vate universities make use of external research supervisors and enroll a large number of doctoral students. These institutions are heading toward becoming doctoral-degree mills.

The main problems of the private universities relate to the de facto management—the trustees of the sponsoring societies or trusts. They control all financial transactions from the purchase of stationery, to purchase of the most sophisticated equipment. They also have a say in the appointment of faculty. Admissions in many universities are manipulated, though they are supposedly made on merit—determined by annual entrance tests, conducted by the university. The attempts of the government to make admissions on the basis of a single national entrance examination have repeatedly failed. Reportedly, large amounts pass under the table in the form of a "capitation fee" that goes not to the institution, but to the sponsors. The tuition fees are high. The fact of the matter is that a student belonging to a family of average means does not get admission to the well-known private universities. Many private universities, though legally not-for-profit, are actually for-profit institutions. For the "haves" private universities provide a solution to the problem of access to higher education. For the "have nots private universities are a social evil responsible for the widening of the economic and social divides.

■

UK Research Excellence: Getting Better All the Time?

Simon Marginson

Simon Marginson is professor of International Education at the UCL Institute of Education, University College London, and joint editor-in-chief of Higher Education. *E-mail: s.marginson@ioe.ac.uk.*

Each half decade, the UK higher education system puts itself through a massive exercise run by the national higher education regulator, designed to catalogue, compute, and judge university research. This time consuming and intensely competitive process, once known as the Research Assessment Exercise, has become the Research Excellence Framework (REF). The results of the first REF were published just before Christmas.

Purposes of Research Assessment

The REF has a number of purposes, not always consistent with each other. It is used to allocate research-specific funding support and to concentrate resources in the highest performing institutions and disciplines, stretching the national research dollar, as far as possible. It shapes the academic labor market, encouraging researchers to shift to high-performing units, and universities to bid for the best researchers. It is also meant to strengthen the focus on high-quality work—researchers submit their four-best publications for evaluation—and to compare UK research against global standards, while at the same time showcasing that same UK research before the world. It also encourages researchers to focus on the economic and social impact of research, as universities are required to submit evidence of such impact.

Any system of research assessment is only partly reliable as an indicator of the real quality of research. Research assessment has a dual character. On one hand it is rooted in material facts and objective methods. On the other hand, it favors some norms, activities, and interests above others—no assessment can cover everything in the same way, each assessment uses specific and partial methods, and the experienced and high-status players are best at gaming the system.

Some aspects of research, such as citations in top journals, are easier to standardize than other aspects, such as the long-term impacts of research on policy and professions. Comparisons between disciplines, between universities with different missions, between experienced professors and early career researchers, and between established ideas and new ideas are all fraught.

The outcome of the UK REF was partly shaped by the universities that selected and fashioned the data for competitive purposes, and the REF's own subject area panels that defined the research judged to be outstanding on a global scale. Precise league table positions in the REF should be taken with a grain of salt.

Measuring Research Impact?

In the UK REF the indicators for "impact," new to the 2014 assessment, are the most vulnerable to manipulation. This is partly because of the intrinsic difficulty of measuring the changes to society, economy, and policy induced by knowledge, especially in the long term. It is also because of the kind of crafted "impact-related" data that were collected during the REF assessment process. A sophisticated industry has emerged in the manufacture of examples of the relevant "evidence" of impact. The REF assessed simulations of the impact of research, rather than actual impact.

At best, it got everyone thinking about real connectivity with the users of research, which was one of the starting points when producing impact documentation. At worst, the measures of "impact" collapsed into a smoke and mirrors exercise, producing data that bear as much relation to reality as the statements of output made by Soviet factories

in response to official targets.

Inevitably, those universities most adept at managing their response to performance measures of all kinds, performed especially well in producing impact documentation. One suspect there was also the "halo" effect, always associated with all measures contaminated by prior reputation. Thus, research at the University of Cambridge was more likely to be seen to have impact precisely because it was from Cambridge.

Measuring the Quality of Outputs

In the REF output quality was measured using a four-star system, producing a ranking based on the average star level of an institution's researchers (the "grade point average"), and another ranking based on the proportion at 4 star level. These assessments of output quality were grounded in considered judgments of real research work, by panels with expertise. But the standardized value of the output indicators, especially as measures of *comparative* quality, are subject to two caveats.

Some aspects of research, such as citations in top journals, are easier to standardize than other aspects, such as the long-term impacts of research on policy and professions.

Between the 2008 Research Assessment Exercise and the 2014 REF, there was a remarkable inflation of the proportion of UK research outputs judged to be "world leading" (rated 4 star) and "internationally excellent" (rated 3 star). Universities could game the assessment by being selective about whose work they included in their REF submission. Including only the best researchers pushes up the grade point average and the proportion of research ranked 4 star. Universities that do this pay a financial price, in that their apparent volume of research is reduced and their subsequent funding falls. Nevertheless, it is good for reputation, which has many long-term spinoffs, including financial benefits.

In 2008, 14 percent of research outputs were judged to be 4 star, and 37 percent were judged to be 3 star, meaning 51 percent of work was in the top two categories. Six years later in 2014, the proportion of work judged to be world leading or excellent had somehow jumped to 72 percent, with 22 percent judged to be 4 star and 50 percent at 3 star. This phenomenal improvement happened at a time when resources in higher education were constrained by historical standards. "It's getting better all the time," as that Beatles song puts it. But *is* UK research getting better?

While real improvement no doubt occurred in at least some fields, the scale and speed of *this* improvement beggars belief. One suspects that it reflects a combination of factors that generate boosterism. Universities have a vested interest in maximizing their apparent quality. Subject-area panels have a vested interest in maximizing the "world-class" character of their fields. UK higher education is competing with other nations, especially the United States, for research rankings, doctoral students, and offshore income. The system, as a whole, benefits from "it's getting better all the time."

The marketing purpose of the REF appears to have overwhelmed its purpose as an assessment of the global position of UK research. This does not impair the other purposes of the REF, including its roles in funding allocation and research concentration, mediating the internal labor market in researchers, and driving performance through competition. But if competition is intensified while the bar is too low, this is more likely to reward competitiveness per se, than reward genuine global-research excellence.

For UK research, grade inflation is a worrying sign of a system becoming complacent about its own self-defined excellence. This is not the best way to drive long-term improvement. Less hubris and more hardnosed Chinese-style realism would serve the United Kingdom better. The next REF should enhance the role of international opinion in the subject panels and place more emphasis on those areas where improvement is most needed.

The next assessment should also require universities to include all of their researchers or, alternately, a fixed proportion, such as the top 75 or 90 percent. With individual institutions pursuing a variety of strategies on inclusion, the REF did not compare like-with-like. This undermines the validity of the REF as a league table of comparative performance, though everyone treats it that way.

For example, the leader on the volume of high quality research was University College London, a large institution that included 91 percent of its researchers. Oxford was second in the volume of high-quality work and did especially well in measures of average researcher quality. It included 87 percent of its researchers in the count. Oxford's great rival, Cambridge included 95 percent of its researchers, generating a grade point average just below Oxford.

Almost certainly, the best 87 percent of Cambridge researchers outdid those at Oxford, but the REF allowed Oxford to game the process so as to present itself as the best research university in the United Kingdom. Meanwhile the University of Cardiff pushed itself up to equal seventh in the land on grade point average by including just 61 percent of its researchers in the count. ■