

How predictable is the REF?

by Blog Admin

April 24, 2012

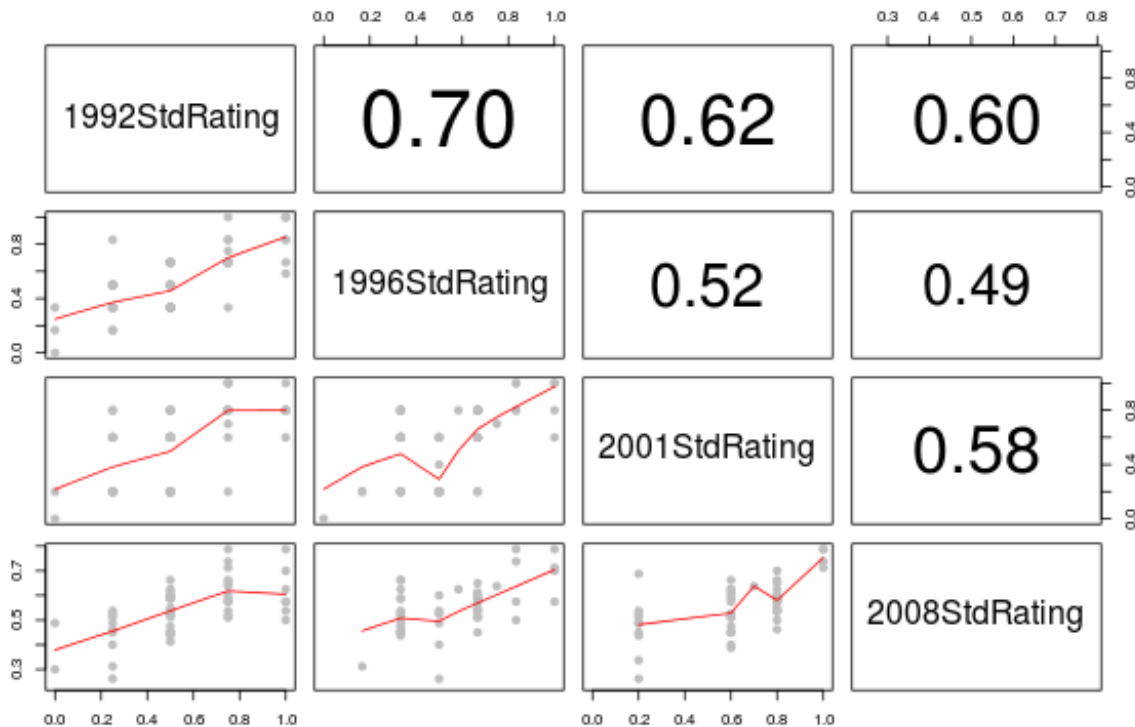
As universities prepare their submissions to the Research Excellence Framework, it's important to know whether the results of the REF could be approximated using other proxy measures. Patrick Dunleavy has [argued vocally](#) for using bibliometrics however, [Chris Hanretty](#) investigates how predictable the REF is — and whether we can generate predictions now based on leading indicators of “research excellence”.



One of the most obvious ways in which to assess the predictability of the REF, or similar exercises, is to check whether performance at time (t) predicts performance at time ($t+1$). The more past performance predicts present performance, the more we can rely on historical allocations, putting our fingers on the scale for departments that have made big improvements.

In order to do this, we must put RAE/REF rankings on a common metric. The current REF uses five grades (Unclassified, 1*, 2*, 3*, 4* — one wonders what extra force the stars add), the 1996 and 2001 RAEs seven grades (1, 2, 3a, 3b, 4, 5, 5*), and the 1992 RAE six (as above, but a single '3' grade in place of 3a and 3b).

If we assume that these grades are all equally spaced — and you might reasonably think that a 2001 '3a' and '3b' are closer together than a 2001 '1' and '2' — then we can convert these grades to a common metric by setting the lowest grade to zero, the highest grade to one, and dividing accordingly.



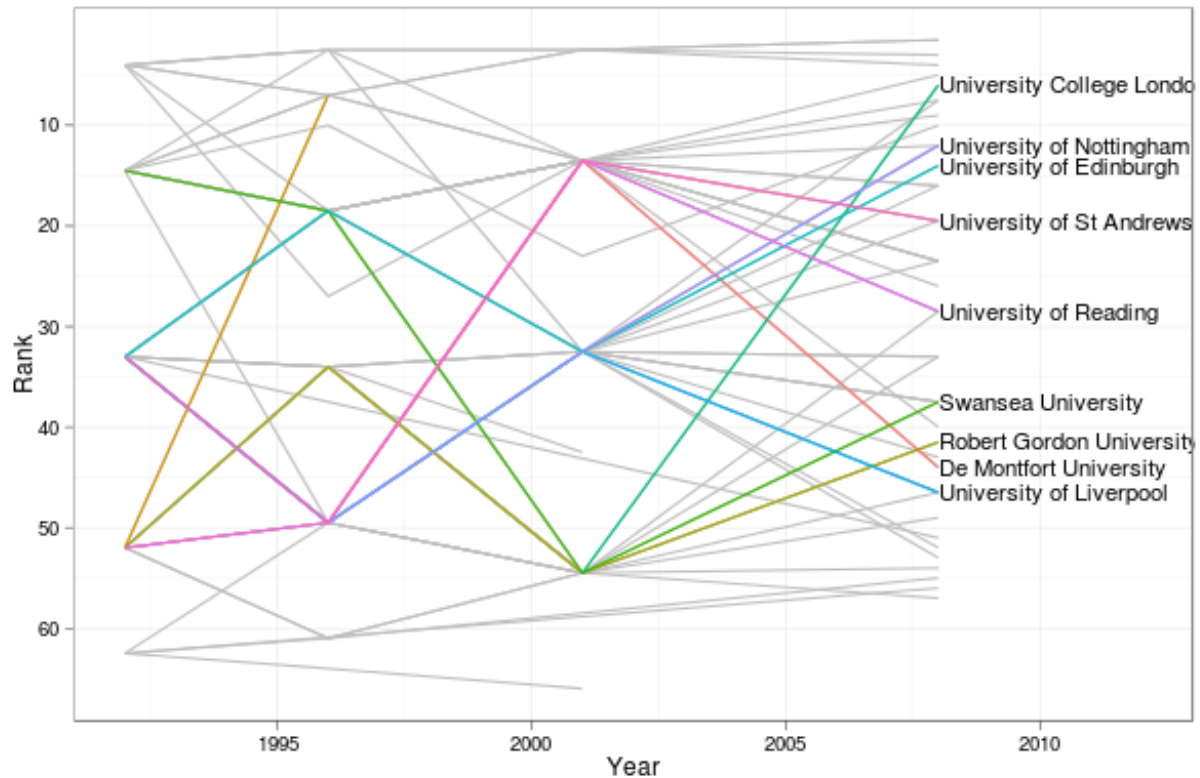
The figure above shows the pairwise correlations between average REF/RAE grades for departments across the four research exercises so far, for departments submitting returns in “Politics and International Relations”. All of the correlations reported are rank correlations, because in 1992 and 1996 the grade the department received was a single grade, rather than an average of grades.

The correlations are all reasonably strong. They wouldn't pass the (somewhat arbitrary) 0.7 rule of thumb

for test-retest reliability — but since we are allowing for department performance to vary over time, that's not a problem.

These correlations are weaker than similar correlations for the REF subject heading “Economics and econometrics”, all of which are greater than 0.7, and one of which (1992 to 1996) is very high indeed, at 0.9.

We can show some of the variability that is present in the data by using a [bumps chart](#). Plotting all of the year-to-year variation in a bumps chart gets messy very quickly, so here I've focused on nine of the more variable departments.



No doubt individuals familiar with the histories of these departments will be able to say what happened such that RAE performance was so volatile. Volatile departments such as these represent exceptions to the general rule of stability across RAE cycles. Whether the new REF, with its additional impact elements, will shake up the rank-ordering of universities substantially, remains to be seen.

This piece was originally published on Chris Hanretty's [personal blog](#) as the first in a series of six blog posts analysing the REF, which can be [found here](#).

Note: This article gives the views of the author(s), and not the position of the Impact of Social Sciences blog, nor of the London School of Economics

No related posts.