reinforcing the behaviours in institutions that funders wish to encourage.
3. The use of metrics in the RAE2008 process has increased over previous assessment exercises and, for engineering subjects (Panel G), is currently at an appropriate level.
4. All of the five proposed metrics based models in the consultation paper are flawed and could lead to unforeseen consequences or behaviours in institutions if adopted
5. The principles developed by the Academy in Measuring Excellence in Engineering Research and The Assessment of Research Quality in Engineering Disciplines should continue to be used in developing the RAE for engineering disciplines as well as other applied science STEM subject areas where appropriate.

6. Over reliance on metrics based on institutions' income from Research Councils and other external sources as a measure of research quality has the potential to reward expensive research rather than good research.
7. The administrative cost and burden of the RAE process is not over burdensome as claimed by HM Treasury and delivers QR funding to institutions for lower overheads as a proportion of funding granted than funding through Research Councils.

In summary, prior to the launch of the next exercise it was clear that

- Peer review was seen as "promoting conservatism, risk avoidance and 'silo' research"
- There was an interest in introducing more bibliometrics in some disciplines but a resistance to doing so in others
- More broadly, there was resistance to a metrics based system and to the use of metrics in allocating funding
- Notwithstanding the objections to peer review noted in point 1 above certain university groups (notably the Russell Group universities) would not "relinquish peer review easily." (Barker, 2007)

In preparation for the next exercise, the four funding bodies first commissioned an initial open consultation on the REF. This consultation was informed by a series of studies that were commissioned by HEFCE to scope the development of introducing more metrics. Specifically, the company Evidence) (2007) produced a bibliometric analysis of interdisciplinary use and Leiden University (CSTW, 2007) looked at

- How far bibliometric techniques can be used to produce appropriate indicators of research quality
- What bibliometric-based indicators would best suit HEFCE's purposes, and how these indicators can be produced and used

The outcomes of the consultation were launched in 2008 (HEFCE, 2008) and concluded, amongst other, that there were concerns about the use of metrics, drawing a clear distinction between science-based research and humanities and social sciences. During 2008-09, HEFCE ran a pilot exercise targeting 22 universities to test the potential for using bibliometric indicators of research quality in the REF.

Following this first consultation by the funding bodies, a series of reports were published or commissioned by HEFCE.

- Interim report on the bibliometrics pilot exercise (HEFCE, 1997)
- Report on the pilot exercise to develop bibliometric indicators for the Research Excellence Framework (HEFCE, 1999)
- Analysis of data from the pilot exercise to develop bibliometric indicators for the REF: The effect of using normalised citation scores for particular staff characteristics (HEFCE, 2011)
- Analysis of data from the pilot exercise to develop bibliometric indicators for the REF (HEFCE, 2011)
- The results of the consultation and the concerns of the participating HEIs with the data collection and validation phase of the pilot are summarised in two reports by Technopolis (Technopolis, 2009): results of round one consultation and results of round two consultation.

Towards the end of 2009, the UK funding bodies issued a second consultation that was informed by the results of the bibliometrics pilot and the outcomes of this consultation contributed to the design of the next exercise. A summary of the proposals made during the second round consultation is publicly available (HEFCE, 2009). This second round of consultation included proposals to assess the impact of research and this was seen as a completely new idea. The introduction of impact was seen by some academics as an attempt to convince the Treasury to abandon the idea of switching to a metrics based system (Sayer D. , 2015) (Jump, 2013). However, the idea of establishing impact was not entirely new, it was already presented as part of the white paper 'Realising our Potential' in 1993 (HM Government, 1993). Moreover, in 2011, RCUK also undertook a review of pathways to impact in connection with its Impact Agenda and published a series of case studies (RCUK, 2011).

HEFCE commissioned an Accountability Review of the RAE 2008 (PA Consulting, 2009). This review estimated that the system cost around £47M to the English HEIs, which was 0.5% of the total funds allocated to research in England. This cost was judged to be proportionate to the rewards of undertaking it (REF, 2011). Independently, Sastry and Bekhradnia (2006) estimated that the RAE may have cost up to £100M.

Along with the RAE Accountability review, a review of the RAE 2008 equality and diversity procedures was published (Equality Challenge Unit, 2009) which included several propositions that were carried forward into the REF. And, HEFCE commissioned RAND to undertake an international review of the experience with capturing research impact (Grant, Brutscher, Kirk, Butler, & Wooding, 2010). This study included an overview of the proposal for a Research Quality and Accessibility Framework (RQF) in Australia (that was never implemented), a framework developed by RAND to capture the impact of research grants, the US PART framework (which is a self-evaluation approach used to assess programme performance across the US federal government) and ERiC which is a framework to assess societal quality in the Dutch higher education system. This material fed into a discussion paper on the accountability burden (HEFCE, 2009), which maintained that expert review should remain the sole means of assessing outputs in a number of disciplines. It would also need to be combined with bibliometrics for other disciplines. There also was a suggestion to reduce the number of outputs from four to three (which was not taken up). HEFCE established Expert Advisory Groups, which convened throughout this period to discuss experience of the 2008 Research Assessment Exercise (RAE) and key features of the REF, the use of bibliometrics, expert review of outputs, environment, taking account of impact in the REF, equalities issues, combining the three elements of assessment and REF panels (HEFCE, 2009).

The idea of assessing impact was not received positively by the entire academic community. The University and College Union (UCU) submitted a petition to withdraw the impact proposals, which 17,500 people signed. In particular, the HEFCE's idea to give the assessment of impact a weight of 25 percent in the REF was found problematic (Universities and College Union, 2009), owing to the difficuties of tracjing impact in the shocial sciences and humanities and in basic research in general, the lack of a proven methodology for tracking impact and the difficulties of attribution and evidencing impact. An academic paper used data from the second consultation mentioned above (Smith, Warda, & House, 2011) and concluded that assessing impact could have a positive influence on the relationship between academia, state and industry although the proposed operationalization of the impact agenda was viewed as flawed.

In 2010 the REF team launched a separate Impact pilot exercise with five UoAs and 29 HEIs with the objective to "test and develop the proposed approach to assessing impact in the REF." The findings of this pilot are published in a report by the chairs of the expert panels to the funding bodies (REF 2014, 2010). Feedback from the 29 HEIs on their pilot impact submissions are summarised in a report by Technopolis (2010). Several exemplary impact case studies were made publicly available. The decision on assessing research impact was announced in 2011 (HEFCE, 2011). Impact would count for 20% of universities' total score. There was an intention to increase this weight in future exercises (eg to 25%) (HEFCE, 2011). UCU remained "disappointed that impact criteria have retained a 20% weighting in the 2014 REF" and comment (University and College Union, 2011)

> *"A more limited 'case study' approach is preferable to requiring all academics to engage artificially with the 'impact' agenda (i.e. as is required by the research councils). However, the current 'impact' proposals are still likely to increase the bureaucratic requirements and 'games-playing' potential of the REF. In addition, we suspect that the REF panels will find it very difficult to make valid and reliable distinctions between the various starred levels for assessing 'impact', i.e. 'outstanding', 'very considerable', 'considerable', 'recognised but modest'.*

Several challenges in defining and measuring impact were outlined: time lag, the developmental nature of impact, attribution, knowledge creep, gathering evidence. See Penfield et al. (Penfield, Baker, Scoble, & Wykes, 2014) for a detailed explanation.

3.8 The Research Excellence Framework, 2014

Major changes	• Introduction of separately assessed Environment and Impact dimensions • Introduction of the use of citation information in certain sub-panels to supplement the peer-review process • Introduction of strengthened measures to promote equality and diversity

The Research Assessment Exercise changed its name to Research Excellence Framework in 2014. Societal impact and environment (in the sense of the institutional environment in which research is done) were introduced in addition to the traditional concern with output.

Building on its original mandate, the REF had three stated purposes (REF, 2011).

1. Inform the selective allocation of research funding to HEIs

   Since the introduction of the first exercise, the larger volume of available grant money is distributed to the HEIs with a larger volume of quality research. By rewarding HEIs based on excellence in research, the funding bodies aim to contribute to the development of a dynamic and internationally competitive research sector that makes a major contribution to economic prosperity, national wellbeing and the expansion and dissemination of knowledge (REF, 2011). This skewed distribution of grant money ensures that the research infrastructure at top-level research institutions is well protected.

2. Provide benchmarking information and establish reputational yardsticks.

   The REF has had a powerful reputational effect on HEIs and, as a result, has incentivised HEIs to maintain and/or improve their ranking amongst HEIs. This is because the outcomes of the exercises are publicly available and used by prospective students, industry, charities and other sponsors of research to shape their judgement of the quality of research throughout the UK. The results are also used internally to gain a strategic view on the performance of faculties and departments (see also Farla & Simmonds (2015)).

3. Provide accountability for public investment in research and demonstrate its benefits.

   The last exercise has highlighted the importance of demonstrating impact in relation to research excellence by the introduction of the Impact case studies. However, it should be noted that HEIs are not held accountable as to how they distribute the funding and whether excellent research, within the HEI, is rewarded.

Numerous changes were introduced in the REF 2014 and these changes and their estimated impact on cost are listed in Figure 11. The figure shows the survey responses of REF research managers and/or pro Vice-Chancellors that were responsible for organising the submission to the REF. These survey respondents were selected based on a representative sample of 20 UK HEIs. Notable changes include the following (REF, 2011) .

- The criterion of societal Impact was introduced into the REF 2014 and this is perhaps the single biggest change in the scope of the exercise since its inception. Impact assessment was based on expert review of case studies that *describe social, economic or cultural impact or benefit beyond academia*. Only impact realised in relation to underpinning research that was of 2-star or higher quality was eligible. Impact was given a weight of 20%, whereas research 'Output' had a weight of 65%, and 'Environment' 15%, applied to all UOAs (REF, 2011). Because it was the first time that such exercise has been done, HEIs had to master the process, and collect the relevant data retrospectively. What was understood by impact and how impact should be measured was unclear (Penfield, Baker, Scoble, & Wykes, 2014) and several REF managers went to workshops (eg organised by HEFCE) to learn more about the Impact agenda to help the submission process along (Farla & Simmonds, 2015). A report by RAND estimated that it cost HEIs £55m to prepare the impact submissions (2015) (Manville, et al., 2015). A separate study by RAND evaluates the Impact assessment process (Manville, et al., 2015).

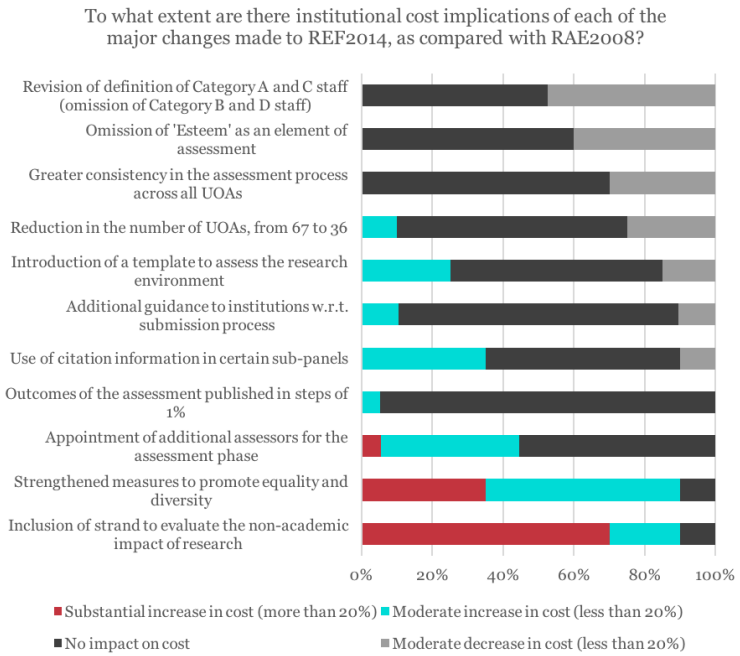
- Use of citation information in certain sub-panels to supplement the peer-review process and provide additional information about the academic significance of research outputs. Some research managers argued that this additional material was useful and others argued that this formed an additional obligation (Farla & Simmonds, 2015).
- The four funding bodies set up an Equality and Diversity Advisory Group to advise on the development of guidance on equality and diversity issues and on special circumstances. This advice was reflected in new measures and codes of practice to promote equality and diversity, increased the cost of submission. Nevertheless, (some) research managers and pro-vice chancellors welcomed the improved equality and diversity agenda (Farla & Simmonds, 2015).

Specific changes to reduce the burden were also introduced.

- Introduction of a template to support the presentation of the research environment and use of Higher Education Statistics Agency (HESA) data in the Environment statement
- Omission of 'Esteem' as an element of the submission process and assessment
- Reduction in the number of UOAs, from 67 to 36

It was expected that some of the reductions in burden would offset the extra work of preparing impact statements and that the costs to funding ratio of the REF would be similar in proportion and reasonable' (REF, 2011).

Figure 11 Changes from the RAE 2008 to the REF 2014



Source: Adapted from (Farla & Simmonds, 2015)

The REF 2014 Accountability Review (Farla & Simmonds, 2015) presents the estimated costs, benefits and burden for HEIs of submitting to the REF. The report sought to establish whether and where there has been any significant change compared with RAE 2008, as detailed in the report of the RAE Accountability Review (PA Consulting, 2009). The review entailed a comprehensive analysis of institutional costs for a sample population of 20 UK HEIs, covering all staff-related costs and related expenditure, both centrally and at the unit of assessment (UOA) level. The total cost to the UK of running REF 2014 is estimated at £246M, out of which £212M is the estimated cost to the UK HE community. This £212M yields a cost per submitted researcher of around £4K, which equates to close to 1% of the submitted researchers' basic salary plus on-costs over a six-year period. It was concluded that the major

cost driver was the REF element on research outputs, which included time spent reviewing and negotiating the selection of staff and publication. The REF 2014 Accountability Review also presents numerous benefits of the REF to the HE community (some of these benefits are already referred to above). However, the increase in estimated costs of the peer review system, as presented in the review, relative to the costs presented in the RAE2008 accountability Review have provided further ground for criticism of the system and of the burden it imposes on researchers.

The accountability review estimated that the £246M total cost for REF 2014 was less than 1% of the value of universities' total research income during the period covered or some 2.4% of the funds distributed via research assessment in the same period. These figures compare with up to 6% for the research councils. However, the relevant cost comparison is with having no REF or using a less expensive approach, such as metrics. At the time of writing, however, while international opinion is that metrics-based alternatives are less expensive, there are few data available to corroborate this idea (Arnold & Mahieu, 2015).

Previous to the release of the REF 2014 Accountability Review, the University and College Union conducted a survey of the experience of 7,000 academics on a range of topics (University and College Union, 2013).

- Transparency in REF selection processes
- Appeals
- Reduced Output Requests and Individual Circumstances
- Equality Principles and potential discrimination
- REF and Performance Expectations
- Impact on career development and terms and conditions
- Variation in Institutional Approaches to REF Performance and Career Detriment
- Probation and Early Career Researchers
- Workload

Overall, the survey results suggest that the REF has a detrimental impact on academics' working conditions and career development. Moreover, the majority of respondents regard the REF as a poor indicator of research excellence and are in favour of a different system for assessing the quality of research (University and College Union, 2013) (UCU, 2008). No alternative system was proposed as part of this report. Around 26% of the respondents are in favour of abolishing it (University and College Union, 2013).

Several studies have been commissioned that use REF impact statements to explore the value of the research on which the REF impact cases are based. One of these studies produced an overview of 6,679 impact case studies that were submitted to the REF 2014. In line with the requirements of submission, each of the case studies aimed to showcase "how research undertaken in the past 20 years has benefited society beyond academia – whether in the UK or globally" (King's College London and Digital Science, 2015). The headline findings are listed below.

- The societal impact of research from UK Higher Education Institutions is considerable, diverse and fascinating
- The research underpinning societal impacts is multidisciplinary, and the social benefit arising from research is multi-impactful
- Different types of Higher Education Institutions specialise in different types of impact
- UK Higher Education Institutions have a global impact
- The quantitative evidence supporting claims for impact was diverse and inconsistent, suggesting that the development of robust impact metrics is unlikely
- The impact case studies provide a rich resource for analysis, but the information is collected for assessment purposes and may need to be aligned for analysis purposes

Another study commissioned by HEFCE looked at the characteristics of departments/units that scored highly in terms of Output and Impact (RAND, 2015). This analysis yields a perhaps not too surprising

set of observations, eg: "High-performing research units receive more income per researcher than the average research unit (2015, p. 6).

The Metric Tide (Wilsdon, 2015) analyses the use and possible limitations of metrics and indicators for research evaluation and concludes, amongst other things, that "within the REF, it is not currently feasible to assess the quality of UoAs using quantitative indicators alone" and advocates the responsible use of metrics. The study has two supplementary reports.

- The Metric Tide: Correlation analysis of REF2014 scores and metrics (HEFCE, 2015) and this study provides further evidence that individual metrics can give significantly different outcomes from the REF peer review process
- The Metric Tide: Literature review (Wouters, et al., 2015) provides an overview of the findings of academic research about a range of indicators that may be useful in research evaluations

Table 1 summarises the main elements of the submissions made by the universities across the period described and how rewards have successively been withdrawn from lower grades. Table 2 shows how the numbers of Units of Assessment have changed over time and summarises the main elements of the submissions required of the universities.

Table 1 Rating scale: thresholds for quality decisions

RSE 1986	RSE 1989	RAE 1992	RAE 1996	RAE 2001	RAE 2001 (HEFCE)	RAE 2008	RAE 2008 From 2012/13 (HEFCE)	REF 2014
Outstanding	5	5	5*	5*	6*	4*	4*	4*
Above average	4	4	5	5	5*	3*	3*	3*
Average	3	3	4	4	5	2*	2*	2*
Below average	2	2	3a	3a	4	1*	1*	1*
	1	1	3b	3b	3a	Unclassified	Unclassified	Unclassified
			2	2	3b			
			1	1	2			

Note: in yellow, the quality scales where funding was allocated

Table 2 Main elements of the submissions and number of Units of Assessment (UoA)

	Main elements of the submission	No. of UoA	Key source
RSE 1986	• Research income and expenditure • Research planning priorities • Five best publication and other output from the previous five years of the unit/department • Descriptions of links to industry	37 (cost centres)	(Bence & Oppenheim, 2005) (Martin & Whitley, 2010)
RSE 1989	• Staff returns • Publication and other output including two publications and other noteworthy publications or public output per member of staff • Total numbers of publications in relation to the number of full-time staff • Data on students and studentships • Data on research grants and contracts • Statement of research plans • General observations on other matters including consultancies and external recognitions	152	(Jones & Sizer, 1990)
RAE 1992	• Overall staff summary • Active research staff return • Two publications and two other forms of output for only the research active staff and other output and summary of publications and other output • Students and studentships • External research income	72	(Universities Funding Council, 1992)

	Main elements of the submission	No. of UoA	Key source
RAE 1996	• Overall staff summary • Research active staff details • Publications and other public output: up to four outputs per selected staff • Research students • Research studentships • External research income • Research environment and plans • General observations and additional information (indicators of research excellence and indicators of peer esteem)	69	(RAE, 1996)
RAE 2001	• Staff details: individual selected as research active and data on other academic staff, postdoctoral research assistants, postgraduate research assistants, other staff • Research outputs: up to four research outputs produced by each member of selected staff • Research students and studentships • Research income • Textual commentary	69	(RAE, 2001)
RAE 2008	• Staff details (research active staff selected) and individual staff circumstances • Research output: up to four research outputs produced by each member of submitted staff (minimum 50%) • Description on research environment and data on research students, research studentships external research income (minimum 5%) and esteem indicators (minimum 5%)	67	(RAE, 2005)
REF 2014	• Staff details (research active staff selected), including individual staff circumstances • Research outputs: up to four research outputs produced by each member of submitted staff (65%) • Impact template and case studies, underpinned by research excellence (20%) • Environment: data on research doctoral degrees awarded, the amounts and sources of external research income and research income-in-kind (15%)	36	(REF, 2011)

4 Effectiveness of the REF

In its original form, the RSE was introduced in order to make the allocation of institutional research funding more selective and transparent to the research community by rewarding quality using publicly-available rules. The transformation of the polytechnics into universities in the early 1990s forced the issue of the affordability of a model where all universities were expected to be research universities. Focusing research resources on excellent research ('selectivity') drove the concentration of resources into a limited number of institutions (Martin & Whitley, 2010). Once achieved, this concentration tended to be self-perpetuating.

4.1 Purposes of the REF

The REF currently has three purposes (REF, 2011).

1. Inform the selective allocation of research funding to HEIs, which involves
   i) Rewarding HEIs based on excellence in research
   ii) Protecting the research infrastructure at top-level research institutes and rewarding 'pockets' of high-quality research. The RAE 2008 changed the quality definition so that the results were presented as quality profiles for each department. This change made it possible to identify and reward 'pockets of high-quality research wherever these may be located' (REF, 2008), allowing funding to be distributed more widely amongst HEIs.
   iii) Contributing to the development of a dynamic research sector
   iv) Contributing to the development of an internationally competitive research sector
2. Provide benchmarking information and establish reputational yardsticks, allowing, amongst others, industry, charities and other sponsors of research to shape their judgement of the quality of research throughout the UK, and potentially leverage additional funding.
3. Provide accountability for public investment in research and demonstrate its benefits

This section focuses on the first two purposes.

4.2 Funding via the REF

Under the dual support system, the UK Higher Education Funding Councils (HEFCs) provide block grant funding for research to HEIs and the UK Research Councils provide grants for specific projects and programmes. The HEFCs are:

- Higher Education Funding Council for England (HEFCE)
- Scottish Higher Education Funding Council (SFC)
- Higher Education Funding Council for Wales (HEFCW)
- Department for Employment and Learning Northern Ireland (DEL)

Block grant funding for research (QR funding) is distributed based on the RAE/REF.

As illustrated in Table 3, the Higher Education Funding Council for England (HEFCE) provides the most research funding as it has the highest number of HEIs. In 2016/17 The volume of QR related funding awarded by HEFCE was £1.578m. This is about 82% of the total HEFCs' expenditure.

Table 3 Overview of annual allocation of QR funding by UK region

Region	Funding Council	Allocation of quality related research funding	Percentage of total quality related research funding
England	HEFCE	£1,578m	82%
Scotland	SFC	£232m	12%
Wales	HEFCW	£71m	4%
Northern Ireland	DEL	£49m	3%

*Data for HEFCE, SFC, and HEFCW are for 2016/17, data for DEL is for 2014/15: HEFCE: http://www.hefce.ac.uk/pubs/Year/2016/201631/, SFC: http://www.sfc.ac.uk/communications/Announcements/2016/SFCAN082016.aspx, HEFCW: https://www.hefcw.ac.uk/policy_areas/research/funding_research.aspx, DEL: https://www.economy-ni.gov.uk/articles/higher-education-quality-related-research-qr-funding

Aside from the difference in the volume of funding allocated between the regions, there are also differences in the *way* QR funding is allocated. HEFCE's allocation of mainstream QR funding takes into account[14]

- **Volume of research**, based on numbers of submitted research-active staff
- **Subject cost weights** reflecting, for example, the fact that laboratory-based research is more expensive than library-based research
- **Quality of research** as measured in the REF
- **London weighting**

'London weighting' is a bonus paid to universities with London post codes and is intended to compensate them for the higher costs of operating in London compared with other parts of the UK. It was introduced in 2002/3 (Table 31) and is delivered as a percentage of the funding univerities obtain based on their quality rating in the REF. For consistency and comparability over time, we retain the original definition of 'mainstream QR' as funding driven directly by the RAE/REF in our analyses below. We show London weighting as a separate component or include it in 'non-mainstream QR'.

HEFCE includes in its calculation of the volume of research the number of research active staff and research students, research assistants, research fellows, and the amount of charitable income submitted[15]. There are also differences in, for example, the way HEIs receive supplementary funding (aside from the Mainstream QR component) to help meet the full economic cost of research funded by charities[16]. SFC has included in its calculation alongside research active staff, research assistants, postgraduate research students, and research grant and contract income[17].

DEL has included research fellows and research assistants in the calculation of volume of research[18] and in 2011/12 DEL allocated some funding held back from the QR funding to create a separate fund to encourage increased participating in the EU framework programme[19].

[14] http://www.hefce.ac.uk/rsrch/funding/mainstream/ See also the guide on how HEFCE allocates funding: http://www.hefce.ac.uk/media/HEFCE,2014/Content/Pubs/2016/201607/HEFCE2016_07.pdf

[15] https://www.hefcw.ac.uk/documents/policy_areas/research/Explanation%20of%20QR%20Funding%20Method.pdf

[16] https://www.hefcw.ac.uk/documents/policy_areas/research/Explanation%20of%20QR%20Funding%20Method.pdf

[17] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/181652/bis-13-545-dual-funding-structure-for-research-in-the-uk-research-council-and-funding-council-allocation-methods-and-the_pathways-to-impact-of-uk-academics.pdf

[18] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/181652/bis-13-545-dual-funding-structure-for-research-in-the-uk-research-council-and-funding-council-allocation-methods-and-the_pathways-to-impact-of-uk-academics.pdf

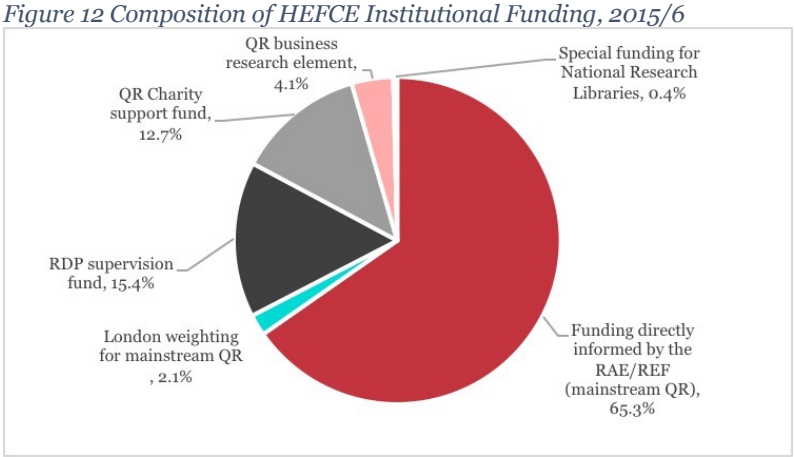
[19] https://www.economy-ni.gov.uk/articles/higher-education-quality-related-research-qr-funding

The incompatibilities among the ways the different HEFCs translate the results of the REF into funding means that they cannot be combined into a single analysis. We have therefore focused on HEFCE's funding, which accounts for over four fifths of the total, in order to be able clearly to describe the principal funding mechanisms and their effects.

The institutional funding that HEFCE provides is is composed of several elements. (Table 31 in the Appendix shows this in detail for the period 1994/5 to 2015/6.) The biggest element is 'mainstream QR', which is the money directly allocated through the REF. The amount of mainstream QR allocated to any Unit of Assessment depends on the number of researchers the university submits to the exercise, the relevant panel's judgement about the quality of the publications submitted and the weights HEFCE applies to different subjects (reflecting the fact that 'hard' science subjects tend to need much more expensive facilities than the 'soft' sciences or the humanities).

In 2015/6, total institutional research funding from HEFCE consisted of the elements shown in Figure 12[20]. Mainstream QR comprised some 65% of the total. Since 2002/3, universities in London benefit from a 'London weighting', intended to compensate for the higher costs of being located in London and which is based on the amount of mainstream QR at the Unit of Assessment level that HEIs located in London receive. London weighting is calculated as 12% for inner London and 8% for outer London of Mainstream quality-related (QR) research funding.[21] The RDP (research degree programme) supervision fund has also in various guises been present since 2002/3, replacing earlier institutional arrangements for funding PhD supervision. In 2015/16 RDP is allocated to universities based on the number of (London-weighted) full-time equivalent eligible students, their subject weights and the London-weighted quality rating of their Unit of Assessment.[22] The Charity Support Fund (from 2006/7) and the so-called QR business research element (from 2007/8) are intended to encourage universities to undertake research funded by business and charities by providing indirect costs to help meet the full economic costs of the work. The amounts individual universities get from these funds depends on the amount of charity and industry funding they obtain[23] [24].

Figure 12 Composition of HEFCE Institutional Funding, 2015/6



Source: HEFCE data

The stable component in the mix is mainstream QR in the narrow sense of the funding most directly informed by the REF (shown as 65.3% of total institutional funding for research in Figure 13). It is

[20] The total recurrent grant in 2015/16 included research funding, knowledge exchange funding and teaching funding.

[21] http://www.hefce.ac.uk/rsrch/funding/mainstream/

[22] http://www.hefce.ac.uk/rsrch/funding/resdegree/

[23] http://www.hefce.ac.uk/rsrch/funding/charity/

[24] http://www.hefce.ac.uk/rsrch/funding/business/

similarly defined across the whole period for which we have funding data[25] and we therefore focus on this type of funding in those of the following analyses that explore changes in allocation to various universities over time.

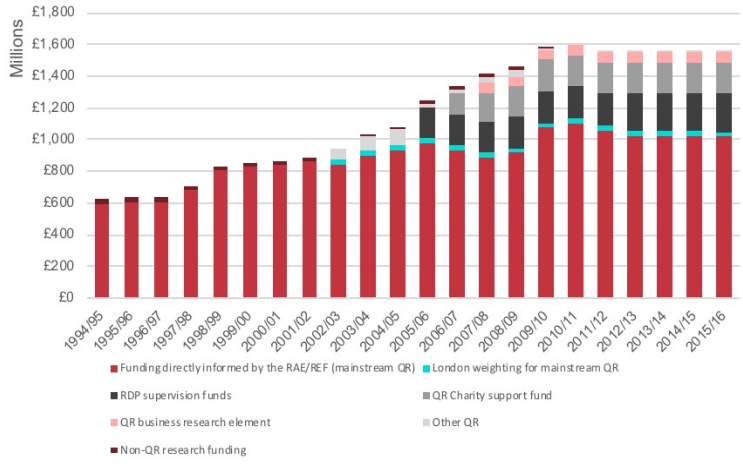
We also present the analysis using the sum of total QR, acknowledging that 'non-mainstream' QR is also influenced by the RAE/REF results. At various points, HEFCE has provided extra quality-based funding for short periods. For example, between 2006/7 and 2008/9, it awarded extra money to the 'best 5-star departments' (Table 31). We have chosen not separately to take account of these in our calculations at the institutional level because these sums are a relatively small components of the total QR funding but this funding is likewise included in the total sum of QR funding.

Figure 12 shows the increase in institutional research funding allocated by HEFCE between academic years 1994/5 and 2015/16 at current prices. (Neither HEFCE nor we have been able to locate data by institution for earlier years, which predate the creation of HEFCE itself.) The Figure shows narrowly-defined 'mainstream' QR, the RDP and then all the remaining components of HEFCE institutional funding for research in a separate block.

The share of HEFCE's institutional research funding budget 'directly informed by the RAE/REF' ('mainstream QR', narrowly defined) has changed considerably. In the academic year 1994/5 it was 94% of institutional funding; in 2001/02, it accounted for 98%; in 2009/10 it accounted for 68%; and in the academic year 2015/16 it provided 65% of the total (£1,017m versus a total of around £1,559m). Taking account of London weighting and the RDP supervision fund, the REF assessment drove 83% of institutional funding in 2015/6.

Narrowly-defined mainstream QR has remained largely stable since 2003/04 at current prices and therefore has been declining in real terms. Since 2004/05, the increase in total institutional research funding allocated to English HEIs is entirely to be attributed to other components of the research funding budget (Figure 13).

Figure 13 Institutional Research Funding Distributed by HEFCE (£m)



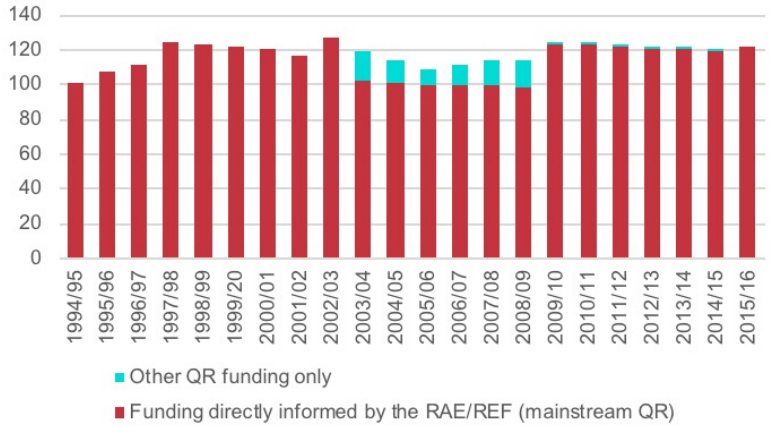
Source: HECFE data (see Table 31 and Table 28)

The total number of universities that receive mainstream QR funding is large (Figure 14). While all HEIs are eligible, the number of HEIs located in England that were actually awarded QR funding more directly informed by the RAE/REF ('mainstream' QR) decreased from 124 HEIs in 1997-98 to 98 in 2008/09.

[25] From 1994/95 – 2001/02 this funding was simply referred to as 'QR' funding and from 2002/03 onwards it was referred to as 'Mainstream QR' Based on our understanding of the data available, until 2001/02 total QR is closest related to what was later referred to as 'mainstream 'QR' as there we no additional pots of quality related research funding, see also Table 31 and Table 33.

Following the introduction of quality profiles in RAE 2008, an additional 25 HEIs were awarded this QR funding. From 2009/10 onwards the number of HEIs receiving this QR funding has remained relatively stable and QR funding has been distributed amongst roughly 122 HEIs. The introduction of new criteria in REF2014 does not seem to have had any effect on the number of institutions receiving QR funding. From 2002/03 – 2014/15 some HEIs did not receive QR funding more directly informed by the RAE/REF (mainstream QR) but did secure other types of QR funding. For example, in 2008/09 (under the RAE 2001) the University of Chester only received QR funding from the business research element and QR Charity support fund. In 2009/10 (under the RAE 2008) it also received QR funding more directly informed by the RAE/REF (mainstream QR).[26]

Figure 14 Number of HEIs Awarded QR Funding by HEFCE

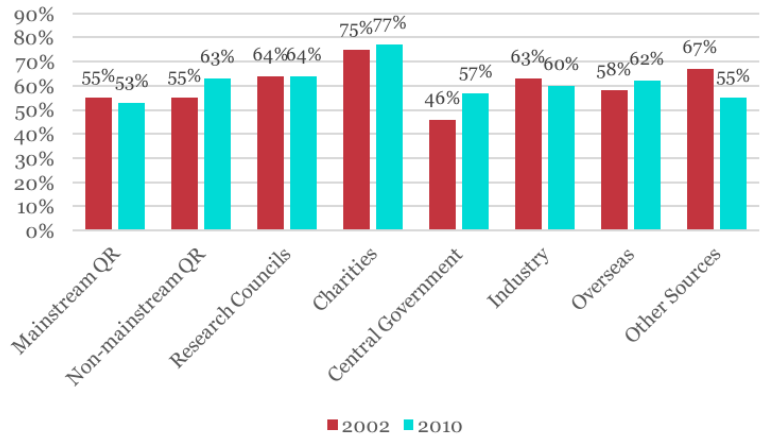


Source: HEFCE data

As illustrated in Figure 15, in 2002, 55% of mainstream QR funding was awarded to the top 10% recipients of QR funding among UK HEIs. In 2010 this proportion was little changed, at 53%. This shows that mainstream QR funding is highly concentrated. However, in 2010, all the other sources of funding, including non-mainstream QR, were even more concentrated.

[26] Data for 2002/03 – 2004/05 is based on the data from the preliminary funding announcement of funding, all other data is based on final allocations made. Unlike the preliminary funding announcement data, 2002/03 – 2004/05 data on final allocations does not include an overview of the broader budgetary envelope. We have not corrected for any discrepancy in the changes of 'mainstream' QR funding announced (in March) and received in the next year.

Figure 15 Proportion of funding from various sources going to the top 10% of beneficiaries for each source, 2002-2010



Source: Based on analysis from (Hughes A. , Kitson, Bullock, & Milner, 2013), using data from HESA Financial statistics

Table 4 shows how much of the total QR funding was allocated to each of four groups of English universities: the Top Ten in terms of the total QR funding they received in 2015/6 (and which are a sub-set of the English Russell Group); the English part of the Russell Group; the English post-1992 universities; and the remaining English universities. The Top-Ten universities consistently take about half the QR funding while the Russell Group as a whole take two thirds or more. The English Post 92 Universities receive about 5%-7% of the funding. This group's share increased following the introduction of 'profiles' in RAE2008.

Table 5 shows how much of the narrowly-defined mainstream QR funding, which is more directly informed by the RAE/REF, was allocated to each of four groups of English universities. Under the RAE and the REF, the share of 'mainstream QR' funding allocated to the Top-Ten universities and Russell Group Universities is slightly lower than the share of total QR funding allocated to the Top-Ten universities and Russell Group Universities, suggesting that these universities are benefitting relatively more from the other post of funding. The English Post 92 Universities received, on average, 8% of the funding, up from 5% under the RAE 2001 and following the introduction of 'profiles' in RAE2008.

Table 4 Percentage of total QR funding allocated, by Top Ten, English Russell Group Universities and English Post-92 Universities.

	RAE 1996	RAE 2001	RAE 2008	REF 2014
Funding year	**1997/98**	**2002/03**	**2009/10**	**2015/16**
Top Ten Universities	47%	50%	44%	52%
English Russell Group Universities	67%	70%	69%	71%
English Post-92 Universities	5%	6%	7%	7%
Remaining English Universities	28%	25%	24%	22%

*The group of Top Ten Universities are selected based on the total QR funding received in 2015/16. *The group of Post-92 Universities include all English HEIs that were given university status in 1992. Data for 1996/97, 1997/98, 2001/02, and 2002/03 includes funding allocated to both UMIST and Manchester University as part of the Russell Group. Source: Analysis by Technopolis data from HEFCE.

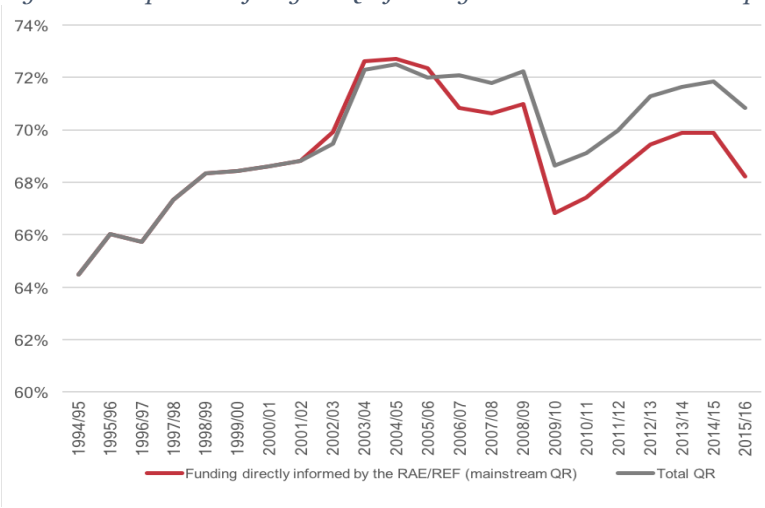
Table 5 Percentage of QR funding more directly informed by the RAE/REF allocated, by Top Ten, English Russell Group Universities and English Post-92 Universities.

	RAE 1996	RAE 2001	RAE 2008	REF 2014
Funding year	**1997/98**	**2002/03**	**2009/10**	**2015/16**
Top Ten Universities	47%	50%	42%	48%
English Russell Group Universities	67%	70%	67%	68%
English Post-92 Universities	5%	5%	8%	9%
Remaining English Universities	28%	25%	25%	23%

*The group of Top Ten Universities are selected based on the total mainstream total QR funding received in 2015/16. The selection is equal to that used for Table 4. *The group of Post-92 Universities include all English HEIs that were given university status in 1992. Data for 1996/97, 1997/98, 2001/02, and 2002/03 includes funding allocated to both UMIST and Manchester University as part of the Russell Group. Source: Analysis by Technopolis data from HEFCE.

During the period 1994/95 – 2015/16, the English Russell Group Universities received on average 69% [64% - 73%] of the 'mainstream' QR funding and 70% of total QR funding. The highest proportion was under the RAE 2001 but the Group's share has remained consistently high across all the exercises for which we have data (Figure 16). In the years 2003/04 – 2008/09 a separate pot of QR funding was allocated to the 'best 5-star' departments (£20.1m-24.6m). HEFCE data show that, on average, 84% of this pot was allocated to the English Russell Group Universities. Moreover, from 2006/07 onwards HEIs were allocated QR charity support funding (£135.7m-£197.8m) and the English Russell Group Universities were awarded, on average, about 82% of this type of funding (see also Table 34).

Figure 16 Proportion of English QR funding awarded to Russell Group Universities, 1999/2000 – 2014/5



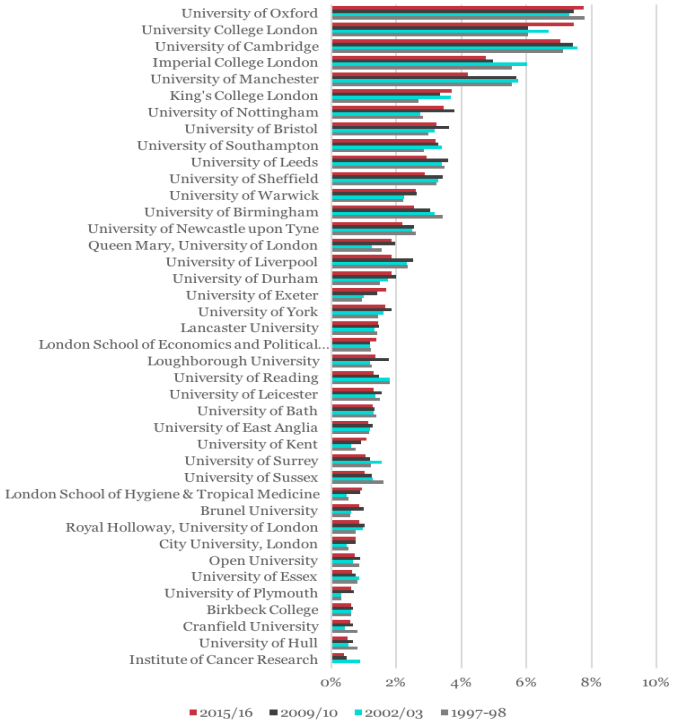
Source: HEFCE data

Figure 17 presents the sum of QR funding awarded to English universities in the last four exercises. It shows only those 40 universities that received at least 0.05% of the total QR distributed in England in 2015/16. Some degree of volatility is visible between exercises. Figure 18 shows graphically the percentage of narrowly-defined mainstream QR funding awarded to English universities in the last four exercises. It shows only the same 40 universities that received at least 0.05% of the total QR distributed in England in 2015/6. As before, some degree of volatility is visible between exercises. The difference between QR and 'mainstream' QR' is substantial for some of the top universities: eg in 2015/16 Oxford university received an additional £60.0m of 'non-mainstream' QR and University College London received an additional £55.7m of 'non-mainstream QR'.

While there are winners and losers in each exercise, the shape of the cumulative concentration curve has remained rather similar (Figure 19 and Figure 20), confirming that the RAE/REF concentrates funding[27]. When looking at the QR funding more directly informed by the RAE/REF allocated under the RAE 2008 and REF 2014, there was a slight reduction in concentration (Gini coefficients of 0.70 and 0.71 respectively).

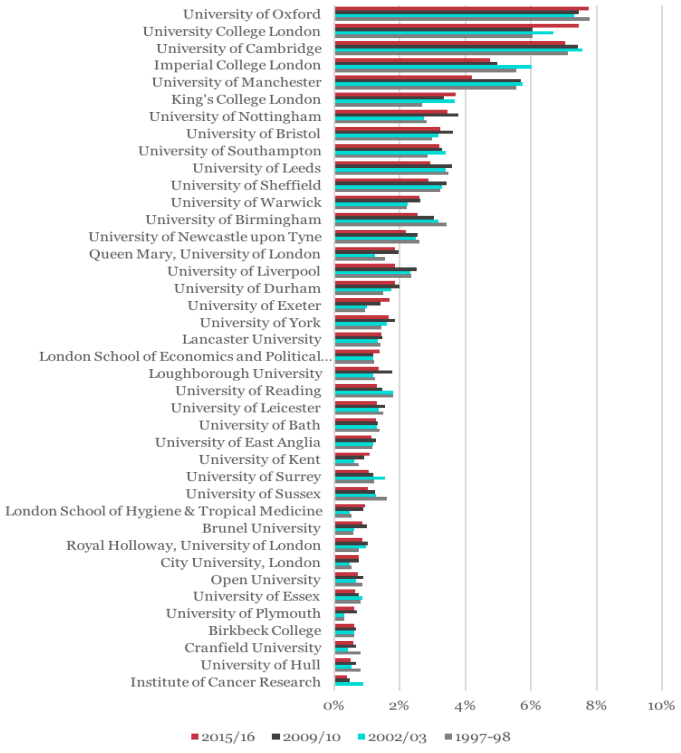
[27] After adjusting for mergers of HEIs, we assume that for each exercise 124 HEIs made a submission. For years where eg only 122 HEIs made a submission we assume an additional two HEIs made a submission but received no funding. This data treatment facilitates presentation of the results but has no meaningful influence on the findings.

Figure 17 Percentage of total QR funding awarded to those universities receiving at least 0.05%o of the English total QR in 2015/16



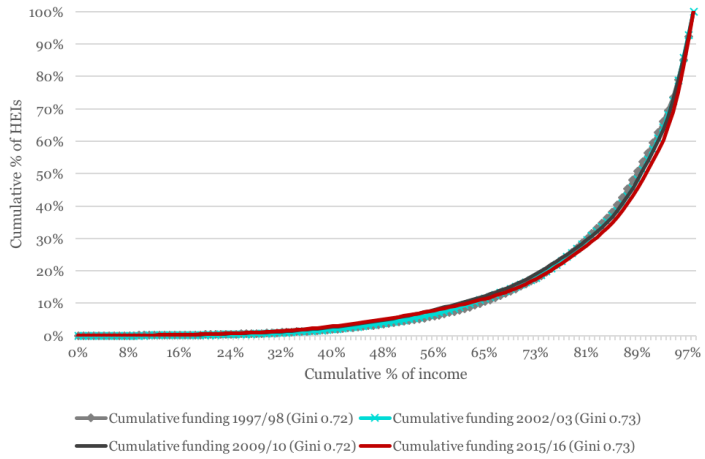
Source: HEFCE data

Figure 18 Percentage of 'mainstream' QR funding awarded to those universities receiving at least 0.05%o of the English total QR in 2015/16



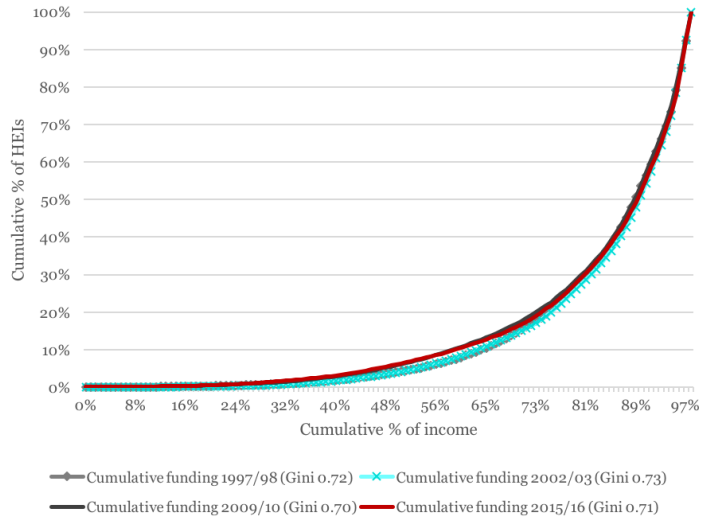
Source: HEFCE data

Figure 19 Distribution of total English QR funding allocated in the last four exercises (Lorenz curve)



Source: HEFCE data

Figure 20 Distribution of English Mainstream QR funding more directly informed by the RAE/REF allocated in the last four exercises (Lorenz curve)



Source: HEFCE data

4.3 Does the REF induce dynamism into the funding system?

In principle, a quality incentive such as QR funding should induce competitive behaviour and hence variations in the way the incentive is distributed. It would be desirable to compare the perturbations in the funding system in the early period of the RSE and RAE with the most recent decade, both to test whether the perturbations have indeed been large and to find out whether the system has tended to settle down into a more stable pattern. Unfortunately, data series going back to the 1980s are not available by institution, so our analysis effectively starts with the RAE of 1996.

Table 6 shows the proportion of total QR that was 'redistributed' (ie gains plus losses at individual institutional level) in the last four exercises. It compares funding in the year immediately after the REF/RAE exercise (ie the first year in which that exercise drives the distribution of funding) with that

in the previous year. The proportions of the total that are reallocated range from 8.6% to 18.6%[28]. Perhaps surprisingly, given the novel elements introduced in the REF, the amount of money reallocated by the REF was less than that in the RAE 2008. The table also shows the percentage of funds reallocated, net of changes in the volume of the total pot of QR funding allocated.

In the RAEs, overall budget increases mitigated the extent of reallocation but there was no budget increase associated with the REF so a greater degree of reallocation was associated with the REF than with its predecessors. The net effect of reallocations 2001/02-2002/03 was as high as 20.6% and the degree of reallocation is in part a result of the creation of additional pots of QR funding in 2002/03, which were allocated based on a different set of criteria (eg the degree of charity funding received).

Table 6 Proportion of total QR funding reallocated following exercises in 1996, 2001, 2008 and 2014 (£millions)

	RAE 1996	RAE 2001	RAE 2008	REF 2014
	1997/98	2002/03	2009/10	2015/16
Total funding awarded in first year post RAE / REF	£684	£940	£1,571	£1,559
Change in volume of total QR funding allocated, compared with last year pre RAE / REF	£82	£72	£134	£1
Average QR funding awarded in first year	£6	£8	£13	£13
Sum of funding reallocated between HEIs in first year	**£112**	**£107**	**£199**	**£134**
Percentage reallocated without taking into account any change in funding allocated	18.6%	12.3%	13.9%	8.6%
Sum of funding reallocated between HEIs, net of any change in total funding	£30	£178	£65	£134
Net effect - Percentage reallocated taking into account the change in funding allocated	4.9%	4.0%	4.5%	8.6%
Number of HEIs awarded QR funding	124	127	124	122

Reallocation is defined as the sum of the gains and losses by individual institutions, compared with the previous year. Funding volumes are rounded to the nearest million. Source: Analysis by Technopolis, data from HEFCE.

Table 7 shows the proportion of the narrowly-defined mainstream QR funding pot that was 'redistributed' (ie gains plus losses at individual institutional level) in the last four exercises. The proportions of the total that are reallocated are considerable, ranging from 10.8% to 20.4%. The Table also shows the percentage of funds reallocated, net of changes in the volume of the total pot of mainstream QR funding allocated.

[28] The analysis controls for mergers or joint submission that impact QR funding allocated in any given year. In the transition from 2001/02 to 2002/03 the following changes took place that are accounted for (aggregated) in the data analysis: The Courtauld Institute of Art made an independent submission in 2002/03. In the previous year, its submission was part of the University of London - Institutes and activities. The University of North London and the London Guildhall University merged and continued under the name London Metropolitan University. Bretton Hall merged with the University of Leeds in 2001, Homerton College was a constituent college of the University of Cambridge and the RCN institute is part of the University of Warwick. In the transition from 1996/97 to 1997/98 the following changes took place: the University of Southampton and the Winchester School of Art merged, the University of Salford and Salford College of Technology merged. Moreover, the UCL - Eastman Dental Institute, UCL - Institute of Child Health, and UCL - Institute of Neurology made separate submissions from UCL. Data for these sets of HEIs are aggregated. Note that some HEIs did not submit to all RAE/REF. Data for 2002/03 – 2004/05 is based on the data from the preliminary funding announcement of funding, all other data is based on final allocations made. Unlike the preliminary funding announcement data, 2002/03 – 2004/05 data on final allocations does not include an overview of the broader budgetary envelope. We have not corrected for any discrepancy in the changes of 'mainstream' QR funding announced (in March) and received in the next year.

Table 7 Proportion of the funding directly informed by the RAE/REF reallocated following exercises in 1996, 2001, 2008 and 2014 (£millions)

	RAE 1996	RAE 2001	RAE 2008	REF 2014
	1997/98	2002/03	2009/10	2015/16
Total funding awarded in first year post RAE / REF	£684	£841	£1,074	£1,017
Change in volume of total QR funding allocated, compared with last year pre RAE / REF	£82	-£27	£154	-£1
Average QR funding awarded in first year	£6	£7	£9	£8
Sum of funding reallocated between HEIs in first year	£112	£93	£188	£124
Percentage reallocated without taking into account any change in funding allocated	18.6%	10.8%	20.4%	12.1%
Sum of funding reallocated between HEIs, net of any change in total funding	£30	£66	£34	£123
Net effect - Percentage reallocated taking into account the change in funding allocated	4.9%	7.7%	3.6%	12.0%
Number of HEIs awarded QR funding	124	127	123	122

Reallocation is defined as the sum of the gains and losses by individual institutions, compared with the previous year. Funding volumes are rounded to the nearest million. Source: Analysis by Technopolis, data from HEFCE

Table 8 provides the equivalent analysis of the churn in the allocations of 'mainstream QR' for the English Russell Group Universities. There also is a fairly high percentage of funding that is reallocated as a result of the REF2014 but the amount of churn is systematically lower for the Russell Group at each exercise than is the case for the English universities as a whole. The corresponding figures for the group of English Universities that were given university status from 1992 (Post 92 Universities) (Table 9) shows that they generally experienced a higher level of churn than the Russell Group, but not always as a high a churn as the set of English universities as a whole.

Table 8 English Russell Group Universities. Proportion of the funding directly informed by the RAE/REF reallocated following exercises in 1996, 2001, 2008 and 2014 (£millions)

	RAE 1996	RAE 2001	RAE 2008	REF 2014
	1997/98	2002/03	2009/10	2015/16
Total funding awarded in first year post RAE / REF	£461	£589	£718	£694
Change in volume of total QR funding allocated, compared with last year pre RAE / REF	£65	-£8	£65	-£18
Average QR funding awarded in first year	£23	£28	£36	£35
Sum of funding reallocated between HEIs in first year	£69	£46	£72	£63
Percentage reallocated without taking into account the increase/decrease in funding allocated	17.5%	7.7%	11.0%	8.8%
Sum of funding reallocated between HEIs, net of any change in total funding	£4	£38	£7	£45
Net effect - Percentage reallocated taking into account the change in funding allocated	1.1%	6.3%	1.1%	6.3%
Number of HEIs awarded QR funding*	21	21	20	20

*Includes all English Russell Group Universities for all years. Data for 1996/97, 1997/98, 2001/02, and 2002/03 include funding allocated to both UMIST and Manchester University, which merged in 2003. Reallocation is defined as the sum of the gains and losses by individual institutions, compared with the previous year. Funding volumes are rounded to the nearest million. Source: Analysis by Technopolis, data from HEFCE

Table 9 Post-92 Universities. Proportion of the funding directly informed by the RAE/REF reallocated following exercises in 1996, 2001, 2008 and 2014

	RAE 1996	RAE 2001	RAE 2008	REF 2014
	1997/98	2002/03	2009/10	2015/16
Total funding awarded in first year post RAE / REF	£35	£44	£87	£87
Change in volume of total QR funding allocated, compared with last year pre RAE / REF	£3	-£6	£51	£19
Average QR funding awarded in first year	£1	£1	£3	£3
Sum of funding reallocated between HEIs in first year	£9	£10	£51	£25
Percentage reallocated without taking into account the increase/decrease in funding allocated	27.4%	20.5%	140.9%	36.6%
Sum of funding reallocated between HEIs, net of any change in total funding	£6	£4	£0	£6
Net effect - Percentage reallocated taking into account the change in funding allocated	18.5%	8.9%	0.0%	8.3%
Number of HEIs awarded QR funding*	34	34	34	34

*The group of Post-92 Universities include all English HEIs that were given university status in 1992. Funding for London Guildhall University and University of North London (now London Metropolitan University) are aggregated for all years. Reallocation is defined as the sum of the gains and losses by individual institutions, compared with the previous year. Funding volumes are rounded to the nearest million. Source: Analysis by Technopolis, data from HEFCE

Table 10 summarises changes in the position of the English universities over time when they are ranked by total QR income. In 2015/16, 47 HEIs increased in rank relative to the previous year while 54 HEIs decreased in rank. The table also shows that under the RAE 2008 an additional 25 HEIs submitted to the RAE for the first time. The overall picture is of considerable movement within the ranking but the numbers of winners and losers staying relatively stable over time. Of course, very small changes in relative performance will induce changes in the rankings.

Table 11 summarises changes in the position of the English universities over time when they are ranked by narrowly-defined mainstream QR income. Thus, in 2015/16, 50 HEIs increased in rank relative to the previous year while 60 HEIs decreased in rank. The table also shows that under the RAE 2008 an additional 25 HEIs submitted to the RAE for the first time. As for the analysis of total QR, the overall picture is of considerable movement within the ranking but the numbers of winners and losers staying relatively stable over time.

Table 10 Changes in rank– according to the total QR funding allocated following exercises in 1996, 2001, 2008 and 2014

	RAE 1996	RAE 2001	RAE 2008	REF 2014
Funding year	1997/98	2002/03	2009/10	2015/16
Number no change	11	14	10	19
Number gaining	43	47	48	47
Number losing	50	50	56	54
Number received no funding in funding year	3	5	0	0
Number received no funding in previous year	22	6	10	2
Total number of English HEIs submitting*	129	122	124	122

Source: Analysis by Technopolis, data from HEFCE, *Total number of English HEIs submitting after correcting for mergers

Table 11 Changes in rank– according to the funding directly informed by the RAE/REF allocated following exercises in 1996, 2001, 2008 and 2014

	RAE 1996	RAE 2001	RAE 2008	REF 2014
Funding year	**1997/98**	**2002/03**	**2009/10**	**2015/16**
Number no change	11	12	12	9
Number gaining	43	50	38	50
Number losing	50	49	48	60
Number received no funding in funding year	3	5	0	0
Number received no funding in previous year	22	6	25	3
Total number of English HEIs submitting*	129	122	123	122

Source: Analysis by Technopolis, data from HEFCE, *Total number of English HEIs submitting after correcting for mergers

A study by UK-IRC (Hughes A. , Kitson, Bullock, & Milner, 2013) concludes that there is a high percentage of variation in the allocation of mainstream QR funding across HEIs. Those with a smaller level of mainstream QR income in 2002/03, on average, received a larger increase of mainstream QR income in 2010/11, consistent with the idea that there was a slight decrease in the concentration of mainstream QR funding as a result of RAE2008.

Table 12 and Table 13 show average, minimum and maximum gains and losses in mainstream/total QR funding as a result of REF2014 in percentage terms. As one might expect, the proportion of funds reallocated is smallest in the top quartile of universities. This comprises the biggest beneficiaries, so the absolute amounts of money involved can nonetheless be significant. The degree of volatility in percentage terms tends to increase further down the rankings, so universities receiving only modest amounts of QR funding are more likely than the big beneficiaries to experience big percentage changes from one period to the next.

Table 12 Summary of percentage changes in total QR funding allocated to English HEIs (2014/15 – 2015/16), by quartile

	Top quartile	Second quartile	Third quartile	Bottom quartile
Average loss	-5%	-13%	-24%	-20%
Average gain	6%	34%	91%	329%
Maximum change	20%	141%	355%	4589%
Minimum change	-17%	-39%	-46%	-32%
Number of HEIs	30	31	30	29*

Source: Analysis by Technopolis, data from HEFCE. *Excluding Writtle College and the University of London Institute in Paris. These two HEIs received no QR funding in 2014/15.

Table 13 Summary of percentage changes in mainstream QR funding allocated to English HEIs (2014/15 – 2015/16), by quartile

	Top quartile	Second quartile	Third quartile	Bottom quartile
Average loss	-8%	-18%	-40%	-25%
Average gain	12%	53%	174%	189%
Maximum change	35%	272%	1257%	864%
Minimum change	-25%	-49%	-70%	-40%
Number of HEIs	30	31	30	28*

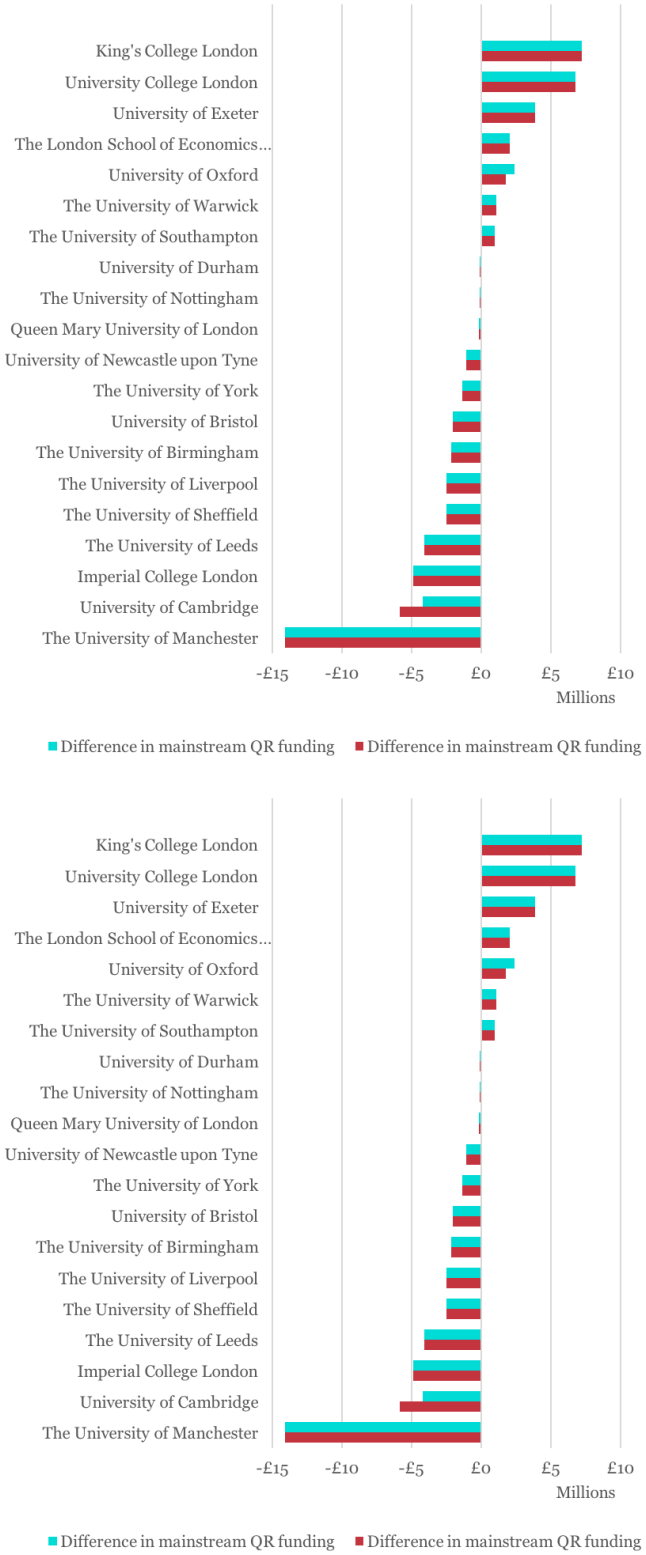
Source: Analysis by Technopolis, data from HEFCE. *Excluding Trinity Laban Conservatoire of Music and Dance, Writtle College, and the University of London Institute in Paris. These three HEIs received no mainstream QR funding in 2014/15.

The ten English universities (all Russell Group Universities) that received the most funding were allocated 52% of total QR funding and 48% of the total volume of 'mainstream' QR funding in 2015/16 (see also Table 4 and Table 5). The biggest changes in the proportion of QR funding awarded to individual institutions occurred among these universities[29]. One dynamic example is University College London which received 6.05% of the funding awarded in 1997/1998, 7.20% of total QR (6.05% of 'mainstream QR) in 2009/2010, and then 8.44% of total QR (7.46% of 'mainstream' QR) in 2015/2016. King's College London's total QR funding rose 1.5 percentage points between 1997/98 and 2015/16. Imperial College London's funding also varied, changing from 5.56% of total QR funding in 1997/98 to 6.38% of total QR (5.00% of 'mainstream' QR) in 2009/2010 and then changing again to 6.04% of total QR (4.76% of 'mainstream' QR) in 2015/2016.

The absolute amounts of money that universities gained or lost in the REF exercise were significant, though the total amount of money reallocated in some earlier RAEs was bigger than that in the REF. The most extreme example is Manchester University, which lost over £14m of total QR ($13.7m of 'mainstream QR'). By way of example, Figure 21 shows how Russell Group universities fared in REF2014, change from 2014/15 to 2015/16.

[29] As explained in the above, data for 2002/03 is based on the data from the preliminary funding announcement of funding, all other data is based on final allocations made.

Figure 21 Change in the distribution of funding to the English Russell Group Universities between 2014/15 and 2015/16



Source: HEFCE data

There are highly-ranked universities that are not part of the Russell Group. For example, the University of Bath ranked joint 14th on grade point average (GPA) in the REF and received 1.4% of the funding share in 1997/98 and a slightly lower share, 1.2% of total QR (1.3% of mainstream QR), in 2002/03. The University of Sussex received 1.6% of the funding in 1997/98 and 1.3% of mainstream and non-mainstream QR in 2002/03. The University of Reading received 1.8% of the funding in 1997/98 and 2002/03. Its 2009/10 share dropped to 1.3% of total QR and for 2015/16 to 1.1% of total QR. A more specialist institution, the Institute of Cancer Research received 0.9% of the QR funding in 2002/03 and as much as 1.2% of total QR in 2009/10 but only half as much, 0.44% of 'mainstream QR' in 2009/10.

A few 'success stories' are apparent, where a university has been able consistently to increase its share of the QR pot over time. For example, the University of Exeter received 1.0% of the funding in 1997/98 and 2002/03, 1.2% of total QR (1.4% of 'mainstream' QR) in 2009/10 and subsequently 1.5% of total QR (1.7% of mainstream QR) in 2015/16. The University of Warwick: improved its share of mainstream QR from 2.2% in 1997/98 to 2.6% in 2009/10. Looking at changes in total QR overtime among the Post 92 Universities we find that

- Anglia Ruskin University (previously Anglia Polytechnic University) experienced an improvement from close to zero percent in 1997/98 and 2002/03 to 0.15% in 2009/10 and 2015/16
- Bournemouth University experienced an improvement from close to zero percent in 1997/98 to 0.19% in 2015/16
- Birmingham City University and the University of East London experienced an improvement from close to 0.1% in 2002/03 and 2009/10 to 0.14% in 2015/16
- The University of Huddersfield experienced an improvement from 0.1% in 1997/98 and 2009/10 to 0.3% in 2015/16
- Liverpool John Moores University experienced an improvement from 0.2% in 1997/98 and 2002/03 to 0.3% in 2008/09 and 2015/16

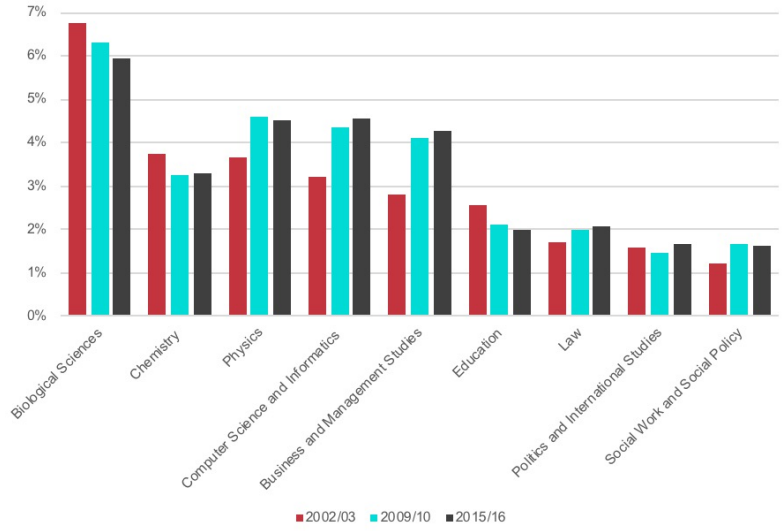
4.4 Movements among disciplines

It is hard to analyse the effects of the REF on the allocation of QR to different disciplines because the disciplinary composition of the panels and therefore the Units of Assessment (UoA) often change over time. However, in a small number of cases, definitions seem to have been stable enough to allow comparisons

HEFCE data suggest that in these cases there has indeed been some reallocation of mainstream QR funding among disciplines. Figure 22 presents the share of funding allocated to a selection of UoAs for different funding years[30]. These UoAs were chosen because their definition appears to have remained stable over time. We find an increase in the share of mainstream QR funding allocated to the UoAs Computer Science and Informatics and Business and Management Studies. There are decreases in the share of funding allocated to other UOAs: Biological Sciences and Education. Physics increased its share from 3.6% (some £15m) in 2002/03 to 4.6% in 2009/10, falling back slightly (about £4m) in the following exercise. In the other subject areas whose definitions appeared stable over time, changes are less pronounced.

[30] This selection of UOSs is chosen because there was no (substantive) name change to the UOA for the RAE 2001, RAE 2008 and the REF2014.

Figure 22 Narrowly-defined mainstream QR funding awarded to disciplines over time, 2002/3 – 2015/6



Source: HEFCE data

4.5 Has the REF affected performance?

Internationally, studies tend to associate PRFS with increased research production and quality (measured as citations) (Moed, 2005) (Butler, 2003) (Jiminez-Contreras, Anegon, & Lopez-Cozar, 2003) (Sivertsen, 2010) (Adams & Gurney, 2014) (Smart, 2013). UK and Australian experience shows that rule changes to incentivise production over quality or vice versa lead to corresponding behaviour changes (Butler, 2010) (Moed, 2005).

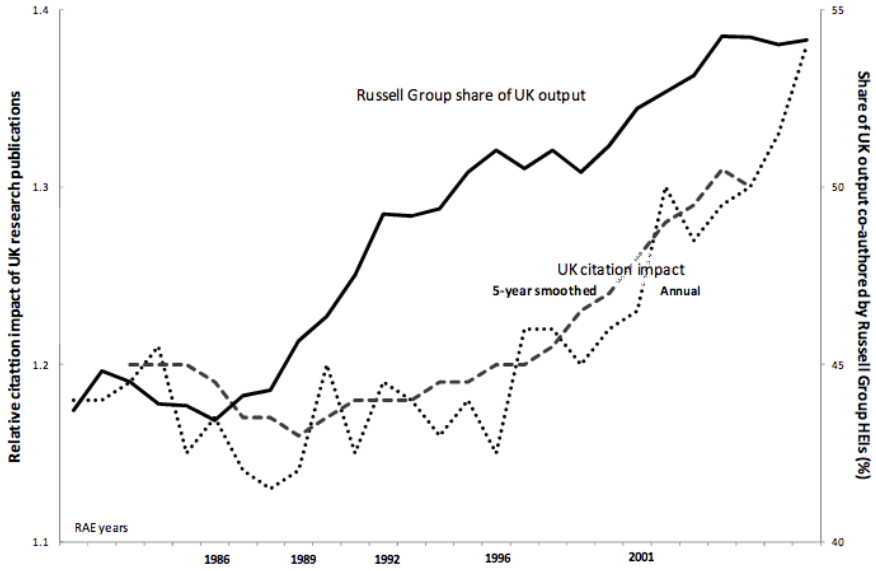
Wang and Hicks (2013) compared trends in publication and Higher Education Expenditure on R&D (HERD) in a handful of countries, searching for structural shifts in publication output. They found one in the UK, associated with the second RSE (consistent with Martin and Whitley's (2010) assertion that the universities did not take the first exercise seriously) but no further structural shifts in the UK thereafter. They also found a discontinuity in Australia but date it to a point before the introduction of the Australian PRFS, when the university system was expanded and the universities were required to report their publications as part of a process of increasing monitoring and quality control. Other discontinuities link to changes in funding or system size except for one in Germany that appears to be associated with a 1993 law allowing universities to make academics' pay dependent upon performance. In the Czech Republic, Vanecek (2013) links a discontinuous increase in publication volume to the introduction of a national evaluation system in 2004 and a context of increasing research funding, rather than to the start of the PRFS in 2008. It seems that PRFS can contribute to structural changes as part of larger processes of output-orientation rather than necessarily being the unique cause of such changes.

The JRC has recently produced two overviews of PRFS in EU member states (Jonkers & Zacharewicz, 2015) (Jonkers & Zacharewicz, 2016). The authors observe that almost all the countries considered have been improving their research performance measures in bibliometric terms. Thus, while all those that have adopted PRFS have improved their performance, those that have not adopted such systems have also improved, so there is not clear relationship between PRFS use and performance. Other factor such as increasing investment in higher education research seem likely to have a positive influence but there is also no simple relationship between these and performance. It is possible that PRFS contribute to improved performance, therefore, but it has not been possible so far to identify their net effect.

For the UK, there is little long-term bibliometric evidence in the literature that can be related to the period of the REF. A study published by HEPI (Adams & Gurney, 2014) looks into whether the RAE/REF have influenced the international competitiveness of the UK research base in terms of citation impact. The authors find that the UK performs above average and that there is an improvement over time, but

cannot unambiguously attribute this improvement to the REF. The dashed line of Figure 23 shows the five-year smoothed averages of the normalised relative citation impact of the UK research base for the years 1981-2007. In the early 1980's, UK research quality was, on this measure, declining although it remained above average (1.0 refers to above the world average). In the years thereafter, there was an increase in the citation impact. The right hand axis of the figure shows the share of UK output that is co-authored by Russell group academics. In 1986 the Russell group staff co-authored around 45% of the UK's research output captured by the Scopus database and by 2005 this percentage increased to close to 55%. Thus, the group of HEIs that benefited most from the RAE/REF in terms of income was also able to increase this aspect of its performance over time.

Figure 23 The relative citation impact of the UK research base (1981-2007)



Source: (Adams & Gurney, 2014)

It is thought that QR funding helps to leverage external funding. If this is the case, we would expect to see a positive relationship between QR funding and other funding sources available to HEIs.

One study has looked at the degree to which there is a relationship between mainstream QR funding and Third Stream Income (TSI) to HEIs (PACEC and University of Cambridge , 2014). In this study, TSI is viewed as a measure of external organisations 'benefits' or 'economic impact'. TSI sums external organisations' willingness to pay for research, consultancy, and other forms of engagement with HEIs[31]. The study finds a positive relation between QR funding per FTE academic staff (2002/03 and 2006/07) and TSI per FTE academic staff generated in the subsequent period (2007/08 and 2011/12). This suggests that HEIs that achieve high scores on the RAE/REF are in a better position to obtain external funding. The study also suggests that there is a *non-linear* correlation between QR funding per FTE academic staff and the other funding sources per FTE academic staff, generated in the subsequent period. For HEIs that have relatively high levels of QR per member of staff, on average, an increase in QR is associated with an increase in TSI. However, for HEIs that have relatively low levels of QR per staff member, on average, an increase in QR is associated with a decrease in TSI (PACEC and University of Cambridge , 2014). The implication seems to be that at reasonable scale, QR funding can be used to improve access to external funding, while at small scale it may substitute for external funding, hence QR funding tends to promote a 'Matthew principle' in which the bigger beneficiaries can build proportionately more capacity than at least some of the smaller ones.

[31] TSI includes income from collaborative research, consultancy, contract research, intellectual property (including sales of shares), facilities and equipment-related services, regeneration and development programmes, and continuing professional development and continuing education.

5 Effects on research

The 2010 OECD report on performance based research funding provides a list of potential effects of PRFS (including the RAE as a prime example) on aspects of research and we use these to structure this Chapter.

- Interdisciplinarity
- 'Blue skies', 'high risk' and 'transformative research'
- Research on the periphery, or non-conventional research
- Applied research
- Researcher autonomy
- Collaboration among researchers (OECD, 2010)

The first two of these areas are of particular relevance in the context of the REF. Despite efforts to facilitate submission of interdisciplinary work, there is still a strong perception that PRFS in general and the REF in particular lead to disciplinary lock-ins and discriminate against interdisciplinary research. PRFS are often seen as discouraging high-risk research, incentivising researchers instead to opt for 'safe' or mainstream topics most likely to guarantee the types of publication needed within the timeframes required to contribute to assessments of institutional performance. According to this argument, a systemic resistance to ground-breaking, transformative research is thereby created. We address applied research below but tackle the issue of researcher autonomy under the heading of effects on researchers and researcher careers in Chapter 6. This Chapter closes with a discussion of collaboration.

5.1 Interdisciplinary research

Interdisciplinarity is viewed as important in research and research funding. The perception is that work at the boundaries between disciplines can be the birthplace of scientific innovation, new disciplines and sub-disciplines. Work of industrial and social relevance may often be interdisciplinary because there are seldom mono-disciplinary answers to industrial or social problems. Research councils therefore tend to promote interdisciplinarity and to implement special procedures or structures into their peer review processes in order to do so. Innovation agencies tend to promote interdisciplinary work while applied industrial research institutes internationally have tended to form polytechnic structures in recent decades as a direct response to their perception that their customers need multi- and inter-disciplinary solutions (Arnold, Barker, & Slipersæter, 2010).

Bibliometric analysis suggests that there is a growing focus on interdisciplinary research in the UK. Similar trends exist elsewhere, and are if anything stronger in emerging economies where the balance of research effort tends to be more applied than in more mature research systems (Elsevier, 2015). The Elsevier report also notes that interdisciplinary research is associated with industry collaboration but also goes on to show that interdisciplinary research is associated with lower levels of international scholarly collaboration and, most notably, with lower overall citation impacts (Elsevier, 2015). Analyses of REF impact case studies show an association between interdisciplinarity and social impact (King's College London and Digital Science, 2015) (Greenhalgh & Fahy, 2015).

The literature is divided on the treatment of interdisciplinary research (IDR) in PRFS. Rafols et al summarise as follows:

> *...the strong impression from qualitative studies that IDR is at a disadvantage in peer review has apparently not been robustly substantiated by quantitative studies. Examining a total of 257 reviews of 38 projects from five somewhat interdisciplinary programmes (e.g. neurobiology) of the US National Science Foundation, (Porter & Rossini, 1985) found a weak but significant correlation between low grades and degree of interdisciplinarity ($r = 0.29$, $p < 0.05$). In contrast, (Rinia, van Leeuwen, van Buren, & Van Raan, 2001), who analysed the evaluation by an international panel of 185 physics programmes in Dutch universities, did not find a bias against [interdisciplinary research]. However, they*

> *did note that IDR tends to be published in journals with a lower citation impact (Rinia, van Leeuwen, van Buren, & Van Raan, 2001) (Rinia, van Leeuwen, & van Raan, 2001)[…] In summary, in contrast to the numerous qualitative studies pointing to clear bias against IDR in evaluation (Travis & Collins, 1991) (Langfeldt, 2006), there are only a few quantitative studies on the subject and these have produced ambiguous and somewhat contradictory results. (Rafols, Leydesdorff, O'Hare, Nightingale, & Stirling, 2012)*

Rafols et al themselves note that there are different types or degrees of interdisciplinarity and suggest that this determines whether or not there is a noticeable level of disadvantage. Where the degree of interdisciplinarity is low, eg where a research project is largely based in the paradigms of one discipline but 'borrows' to some extent from another, there is little evidence of disadvantage; in fact, such work has been observed to perform better than purely mono-disciplinary work in bibliometric terms. However, for strongly interdisciplinary work, where the primacy of one particular discipline is not evident and a whole new sub-field with new paradigms emerges, negative effects are common (Rafols, Leydesdorff, O'Hare, Nightingale, & Stirling, 2012).

While the REF and other PRFS may be quite explicit in their treatment of interdisciplinary research, they are nonetheless constructed around disciplines: interdisciplinary research is the oddity, spanning fields that often have different quality norms, making assessment that starts from a disciplinary perspective or using disciplinary journal classifications cognitively difficult. How an interdisciplinary research output might fare when presented to a panel composed of different disciplinary orientations is therefore hard to predict. It might find more favour with one disciplinary perspective than with another or, given its often intermediate character, it may fail to fully satisfy the established markers of quality in either perspective (Griffin, Medhurst, & Green, 2006). Several authors make the point that there are still effectively no established interdisciplinary indicators or criteria for quality (Wagner, et al., 2011),[32] and that instead a coming-together of mono-disciplinary perspectives is typically relied on to reach a judgement (Strang & McLeish, 2015).

Whatever concerns exist about the treatment of interdisciplinary outputs, there is no evidence to suggest that they are unfairly treated within the REF. The summary of REF2014 explicitly notes that

> *Panels observed increasing trends towards inter- and multi-disciplinary research and recognised excellence in these outputs. Indeed, interdisciplinary research scored equally highly as outputs from a single discipline.*[33]

In practice it is hard to test a hypothesis that researchers are shying away from interdisciplinary work[34]. However, there is emerging evidence about submissions to REF. Elsevier's report allocates an interdisciplinary research score to outputs based on the distance between the fields of sources cited in each output. The authors use this distance to help them rank outputs in terms of their level of interdisciplinarity. It then considers the UK share of the world's top interdisciplinary research outputs, and contrasts it to the share of REF outputs.

> *Around 6.4% of the REF submitted publications belong to the world's top 10% most [cited] [interdisciplinary research]. This percentage is higher in wider UK publications in Scopus (8.4%) and all publications in Scopus (10%). (Elsevier, 2015)*

---

[32] Laudel and Origgi note that some efforts have been made internationally, but that these are rarely institutionalized. They suggest that consistency with previous research, balance between interdisciplinary perspectives and potential effectiveness in terms of the challenge to be addressed might be useful starting points for criteria. (Laudel & Origgi, 2006)

[33] http://blog.hefce.ac.uk/2014/12/18/research-excellence-in-numbers/

[34] There is some survey-based evidence that researchers choice of topics has shifted away from these areas (Joynson & Leyser, 2015) but this is not corroborated by many other sources.

In other words: 8.4% of UK-based articles on Scopus belong to the global top 10% in terms of interdisciplinarity, whilst only 6.4% of REF-submitted outputs fall into this group. Assuming interdisciplinarity and quality are independent of each other, the presence of interdisciplinary research in the REF is only around three quarters of what one would expect it to be.

At the level of submission to the REF, therefore, interdisciplinary research does appear to fare less well than would be desirable. In the absence of clear criteria and a replicable process to gauge the quality of an interdisciplinary output, there is an inherent risk attached to submitting it to the REF. Unpredictable combinations of different disciplinary perspectives instead of established, stand-alone interdisciplinary standards and criteria drive this unpredictability to some extent. However, it is compounded by the fact that often very small numbers of people will have input into the judgement of an interdisciplinary output. Hence, it is hard for university research managers to predict how interdisciplinary research will be rated in the REF.

In mono-disciplinary research, in contrast, an institution can assemble experts whose understanding and definition of quality are likely to correspond to that of the panellists in that discipline's UoA, thus creating a basis for predicting whether or not an output will do well.[35] These considerations help explain the growing use of journal impact factors and lists of approved or 'diamond' journals by universities in research management, with researcher recruitment and promotion tied to such performance. To the extent that publication in high-status, high-impact journals is taken as an indicator of quality, interdisciplinary research is likely to be seen as of lower quality because of the more fragmented pattern of publication and the tendency for interdisciplinary research journals to have lower impact factors, as Rafols et al (2012) illustrate in the case of innovation research. (In this respect, it is worth recalling that impact factors result not only from the quality of articles published in a journal but also the number of people who read them.)

In indicator-based assessment systems, it has been noted that interdisciplinary research is most clearly disadvantaged, for reasons already noted in the above Elsevier report (Rafols, Leydesdorff, O'Hare, Nightingale, & Stirling, 2012) (Roessner, Porter, Nersessian, & Carley, 2013) (Roessner, Porter, Nersessian, & Carley, 2013).

Rafols et al's study comparing the interdisciplinary field of Innovation Studies with the more established field of business and management studies, into which IS departments are often integrated points to the role of the list of four-star journals rated by the Association of Business Schools (ABS). Publication in these journals is becoming a critical marker for career prospects in business and management studies; yet, the analysis shows that outputs from IS rarely feature in these journals, leaving IS scholars within business and management departments at a comparative disadvantage and threatening the growth of this interdisciplinary field as a whole. The authors conclude:

> *…when journal rankings are used to help determine the allocation of esteem and resources, they can suppress forms of interdisciplinarity that are otherwise widely acknowledged to be academically and socially useful. Important implications arise, both for research evaluation in the specific fields in question, as well as for wider investigations to inform the more general governance of science and technology using metrics to capture multidimensional qualities that cannot be intrinsically reduced to a single indicator. (Rafols, Leydesdorff, O'Hare, Nightingale, & Stirling, 2012)*

The use of proxies such as journal esteem or impact factor is more widely documented in the literature. Other academic studies demonstrate evidence of this (TaylorJ, 2011), and a recent UCU survey found that 27.7% of respondents noted that journal rankings or impact factors had been used to take decisions about REF submissions (University and College Union, 2013).

---

[35] This can take the shape of internal peer review, a full mock-REF at departmental or institutional level, or a range of other techniques, often also including metrics, see (UCU, 2013)

Over time, considerable efforts have been made in REF panels and processes to address interdisciplinary research. There is no evidence that suggests that the REF processes do this badly. Therefore, if indeed interdisciplinary research is disadvantaged it appears to be because of universities' submission behaviour. If it is hard in a REF or Research Council panel to assess interdisciplinary research, it is doubly so for university research managers to anticipate the way a REF panel may treat it. The risk-minimising strategy is therefore to prefer mono-disciplinary work for REF submissions.

5.2 Transformative research

Especially over the last 20 years or so, research funders have become increasingly anxious that their peer review processes tend to squeeze out proposals that are riskier, more translational, more radical or potentially disruptive than the norm. We use the term 'transformative research' as a shorthand for these kinds of research.

The idea 'transformative research' originates from Kuhn's (1970) notion of revolutionary science and competing paradigms. Like most philosophy of science in the first two-thirds of the Twentieth Century, Kuhn's work is essentially about physics. Subsequently, the idea has been applied to the sciences more generally. A discourse has become established around research that challenges some form of a 'mainstream'. Terms used to describe these activities include 'revolutionary science' (Kuhn, 1970), blue-sky research (Kilduff, Mehra, & Dunn, 2011) or transformative research (Mertens, 2009). The term 'transformative research' emerged from a debate among US researchers and funders (especially the NSF) about how to assess and fund high-risk research (National Science Board, 2007).

Transformative research tends to struggle in peer review, be that for grant funding or for journal publication (Wessely, 1998) (Horrobin, 1996) (Roy, 1985) (Lakatos & Musgrave, 1970) (National Science Board, 2007) (Arnold, et al., 2013) (Luukkonen, Stampfer, & Strassnig, 2015). This is often due to risk-averseness, (Chubin & Hackett, 1990) (Wagner & Alexander, 2013), (Hävrynen, 2007) (Luukkonen, 2012). Laudel and Gläser (2014) argue that the often ambitious and less predictable nature of transformative research results in longer lead times and greater need for follow-up funding than 'normal science'. Others – in particular research funders themselves – have drawn links between transformative research and early career researchers. Those who have not spent an entire career socialised into existing paradigms are more likely to break away from them; yet they are also typically at a disadvantage in peer review processes, especially where these take account of applicants' track record and experience (Roy, 1985) (Travis & Collins, 1991).

Some funders have introduced specific funding schemes for transformative research. These include most notably the ESRC and EPSRC in the UK and the NIH and NSF in the US.[36] Much of the literature on transformative research and the RAE/REF is qualitative, as the notion of transformative research is arguably less quantifiable than interdisciplinary research. Nevertheless, the available evidence allows us to identify two elements of the RAE/REF that likely discourage TR, which we take here to mean any forms of non-conventional research besides interdisciplinary research.

The element of controlling risk that exists in ex ante grant funding reviews of course does not exist in ex-post research assessments: the risk involved in conducting the research has already been taken (Klein, 2006). Hence, if transformative research is disadvantaged in the REF, it is more likely to be

[36] ESRC: http://www.esrc.ac.uk/news-events-and-publications/news/news-items/the-transformative-research-scheme-outcomes-of-the-third-round/

EPSRC: https://www.epsrc.ac.uk/funding/howtoapply/routes/network/ideas/

NSF: https://www.nsf.gov/about/transformative_research/definition.jsp

NIH: https://commonfund.nih.gov/tra

because universities are reluctant to conduct or submit it than because the REF processes cannot value it[37].

Laudel and Gläser note that transformative research can often be significantly harder to plan than more traditional research projects. Transformative research may require more flexible timeframes and resources. Budgeting and timetabling may be done robustly when there is substantial existing knowledge about the chosen methods and fields of application, but in the absence of such knowledge, parameters are harder to set (Laudel & Gläser, 2014).

The requirement, derived from the REF, for researchers to produce a minimum number of excellent outputs within the current window of assessment discourages risk-taking and longer-term research. Particularly given a growing interest in how to encourage high-risk, transformative research (Luukkonen, Stampfer, & Strassnig, 2015) (Laudel & Gläser, 2014), the suggestion that the presence of the REF makes it risky for the individual researcher to conduct long-term, risky or radical research needs to be taken seriously. As with interdisciplinary research, the issue is not in the REF process itself but in the way the REF encourages universities and researchers to behave.

We conclude that while the REF is probably equipped to value both interdisciplinary and transformative research, its presence in the research system discourages the performance of both. Of course, many people do both in any case – but it seems that the way universities manage research in order to obtain REF funding reduces the volume of each, compared with what researchers would otherwise do.

5.3 Non-conventional research

In non-Anglophone countries, there is a tendency to push research into Anglophone publication channels as these are the ones best covered by the bibliometric databases (Scopus and Web of Science), often at the expense of nationally or locally relevant research, and cementing a peripheral status for non-Anglophone countries (Stöckelova, 2012) (Meriläinen, Tienari, Thomas, & Davies, 2008).

Lee and Harley (1998) conducted research on the distinction between mainstream economics, driven largely by mathematical models, and heterodox economics which departs considerably from the established norms of the discipline. Critical in Lee's analysis is the presence of so-called diamond journals in the discipline. Publication in these journals has since the introduction of the REF become a requirement for recruitment by an increasing number of institutions. Publications in these elite journals consistently make up the most highly rated outputs in the RAE. The diamond journals all have a focus on mainstream economics, methodologies and approaches, so that alternative approaches have been squeezed out, resulting in a decrease of heterodox economics across institutions, which in turn has not only affected the breadth of approaches in research, but also undergraduate teaching curricula. Rafols et al similarly point to the role of preferred lists of journals in business studies in making it hard for research in innovation studies to be recognised (Rafols, Leydesdorff, O'Hare, Nightingale, & Stirling, 2012).

5.4 Applied research

Applied research has been identified as likely to be disadvantaged by performance-based research funding systems. The OECD report notes:

> *This is of particular concern in PRFS systems based primarily on quantitative indicators, as much of this research is published in the "grey literature" that falls outside the ambit of standard performance indicators. (OECD, 2010)*

The focus of the UK system on peer review means that this problem is not likely to be especially pertinent and since 2014 the REF's impact dimension mitigates this issue further, as applied research is thereby

[37] That said, a commentator on an earlier draft of this paper pointed out that the REF does not have a category of 'game-changing research'. The highest praise it can bestow is to describe research as 'world-leading', which may nonetheless be research that is very conventional

explicitly encouraged. The use of different panels for different subjects means that the opportunities directly to compare basic and applied fields are reduced, reducing the risk that one form of research is judged by the standards of another – although this risk is not completely eliminated (Rafols, Leydesdorff, O'Hare, Nightingale, & Stirling, 2012). University career incentives to publish in indexed journals, however, may influence researchers' choice of topics.

5.5 Collaboration among researchers

Rising international collaboration in research is an established trend, which goes well beyond Europe and the Framework Programme. For example, an analysis of Swedish researchers' international co-publication patterns showed there was no difference in co-publication patterns between Framework Programme participants and non-participants (Arnold, et al., 2011).

There is a widely held view that fostering collaboration – including in the shape of co-authorship – is of strategic importance. Figure 9 in Section 2.4 of this report illustrates the large and growing importance of co-publication in UK research, almost doubling in the past decade. Elsewhere, Adams compares the increase in international co-publication as a share of total research outputs with five other countries (US, Switzerland, China, South Korea, Brazil) and notes that the rate in the UK is in fact amongst the strongest of these. International co-publication is increasingly associated with higher citation and, by proxy, with higher quality (Adams J. , 2013).

There appears to be no general evidence about whether PRFS *per se* affect national or international collaboration, suggesting that any influence of a PRFS is likely to be caused by the specific way in which it treats collaboration. Both for Norway (Bloch & Schneider, 2016) and Morocco (Bouabid, 2014) there is evidence that the introduction of metrics-led research assessment has led to higher levels of collaboration, specifically international co-publication in the case of Morocco.

The OECD notes that the principal choice for PRFS designers in the area of co-authorship is between giving each institution or author full credit for outputs to which they contribute, or only to give fractional credit (OECD, 2010). The former is intuitively more conducive to collaboration.

> *Concerns about the effect of fractional counts on collaborative activities led Australian governments to reject this methodology in favour of whole counts. Norwegians appear less concerned and believe their use of fractional counts has not resulted in a decline in collaborative activities (Schneider, 2009, p. 372). Schneider believes that "'invisible colleges' and social networks within research specialties have eventually ensured collaboration". It is believed that the dependence of research on collaboration will counteract any adverse behaviour that might result from the funding model. (OECD, 2010)*

Studies in New Zealand (Edgar & Geare, 2010) and the UK (Henkel, 1999) point to the tension between PRSFs' tendency to encourage competition among individual researchers and the benefits of collegiality in research. This tends to confirm the idea that any effects on collaboration are determined by the detail of how particular PRFS are designed.

One specific effect of the REF, however, relates to cooperation with developing countries. Jeffery (Jeffery, 2014) notes that in North-South collaborations UK authors often prioritise aspects of research most conducive to a suitable REF submission.

6 Effects on researcher careers

Earlier Chapters imply that effects of the REF and other PRFS on individual researchers are mediated through the mechanism of careers. This Chapter starts with the observation that a PRFS induces focus on research outputs rather than the wider process of research and that this filters through to the way researchers have to behave in order to obtain career success. It then looks at the way the influence of the REF is mediated through careers. The third section looks at what the researchers themselves say. The fourth explores the way the REF seems over time to have redefined the academic role, reducing the importance of teaching and increasing that of research in defining academics' careers and legitimacy. The final section looks at effects of the REF on academic mobility.

6.1 The output imperative

PRFS normally analyse **outputs** of research and therefore incentivise increases in the number of outputs produced. This can lead to 'salami slicing' and a reduction in average quality (Butler, 2003) unless the incentives for quantity are also complemented by requirements for quality.

The REF and its predecessors conform to this norm of output focus and have tended therefore to draw attention away from other aspects of the academic career, reducing academics' ability to choose between being researchers or scholars and teachers (Elton, 2000). The older literature describes a growing focus on the production of particular types of output with a corresponding de-emphasising of academics' roles as teachers. More recent literature discusses the pressure on academics to be 'research active' (Sikes, 2006) (Elton, 2000) (OECD, 2010).[38] Henkel notes:

> *"Where 'research inactivity' was once tolerated and sometimes needed as part of the departmental economy, it is now perceived as a liability" (Henkel, 1999)*

Authors also note that the definition of research has been narrowed and standardised as a result of the REF (de Boer, et al., 2015). Survey data highlight a widespread perception that career development activities and non-publication activity (professional education, training and supervision, mentoring, training, teaching, peer review, university administration, public engagement and contributing to the work of national bodies and policy makers) are undervalued in the REF (Joynson & Leyser, 2015). Although these factors were included in the 'environment' element of REF2014, it is not yet clear whether this will reduce the focus on published outputs at the level of individual researchers and their careers.

Researchers' understanding is that successful publication rather than conducting research, is the objective of the job (Bence & Oppenheim, 2005). Moreover, as the outputs need be of a certain scientific quality, they need to appear in very particular publication channels, especially in high-impact factor journals, which have increased significantly as a proportion of overall research outputs both in the UK and elsewhere (OECD, 2010). Although the REF itself does not require researchers to use these specific channels, their employers increasingly do, in the expectation that this will maximise returns from the REF (Farla & Simmonds, 2015). Publication in an often pre-selected list of high-status journals therefore becomes a critical part of researchers' activities (Rafols, Leydesdorff, O'Hare, Nightingale, & Stirling, 2012) (Joynson & Leyser, 2015). This focus on journals has become especially problematic in fields where journal publication has traditionally not been central. For instance, in the field of medieval studies, the it has shifted attention away from updating editions, so that out-of-date editions have become more common (Maude, 2014).

The focus on high-impact-factor journals has also tended to displace publication for non-academic audiences.

[38] The notion of being non-research active has also been highlighted as having negative career consequences in New Zealand, see (Middleton, 2008)

> *"...since publications only counted for the RAE if they were in refereed journals, publication became more biased towards these and away from professional and more popular journals. This in turn resulted in a reduction of dissemination to potential users of research..." (Elton, 2000)*

Hicks has also noted that what she calls 'enlightenment literature', ie pieces for wider circulation and dissemination, has suffered and been reduced in the social sciences and humanities as a result of performance based research funding systems, a problem not only visible in the UK but also frequently in the international context (OECD, 2010) (Hicks, 2013).

6.2 The career as a pathway to impact for the REF

The previous Chapter indicated that effects of REF on the character of research tend to result not only from the operation of the REF itself but also from the way universities manage and decide to submit research to it. The fact that REF submissions are generated in a centralised process means that the institution is more involved and impinges more on the individual's career and activities than it did in the past, when the department was the main context (Henkel, 1999). The individual researcher is therefore obliged to demonstrate value not only within the context where research is undertaken but also at a higher organisational level that is less aware of that departmental or disciplinary context (Agyemang & Broadbent, 2015) (OECD, 2010) (Leisyte & Westerheijden, 2014).

From the university perspective, obtaining a high REF score and therefore high funding depends on recruiting and retaining academics capable of producing outputs that will be given high marks in the REF and then encouraging them to do so. Since university recruiters and research managers cannot predict what REF scores individual researchers will obtain, they work with the available proxies. Do the academics have a history of achieving high REF scores? Do they publish in highly-rated journals, which are likely to carry the kind of articles that REF panels will in turn rate highly? In many cases, universities also use internal or external peer review to help them decide which outputs to submit to the REF process (Farla & Simmonds, 2015). These considerations underpin universities' heavy reliance on REF-related criteria in recruitment, appraisal, promotion and career development.

The REF is in this respect not unique. The influence of PRFS more generally on behaviour appears to be strongly mediated by the way researcher careers are managed and, correspondingly, the extent to which researchers conform to the demands of university Human Resource (HR) management. Thus, almost half of department heads in Norway make use of the Norwegian publication indicator (used in the in Norwegian PRFS) for recruitment and promotion purposes, and almost 90% use it for monitoring departmental activity (Aagaard, Bloch, & Schneider, 2015). Institutional incentive systems linked to the national assessment exercise are also used in Italy (Abramo, D'Angelo, & Di Costa, 2011). In the Spanish system, performance in the *sexenios* influences the award of tenure and eligibility to become part of the panel that grants tenure as well as influencing competitive grant funding and has been shown to drive individual performance (Jiminez-Contreras, Anegon, & Lopez-Cozar, 2003). A change in German law to allow performance to drive salary had a similar effect. Evaluations consistently identify changes in research management as an immediate effect of PRFS (Butler, 2003; Martin & Whitley, 2010) and these in turn affect staff appraisal and career development. Other studies demonstrate the link between individuals' REF performance and the extent to which they continue to be allowed to do research (Sikes, 2006) and at the upper end of the seniority ladder that considerations based on the REF have resulted in forced early retirements (Bence & Oppenheim, 2005).

The idea that the career is a central driver of researcher performance is supported by an unrelated exercise aiming to understand the effect of incentives on publication behaviour. It looked at submissions to *Science* and showed that while cash incentives produced the greatest number of articles submitted, it was career-related incentives that drove increases in quality and hence the proportion of articles accepted for publication (Franzoni, Scellato, & Stephan, 2011).

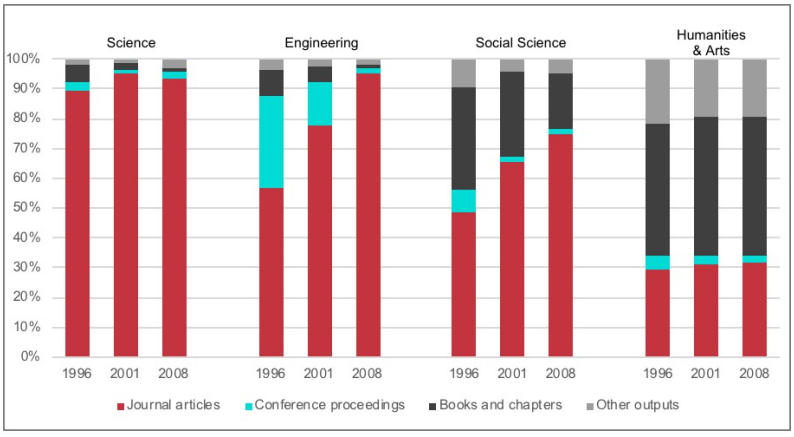
Adams and Gurney's analysis of RAE submissions clearly indicates the influence of research management on the selection of publications to submit, and implies that management resorts to various inappropriate proxies for quality that push researchers towards producing material that meets the

expectations of research management rather than the needs of research. (Note that their use of 'researcher' in this quote is inappropriate: research managers not individual researchers make REF submissions.)

> *... researchers submit journal articles in preference to the outputs that elsewhere they say are central to their field, they skew their selection to high-impact journals, and they submit pieces from such journals even when they are not well cited and, sometimes, not even research papers. We suggest that this is because they believe that the 'brand' of a journal known to have high average impact is a better proxy 'signal' in place of real evidence of excellence. ... The articles submitted in RAE2008 are a concentrated selection from journals with relatively high citation impact – either globally or more specifically in their field. For some journals, such as Nature, there were more RAE records than papers with UK addresses. Some researchers submitted editorial, letters and other ephemera – often not cited at all – instead of substantive research contributions. Why? Because the kudos of the well-cited journal was a marketing signal outweighing the individual item and outweighing even the opportunity to submit better cited papers from less prestigious journals. (Adams & Gurney, 2014)*

Figure 24 shows the authors' analysis of the publication channels used in REF submissions over time. It indicates a sharp increase in journal publication in engineering and the social sciences. The natural sciences begin the period using journals as their primary communications medium while the growth in the use of journals in the humanities is limited, maintaining their traditional focus on books and other outputs.

Figure 24 Submissions to the REF classified by area of research and type of publication, 1996 – 2008



Source: based on (Adams & Gurney, 2014)

We have not been able to identify a systematic account of the techniques university research managers use in deciding what outputs to submit to the REF. The University and College Union's 2013 survey of its membership suggests that whilst peer review dominates these processes, there is also considerable use of journal rankings, publication type and citation data (Table 14), all of which are known to be problematic measures of quality, particularly at the level of the individual researcher or research output (Eyre-Walker & Stoletzki, 2013) (Seglen, 1997). However, since the unit of analysis in the survey is the individual academic rather than the university, we cannot be sure that it accurately reflects the balance of practice among institutions.

Table 14: UCU members' view of methods used for selecting outputs to be submitted to the REF

Answer Options	Response Percent
Internal peer review	65.6%
External peer review	49.4%
Mock REF exercise	41.9%
Assessment by senior manager/REF-lead	47.7%
Journal rankings/impact factors	27.7%
Type of publication/output	25.3%
Citation data	9.8%
Don't know	18.0%
Other (please specify)	3.0%
Answered Question	*6,902*

"Methods used in making decisions on inclusion in REF submission"
Source: (University and College Union, 2013)

6.3 What the researchers say

Consultations with the academic community in the UK about research assessment and the REF tend to produce agreement that the REF functions as a performance management tool. Institutional comment (ie from research managers who use the REF as a management tool) tends to be positive while, unsurprisingly, comment from individual academics (on whom the tool is used) tends to be negative (Nielsen, Kolarz, Farla, Blessing, & Simmonds, 2016).

The University and College Union (UCU), an academic trade union, did an extensive survey in 2013 of members whom it could identify as academics or researchers (University and College Union, 2013). It sent out 46,225 questionnaires and obtained a 15% response rate. There was little difference between the responses of people who expected to be included in REF 2014 and those who did not.

While 71% of respondents believed that some kind of assessment of individuals' research outputs was necessary, a small majority (51%) felt that the REF should be replaced by a better method while 36% had no view on the subject. Only 16% felt that the REF and its predecessors had resulted in an increase in the quality of UK academic research and only 20% felt that the REF was a good indicator of the quality of research being undertaken in the universities. Substantial proportions of respondents said that failure to meet institutional expectations regarding the REF would result in denial of progression or promotion and could result either in them being transferred to teaching-only positions or to losing their job (Table 15).

Table 15: Perceptions and concerns about not fulfilling institutional expectations

Answer Options		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I would be concerned about my future career prospects if I was not included in my institution's REF submission	All respondents	3.5%	9.9%	15.0%	35.1%	36.5%
	Lecturers	1.9%	5.4%	11.2%	37.3%	44.1%
	All – not in REF	3.8%	13.8%	20.1%	34.7%	27.6%
	Lecturers – not in REF	2.1%	6.8%	17.5%	40.5%	33.1%
If I do not perform to institutional expectations in relation to the REF, it is likely that I will not be supported by my institution to undertake research in the future	All respondents	7.5%	17.5%	30.2%	29.2%	15.7%
	Lecturers	5.6%	15.3%	31.0%	30.3%	17.7%
	All – not in REF	4.7%	13.9%	29.0%	28.4%	24.1%
	Lecturers – not in REF	4.2%	10.4%	28.8%	30.3%	26.4%
If I do not perform to institutional expectations in relation to the REF, it is likely that I will be transferred to a teaching only/teaching focused contract	All respondents	12.1%	29.0%	37.3%	15.0%	6.6%
	Lecturers	8.0%	25.6%	40.1%	17.7%	8.6%
	All – not in REF	9.0%	24.9%	38.9%	16.8%	10.5%
	Lecturers – not in REF	4.8%	18.5%	43.8%	20.2%	12.8%
If I do not perform to institutional expectations in relation to the REF, I am concerned I will lose my job	All respondents	20.9%	31.8%	23.9%	16.1%	7.2%
	Lecturers	14.8%	30.2%	24.2%	20.6%	10.2%
	All – not in REF	13.6%	30.2%	26.5%	19.8%	9.8%
	Lecturers – not in REF	8.6%	30.1%	22.3%	25.6%	13.4%

Source: (University and College Union, 2013)

Respondents felt that the REF impinged in substantial ways on their use of time and their ability to do their jobs as academics. Women experienced slightly more difficulties than men.

- Over 67% of all respondents (and 73% of women responding) disagreed/ strongly disagreed that they were able to undertake the necessary work to produce the required REF outputs without working excessive hours
- Over 52% of all respondents (and over 57% of women responding) agreed/ strongly agreed that pressure within their institution to meet institutional expectations in relation to the REF had made it difficult to manage their workload
- Nearly 29% of all respondents (and close to 31% of women responding) agreed/ strongly agreed that they did not have enough time to prepare their teaching because of the need to focus on their REF outputs (over 36% of all respondents and 32% of women disagreed/strongly disagreed)
- Nearly 27% of all respondents (and close to 28% of women responding) agreed/strongly agreed that they did not have enough time to provide feedback and support to students because of the need to focus on their REF outputs (close to 40% of all respondents and over 36% of women disagreed)
- Over 60% of all respondents (and over 63% of women responding) agreed /strongly agreed that pressure to meet expectations in relation to the REF had increased their stress levels
- Over 34% of respondents agreed/strongly (and close to 39% of women responding) agreed that pressure to meet expectations in relation to the REF had negatively impacted on their health (University and College Union, 2013)

The survey suggested that more junior staff were more likely than their senior counterparts to receive warnings of sanctions the university might take if their performance was unsatisfactory.

Similar observations have been made elsewhere (Henkel, 1999). Brinn et al have similar findings from conducting a survey based on data from over 700 research active academics in finance and accounting in 2001. Junior researchers in particular associated the RAE (REF) with adverse career consequences (Brinn, Jones, & Pendlebury, 2001).

Both the link between the REF and career consequences, as well as the greater negative effects on more junior researchers, are long-standing phenomena. Even the most recent literature still notes the critical significance of the REF on these most evident markers of career progression (Gray R. , 2015).

Joynson and Leyser have recently conducted research on the culture of scientific research in UK academia based on 970 responses to an online questionnaire run by the Nuffield Council for Bioethics and reflect more broadly on the criteria used to decide on promotion and career progression, of which the REF is clearly a significant part. They note that

> *… fifty-four per cent of respondents think the way scientists are assessed for promotion during their career is having a negative or very negative effect overall on scientists in terms of encouraging the production of high quality science, compared to 22 per cent who think it is having a positive or very positive effect. (Joynson & Leyser, 2015)*

Several studies have pointed out a gender imbalance: a slightly higher proportion of female than male respondents report that they had been warned about the career consequences of inadequate REF performance. A further imbalance is connected to the fact that

> *The results in 2001 demonstrated that while 64% of men were chosen for submission for RAE, only 46% of women … (Broadbent, 2010)*[39]

6.4 Teaching versus research

It has been evident since the early years of the REF that applying incentives and attention to research while not at the same time doing something equivalent to teaching would result in a redefinition of the role and importance of research in academic careers (Brinn, Jones, & Pendlebury, 2001). As a consequence, the division of labour between teaching and research has become more stark (Henkel, 1999) (Sikes, 2006) (Arnold & Mahieu, 2015) with the research role becoming dominant and less priority being given to teaching and other duties such as administrative tasks (which are rarely mentioned in the literature) (Court, 1999). A clear hierarchy of esteem privileging research over teaching has been the result (Broadbent, 2010), as has a lack of reward for teaching across large parts of the system (Brinn, Jones, & Pendlebury, 2001).

Based on survey research, Court identifies the rise of a 'research imperative', leading to a standardisation and narrowing of the academic career path, with teaching and administration deemed largely insignificant for career progression, and over-emphasis on research widely acknowledged, alongside a growing notion of teaching-focused staff as second-class citizens (Court, 1999). Based on a large programme of interviews, Henkel describes a shift from more pluralistic groups of academics, where emphases between teaching, research and other tasks were fluid, to a more rigid structure.

> *"Division of labour within departments was evolutionary and adaptive to individual variations, within an overall ethos of collegial governance and largely unspoken assumptions of academic community in an elite system. To some extent, it relied on some members not being particularly active researchers. […] The RAE brought with it new and more formal demarcations between researchers and teachers. Judgements of the quality of a department's research are made in the*

[39] In terms of ratings other than inclusion, there are likewise gender imbalances (Brooks, Fenton, & Walker, 2014)

> *RAE only by reference to those members designated 'research active'." (Henkel, 1999)*

The fact that all universities submit returns to the REF means that this logic extends right across the university sector – including to those newer universities that in practice tend to focus on teaching (Elton, 2000). With the RAE a critical source of university research funding, the remit of academics across the system was thereby pushed towards research, away from teaching (Brinn, Jones, & Pendlebury, 2001).

A clear hierarchy of esteem privileging research over teaching has been the result (Broadbent, 2010). In this new situation, Elton notes that departments increasingly made use of teaching assistants, part-time and temporary staff to relieve researchers of their teaching duties (Elton, 2000). A comparison with four other countries suggests that this re-delineation of duties is particularly noticeable in the UK (Key Perspectives Ltd, 2009).

Whilst, as the 1992 RAE results show, there was at least some degree of high quality research being conducted at all institutions, many academics – especially at the post-1992 institutions – did not view their career as being strictly or even primarily geared towards the production of research. In post-92 universities, identities were built around teaching, which led to a need for individuals to re-evaluate identities and missions (Kogan & Henkel, 1998). Pushing academics towards a 'researcher' identity, even if they did not see themselves as such before, also meant that many found themselves with a newly defined remit that did not reflect the reasons they entered the profession, as Sikes' extensive programme of interviews found (Sikes, 2006). Reorientation of their remit was particularly challenging for academics who saw teaching as their central role (Henkel, 1999).

The extent of synergy between university teaching and research is not well explored but the idea of research-based or research-informed teaching is something of a mantra. Young academics in particular often have interest and ambitions in both areas (Kogan & Henkel, 1998). In the hard sciences and economics this synergy is believed by some to apply mainly at the postgraduate level but there is a prevalent understanding in the humanities and social sciences that teaching benefits from being underpinned by excellent research (and delivered by the individuals conducting it), whilst teaching itself also informs and allows expansion of research horizons (Kogan & Henkel, 1998) (Henkel, 1999; Sikes, 2006).

The UCU survey (University and College Union, 2013) shows that academics believe that increased teaching loads or the removal of the right to research is a likely sanction against those who do not perform in REF terms. This raises a problem of path-dependency – once the resources needed to do research have been removed, it is difficult to improve one's research performance and move back 'on track' in a research career – as is implicitly recognised in some of the more recent arrangements that take account of individual researchers' special circumstances in selecting what outputs are to be submitted. The inability to continue pursuing or move back towards research following an increase in teaching loads after non-submission for RAE has also been documented by Henkel, based on a broad programme of qualitative research across subjects and institutions (Henkel, 1999).

6.5 Mobility

The reasons why researchers choose to change institutions are not systematically understood (Fernandez-Zubieta, Geuna, & Lawson, 2015) and the REF is unlikely to be more than one among several drivers. However, national research assessment systems in general, and the REF in particular, foster transfer markets for academics, where those who produce highly rated outputs can find themselves in high demand. Others who do not rate well in the PRFS are likely to be disadvantaged in comparison (OECD, 2010). In effect, the 'labelling' done by the PRFS adds to the information otherwise available to potential employers through CVs, track records and bibliometric indicators like h-indices.

In this sense, performing well in the REF can become an asset through which transfer to more prestigious positions or institutions becomes possible (de Boer, et al., 2015) (Brinn, Jones, & Pendlebury, 2001) (Leisyte & Westerheijden, 2014) (OECD, 2010). We have not been able to find systematic work relating RAE performance and academic pay, although there are plenty of anecdotes

supporting the idea that it does. However, Leisyte and Westerheijden (2014) show that between 2002 and 2006 the number of academics in the UK earning more than £100,000 increased by 169%.

Brinn et al collected survey data on the issue of positive vs negative impacts of the RAE on researcher mobility. Whilst overall responses tend more towards the positive, the data also show divisions between senior and junior academics, as well as between post-1992 and older universities. Whilst more senior researchers and those based at 'old' institutions largely tend towards seeing a positive impact on mobility, the views in the other groups are far more cautious, and tend further towards negative impact among the 'new' universities. Mobility is often very closely mediated through RAE scores, meaning for instance that authoring an article that received a 5-star rating makes the author attractive to 4-star rated department (Brinn, Jones, & Pendlebury, 2001).

Court (1999) established through a survey of 2000 AUT[40] members (561 or 28% responded) in 1999 that over 30% of academics believed that normal appointment procedures had been circumvented in order to bring in 'star' academics to boost RAE ratings.

An important issue not treated in the literature is the influence of PRFS in general and of the REF in particular on the circulation of manpower from industry into the academic system. This seems likely to be constrained by the fact that the REF (and therefore university HR policies) value only academic and not industrial track records.

[40] The Association of University Teachers (AUT) later merged with NATFHE the University and College Lecturers' Union to become the University and College Union, (UCU) referred to earlier

7 International practice in performance-based research funding systems (PRFS)

In this Chapter, we provide an overview of international practice in performance-based research funding systems (PRFS) covering 9 countries: Belgium/Flanders, Czech Republic, Denmark, Finland, Italy, New Zealand, Norway, Sweden, UK. Where relevant, we cover two additional countries where research assessment at the national level is not linked to funding, namely Australia and the Netherlands.

The Chapter is based on our research on PRFS in the "Measuring scientific performance for improved policymaking" study commissioned by the STOA office in the European Parliament (Mahieu, Arnold, & Kolarz, 2013) and a study designing a new R&D evaluation and institutional funding system for the Czech Republic, commissioned by the Ministry of Education, Youth and Sports (Arnold & Mahieu, 2015). In these studies, we conducted an extensive analysis of the scientific literature, relevant policy documents and information from the national organisations organising PRFS. Where necessary, information about individual countries has been updated for this report as some PRFS have changed over time.

Performance-based research funding systems are defined in different ways in the literature, depending on the scope of investigation. In this Chapter we follow the definition of PRFS as "national systems of research output evaluation used to distribute research funding to universities" (Hicks, 2012). A PRFS consists therefore of two core components: evaluation (or 'assessment'); and the use of assessment results in a funding formula.

PRFS are essentially policy instruments. Policy objectives directly guide the parameters used for the research assessment as well as the funding tools used and the proportion of universities' overall and institutional funding that is governed by the PRFS. The policy context of the PRFS is therefore a critical factor that drives - and explains - differences among the performance-based research funding systems used in international practice.

We discuss PRFS in the countries considered from a research governance perspective in Section 7.1, setting out the types of policy objectives most commonly addressed in international practice, their influence on the institutional funding system, and the uses of the evaluation results. In Section 7.2 we describe the different models for research assessment that are used in PRFS and their key design elements. Section 7.3 sets out the influence that the PRFS in the different countries have on institutional funding for research and the universities' institutional funding overall. In Section 7.4 we position the RAE/REF in the international context.

7.1 PRFS and their function in research governance

Performance-based research funding systems are intrinsically linked to the structure of the research system, the national approach to research governance and the policy context. As a consequence, PRFS are country-specific and respond to specific policy objectives, even though some common characteristics can be identified.

Historically, many PRFS can be seen as part of the New Public Management model, under which governments have launched major reforms in an effort to make public management more efficient and accountable. In the field of research, more autonomy has been granted to the research-performing institutions in exchange for accountability. Policymakers correspondingly steer research using funding as an incentive rather than controlling it through detailed prescription. Different approaches emerged. In some countries, performance contracts have been established, linking institutional funding to commitments about future performance. In others, PRFS were set up, acting as a reward for past performance.

The rationale for PRFS in Norway

In the beginning of the 2000s, evaluation studies had repeatedly pointed to failures in the system in terms of a lack of critical mass, lack of mobility, scientific inbreeding, and fragmented relationships in the national research system. Higher education institutions were operating in traditional ways, often separately from the applied research institutes, and with limited input or interaction with Norwegian or international industry.

In 2002, the Norwegian government tackled these systemic issues through two major policy interventions: on the one hand, the public research actors were granted a higher level of autonomy; on the other hand, some of the institutional funding for research was allocated in competition. More open competition for funding based on quality and relevance was expected to lead to improved quality and a more 'dynamic' division of labour in the research system. A performance-based funding model for the higher education sector was therefore introduced and fully implemented in 2006. Key objectives were to boost excellence in research and to act as an incentive for the higher education sector to look for external funding. A similar model was piloted also for the research institute sector in 2009, in the context of a broader overhaul of the institute funding system.

Source: (Arnold & Mahieu, 2015)

Policy objectives directly influence core design elements of the system such as the focus and scope of the assessment, the choice of assessment criteria and their indicators, and the choice of funding criteria and mechanisms. Purposes vary (Table 16) but four main categories of policy objectives can be observed.

1. To enhance the quality of research and the country's research competitiveness
2. To steer behaviour in order to tackle specific failures in the research system
3. To strengthen accountability
4. To provide strategic information for research strategy at institutional and/or national level

All countries use PRFS in an effort to increase the quality of research. Most also aim to trigger other behaviours, in line with policy priorities or a perceived need for change in the national research system. Objectives include: fostering critical mass; enhancing research-industry collaboration and the knowledge transfer; identifying or directing funding toward areas of research strength and emerging areas of research excellence; and strengthening the international competitiveness of research. (NZ Ministry of Education, 2012). Some also seek more accountability-related objectives, to stimulate efficiency in research activity and to demonstrate that investment in research is effective and delivers public benefits (Abramo, D'Angelo, & di Costa, 2011).

While rewarding of performance is an intrinsic characteristic of all PRFS, only in few countries does this selective distribution of funding aim also at a concentration of resources. This was originally the case in the United Kingdom. It constituted part of a government response to limited resources and the increasing costs of research (see Section 3.1, above). The goal was to maintain research excellence but in fewer places (OECD, 2010). Sweden has also been looking to concentrate resources through a revision of the national PRFS, trying to foster more peaks of excellence in a system whose average level of quality is already quite high (Swedish Research Council, 2015). However, the proposal was dropped from the subsequent Research Bill so the old system has been retained.

Other countries, instead, aim to strengthen research capacity in the weaker parts of the system rather than focusing money on 'excellent' researchers or institutions. The main purpose of the performance-based allocation system in Norway, for example, is to enhance the quality of research by motivating institutions to increase their research activities and by distributing resources according to research results. The main winners have been the provincial institutions (Aagaard, Bloch, & Schneider, 2015). The Italian system tries to foster excellence without changing the regionally distributed structure of the university system (Abramo, Cicero, & D'Angelo, 2011), and Belgium (Flanders) established its performance-based funding mechanism with a clear intent to distribute research funding on a wide basis (Mahieu B. , 2015).

Table 16 Main PRFS policy objectives of PRFS in the comparator countries

	Quality of research	Systemic factors	Accountability	Strategic intelligence
Australia	Identify excellence across the full spectrum of research performance	Enable comparisons of research, nationally and internationally, for all discipline areas	Give government, industry, business and the wider community assurance of the excellence of research conducted in higher education institutions	Provide a national stocktake of discipline level areas of research strength institutions Identify emerging research areas and opportunities for development
Belgium /Flanders (BOF)	Stimulate scientific performance and the quality of research	Create an incentive for technology transfer	Make quality visible	
Czech Republic	Reward research quality		Make quality visible to the national and international public	
Denmark	Improve quality, increase productivity and enhance efficiency in HE research	Sharpen international profile and international competitiveness of HEI		
Finland	Strengthen research quality	Sharpen international profile and international competitiveness of HEI	Demonstrate to the public that research funding is spent optimally	Assist the institutions to fulfil strategic goals and priorities
Italy (VQR, 2011)	Provide an objective and rigorous assessment of research	Define a national ranking per scientific area and 'structure' typology		Allow for a comparison of the national research quality with the quality in the major industrialised countries
Netherlands (2015)	Reveal and confirm the quality of research		Reveal and confirm the relevance of the research to society	Improve quality and relevance of research where necessary
New Zealand (2013)	Increase the quality of basic and applied research Support world-leading research-led teaching and learning		Provide robust public information to stakeholders about research performance within and across tertiary education organisations	Assist New Zealand's tertiary education organisations to maintain and lift their competitive research rankings relative to their international peers
Norway	Strengthen research capacity	Enhance co-operation among the different actors in the research system and knowledge transfer		
Sweden	Assess the quality of research	Stimulate HEI's to find a profile where they have a competitive advantage, which will help a clearer division of roles between HEI's, and increased specialisation		
UK (REF, 2014)	Reward research excellence		Produce evidence of the benefits of public investment in research	Provide benchmarking information and establish reputational yardsticks, for use within the higher education sector and for public information

Sources: (OECD, 2010) (Mahieu & Arnold, 2015)

National research assessment is not always linked to funding distribution. Countries that aim predominantly to allocate resources based on past performance use a PRFS. Others (in this report, Australia and the Netherlands) focus on informing research policies and institutional strategies but base the allocation of institutional funding on performance agreements between the universities and the responsible agency or ministry. An example of such an evaluation system is the Standard Evaluation Protocol, adopted in the Netherlands.

The Standard Evaluation Protocol (SEP)

In the Netherlands, the national research assessment is conducted in a bottom-up manner, guided by an evaluation protocol that was established jointly by the Royal Dutch Academy of Science (KNAW(and the research council (NWO) and the Universities. By law, all universities are obliged to conduct this 'bottom-up' evaluation every six years. The research organisations themselves are responsible for the assessment and research units have a significant influence on the process.

The main aim of research assessments described in the SEP is: "*to reveal and confirm the quality and the relevance of the research to society and to improve these where necessary*". Specific aims for various target groups are also further specified: researchers need to know the quality and societal relevance and their unit's strategy, and how these aspects can be improved; boards of institutions wish to track the impact of their research policy; government wants to know the assessment outcomes in relation to the institution's accountability and the government's efforts to support an outstanding research system; society and private sectors seek to solve a variety of problems using knowledge that research delivers.

The boards of the universities and institutes are responsible for planning the assessments. The SEP provides guidelines on review procedures, criteria and scoring in order to ensure coherence among assessments. However, guidelines are broad enough to ensure applicability to all disciplines, and the exact terms of reference are established by the boards, using the research units' strategy and targets as guiding principles. The boards are also free to define the level of aggregation of the evaluations, and in consultation with the research units under evaluation they establish a Peer Review Committee and appoint its members. The boards are also responsible for following up the review committee's recommendations, with no specific guidelines being provided by the SEP. The committee's recommendations are to be used to improve the research at the institutions and do not directly affect external funding.

Based on (Mahieu B. , 2015)

7.2 Key design choices for evaluation in PRFS

The methodologies adopted for research assessment differ, reflecting different government priorities. Policy objectives drive the approach to assessment, the number and type of indicators selected, and the relative weight placed on each indicator or assessment criterion in the construction of the final score for the allocation of the institutional funding.

Looking at the various designs used in the PRFS studied, we have assembled a list of parameters available to PRFS designers. Table 13 lists these and they are discussed in more detail below. Many of these elements are interlinked. In particular, the overarching model chosen for the assessment (ie peer review, bibliometrics or a combination of both), drives the granularity and the periodicity of the PRFS.

Table 17 Key design parameters for the assessment component in PRFS

Key design parameter	Variations
Model used for the assessment of research quality	• Peer review-based • Informed peer review • Mix of peer review & bibliometrics • Metrics-based
Scope of research activity included	• Research • Innovation • Societal relevance
Type of indicators	• Output indicators • External funding indicators • Systemic indicators • Outcome/impact indicators
Assessment criteria in peer review-based systems	• Quality of outputs • Relevance of research activities • Institutional environment • Esteem measures
Granularity	• Units of analysis (grouping of scientific disciplines) • Inclusion of individual staff (inclusive/exclusive)
Periodicity	• Annual • Longer time frames

Based on (Arnold & Mahieu, 2015)

7.2.1 Overall 'model' used for the assessment of research quality

While all PRFS aim to assess the quality of research, Hanne Foss Hansen notes in her discussion of performance indicators that the *concept of performance remains ambiguous* (OECD, 2010). There is an ongoing debate in the academic community about the definition of quality in research outputs and the gap between outputs and their effects.

Closely related to this difficulty of defining what constitutes 'quality' or 'excellence' in research is the discussion about the choice of the methods to assess research performance, ie peer review or metrics.

Peer review versus bibliometrics

The choice between metrics or peer review is contentious. On the one hand, metrics-based systems typically encounter criticism from the research community on the grounds that metrics provide imperfect measures of quality. The community tends to prefer peer review, thanks to its flexibility and

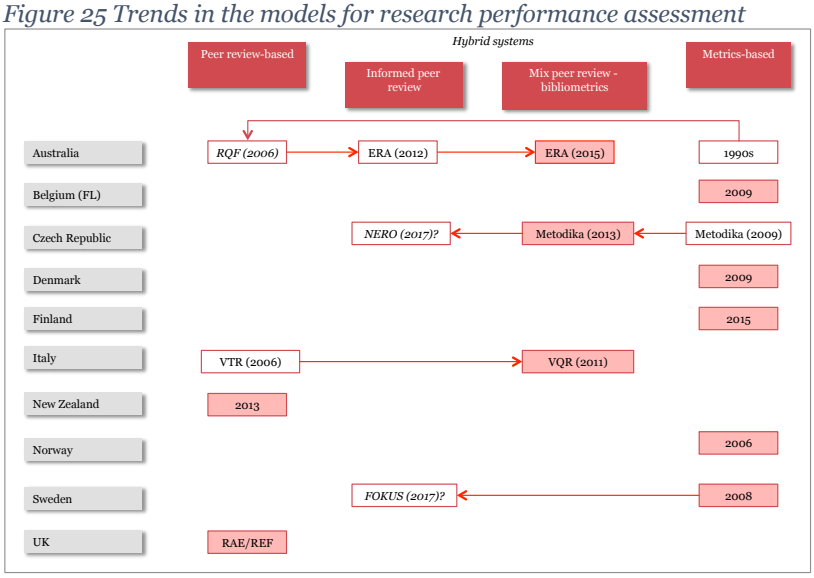
its ability to assess a wider range of research outputs and research-related activities. PRFS that use peer-review are more comprehensive and appear to have greater credibility and sector buy-in, meaning that PRFS have status incentives as well as funding incentives (NZ Ministry of Education, 2012). The recent proposal for a new PRFS in Sweden involved a shift from metrics to peer review (Swedish Research Council, 2015).

Cost, however, is a major factor. PRFS that rely entirely on metrics are generally considered to be less expensive to administer and less compliance-heavy than systems that use peer reviews, which are seen as cost-intensive and time consuming. Our sample of metrics-based PRFS suggests that policymakers generally fail to adopt many of the more sophisticated indicators the bibliometricians can provide. Use of journal impact factors is widespread, despite the growing understanding that these are inappropriate as indicators of the quality or impact of individual articles.

Current international practice

PRFS are dynamic. A historical analysis of the PRFS in scope to this study indicates a continuing search for improvement (Figure 25). Factors leading to change include pressure from the research community, developments in evaluation methodologies and concepts and a search for an improved cost-benefit ratio. In Italy and Sweden, a major driver was an envisaged increase in the proportion of institutional funding to be allocated based on the assessment results and a concomitant desire for more reliable assessment methods (Mahieu B. , 2015).

Figure 25 Trends in the models for research performance assessment



Notes: The Australian Research Quality Framework (RQF) that was proposed in 2006 was an RAE-like system, based on peer review panels, but also including end-user assessments of impact on the economy and society. The system was criticised for being costly and non-transparent and when a new government took over in 2007, it was never implemented. Both in Sweden and the Czech Republic, the PRFS system is currently being reviewed (Mahieu B. , 2015)

Current international practice is as follows.

The UK and New Zealand are the only two countries relying close-to-uniquely on peer review. The UK REF allowed the use of informed peer review in cases where panels desired it by accessing citation counts and contextual analysis to help clarify citation behaviour and patterns in the relevant field. Use of journal impact factors and other bibliometric indicators not supplied through the REF administration was forbidden (HEFCE, 2012) (HEFCE, 2012a). One panel (Computer Science and Informatics) had planned to use Google Scholar data as a way to capture more of the conference activity that is central to the way

that field communicates, but was defeated (according to HEFCE) by its by inability to harvest the needed information from Google[41]

Belgium, Denmark, Finland and Norway use bibliometrics for the assessment of research quality in the PRFS. All of them use the "Norwegian model" for the bibliometric indicators, extending the data coverage of the international, commercial bibliometric databases (Web of Science and Scopus) by means of country-specific publication databases. It should be noted that in the Scandinavian countries, the PRFS is complemented with system-based or discipline-based national evaluations using informed peer review that have a more formative character.

An increasing number of systems combine the two approaches. The 'mixed peer review-bibliometrics' model uses both bibliometrics and peer review. In Italy for example, the latest VQR (2011) used an informed peer review process for the greater part of the funding, based on outputs submitted by the research organisations as well as a self-evaluation, but complements this with indicators to allocate the balance. The results of the peer review and bibliometric exercises are used separately to allocate units of assessment into broad quality bands, and the combination of these bands with the volume of output then drives the funding provided. In Australia the ERA 2015 uses a broad range of assessment tools, including bibliometric and non-bibliometric indicators, as well as peer review. Crucially, these are not all used equally across all disciplines. Citation analysis is used more extensively in the sciences and peer review is used more extensively in social sciences, humanities and computing.

The 'informed peer review' model uses bibliometrics to inform the peer reviewers - to varying degrees and at the peers' discretion. This exploits the ability of indicators to represent large sets of data while exploiting the ability of peers to make more qualified judgments about excellence, coherence and other qualitative aspects that cannot be achieved through indicators alone.

In the Netherlands, Sweden and Austria, peer-review based evaluations are organised at the institutional level with no direct linking to funding.

Internationally, the trend is away from UK-style peer-review based exercises and towards greater use of bibliometrics and simple indicators. But this is mostly happening in the context of systems that (re)allocate a lower proportion of institutional funding than the REF, so that they can afford to be a little cavalier with their judgements. If (as in the UK) most of the institutional funding is at stake, then it will clearly be more heavily contested and the academic community will itself drive up the costs (including to itself) associated with precision.

7.2.2 Scope of research activity included

Over time, there is a clear trend to increase the scope of research assessment in the context of PRFS. While early PRFS focused only on scholarly outputs, the current trend is to encompass also aspects of innovation and the universities' 'third mission' of knowledge exchange with society. Increasingly, assessments address research performance also in terms of its impacts on research, innovation and society at large (Figure 25).

Most PRFS, no matter whether the assessment is peer-review or metrics-based, use indicators beyond academic outputs to measure performance. In practice, however, PRFS have not attempted a heavily metrics-based treatment of wider societal, cultural or economic impacts of research. Impact presents a significant challenge to research assessors, primarily because there often is a long time delay between publication and any social impact. Impacts are therefore predominantly assessed indirectly, ie by using proxy indicators such as the universities' capacity to gain external research funding (see Section 7.2.3, below).

The UK REF 2014 stands out as the first major concerted attempt – in a PRFS - to demonstrate research impact in a systematic way across all disciplines (by means of narratives). Other national evaluation systems that are not linked to institutional funding have made similar attempts. An example is the Dutch

[41] http://www.ref.ac.uk/about/guidance/citationdata/googlescholar/#d.en.78940

Standard Evaluation Protocol where the self-evaluation component of the exercise requires the universities to provide narratives on the societal relevance of their research activities and outputs, albeit not at the rigorous and detailed level of the REF 2014. A similar approach is being attempted also in some disciplinary evaluations organised at the national level in Norway.

7.2.3 Indicators used

Indicators used in the context of PRFS can be grouped into three categories: indicators directly assessing research outputs; external funding indicators; and systemic indicators (Table 18). The two last categories act as proxies.

Table 18 Indicators used in PRFS

		Belgium /FL (2009)	Czech Republic (2013)	Denmark (2009)	Finland (2015)	Italy (VQR, 2011)	New Zealand (2003)	Norway (2006)	Sweden (2008)	UK (REF 2014)
Output indicators	Academic outputs	√	√	√	√	√	√	√	√	√
	Non-academic outputs		√			√			√	√
	Innovation-related outputs (IPR)		√			√				√
External funding indicators	Competitive funding / national			√	√	√	√	√	√	√
	Competitive funding / international			√	√	√	√	√	√	√
	Contract research funding			√		√	√		√	√
	Non-competitive funding			√		√	√			√
Systemic indicators	Esteem (conferences, editorships, rewards etc)					√				√
	Collaborations / national					√		√		√
	Collaborations / international					√				√
	International mobility				√	√				√
	Collaboration research-industry					√				√
	PhD recruitment/awarding	√		√		√	√	√		√
Outcomes/impact indicators	Academic impacts (citations)	√	√			√			√	
	Socio-economic outcomes/impacts (eg spin-offs)					√				√

The degree of emphasis given to each of these indicator classes depends on the function of the evaluation and its policy objectives.

- The 'direct' assessment of research productivity and research quality or excellence is a topic of importance in all PRFS but it is particularly stressed in those research systems where the lack of productivity is cause for concern (eg Italy, the Czech Republic) or where the need is felt for a focus on excellence (eg Sweden, the UK)
- Several PRFS directly assess also the productivity, quality and value of non-academic outputs and innovation-related ones such as patents
- Most countries use the universities' capacity to obtain external funding as a proxy indicator for research quality and/or relevance. To an extent, this compensates for the difficulty of assessing the socio-economic impacts of research. Competitive project funding is considered an indication of the

university's scientific competitiveness and thus acts as a proxy indicator for *research quality*; contract research funding stands (also) as an indication of the university's research alignment with the needs of society (as expressed by industry or the public sector) and/or national research priorities and acts as a proxy indicator for *impact on innovation*

- Several countries also use a series of systemic indicators assessing the universities' broader research *activities*, as proxy indicators for quality or relevance for the national research or innovation system. Most of these treat the number of PhDs awarded as an indicator of the university's contribution to the strength of the national research system. Indicators related in the 'pathways to impact' (research esteem, collaborations, etc) relate to the dynamics and interactions between the various elements of a research system (actors, context, infrastructures, etc). These are focus points especially for those systems where a major intent is to *steer* research behaviour in order to overcome specific systemic failures (eg in Norway an enhanced institute-HEI collaboration), enhance internationalisation (eg in Finland and Norway) or foster the socio-economic relevance of research (eg enhancing science-industry collaboration in Italy)

In a few countries, citations are considered to provide proxies for impact on the research base. The use of indicators to assess socio-economic impacts is rare, perhaps because these are easily 'gamed' (Good, Vermeulen, Tiefenthaler, & Arnold, 2015)

In Table 15, above, the UK REF and the Italian VQR stand out for the breadth of the indicators they use, covering all indicator categories and focusing more than other PRFS on the use of systemic indicators.

In addition, the UK REF pays considerable attention to the specifics of the different scientific disciplines. For example, the sub-panels are allowed to identify the types of academic and non-academic outputs they considered to be relevant to their discipline. In addition to papers in peer-reviewed journals, conference proceedings and books, the panels indicated *sixteen* other types of academic output that could be submitted, ranging from technical reports to textbooks. They also accepted the submission of in total *nineteen* types of non-academic outputs. These included nine types of 'digital artefacts' (such as software, archives, films etc.), seven types of 'physical artefacts' (eg new materials or prototypes), and three types of 'temporary artefacts' (exhibitions, performances, and 'additional' outputs).

In relation to the indicators related to the REF's 'environment' criterion, the sub-panels could formulate field-specific variations of the information requested. The indicators for this assessment criterion were grouped into six categories: research strategy; environment; people; research students; income, infrastructure and facilities; and collaboration and contribution to the discipline or research base. This included several 'esteem' indicators, and in some sub-panels, information related to inter-disciplinary research. The emphasis was on narratives and descriptions rather than quantitative data.

7.2.4 Assessment criteria in peer review-based systems

Metrics-led research assessments constitute a form of assessment that is numerical and measurable from the point of observation onwards. Peer review requires at some point a transition from qualitative and often holistic expert observation to numerical grading or ranking. This transition requires trust in the expertise and judgement of the reviewers. Typically, complex guidelines and criteria are put in place to minimise the role of subjectivity and systematise peers' qualitative judgements in a detailed and rigorous fashion. Broadly, the processes are the similar in the different countries; approaches differ only at a detailed level.

Four of the PRFS we consider in this report build at least partly upon input from peer reviewers, ie the Czech Metodika, the Italian VQR, the PBRF in New Zealand, and the REF in the UK. Two other peer review-based national evaluation frameworks, the Australian ERA and the Dutch SEP, do not inform funding allocations.

In the Czech Republic and New Zealand, the panels focus only on assessing the quality of the research outputs. In Italy, the UK, the Netherlands and Australia, instead, the evaluation panels consider various aspects of research activities. The assessment criteria against which panel members are expected to formulate a final judgement are broadly similar. The 'pure' peer review-based systems (SEP and REF)

consider the institutional environment (based upon narratives and self-assessments by the universities) while the mixed peer review/metrics systems (ERA and VQR) focus only on a range of esteem measures.

Table 19 Assessment criteria in national evaluation frameworks involving peer review

	Australia – ERA 2015	**Netherlands – SEP 2015**	**Italy – VQR 2011**	**UK REF 2014**
Outputs	Volume and activity; publishing profile; peer review; citations; research income	Research quality	Originality & innovation	Originality, significance and rigour
Relevance/ impact	Applied measures (IPR & research commercialisation)	Relevance to society	Relevance for the advancement of knowledge & social benefits	Reach and significance
			Technology transfer activities and (potential) socio-economic fallouts	
Environment		Viability		Vitality and sustainability
Esteem	Esteem measures (at eligible researcher level)		Internationalisation and/or international standing	

A key issue in a PRFS is the weighting that is given to the different assessment criteria for the definition of the final score. These weightings depend on the policy priorities and objectives of the PRFS. We cover this in Section 7.3.3, below.

7.2.5 Granularity

7.2.5.1 Unit of analysis

The unit of analysis in an evaluation can be the individual researcher, a research group (field defined), the faculty, the department, or the institution. It is a fundamental component in the design of research assessment systems. The selection of the most appropriate unit of evaluation is to a large extent dependent on the purpose of the research assessment: in evaluations that do not drive funding, the deciding factor is the most suitable level for the collection of the information that is required for policymaking or governance; in PRFS it is driven by the level at which the funding is allocated.

There are various sets of considerations that leave most systems with the task of balancing emphasis on institutions on one hand and scientific field or research group on the other. A major factor in this regard is the evaluation model. In practice, the tension between complexity and practicality means that while research groups are theoretically the ideal unit of evaluation, departments or universities are usually the focus of PRFS (OECD, 2010).

In general, the methods used in assessment tend to correlate with the choice of focus: peer review is used for departmental evaluations, while metrics-based systems are used for university-level evaluations. In fact, bibliometrics and panel evaluations are variably flexible from this perspective.

- **Metrics-only systems** can collect data about outputs at the level of the individual researcher, as is most obviously done in the countries that operate a national research information system for this purpose, such as Norway and the Czech Republic. To our knowledge, nobody allocates institutional funding at this individual level. Normally, outputs are aggregated to the *organisational level* and used to determine the institutional funding for the research organisation as a whole. It is possible to aggregate results also to the level of individual groups or faculties (and some research-performing organisations appear to run shadow systems in order to do this). For the external funders directly to allocate institutional funding to intra-organisational entities would challenge the principle of the autonomy of universities and other research organisations, so this is not done. As a result, the principles and patterns of allocation of institutional research funding within universities are very much a black box – leaving significant uncertainty about how PRFS actually impact aspects of day to day operation of research. A rare insight is provided by a study of internal allocation patterns in Swedish universities, which reveals diversity of behaviour, not all of which is consistent with attaining the goals of the PRFS (Fridholm & Melin, 2012).

- **Peer review systems** could have this same flexibility only if they had the capacity to assess all the output produced by each individual researcher – which is impossibly resource intensive and is rarely done in PRFS. To our knowledge, the only country adopting such system is New Zealand. It is also done in specialised systems that assess individual researchers as part of a promotion or award process (as in current Spanish and German systems). Assessment at the *departmental or discipline level* is most common under peer-review systems. This recognises the collegial nature of much research activity and the importance of collaborative research arrangements (NZ Ministry of Education, 2012)

7.2.5.2 The inclusion of individual staff

Assessment systems take different approaches to the question whether and how individual researchers should submit a selection of their research outputs for review. Broadly, there are two approaches.

- The evaluation is comprehensive, i.e. all researchers at an institution are included. Systems that take this approach typically specify clear inclusion criteria, including most often a minimum level of professional attachment to the institution (eg at least 20% full-time equivalent contract in New Zealand). This approach allows a relatively representative overview of the outputs, quality and/or impact of research within the unit of analysis. Table 20, below, shows that this approach is taken in Australia, Finland, Italy and New Zealand
- In the UK RAE/REF, the universities are expected to identify a smaller selection of their 'best' researchers who will then submit their work. This reduces the burden on the evaluators, as the overall amount of submitted work is smaller. This approach does not give a representative view of all research activity that has occurred in an evaluated unit but instead indicates the maximum standard of which that the unit is capable, in the view of those who prepare the submissions

There are weaknesses in both approaches: comprehensive inclusion of staff may for instance obscure the presence of a select few outstanding researchers in a unit whose level is otherwise ordinary, whilst selection of the best examples may obscure that relatively poor quality research is also being done.

Some evaluation systems avoid bringing the level of assessment down to the individual researcher level. Research outputs are submitted for peer review based upon a selection of the work at the level of research group or department, not at the level of the individual researcher.. Even though the assessment of individuals is not an objective in any of these cases, the inclusion of results at the individual level has effects on career prospects and the R&D system as such. The UK REF 2014 considered the work of research groups rather than individual researchers. Nevertheless, it still required the 'nomination' of individual researchers, so the effects of this change remain to be seen.

Table 20 Individual researchers included in peer review-based evaluations

	Australia	UK	Italy	New Zealand
All academics	Yes		Yes	Yes (for submission, not evaluation)
Selected sub-group of academics		Yes		
Criteria for inclusion	There are detailed eligibility criteria although the objective is comprehensive submissions	Units select the academics to be included. Beyond that: 'Category A' specifications must apply, special circumstances for early career researchers; minimum 0.2 FTE	Academic staff (gradations apply depending on rank/ seniority)	All eligible academics are included in the assessment of the TEO. The EO only excludes the academics (with low score) that do not add value to the overall score of the university department

7.2.5.3 Handling the differences among the scientific disciplines

Differences among fields present a major challenge for research assessment. Two models emerge in international practice.

- Some bibliometric-based systems try to overcome field differences in publication patterns by introducing a system of weights that balance the differences in publication patterns and by presenting scholarly publications with complete data from research information systems, in order

to compensate for the differences in the coverage of scientific fields in the commercial data sources, ie Web of Science and Scopus

- Peer review-based systems like the UK RAE/REF, instead, solve the issue by using discipline-based panels and units of assessment

In peer review based systems this generally means that different fields are not put into direct competition with each other at the level of assessment, though the funding formulae may do so. Mechanisms are put in place which aim to ensure that peer review panels all use assessment scales that have the same meanings, irrespective of discipline. Putting different fields into competition within metrics-based approaches means that the designer of the assessment has to produce some sort of bibliometrics-based technique for comparing across fields without, in reality, having a deep understanding of what these inter-field differences are.

A topic closely linked to the choice of method is the ability suitably to assess and reward interdisciplinary research. Whichever approach is used, assessment systems – like wider evaluation systems or systems for deciding on the quality of research proposals in research councils and other funding agencies – struggle with interdisciplinary research. Peer panels may have to cooperate to address it. Indicator-based approaches do not have a good frame of reference for interdisciplinarity precisely because the statistical pictures they use to describe 'good' research are defined within disciplines. As in other situations where it is hard to define very precise rules, there is virtue in using (peer) judgement to find solutions to these problems.

- Overall, interdisciplinary researchers often publish in a range of journals and their outputs are spread over several different fields, thus weakening a university's claim to have concentrations of excellent research. A risk is that more narrowly specialised research is encouraged at the expense of cross-disciplinary innovation.
- Interdisciplinary work also poses a challenge in peer-review based systems as peer review requires grouping panellists into areas of expertise. Here the structuring function of disciplines required for peer review clashes directly with work that seeks to transcend these structures.

In most of the discipline-based evaluations (in the UK, Australia and NZ), interdisciplinarity is tackled by defining a broader disciplinary focus for the panels, thus allowing for the consideration of interdisciplinary research among sub-disciplines, and if necessary via cross-referrals among different panels. In Australia, submissions at area rather than field level are allowed, thus making research from a mix of reasonably closely related fields un-problematic. Panel members may be assigned between the various panels to bring appropriate expertise to bear on the evaluation.

This approach goes some way to tackle the problem of assessing interdisciplinary research. However, simply conducting reviews according to a 'mix' of the contributing fields does not recognise that such research may not just combine but transcend established norms in any of those fields. Moreover, it does not solve the issues around metrics use for interdisciplinary work discussed above, at least where metrics are designated as an assessment tool in contributory fields.

This problem has been tackled in the case of Italy's VQR: here, every research output submitted needs to be accompanied by several pieces of information, including a statement indicating whether it should be regarded as a product of an established discipline, a newly emerging area or that is has an inter-disciplinary character. In addition to 14 disciplinary categories, the VQR has six additional inter-disciplinary ones. Though the VQR has a broad range of assessment tools, in the inter-disciplinary sectors peer review is preferred to metrics owing to the limited availability of bibliometric indicators and the other known difficulties in assessing research in emerging and inter-disciplinary areas.

7.2.6 Periodicity

The frequency with which research activities are assessed varies (Table 21) tends to be influenced by system design and cost considerations. Collecting quantitative information is generally easier than peer-review evaluation, so information measured by metrics is collected more frequently – in contrast to peer review-based systems which are more resource and time intensive and therefore can only be conducted occasionally.

As a consequence, the information upon which PRFS are based is more up-to-date in metrics-based systems. The shorter time lag between evaluation and funding allows these systems to be more responsive to changes in policy objectives and in the research system at large. In countries with PRFS that have a more extended interval such as the UK REF, policymakers can only periodically use the PRFS to understand the research performance of the research organisations, though they may choose to use other approaches to provide more timely information.

Table 21 Periodicity of PRFS

Country (PRFS, start date)	PRFS type	Years between exercises
Australia (ERA, 2010)	Mix peer review/metrics	2
BE – Flanders (BOF, 2003)	Metrics	1
Czech Republic (Metodika, 2013)	Mix peer review/metrics	1
Denmark (2012)	Metrics	1
Finland (2015)	Metrics	1
Italy (VQR, 2011)	Mix peer review/metrics	3
New Zealand (2013)	Peer review	3, 6
Norway (2006)	Metrics	1
Sweden (2008)	Metrics	1
UK (RAE, 1986)	Peer review	5 - 7

7.3 PRFS driving (part of) institutional funding

The proportion of institutional funding affected by the research evaluation is a key aspect of PRFS. Specific policy objectives guide this proportion and the nature of the formula used for the funding allocations. These funding principles evolve over time and influence the design of the evaluation component. PRFS driving a large proportion of institutional funding require greater robustness and community buy-in than those that inform only a small proportion.

PRFS are components of the overall research funding system. A key policy question is what proportion of that funding should be awarded in competition (Section 7.3.1). Second, a PRFS is only one of the tools available for institutional funding. Others include unconditional block funding, formulae that inform the institutional funding for teaching, and performance-based funding based upon a performance agreement (Section 7.3.2). Policy decisions define the shares of these funding components. Finally, policy objectives also determine the weights allocated to the assessment criteria and indicators used in the PRFS, thus determining their influence on universities' institutional funding (Section 7.3.3).

7.3.1 Share of competitive versus institutional funding

Historically, the trend has been for countries to move from 100% institutional funding of research to a mix of institutional and project funding and then finally to making part of the institutional funding performance-based. In the 1990s, higher education funding for research divided in many countries between 20-30% project funding with institutional funding making up the balance, but the share of project funding has risen since then (van der Meulen, 2001) (Lepori, et al., 2007).

Different countries choose different balances between institutional and project-based funding of research. In principle, the project-based component involves competition and therefore induces quality as well as providing a way to reallocate resources towards excellent research. The institutional component provides stability: it makes it possible to take decisions about things like the number and size of research-performing entities (for example, it may not be desirable for all research in a field to be done in one place), supports the provision of infrastructure and enables exploration and capacity-building by providing resources that are sheltered from competition. There appears to be no 'ideal' mix;

rather, the objective is to maintain a balance between competition, stability and restructuring that ensures quality while keeping the research system up to date and relevant.

Table 22 shows the ratio of institutional funding for research to project funding for research in the universities of thirteen countries. The UK, French, Flemish and Irish systems stand out as providing 50% or less of institutional funding for research to the universities.

Table 22 Share of GUF versus direct government funding of R&D expenditure in the higher education sector, 2009 (in millions of €)

	PRFS used	Government sector funding	General university funds (GUF)	Government competitive research project funding
Austria		1,669	76%	24%
Belgium / Flanders	√	1,117	36%	64%
Denmark	√	1,653	72%	28%
Finland	√	1,033	58%	42%
France		7,972	50%	50%
Germany		7,575	71%	29%
Iceland		55	51%	49%
Ireland		704	31%	69%
Italy	√	5,204	85%	15%
Norway	√	1,380	73%	27%
Spain	√	3,012	66%	34%
Sweden	√	2,041	57%	43%
Switzerland		2,000	82%	18%
United Kingdom	√	5,545	48%	52%

Source: Statistics from Eurostat. The countries covered in this chapter are highlighted in grey

7.3.2 PRFS as a component of the institutional funding systems

PRFS are only one of the tools available for providing universities with institutional funding. Others are negotiated or historically determined block grants, formulae for teaching funding, and performance contracts. The use of different components in the institutional funding mix makes it possible to combine elements of continuity and stability with incentives for desirable changes in behaviour. Performance contracts bring in the possibility of developing and reinforcing future elements of institutional strategy.

Table 23 shows that in many institutional funding systems a considerable proportion remains unconditional. PRFS are generally seen as useful ways to add an element of competitive pressure to institutional funding – but only at the margin so most PRFS provide only a small component of the overall funding for research and higher education. Only Finland, Denmark and the UK allocate more than 20% of the institutional funding through PRFS. Our impression is that the systems that only drive a small part of the institutional funding can nonetheless be effective at changing behaviour and performance. In the case of Norway there is clear evaluation evidence to this end (Bloch & Schneider, 2016).

Table 23 Current university institutional funding systems

Country	Block grants (teaching & research)	Formula teaching funding	PRFS	Performance agreements (PA)	Comments
BE – Flanders (2011)	17%	69%	13%		
Czech Republic (2013)	80%	13%	7%		
Denmark (2012)	60%	18%	22%	√	PA has no link to the block grant
Finland (2013)	25%	31%	26%	√	Targets for the block grant are set in the PA
Italy (2014)	82%	8%	10%	√	PA drives 1% of the block grant
New Zealand (2003)	90%		10%		
Norway (2015)	60%	25%	15%	√	PA drives 5% of the block grant
Sweden (2012)	90%		10%	√	PA has no link to the block grant
UK (England) (2013)	13%	51%	23%		

Source: Relevant national authorities

As shown in the Table above, in most cases performance contracts are linked to the block grant setting targets and driving (part of) the funding. Finland provides a prime example for this model of governance.

Finland's' model of governance

In Finland, the institutional funding for teaching and research is determined by a comprehensive development plan established by the government every four years covering the institutional funding for both teaching and research. In this model, performance contracts and the funding formula with performance indicators for teaching and research (ie the PRFS) act as complementary tools. At the beginning of every three-years term, negotiations are conducted and performance agreements signed between the universities and the ministry establishing operational and qualitative targets as well as the required resources. At the same time, monitoring and evaluation indicators are defined.

In 2015, the institutional funding was calculated based on a set of 17 indicators divided into the three areas: education, research and strategies/national tasks (Table 24). Each of the indicators contributes to one or more of three policy objectives: efficiency, quality and internationalisation.

Source: (Claeys-Kulik & Estermann, 2015)

Table 24 Institutional funding system in Finland

	Efficiency		Quality		Internationalisation	
Education (41%)	Completed graduate degrees	14%	Student Feedback	3%	Completed higher degrees of foreigners	1%
	Completed undergraduate degrees	6%	Graduates in employment	1%	International student exchange	2%
	The percentage of students who complete at least 55 credits	12%				
	Credits Rewarding (no degree-conferring) studies	2%				
Research (34%)	Completed PhD degrees	9%	Scientific publications	13%	Competitive Research Funding / International	3%
			Competitive Research Funding / national	6%	Completed PhD degrees of foreigners	1%
					Foreign academic staff	2%
Strategies / national tasks (25%)	Strategy-based financing					10%
	National features / functions					7%
	Subject specific financing - applies particularly for expensive disciplines such as engineering, science, medicine and art					8%

Source: (Claeys-Kulik & Estermann, 2015)

7.3.3 Policy objectives directing the institutional funding

The weights attributed to the indicators and assessment criteria used in the assessment component of the PRFS ultimately define the funding allocations. Table 25 shows the variations in the weights allocated to the different indicators, illustrating the influence of the policy objectives for the PRFS set out in Section 7.1, above.

The PRFS in the UK, New Zealand and the Czech Republic stand out for the high weight attributed to the results of the research quality assessment, while the systems in Finland, Norway and Belgium (Flanders) give high importance to the effects of the activities on the research system. Italy stands out for its focus on innovation-related outputs and activities. Denmark and Sweden give higher than average weight to the ability of the universities to attract external funding for research.

The usefulness of the (international) trend towards using external research funding levels to drive part of institutional funding is an unresolved issue. This seems likely to reinforce existing resource hierarchies or even further concentrate resources, potentially at the expense of wider capacity building in the research system. On the other hand, it does provide a way to channel funding for the overhead and infrastructural costs of external research towards those universities that incur them.

Table 25 Weights allocated to indicators and assessment criteria for funding within the national PRFS

	Research quality	Institutional/national research system			Socio-economic environment		External funding	Other
		Internationalisation	PhD students / completed	Environment	Innovation outputs/ activities	Impact		
BE – FL (2011)	39%		35%					27%
Czech Rep.(2013)	85%				15%			
Denmark (2012)	45%		18%				36%	
Finland (2013)	38%	9%	26%				26%	
Italy (2014)	31%	12%	6%		38%		6%	6%
New Zealand (2003)	60%		25%				15%	
Norway (2014)	30%	20%	30%				20%	
Sweden (2012)	50%						50%	
UK (England) (2014)	65%			15%		20%		

Source: Relevant national authorities

Table 26 shows the proportion of universities' institutional funding for teaching and research influenced by the PRFS and the weights allocated to individual indicators and assessment criteria. The UK stands out as the country where the assessment of research quality influences more than 10% of the universities' overall institutional funding. This is the result of the high weight set on the research quality criterion in the RAE/REF (Table 25), combined with the high share of overall institutional funding guided by the PRFS (Table 22). In addition, one should take into account that in the RAE/REF, the purpose of selectivity and fostering 'excellence' has led to the use of a *non-linear* calculation of the institutional funding allocations, skewing rewards towards those with the strongest scores. The other PRFS have a linear relationship between the production of quality publications or any other indicator/assessment criterion and monetary rewards.

Table 26 Proportions of overall institutional funding driven by PRFS indicators and assessment criteria

	Research quality	Institutional/national research system			Socio-economic environment		External funding	Other	Sum
		Internationalisation	PhD students / completed	Environment	Innovation outputs/ activities	Impact			
BE – Flanders (2011)	5%		5%					3%	13%
Czech Republic (2013)	6%				1%				7%
Denmark (2012)	10%		4%				8%		22%
Finland (2013)	10%	2%	7%				7%		26%
Italy (2014)	3%	1%	1%		4%		1%	1%	11%
New Zealand (2003)	6%		2%				2%		10%
Norway (2014)	5%	3%	5%				3%		16%
Sweden (2012)	5%						5%		10%
UK (England) (2013)	15%			3%		5%			23%

Source: Relevant national authorities

7.4 The REF in international context

The analysis set out above shows that the REF performance-based research funding system differs from the PRFS in other countries in several aspects.

Like other systems, the REF aims to foster quality in research as well as societal relevance. However, the emphasis on research quality is particularly strong. The REF 2014 attributed a weight to the quality of research outputs in the funding formula that was considerably higher than in most of the other countries and second only to the Czech Republic. The high weight given to research quality in the REF combined with the great influence of this factor on institutional funding for universities means that the UK stands out as the country where the assessment of research outputs influences more than 10% of universities' overall institutional funding. The average in the other countries is around 5%.

The REF is an integral part of a research funding system in the UK that is strongly competition-based. The UK is among the few countries where institutional funding accounts for less than 50% of overall public funding for research. The overall institutional funding for UK universities (for teaching and research combined) contains a particularly high share of funding driven by the PRFS. Only in Finland and Denmark does the PRFS govern such a large proportion of institutional funding overall. In contrast to the UK, however, these countries allocate a high share of the funding through block grants (supporting stability and longer-term strategies). They also make use of performance agreements, adding a prospective element to the backwards-looking role of the PRFS.

The strong influence of the REF on the universities' funding brings a particular requirement for its assessment system to be accepted as legitimate by the research community. It is one of the few PRFS where peer review is central to the assessment exercise. Panels in the REF are informed in their judgement by a particularly large amount of information, including both quantitative data and, especially, narratives. Only the Italian VQR uses a similarly broad range of indicators.

The REF 2014 also involves the first major attempt in a PRFS to demonstrate research impact in a systematic way across all disciplines. Through the organisation of its panel system it pays a particularly high level of attention to field-specific characteristics and interdisciplinary research.

8 Research Information Systems (RIS) and performance-based funding

Research information systems (RIS) play an increasingly important role for the governance of research at the institutional, funding body and national levels. Here we focus on the use of research information systems for the collection of information in the context of performance-based research funding systems. A specific point of interest is the extent to which current RIS in the UK can act as sources of information for the REF exercise.

In universities and research institutions the principal motivation for the use of RIS is a need for a consolidated base of information on the organisation's structure, staff, research activities, and research outputs to support reporting both inside and outside the organisation.

A research funder's research information system serves to handle the processes of receiving, evaluating (ex-ante) and selecting proposals to fund, grant administration, interim and final reporting and ex-post evaluation. It can also provide a consolidated base of information about the activities of the funder and the outcomes of funded work, supporting informed strategic planning as well as providing accountability.

At the national level, governments increasingly develop national research information systems that would constitute a consolidated basis for the collection and analysis of information related to research performance in the country as a whole and the identification of short, medium or long-term impacts. Research assessment and in particular the need for information informing performance-based research funding systems (PRFS) is among the drivers for government research information systems.

In this Chapter, we first provide a brief overview of the use of national RIS for the purpose of research performance assessments and funding allocations in other European countries. In the second section we look into the (potential) use of RIS data for the REF.

This Chapter is based on research conducted in the context of a study for the European Parliament (Mahieu, Arnold, & Kolarz, 2013) and a study designing a new R&D evaluation and institutional funding system for the Czech Republic, commissioned by the Ministry of Education, Youth and Sports (Arnold & Mahieu, 2015), as well as on additional desk research done for the purposes of this study.

8.1 National Research Information Systems in international practice

8.1.1 The use of RIS as sources of information for PRFS

National policymakers rely on data from research performing organisations and research funding bodies for building information to inform research strategies or decide on funding allocations. Historically, this data flow was – and in some cases still is – provided via ad-hoc evidence collection. Increasingly, however, there is a role for data stored in various research information and management systems.

Sources of information (beyond ad-hoc evidence collection) used for decision-making at the national level include the following.

- National statistical databases, collecting information on research staff and students by means of census / statistical surveys
- Research information systems set up by funding agencies, providing information on research outputs and outcomes/impacts (based upon reporting by the beneficiaries) in combination with information on the agencies' activities and their focus and level of research funding
- Bibliographic datasets, ie national databases holding meta-data on academic publications and/or acting as national open repositories. Most often, these datasets are linked to the commercial bibliometric sources (Scopus, Web of Science) and/or approved lists of national-language journals and publication channels, the latter addressing issues in the social sciences and humanities where the commercial databases have poor coverage

- Institutional RIS of the research performers, holding information on research staff, students, research outputs and outcomes as well as the broader research activities. These RIS also tend to be linked to (institutional or national/ disciplinary) repositories

In Chapter 7, we indicated that PRFS typically make use of data related to research outputs (academic, non-academic and innovation-related outputs such as IPR and spinoffs); data on external funding; and data on 'broader research activities' such as collaborations, international mobility, and esteem factors such as conferences, editorships, prizes etc. In some cases, these data related to performance are complemented by data on the size of the research-performing organisation such as the number of staff, of PhD students etc.

In, Table 27 below, we list the indicators covered in the PRFS considered for this study and indicate whether the information against these indicators were collected ad-hoc (cells highlighted in yellow) or derived from data stored in national research information systems (orange). Blank cells relate to data not used in the specific PRFS shown.

Table 27 Use of RIS versus ad-hoc evidence collections in European PRFS

	Belgium /FL (2009)	Denmark (2009)	Finland (2015)	Norway (2006)	Sweden (2008)	Czech Republic (2013)	Italy (VQR, 2011)	UK (REF 2014)
INPUT DATA								
Research staff								Selected
PhD students (recruitment/awarding)								
OUTPUT DATA								
Academic outputs								Selected
Non-academic outputs								Selected
Innovation-related outputs (IPR, spinoffs etc)								Selected
DATA ON FUNDING								
External funding								
DATA ON BROADER RESEARCH ACTIVITIES								
Collaborations (national/international)								Selected
International mobility								
Esteem factors (conferences, editorships, prizes)								Selected
Research infrastructure								

While input data as well as data on external funding are collected in close to all cases from national sources (eg databases held at the statistical agencies or Ministries), for the remaining data the metrics-based PRFS tend to use of data stored in information systems while the peer review based systems (Italy and the UK) tend to use ad hoc data collection methods. This appears to be because in metrics-based systems most often the assessment is based on *quantity* (eg the number of interactions or level of funding, in comparison to the other organisations active in similar fields of research). Quality measures exist only in the case of academic outputs and relate to the level of citations as proxy indicators for quality (in the Belgian and Swedish systems) or the quality categorisation of journals and publishing houses (in Denmark, Finland and Norway). In peer review based evaluations, the quality and potential effects are

of primary interest. Especially in the case of the UK, this is obtained in the form of text descriptions or narratives; quantitative data merely put this information into context, allowing 'informed' peer review.

8.1.2 The trend towards comprehensive national research information systems

In recent years an increasing number of initiatives have been launched in Europe that interlink research information and management systems, publication databases and (national) research evaluation datasets. The trend is towards the creation of 'comprehensive' national RIS, linking the datasets on research funding, programmes, outputs and outcomes as well as broader research activities, together with datasets containing information on research staff and students, bibliographic databases and repositories.

These initiatives aim to improve the availability of information on research, its outputs, short- and longer-term impacts. Comprehensive research information systems set up at the national level can act as platforms to create visibility for research outputs and outcomes in society, no matter the source and/or funding agency. An example is the FRIS system in Belgium (Flanders).

Key drivers for the creation of these national RIS systems are the need for greater quality and efficiency in data collection and the need for a more comprehensive view of national research production and its impact. Significant benefits appear to arise.

- Integrated information systems that enable one-time-only data entry address the issue of data quality and the burden imposed on the research communities, creating cost savings for both funding agencies and the research organisations
- The use of compatible web-based systems, integration of the RIS used by funding agencies and research organisations and improved linkages with other systems allows outputs to be reported and captured at any point
- National RIS linked to repositories provide opportunities to reduce reliance on bibliometrics and find alternative methods and tools for the assessment of research performance.
- National RIS make better strategic intelligence available to both research governance and the research performing organisations

In most countries, the development of comprehensive national RIS, is a relatively recent phenomenon. It is therefore no surprise that they have achieved varying levels of maturity (Table 28).

Typically, the major challenges encountered during the development of these RIS included the investment needed, establishing consensus among the stakeholder communities involved (eg the funding bodies and research organisations) on common routines and data semantics. Challenges of a technological nature included lack of interoperability among existing systems and the need to agree unique identifiers for researchers, organisations and research outputs.

There are only a few national systems sufficiently advanced that they encompass all research funding and project data as well as information on research outputs and effects, and research staff and students. The most visible ones are CRIStin (Norway) and FRIS (Belgium/Flanders).

Some of these initiatives integrate or interlink institutional and public agency information systems; some interlink data on research outputs with commercial bibliometric sources or open access repositories (e.g. the Norwegian CRIStin) aiming directly to harvest publication data from open repositories; others such as FRIS in the Flanders (Belgium) are considering links to patent databases.

Table 28 Examples of comprehensive national research information systems in Europe

	Belgium (Flanders) - FRIS	Czech Republic - VAV	Norway - CristIN	Sweden - SweCRIS
Format	CERIF	Not CERIF compliant	CERIF	CERIF
State	Ongoing construction	Ongoing construction	Ongoing construction	Development stage
Scope	All universities & research institutes – in future also private sector	All universities & research institutes	All universities, university colleges, research institutes and the regional health authorities	All universities, university colleges and research institutes
Source	Automated harvesting from /to institutional systems at the universities	Direct input by funding agencies & researchers	Direct input by researchers with automated updates in the institutional systems	Linking to RIS of all external (public) funders
Future intentions	Central service on journals and articles metadata Real integration of data Expansion of output information in content (patents, publications, etc.)	Linking with bibliometric databases	Databases on research outputs such as patents Data on EU-funded projects Yellow pages" of Norwegian researchers Invite companies to present their research activities Incorporate Health Research Classification System	Integration with PRISMA (system under construction of the three major research funders on researchers and research projects, including CV database Integration with. SwePub - database of the National Library of Sweden publications, harvested from the university databases and citations in Web of Science (ISI) National system for information to media/public and the website forskning.se which targets the public

So far, Flanders, Denmark and Norway are the places in Europe where the concept 'one input, multiple uses' is fully applied. In most of the other countries, major investments are being made in interoperability so that this objective can be reached. The use of CERIF, the European standard for Common Research Information Systems (CRIS), as the data model for the research information systems is critical to achieving interoperability.

8.2 The potential use of RIS data for the REF

8.2.1 The RIS eco-system in the UK

Overview of the current situation

The Higher Education Statistics Agency (HESA) and the Research Councils UK (RCUK) are the two main actors in the UK system for the collection of information on Higher Education Institutions (HEI), research activities and their outcomes.

HESA is the official body responsible for collecting and publishing data collected directly from the HEIs in the UK. Its Higher Education Information Database for Institutions (HEIDI) provides a front end to a range of data sets – covering both HESA and non-HESA HEI data sources – to HEIs, funding bodies and government departments. It provides a one stop shop for information including student and staff records, finance returns, destinations surveys, Estates Management Statistics, and HE Business and Community Interaction surveys (HE-BCI), the latter providing information about revenue generated from research and enterprise creation. HEFCE additionally collects and published university cost data in connection with the transparent approach to costing (TRAC) required in order to enable a standardised approach to calculating the full economic cost of university research.

The Research Councils have invested considerable resources during the last decade in developing a national RIS for the collection of information on all RCUK-funded research activities and their outputs and outcomes/impacts. The system consists of a Joint electronic Submissions system (Je-S) for proposals as well as a common online platform for the submission of information on outputs and outcomes, Researchfish.

Researchfish makes it possible to link individual research outcomes with the information systems of multiple funders. RCUK-funded researchers are expected to log the outcomes of their funded projects on the system; they submit an up-to-date record of the outcomes of each award annually during the

period of the award and for about 5 years after it ends. Currently, no bulk uploads are possible nor is there the facility to directly link Researchfish with institutional RIS.

The platform also collects information also for other funders in the UK, including (mainly health-related) non-profit organisations, and other public and private research funders such as DFID. As of 2015, Researchfish captures information for 79 research funders (74 registered in the UK, 5 overseas) (Hinrichs, Montague, & Grant, 2015). In theory, the system can hold information on outcomes of research funded by any fund provider. However, it is unclear to what extent researchers make use of this facility and therefore to what extent Researchfish holds information on outcomes from the entirety of the academic research activities conducted in the UK.

Another online platform that is related to RCUK-funded research is the UK Gateway to Research (GtR), developed by the Research Councils. As part of the Innovation and Research Strategy of the Government's Department for Business, Innovation and Skills (BIS), the Gateway to Research holds regularly updated data from the seven Research Councils as well as Innovate UK and the National Centre for the Replacement and Reduction of Animals in research (NC3Rs). It provides selected information about people and organisations involved, as well as about publications and outcomes from RCUK-funded research projects. Grant information comes from the research councils' and Innovate UK's grant systems.

Innovate UK funds collaborative projects involving HEIs and regularly publishes information on the projects and beneficiaries online. However, the data it provides in a RIS format are only administrative, ie a listing of projects, beneficiaries and their categories, and the amount of funding awarded. Outcomes are not recorded.

Future developments

The fragmentation of the reporting infrastructure for the HEIs in the UK is well known (Scoble & Jones, 2013) and was remarked upon by respondents to the consultation in connection with the 2016 While Paper (BIS, 2016). Several initiatives have been launched in recent years (in particular by HESA, RCUK and JISC) to try and reduce the burden this imposes on the HEIs.

An improved and harmonised data collection process where data are entered once and used many times is clearly desirable, but for the present there is no project in progress to achieve this.

- The 2015 Kings College report on Researchfish concludes that there is a significant need for improvement to the system. The quality of the data submitted is an issue as well as the ease in using the data for assessment and reporting purposes. Many research organisations also indicated the need for better interoperability with their institutional research information systems and expressed a desire for functionality that would enable them to automate data input into Researchfish (Hinrichs, Montague, & Grant, 2015)
- The JISC/RCUK *Overview of Systems Interoperability Project* (OSIP) project focused on identifying the next steps in order to maximise the interoperability of the research information held, and systems used, by RCUK. In its final report, the project team lists a number of future actions endorsed by the RCUK Research Funding Programme (amongst which the use of ORCID for researcher identification), as well the need for support from RCUK for setting up a high-level UK Research Information Steering Group (Lyne, Grout, & Notay, 2015)
- The JISC-funded *UK Research Information Shared Service* (UKRISS) project investigated the reporting of research information across the UK HE sector and assessed the feasibility of a national infrastructure based on CERIF that would increase efficiency, productivity and reporting quality (Waddington, 2015)

  The project developed a core reporting profile that would enable harmonised reporting about RCUK-funded research, taking into account the information needs of the HE-BCI survey and the REF. It argued that the considerable costs that might be imposed upon small institutions and charities would be outweighed by the benefits of harmonisation

The need for a comprehensive view of the whole research landscape in the UK is increasingly recognised. In 2014 the Council for Science and Technology (CST) launched the crowd-sourced Science Landscape project. It led to the establishment of the High-Level Landscape Mapping Group that aims to build a comprehensive and reliable picture of the whole R&D landscape in the UK, creating a stronger evidence base for future decision-making. To this end, the High-Level Group considered the need for a full exploitation of current RIS and datasets documenting capability and investment in UK research and innovation. It will develop a 5-year plan setting out the steps required to reach this goal. The High-Level Group involves representatives of the Research Councils, Government Departments and Public Agencies (Elliott & al, 2016).

8.2.2 The use of RIS data in REF 2014

In the REF 2014, the submitting institutions were requested to submit (in addition to the impact case studies)

- Information on staff in post on the census date, selected by the institution to be included in the submission; individual staff circumstances for those submitting fewer than four outputs; and details of Category C staff
- Details of research outputs that selected staff have produced during the publication period
- A completed template describing the submitted unit's approach to enabling impact from its research, and case studies describing specific examples of impacts underpinned by the unit's research
- A completed template describing the research environment, including data about research doctoral degrees awarded, research income and research income-in-kind

The focus of the REF on selected research staff only, combined with the emphasis on text-based descriptions and narratives, meant that quantitative data were requested only for the assessment of the 'research environment'. HEFCE provided the HEIs with information on these indicators based on data stored in other research information systems.

- Data on research doctoral degrees awarded and amounts and sources of external research income based on information stored in the Higher Education Statistics Agency (HESA)
- Data on research income in kind (ie the estimated value of facility time allocated through peer review and used by researchers at submitted units), based on records held by the UK Research Councils and relevant health research funding bodies

We could find no indication of universities using other information stored in HEIDI or other HESA datasets in the context of the REF, while several HEIs indicated that they used information stored in Researchfish as a basis for their REF submission, especially for the drafting of their impact case studies.

8.2.3 REF-relevant data collected in other RIS

The two sources of information on research that could be of direct relevance for future REF exercises are the RCUK's Researchfish and the HESA databases. The information stored in the Innovate UK RIS is of very limited immediate value for the REF. Table 29, below, lists the indicators against which the REF, Researchfish and HESA collect information.

Table 29 Research information collected by the REF, Researchfish and HESA

		REF	RCUK/RF	HESA / HE-BCI
INPUT				
External funding	Research grants and contracts			X
	Follow-up funding		X	
	Funding from regeneration and development programmes			X
OUTPUTS				
Research outputs	Academic outputs	Selected	X	
	Digital artefacts	Selected	X	
	Temporary artefacts (exhibitions etc)	Selected	X	
	Physical artefacts (new materials, products, processes etc)	Selected	Medical	
	Research Tools & Methods		X	
	Research Databases & Models		X	
IPR	Patents	Selected	X	X
	Published patent applications	Selected	X	X
	Other forms of IPR	Selected	X	X
ENVIRONMENT				
Research strategy	Evidence of the achievement of strategic aims for research	Narrative		
Environment	Evidence of strong research plans	Narrative		
People	Evidence on staffing strategy and staff development	Narrative (& data)		
Research students	Evidence on the training and supervision of postgraduate research (PGR) students	Narrative & data		Data for the REF
Income, infrastructure and facilities	Information about research income, infrastructure and facilities	Narrative & data	Data for the REF	Data for the REF
Collaboration and contribution to the discipline or research base	Contributions to the wider research base, including work with other researchers outside the submitted unit whether locally, nationally or internationally; support for research collaboration; and interdisciplinary research	Narrative (& data)	X	X
BROADER RESEARCH ACTIVITIES				
Esteem indicators	Research conferences / workshops		X	
	Awards and Recognition		X	
OUTCOMES & IMPACTS				
Knowledge & technology transfer	Research Tools & Methods		X	
	Research Databases & Models		X	
	Spin-offs		X	X
	Influence on Policy		X	
Engagement Activities	Communication of research beyond the normal peer group		X	
	Social community and cultural engagement (Public lectures; Performance arts; Exhibitions; Museum education; Other)			X
	Business and community services (courses for business and the community)			X

Our conclusions on the usability for the REF of the data collected in the RIS and the implication for the design parameters are as follows.

- *The model used*. The Researchfish approach for the collection of data on research performance is similar to the one used for the REF. There is an equally strong emphasis on qualitative information, to be provided in the form of descriptions or narratives and there is no direct use of bibliometrics. However, Researchfish also requires systematic input of quantitative data and responses that can be analysed statistically (eg Yes/No answers and categorisation lists). A similar approach is taken only by some panels in the REF and only for some of the indicators. HESA, instead, focuses exclusively on the collection of quantitative data
- *The scope of the research activity included*. The Researchfish approach has strong similarities with the REF one. Research is at the core of the exercise. However, Researchfish devotes more attention to the relevance of the research activities for industry and society, a feature that is present at the same level only in some REF panels. In the REF, this aspect is covered through the impact case studies.
- *Type of indicators used*. While in many cases the REF and RCUK, and even HESA, focus on the same indicators, there are also differences owing to the concern in the REF with the requirements of individual panels, which can tailor indicators to their scientific disciplines in order to address as much as possible the differences in research practice. This flexibility in the REF means that there is a long list of different non-academic outputs that could be submitted to the panels as well as field-specific differences in the expected content of narratives related for example to contributions to the wider research base. RCUK uses a smaller number of standardised indicators, including discipline-specific indicators. There is a considerable list of possible non-academic outputs, however, and there is clearly scope to align the data collection needs of the REF and Researchfish if desired

  HESA/HE-BCI collects data on innovation-related outputs and outcomes, such as IPR and spin-offs, as well as the focus and volume of contract research and consultancy contracts at a level of detail that is not present in the REF or Researchfish
- *Granularity of the exercise*. While the REF has an exclusive focus on the work of researchers producing excellent research, for RCUK the primary focus is on the RCUK-funded project and not on the work of individual researchers. Researchers that have not been involved in RCUK-funded projects will therefore not appear in the RCUK RIS
- *The periodicity of the exercise*. A major distinction is the time frame of the REF versus the Researchfish and HESA collection activities. While the information in relation to RCUK-funded projects and the HEI activities in general are updated on an annual basis, information deriving from the REF is static for a period of 5 years

8.2.4 Conclusions

The analysis set out above shows that there currently is a lack of comprehensive data on research and its outputs and impacts in the UK, owing to the silo approach triggered by the dual support system. HESA collects information on the HEIs' activities as a whole, ie both teaching and research. A clear distinction cannot in all cases be made between the two.

This lack of comprehensive data inhibits the use of most of the research performance data currently stored in the national RIS for the REF exercise, no matter what form it takes.

We understand that the problem is being recognised among policy makers at various levels in the research governance system. No doubt a great deal can be accomplished by linking existing systems. However, the complexity of the matter and the costs involved imply that there will not be a short-term solution. A similar conclusion can be reached in relation to the direct use of the data stored in the institutional RIS.

9 Bibliography

Agyemang, G., & Broadbent, J. (2015). Management control systems and research management in universities. *Accounting, Auditing & Accountability Journal, 28*(7), 1018-1046.

Aagaard, K., Bloch, C., & Schneider, J. (2015). Impacts of performance-based research funding systems: The case of the Norwegian Publication Indicator. *Research Evaluation, 24*, 106-117.

Abramo, G., Cicero, T., & D'Angelo, C. (2011). The dangers of performance-based research funding in non-competitive higher education systems. *Scientometrics, 87*, 641–654.

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2011). National research assessment exercises: the effects of changing the rules of the game during the game. *Scientometrics, 88*(1), 229-238.

Abramo, G., D'Angelo, C., & Di Costa, F. (2011). National research assessment exercises: a comparison of peer review and bibliometric rankings. *Scientometrics, 89*, 929-941.

ABRC. (1987). *A strategy for the science base. A discussion document prepared for the Secretary of State for Education and Science by the Advisory Board for the Research Councils*. London: HMSO.

Adams, J. (2013). The fourth age of research . *Nature, 497*, 557-560.

Adams, J., & Gurney, K. (2014). *Funding selectivity, concentration and excellence - how good is the UK's research?* . HEPI.

Adams, J., & Gurney, K. A. (2014). *Evidence for excellence: has the signal overtaken the substance? An analysis of journal articles submitted to RAE2008*. London: Digital Science.

Adams, J., & Gurney, K. A. (2016). *The Implications of International Research Collaboration for UK Universities Research assessment, knowledge capacity and the knowledge economy*. London: Digital Science.

Arnold, E., & Mahieu, B. (2015). *R&D Evaluation Methodology and Funding Principles*. Prague: Ministry of Education, Youth and Sport.

Arnold, E., Åström, T., Boekholt, P., Brown, N., Good, B., Holmberg, R., . . . van der Veen, G. (2011). *Ioacts of EU Framework Programmes in Sweden, VA 2008:11*. Stockholm: Vinnova.

Arnold, E., Barker, K., & Slipersæter, S. (2010). *Research Institutes in the ERA*. Brussels: European Commission.

Arnold, E., Good, B., & Segerpalm, H. (2008). *Effects of Research on Swedish Mobile Telephone Developments: The GSM Story, VA2008:04*. Stockholm: VINNOVA.

Arnold, E., Luukkonen, T., Boekholt, P., Nooijen, A., Jávorka, Z., & Zuijdam, F. (2013). *Evaluation of the Academy of Finland*. Helsinki: Ministry of Education and Culture.

Barker, K. (2007, March). The UK Research Assessment Excercise: the evolution of a national research evaluation system. *Research Evaluation*, 3-12.

Bekhradnia, B. (2006). *Evaluating and funding research through the proposed "Research Excellence Framework"*. London: HEPI.

Bence, & Oppenheim. (2005). The evolution of the UK's research Assessment Exercise: publications, performance and perceptions. *Journal of Educational Administration and History, 37*(2), 137-155.

Bence, V., & Oppenheim, C. (2005). The Evolution of the UK's Research Assessment Exercise: Publications, Performance and Perceptions. *Journal of Educational Administration and History, 37*(2), 137-155.

BIS. (2011). *International Comparative Performance of the UK Research Base - 2011*. A report prepared by Elsevier for the UK's Department of Business, Innovation and Skills (BIS.

BIS. (2013). *International Comparative Performance of the UK Research Base – 2013*. A report prepared by Elsevier for the UK's Department of Business, Innovation and Skills (BIS).

BIS. (2016). *Fulfilling Our Potential: Teaching Excellence, Social Mobility and Student Choice. Summary of Consultation Responses*. London: Department of Business, Innovation and Skills.

Bloch, C., & Schneider, J. (2016). Performance-based funding models and researcher behavior: An analysis of the influence of the Norwegian Publication Indicator at the individual level. *Research Evaluation, 25*(3).

Boer de, H., & File, J. (2009). *Higher Education Governance Reforms across Europe: Report of the MODERN project*. Brussels: ECSMU.

Boer de, H., Jongbloed, B., Bemmeworth, P., Cremonini, L., Kolster, R., Kottmann, A., . . . Vossensteyn, H. (2015). *Performance-based funding and performance agreements in fourteen higher education systems*. Report for the Ministy of Education, Culture and Science. Enschede: CHEPS.

Bolton, P. (2012). *Education: Historical statistics* . Libary House of Commons.

Bouabid, H. (2014). Science and technology metrics for research policy evaluation: some insights from a Moroccan experience. *Scientometrics, 101*, 899-915.

Brinn, T., Jones, M. J., & Pendlebury, M. (2001). The impact of research assessment exercises on UK accounting and finance faculty. *British Accounting Review, 33*(3), 333-355.

Broadbent, J. (2010). The UK Research Assessment Exercise: Performance Measurement and Resource Allocation. *Australian Accounting Review, 20*(1), 14-23.

Brooks, C., Fenton, E. M., & Walker, J. T. (2014). Gender and the evaluation of research. *Research Policy, 43*(6), 990-1001.

Butler, L. (2003). Explaining Australia's increased share of ISI publications – the effects of a funding formula based on publication counts. *Research Policy, 32*(1), 143-155.

Butler, L. (2010). *Impacts of Performance-Based Funding Systems: A review of the concerns and evidence*. Paris: OECD.

Chubin, D. E., & Hackett, D. J. (1990). *Peerless Science: Peer review and US science policy*. Albany: SUNY Press.

Claeys-Kulik, A.-L., & Estermann, T. (2015). *DEFINE Thematic Report: Performance-based Funding of UNiversities in Europe*. Brussels: European University Association .

Court, S. (1999). Negotiating the Research Imperative: The Views of UK Academics on their Career Opportunities. *Higher Education Quarterly, 53*(1), 65-87.

CSTW. (2007). *Scoping study on the use of bibliometric analysis to measure the quality of research in UK higher education institutions*. Leiden: CSTW, Leiden University.

Dearing. (1997). *Higher Education in the learning society*. London: HMSO.

Department for Education. (2006). *Reform of Higher Education Research Assessment and Funding - Summary results*. London: Department for Education.

DES. (1985). *The Development of Higher Education into the 1990s, Cmnd. 9524*. London: HMSO.

Edgar, F., & Geare, A. (2010). Characteristics of High- and Low-performing University Departments as Assessed by the New Zealand Performance Based Research Funding (PBRF) Exercise. *Australian Accounting Review, 20*(1), 55-63.

Elliott, J., & al. (2016). *The UK Knowledge and Research Landscape: A report on available resources*. High-Level Landscape Mapping Group. STC.

Elsevier. (2015). *A review of the UK's interdisciplinary research using a citation-based approach*. A report by Elsevier to the UK HE funding bodies and MRC.

Elton, L. (2000). The UK Research Assessment Exercise: Unintended Consequences. *Higher Education Quarterly, 54*(3), 274-283.

Equality Challenge Unit. (2009). *The impact of the process to promote equality and diversity in the Research Assessment Exercise 2008*. London: Equality Challenge Unit.

Evidence. (2007). *Bibliometric analysis of interdisciplinary research*. Leeds: Evidence.

Eyre-Walker, A., & Stoletzki, N. (2013). The Assessment of Science: The Relative Merits of Post-Publication Review, the Impact Factor, and the Number of Citations. *PLoS Biology, 11*(10), 1-8.

Farla, K., & Simmonds, P. (2015). *REF 2014 Accountability Review: Costs, benefits and burden*. Bristol: HEFCE.

Franzoni, C., Scellato, G., & Stephan, P. (2011). Changing incentives to publish. *Science*(333), 702-703.

Fridholm, T., & Melin, G. (2012). *"Med glädje, men inte med lätthet" –om högskolans fördelning av de direkta statsanslagen för forskning Distribution of governmental direct appropriations to research at Swedish universities* . Stockholm: Swedish Employers Federation.

Funding Councils. (1997). *Conduct to the Exercise: RAE Manager's Report*. Bristol: HEFCE.

Gazni, A., & Thelwall, M. (2014). The citation impact of collaboration between top institutions: a temporal analysis. *Research Evaluation*(1), 1-11.

Geuna, A., & Martin, B. R. (2003). University Research Evaluation and Funding: An International Comparison. *Minerva, 41*(4), 277-304.

Good, B., Vermeulen, N., Tiefenthaler, B., & Arnold, E. (2015). Counting quality? The Czech performance-based research funding system. *Research Evaluation, 24*(2), 91-105.

Grant, J., Brutscher, P., Kirk, S. E., Butler, L., & Wooding, S. (2010). *Capturing Research Impacts: A review of international practice*. Cambridge: RAND.

Gray, R. (2015). Has the Research Excellence Framework killed creativity? *Journal of Psychiatric and Mental Health Nursing, 22*(3), 155-156.

Greenhalgh, T., & Fahy, N. (2015). Research impact in the community-based health sciences: an analysis of 162 case studies from the 2014 UK Research Excellence Framework. *BMC Medicine, 13*(232).

Griffin, G., Medhurst, P., & Green, T. (2006). *Interdisciplinarity in Interdisciplinary Research Programmes in the UK*. Hull: Research Integration, University of Hull.

Hävrynen, M. (2007). *Breakthrough Research: Funding for High-Risk Research at the Academy of Finland, 8/07*. Helsinki: Academy of Finland.

HEFCE. (1997). *Impact of the 1992 RAE on higher education institutions in England*. Bristol: HEFCE.

HEFCE. (1999). *Higher Education in the Learning Society*. Bristol: HEFCE.

HEFCE. (2009). *The Research Excellence Framework: A brief guide to the proposals*. Bristol: HEFCE.

HEFCE. (2011). *Decisions on assessing research impact*. Bristol: HEFCE.

HEFCE. (2012). *Panel criteria and working methods* . Bristol: HEFCE.

HEFCE. (2012a). *Assessment framework and guidance on submissions (updated to include addendum published in January 2012)*. Bristol: HEFCE.

HEFCE. (2015). *The Metric Tide: Correlation analysis of REF2014 scores and metric*. Bristol: HEFCE.

Henkel, M. (1999). The Modernisation of Research Evaluation: The Case of the UK. *Higher Education, 38*(1), 105-122.

Hicks, D. (2012). Performance-based university research funding systems. (Elsevier, Ed.) *Research Policy, 41*, 251-261.

Hicks, D. (2013). One size doesn't fit all: On the co-evolution of national evaluation systems and social science publishing. *Confero, 1*(1), 67-90.

Hinrichs, S., Montague, R., & Grant, J. (2015). *Researchfish: A forward look - Challenges and opportunities for using Researchfish to support research assessment*. The Policy Institute at Kings College. Researchfish.

HM Government. (1993). *Realising our Potetial: A Strategy for Science, Engineering and Technology, CM 2250*. Lodon: HMSO.

HM Treasury. (2006). *Science and Innovation Investment Framework 2004-201: next steps*. London: HM Treasury.

Horrobin, D. F. (1996). Peer review of grant applications: a harbinger for mediocrity in clinical research? *The Lancet, 348*, 1293-1295.

Hughes, A., Kitson, M., Bullock, A., & Milner, I. (2013). *The Dual Funding Structure for Research in the UK: Research Council and Funding Council Allocation Methods and the Pathways to Impact of UK Academics*. UK-IRC. Cambridge and London: UK-IRC.

Jarrett, S. (1985). *Report of the Steering Committee for Efficiency Studies in Universities ('The Jarratt Report')Report'), report to the Committee of Vice-Chancellors and Principals (CVCP) and to the University Grants Committee*. London: CVCP.

Jeffery, R. (2014). Authorship in multi-disciplinary, multi-national North-South research projects: issues of equity, capacity and accountability. *Compare: A Journal of Comparative and International Education, 44*(2), 208-229.

Jiminez-Contreras, E., Anegon, F., & Lopez-Cozar, E. (2003). The evolution of research activity in Spain – The impact of the National Commission for the Evaluation of Research Activity (CNEAI). *Research Policy, 32*, 123-142.

Jones, P., & Sizer, J. (1990). *The Universities Funding Council 1989 Research Selectivity Exercise*. Retrieved from http://www.bzh.bayern.de/uploads/media/4-1990-jones-sizer.pdf

Jonkers, K., & Zacharewicz, T. (2015). *Performance based funding: A comparative assessment of their use and nature in EU Member States* . Brussels: JRC Science for Policy Report.

Jonkers, K., & Zacharewicz, T. (2016). *Research Performance Based Funding Systems: a Comparative Assessment*. Brussels: JRC.

Joynson, C., & Leyser, O. (2015). The culture of scientific research. *F1000Research, 4*(66), 1-11.

Jump, P. (2013, October 17). Evolution of the REF. *THES*.

Karlsson, S., & Persson, O. (2012). *The Swedish Production of Highly Cited Papers, 5:2012*. Stockholm: Swedish Research Council.

Key Perspectives Ltd. (2009). *A comparative review of rsearch assessment regimes in five countries and the role of libraries in the Research Assessment process*. OCLC Research.

Kilduff, M., Mehra, A., & Dunn, M. B. (2011). From Blue Sky Research to Problem Solving: A Philosophy of Science Theory of New Knowledge Production. *Academy of Management Review, 36*(2), 297-317.

King's College London and Digital Science. (2015). *The nature, scale and beneficiaries of research impact: An initial analysis of Research Excellence Framework (REF) 2014 impact case studies*. London: King's College London.

Klein, J. T. (2006). Afterword: the emergent literature on interdisciplinary and transdisciplinary research evaluation. *Research Evaluation, 15*(1), 75-80.

Kogan, M., & Henkel, M. (1998). The impact of policy changes on the academic profession in England. *European Review, 6*(4), 505-523.

Kuhn, T. (1970). *The Structure of Scientific Revolutions*. Chicago: Chicago UP.

Lakatos, I., & Musgrave, A. (1970). *Criticism and the Growth of Knowledge*. Cambridge: Cambridge UP.

Langfeldt, L. (2006). The policy challenges of peer review: managing bias, conflict of interests and interdisciplinary assessments. *Research Evaluation, 15*(1), 31-41.

Laudel, G., & Gläser, J. (2014). Beyond breakthrough research: Epistemic properties of research and their consequences for research funding. *Research Policy, 43*(7), 1204-1216.

Lee, F., & Harley, S. (1998). Peer review, the Research Assessment Exercise and the demise of non-mainstream economics. *Capital & Class, 22*(3).

Leisyte, L., & Westerheijden, D. (2014). *Research Evaluation and its Implications for Academic Research in the United Kingdom and the Netherlands*. Dortmund: Zentrum für Hochschulbildung & TU Dortmund discussion papers.

Lepori, B., van den Besselaar, P., Dinges, M., Potì, B., Reale, E., Slipersæter, S., . . . van der Meulen, B. (2007). Comparing the evolution of national research policies: what patterns of change? *Science and Public Policy, 34*(6), 372-388.

Luukkonen, T. (2012). Conservatism and risk-taking in peer review: Emerging ERC practices. *Research Evaluation, 21*(1), 48-60.

Luukkonen, T., Stampfer, M., & Strassnig, M. (2015). *Evaluation practices in the selection of ground-breaking research proposals*. Vienna: WWTF Working Paper: 13-08-2015.

Lyne, I., Grout, C., & Notay, B. (2015). *Overview of Systems Interoperability Project - Report to the RCUK Research Funding Programme Board*. RCUK. JISC.

Mahieu, B. (2015). *R&D Evaluation Methodology and Funding Principles Background report: Evaluation systems in international practice (country analyses)*. Prague: Czech Ministry of Education, Youth and Sport.

Mahieu, B., & Arnold, E. (2015). *R&D Evaluation methodology & Funding principles - The R&D Evaluation Mthodology (Final Report 1)*. Technopolis Group. Prague: Ministry of Education, Youth & Sports.

Mahieu, B., Arnold, E., & Kolarz, P. (2013). *Measuring scientific performance for improved policymaking - Literature review*. Technopolis Group. Brussels abd Strasbourg: STOA, European Parliament,.

Manville, C., Guthrie, S., Henham, G. S., Garrod, B., Sousa, S., Kirtley, A., & Ling, T. (2015). *Assessing impact submissions for REF 2014: An evaluation*. Cambridge: RAND.

Manville, C., Jones, M., Frearson, M., Castle-Clark, S., Henham, G. S., & Grant, J. (2015). *Preparing impact submissions for REF 2014: An evaluation. RAND*. Cambridge: RAND.

Martin, B., & Whitley, R. (2010). The UK Research Assessment Exercise: a Case of Regulatory Capture? In R. Whitley, J. Glaser, & L. Engwall, *Reconfiguring knowledge production : changing authority relationships in the sciences and their consequences for intellectual innovation* (pp. 51-80). Oxford University Press.

Maude, K. (2014). Citation and marginalisation: the ethics of feminism in Medieval Studies. *Journal of Gender Studies, 23*(3), 247-261.

Meriläinen, S., Tienari, J., Thomas, R., & Davies, A. (2008). Hegemonic Academic Practices: Experiences of Publishing from the Periphery. *Organization, 15*(4), 584-597.

Mertens, D. (2009). *Tranformative Research and Evaluation*. London: Guildford Press.

Moed. (2005). *Citation Analysis in Research Evaluation*. (Springer, Ed.)

Moed, Glanzel, & Schmoch. (2004). *Handbook of Quantitative Science and Technology Research*. (Kluwer, Ed.)

National Committee of Inquiry into Higher Education (Dearing Report). (1997). *Report 6 Widening participation in higher education for students from lower socio-economic groups and students with disabilities*. London: HMSO.

National Science Board. (2007). *Enhancing Support of Transformative Research at the National Science Foundatio*. Washington DC: National Science Board.

Nielsen, K., Kolarz, P., Farla, K., Blessing, V., & Simmonds, P. (2016). *Independent review of the Research Excellence Framework (REF): Synthesis of responses submitted to the REF Review Call for Evidence and follow-up interviews*. London: Departmnt of Business, Innovation and Skills.

Nowotny, H., Scott, P., & Gibbons, M. (2003). Introduction: `Mode 2' Revisited: The New Production of Knowledge. *Minerva, 41*(3), 179-194.

NZ Ministry of Education. (2012). *An international comparison of performance-based research funding systems (PBRFS)*. New Zealand.

OECD. (2010). *Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings*. Paris: OECD Publishing.

Otley, D. (2010). Research Assessment in the UK: An Overview of 1992–2008. *Australian Accounting Review, 20*(1).

PA Consulting Group. (2009). *RAE 2008 Accountability Review*. London: PA Consulting.

PACEC and University of Cambridge . (2014). *A Review of QR Funding in English HEIs. Process and Impact* . HEFCE 2014.

Page, E. S. (1997). *Data Collection for the 1996 Research Assessment Exercise: Review*. Bristol: HEFCE.

Parliamentary Office of Science and Technology. (1997). *Striking a Balance - The Future of Research Dual Support in Higher Education"*. London: POST.

Penfield, T., Baker, M. J., Scoble, R., & Wykes, M. C. (2014). Assessment, evaluations, and definitions of research impact: A review. *Research Evaluation, 23*.

Porter, A. L., & Rossini, F. (1985). Peer review of interdisciplinary research proposals. *Science, Technology and Human Values, 10*, 33-38.

Price, D. d. (1963). *Little Science, Big Science*. New York, New York, USA: Columbia University Press.

RAE. (1996). *RAE96 2/95 Guidance on submissions* . Retrieved 2016, from RAE 1996: http://www.rae.ac.uk/1996/c2_95.html

RAE. (2001). *Section 3: Information Requirements: notes and definitions*. Retrieved 2016, from RAE 2001: http://www.rae.ac.uk/2001/pubs/2_99/section3.htm

RAE. (2005). *RAE2008 Guidance on submissions*. Bristol: HEFCE.

RAE. (2009). *RAE 2008 Managers report*. Bristol: HEFCE.

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management. *Research Policy, 41*, 1262-1282.

RAND, P. I. (2015). *The Policy Institute at King's College London and RAND Europe (2015). Characteristics of high-performing research units: A preliminary analysis Research Report 2015/02 Prepared for the Higher Education Funding Council for England (HEFCE)* . London: King's College London.

RCUK. (2011). *Pathways to Impact*. Swindon: RCUK.

REF. (2008). *Research Assessment Exercise 2008: the outcome*. The funding bodies.

REF. (2011). *REF2014 Assessment framework and guidance on submissions*. Bristol: HEFCE.

REF. (2011). Sector Impact Assessment.

REF 2014. (2010). *Research Excellence Framework impact pilot exercise: Findings of the expert panels. A report to the UK higher education funding bodies by The chairs of the impact pilot panels*. Bristol: HEFCS.

REF2014. (2014, December 18). *Comparison with 2008 RAE results*. Retrieved 03 03, 2016, from http://www.ref.ac.uk/results/analysis/comparisonwith2008raeresults/

Rinia, E. J., van Leeuwen, T. N., & van Raan, A. F. (2001). New light on old boys: cognitive and institutional particularism in the peer review system. *Scientometrics, 53*(2), 241-248.

Rinia, E. J., van Leeuwen, T. N., van Buren, H. G., & Van Raan, A. F. (2001). Impact measures of interdisciplinary research in physics. *research Policy, 30*(3), 357-361.

Roberts, G. (2003). *Review of research assessment Report by Sir Gareth Roberts to the UK funding bodies* . London: RA.

Roessner, D., Porter, A. L., Nersessian, N. J., & Carley, S. (2013). Validating indicators of interdisciplinarity: linking bibliometric measures to studies of engineering research labs. *Scientometrics, 94*, 439-468.

Rosenberg, N. (1976). *Perspectives on Technology*. Cambridge: Cambridge University Press.

Roy, R. (1985). Funding Science: The Real Defects of Peer Review and an Alternative to It'. *Science, Technology and Human Values, 10*(3), 73-81.

Royal Academy of Engineering. (2006). *Reform of Higher Education Research Assessment and Funding*. London: Royal Academy of Engineering.

Sastry, T., & Bekhradnia, B. (2006). *Using metrics to allocate research funds A short evaluation of alternatives to the Research Assessment Exercise*. Oxford: Higher Education Policy Institute (HEPI).

Sayer, D. (2015). *Rank Hypocrisies. The insult of the REF.* London: Sage Swifts.

Scoble, R., & Jones, R. (2013). *UK Research Information Shared Service (UKRISS) project - Landscape study report*. JISC.

Seglen, P. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ, 314*, 314-497.

Sikes, P. (2006). Working in a 'new' university: in the shadow of the Research Assessment Exercise? *Studies in Higher Education, 31*(5), 555-568.

Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. *ISI Newsletter, 6*(1), pp. 22-28.

Smart, W. (2013). *Analysis of the Impact of the PBRF. Interim Findings.* Wellington: Ministry of Education.

Smith, S., Warda, V., & House, A. (2011). 'Impact' in the proposals for the UK's Research Excellence Framework: Shifting the boundaries of academic autonomy. *Research Policy, 40*.

Stöckelova, T. (2012). Immutable Mobiles Derailed: STS, Geopolitics, and Research Assessment. *Science, Technology, & Human Values, 37*(2), 286-311.

Stern, N. (2016). *Building on Success and Leaning from Experience: An Independent Review of the Research Excellence Framework*. London: Department for Business, Energy and Industrial Strategy.

Strang, V., & McLeish, T. (2015). *Evaluating Interdisciplinary Research: a practical guide* . Durham: Durham University & Institute of Advanced Study.

Swedish Research Council. (2015). *Research Quality Evaluation in Sweden – FOKUS.* Stockholm: Vetenskapsrådet.

TaylorJ. (2011). The assessment of research quality in UK. Universities: peer review or metrics? *British Journal of Management, 22*, 202-217.

Technopolis. (2009). *Identification and dissemination of lessons learned by institutions participating in the Research Excellence Framework (REF) bibliometrics pilot*. Bristol: HEFCE.

Technopolis. (2010). *REF Research Impact Pilot Exercise Lessons - Learned Project: Feedback on Pilot Submissions*. Bristol: HEFCE.

Technopolis. (2015). *R&D Evaluation Methodology and Funding Principles Background report 6: R&D governance and institutional funding in international practice.* Prague: Czech Ministry of Education, Youth and Sports.

Travis, G., & Collins, H. M. (1991). New light on old Boys: Cognitive and Institutional Particularism in the peer review system. *Science, Technology and Human Values, 16*(3), 187-97.

Universities and College Union. (2009). *Response to the Research Excellence Framework: Second consultation on the assessment and funding of research*. London: UCU.

Universities Funding Council. (1992). *Research Assessment Exercise 1992: The Outcome*. Retrieved 2016, from RAE 1992: http://www.rae.ac.uk/1992/c26_92.html#annexc%20research%20Assessment%20Exercise%201992:%20the%20Outcome

University and College Union. (2008). *Response to the Research Excellence Framework Consultation on the assessment and funding of higher education research post-200*. London: University and College Union.

University and College Union. (2011). *UCU response to the HEFCE REF consultation Consultation on draft panel criteria and working methods*. London: University and College Union.

University and College Union. (2013). *The Research Excellence Framework (REF) - UCU Survey Report*. London: University and College Union.

van der Meulen, B. (2001). *The Asutrian Science Fund: Governance and Processes*. Enschede: Universiteit Twente.

Vanecek, J. (2013). he effect of performance-based research funding on output of R&D results in the Czech Republic. *Scientometrics, 10*, 1-25.

Waddington, S. e. (2015). *UK Research Information Shared Service (UKRISS) Project - Final report*. JISC.

Wagner, C., & Alexander, J. (2013). Evaluating transformative research programmes: A case study of the NSF Small Grants for Exploratory Research programme. *Reseatch Evaluation, 22*(3), 187-197.

Wagner, C., Roessner, D., Bobb, K., Klein, J. T., Boyack, K., Keyton, J., . . . Börner, K. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics, 165*, 14-26.

Wang, J., & Hicks, D. (2013). Detecting structural change in university research systems: A case study of British research policy. *Research Evaluation, 22*, 258-268.

Weinberger, H. (1997). *Nätverksentreprenören: PhD Thesis*. Stockholm: Royal Institute of Technology (KYH).

Wessely, S. (1998). Peer review of grant applications: what do we know? *The Lancet, 352*, 501-505.

Wilsdon, J. (2015). *The Metric Tide: Report of the independent Review of the Role of Metrics in Research Assessment and Management*. Bristol: HEFCE.

Wooding, S., & Grant, J. (2003). *Assessing Research: the researchers' view*. Cambridge: RAND.

Wouters, P., Thelwall, M., Kousha, K., Waltman, L., de Rijke, S., Rushforth, A., & Franssen, T. (2015). *The Metric Tide: Literature Review: Supplementary Report I to the Independent Review of the role of Metrics in Research Assessent and Management*. Bristol: HEFCE.

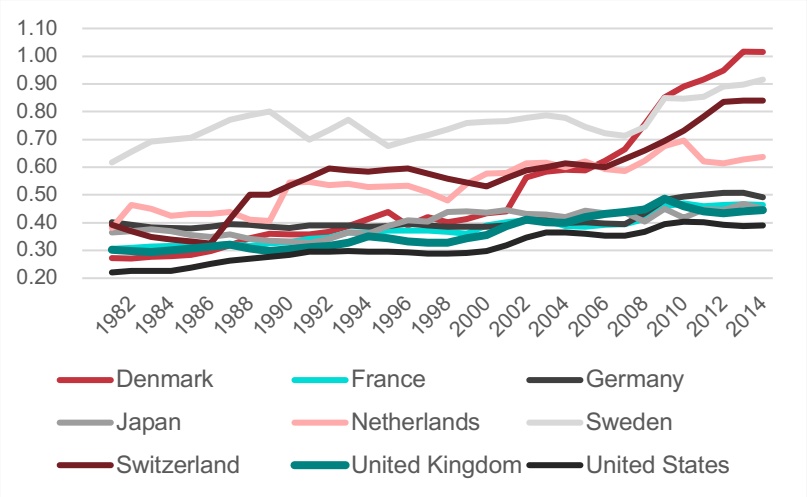
Appendix A Additional material about the UK context

Table 30 Structure of GDP in the UK and comparator countries

Country	Agriculture	Industry	Services
Denmark	4.5%	19.1%	74.6%
France	1.9%	18.3%	79.8%
Germany	0.8%	28.1%	71.1%
Japan	1.2%	27.5%	71.4%
The Netherlands	2.8%	24.1%	73.2%
Sweden	1.8%	26.9%	71.3%
Switzerland	1.3%	21%	78.3%
UK	0.7%	21%	78.3%
USA	1.12%	19.1%	79.7%

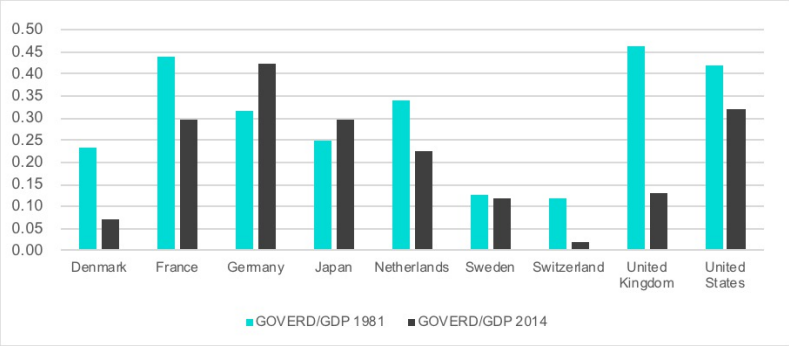
Source: IMF

Figure 26 Higher education expenditure on R&D as a percentage of GDP, 1981-2014



Source: OECD, MSTI, accessed 21/5/16

Figure 27 Government expenditure on R&D as a percentage of GDP



Source: Data from OECD, MSTI, accessed 21/5/16

Table 31 Composition of HEFCE QR funding 1994/5 to 2015/6 (£m in current money)

	Year	Total QR Funding	Funding directly informed by the RAE/REF (mainstream QR)	London weighting for mainstream QR	RDP supervision fund	QR Charity support fund	QR business research element	Special funding for National Research Libraries	RDP supervision funds for departments rated 4 and above	Transitional special funding for research libraries	Supplement for 'best 5 star' departments	Supplementary allocation to maintain funds in real terms	Transitional RDP funds for depts rated 3a in non-capability fund subjects	Supplement for 5*s in 1996 and 2001 RAE	PGR supervision allocation
REF 2014	2015/16	£1,559	£1,017	£33	£240	£198	£64	£7							
	2014/15	£1,558	£1,018	£32	£241	£198	£63	£6							
	2013/14	£1,557	£1,018	£32	£241	£197	£63	£6							
	2012/13	£1,557	£1,018	£32	£241	£197	£63	£6							
	2011/12	£1,558	£1,053	£32	£205	£198	£64	£6							
	2010/11	£1,603	£1,097	£33	£205	£198	£64	£6							
RAE 2008	2009/10	£1,571	£1,074	£32	£203	£192	£63	£6							
	2008/09	£1,436	£920	£29		£185	£62	£6	£199		£25	£12			
	2007/08	£1,391	£892	£28		£180	£60		£193	£2	£25	£10	£2		
	2006/07	£1,319	£936	£29		£136			£188	£2	£24		£3		
	2005/06	£1,228	£981	£33					£184	£2			£5	£24	
	2004/05	£1,061	£932	£32										£24	£73
	2003/04	£1,020	£898	£31									£1	£20	£70
RAE 2001	2002/03	£940	£841	£29									£2		£67
	2001/02	£868	£868												
	2000/01	£847	£847												
	1999/00	£835	£835												
	1998/99	£804	£804												
RAE 1996	1997/98	£684	£684												
	1996/97	£602	£602												
	1995/96	£602	£602												
RAE 1992	1994/95	£591	£591												

Source: data from HEFCE. *For 2004/05 funds for the best 5-star departments includes, where appropriate, funds to maintain funds at the 2003-04 level. Data excludes 'Model fundable home fee-paying research students in 3b rated departments (2002/03) which was less than £1m.

Table 32 Composition of HEFCE QR funding to English Russell Group Universities 1994/5 to 2015/6 (£m in current money)

	Year	Total QR Funding	Funding directly informed by the RAE/REF (mainstream QR)	London weighting for mainstream QR	RDP supervision fund	QR Charity support fund	QR business research element	Special funding for National Research Libraries	RDP supervision funds for departments rated 4 and above	Transitional special funding for research libraries	Supplement for 'best 5 star' departments	Supplementary allocation to maintain funds in real terms	Transitional RDP funds for depts rated 3a in non-capability fund subjects	Supplement for 5*s in 1996 and 2001 RAE	PGR supervision allocation
REF 2014	2015/16	£1,104	£694	£23	£167	£164	£50	£6							
	2014/15	£1,120	£712	£22	£170	£161	£49	£6							
	2013/14	£1,116	£712	£22	£169	£159	£48	£6							
	2012/13	£1,110	£707	£21	£165	£162	£48	£6							
	2011/12	£1,090	£721	£21	£134	£162	£47	£6							
	2010/11	£1,108	£740	£22	£134	£162	£45	£5							
RAE 2008	2009/10	£1,078	£718	£21	£133	£156	£44	£5							
	2008/09	£1,038	£653	£20		£150	£43	£5	£142		£20	£5			
	2007/08	£998	£630	£19		£146	£42		£136	£1	£20	£3	£0		
	2006/07	£951	£663	£20		£113			£132	£1	£20		£1		
	2005/06	£884	£710	£23					£128	£1			£1	£20	
	2004/05	£770	£678	£23										£20	£49
	2003/04	£737	£652	£22									£0	£17	£46
RAE 2001	2002/03	£653	£588	£20									£0		£44
	2001/02	£597	£597												
	2000/01	£581	£581												
	1999/00	£571	£571												
	1998/99	£550	£550												
RAE 1996	1997/98	£461	£461												
	1996/97	£396	£396												
	1995/96	£397	£397												
RAE 1992	1994/95	£381	£381												

Source: data from HEFCE. *For 2004/05 funds for the best 5-star departments includes, where appropriate, funds to maintain funds at the 2003-04 level. Data excludes 'Model fundable home fee-paying research students in 3b rated departments (2002/03) which was less than £1m.

Table 33 Composition of HEFCE non-QR research funding 1994/5 to 2015/6 (£m in current money)

	Year	Total non-QR Funding	Transitional QR funding	Capability fund	GR	DevR	CR
REF 2014	2015/16	£-					
	2010-11-2014/15	£-					
RAE 2008	2009/10	£12	£12				
	2008/09	£22		£22			
	2007/08	£22		£22			
	2006/07	£22		£22			
	2005/06	£22		£22			
	2004/05	£18		£18			
	2003/04	£18		£18			
RAE 2001	2002/03	£-					
	2001/02	£20			£20		
	2000/01	£20			£20		
	1999/00	£20			£20		
	1998/99	£20			£20		
RAE 1996	1997/98	£20			£20		
	1996/97	£36			£20	£16	
	1995/96	£36			£20	£16	
RAE 1992	1994/95	£36			£10	£16	£10

Source: data from HEFCE. **CR** (contract-related research component of funding) was a former component of research funding based on the amount of contract income secured by the institution. This component was replaced by GR funding. **GR** is a variant of the former CR and aimed to encourage institutions to conduct collaborative research programmes, the outcomes of which could be used in a diversity of applications. GR funding allocations are made in proportion to each institution's GR qualifying income. Qualifying income included the amount received from users of research, for collaborative projects where the institution retains the intellectual property and publication rights. **DevR** aimed to encourage the development of research activity by rewarding evidence of research potential in institutions which do not receive substantial research funds.

Table 34 English Russell Group QR funding as a percentage of total QR funding

	Year	Funding directly informed by the RAE/REF (mainstream QR)	London weighting for mainstream QR	RDP supervision funds	QR Charity support fund	QR business research element	Other QR	Total
REF 2014	2015/16	68%	72%	69%	83%	79%	90%	71%
RAE 2008	2014/15	70%	70%	71%	81%	77%	90%	72%
	2013/14	70%	70%	70%	81%	76%	90%	72%
	2012/13	69%	68%	69%	83%	75%	90%	71%
	2011/12	68%	66%	65%	82%	73%	90%	70%
	2010/11	67%	65%	65%	82%	71%	82%	69%
	2009/10	67%	65%	66%	81%	71%	82%	69%
RAE 2001	2008/09	71%	69%	71%	81%	70%	71%	72%
	2007/08	71%	69%	70%	81%	70%	68%	72%
	2006/07	71%	69%	69%	83%		83%	72%
	2005/06	72%	71%	69%			84%	72%
	2004/05	73%	71%				71%	73%
	2003/04	73%	71%	2%			70%	72%
	2002/03	70%	69%	2%			66%	69%
RAE 1996	2001/02	69%						69%
	2000/01	69%						69%
	1999/00	68%						68%
	1998/99	68%						68%
	1997/98	67%						67%
RAE 1992	1996/97	66%						66%
	1995/96	66%						66%
	1994/95	64%						64%

Source: data from HEFCE. **RDP supervision funds** includes 'RDP supervision fund', 'RDP supervision funds for departments rated 4 and above', and 'Transitional RDP supervision funds for departments rated 3a in non-capability fund subjects'. **Other QR** includes 'Special funding for National Research Libraries', 'Transitional special funding for research libraries', 'Supplement for 'best 5 star' departments', 'Supplementary allocation to maintain funds in real terms', 'Supplement for 5*s in 1996 and 2001 RAE', 'PGR supervision allocation' and 'Model fundable home fee-paying research students in 3b rated departments'.

technopolis |group| United Kingdom
3 Pavilion Buildings
Brighton BN1 1EE
United Kingdom
T +44 1273 204320
E info@technopolis-group.com
www.technopolis-group.com