

Analysis of data provided by Professor Paul Ewart

Daniel Lunn

May 16, 2017

Department of Statistics
University of Oxford

Introduction

The data comprise figures taken from *Annexe F* of *EJRA: Review Working Group Report* and answers to three specific questions have been requested. These are:

1. Do the data provide evidence that the introduction of the EJRA in 2011 has had the effect of improving the overall proportion of females in the Academic and Academic-related staff groups subject to retirement under the EJRA (The EJRA population)?
2. Do the data provide evidence that the introduction of the EJRA in 2011 has had the effect of improving the overall proportion of females in the Associate Professor grade?
3. Do the data provide evidence that the introduction of the EJRA in 2011 has had the effect of improving the overall proportion of females in the Statutory Professor grade?

There was also the supplementary question:

In addition to these specific questions I would be grateful if the data showing the proportion of Associate Professors in each age band can be analysed to give any information about the age at which people are appointed to this grade?

Statistical analysis

Although the questions specifically concern Academic and Academic-related, Associate Professor and Statutory Professor grades, RS IVs have also been included in the analysis.

In order to cater for lack of independence of year-by-year proportions log-odds growth models with variance weighted according to sample size were fitted to the years 2006 - 2015; the response variable for each grade was the log of the Female/Male ratio. An indicator variable (EJRA) was included to enable adjustment for a change in growth rate from 2012 - 2015.

Models were fitted to each of the four grades. Formal Shapiro-Wilk tests for normality of the model residuals were carried out with resulting p -values for Statutory Professor ($p = 0.3119$), Associate Professor ($p = 0.4762$), RSIV ($p = 0.3508$), Academic & Academic-related ($p = 0.5349$); the model residuals showed no evidence of heteroscedasticity or serial correlation and therefore there was no reason to doubt the adequacy of the fitted models.

Statutory Professor

The fitted model is shown in the table below. The variable Year:EJRA is the additive adjustment to the coefficient of Year to cater for a change over the post 2011 period; thus a value different from zero would show a change and the p -value produced is for a test of the null hypothesis that there is no change.

<i>Coefficient</i>	<i>Value</i>	<i>Std.Error</i>	<i>t-value</i>	<i>p-value</i>
(Intercept)	-77.89118	34.94054	-2.2292496	0.0610
Year	0.03769	0.01740	2.1664076	0.0670
Year:EJRA	0.00000	0.00005	0.0034333	0.9974

There is very weak evidence of a steady growth in the Female/Male ratio ($p = 0.0670$), but there is no evidence to support a change in growth-rate post 2011 ($p = 0.9974$). The Year coefficient for $\log(\text{Female/Male})$ was 0.03769, giving a multiplicative annual growth of $\exp(0.03769) = 1.0384$ or a compound 3.84% growth per annum which does not change from 2006 - 2015.

Associate Professor

<i>Coefficient</i>	<i>Value</i>	<i>Std.Error</i>	<i>t-value</i>	<i>p-value</i>
(Intercept)	-42.66764	5.566648	-7.664872	0.0001
Year	0.02071	0.002772	7.470826	0.0001
Year:EJRA	0.00000	0.000008	0.127355	0.9022

Here there is very strong evidence of a steady growth in the Female/Male ratio ($p = 0.0001$), but again no evidence of any change post 2011 ($p = 0.9022$). The annual growth rate of $\log(\text{Female/Male})$ was 0.02071 giving a multiplicative annual growth of $\exp(0.02071) = 1.0209$ or a compound 2.1% growth per annum which is unchanged from 2011 - 2015.

RS IV

<i>Coefficient</i>	<i>Value</i>	<i>Std.Error</i>	<i>t-value</i>	<i>p-value</i>
(Intercept)	19.252244	26.265162	0.7329955	0.4874
Year	-0.010089	0.013074	-0.7716416	0.4656
Year:EJRA	-0.000033	0.000035	-0.9337593	0.3815

There is no evidence of any annual growth in Female/Male ratio ($p = 0.4656$) and no evidence of change post 2011 ($p = 0.3815$).

Academic & Academic-related

<i>Coefficient</i>	<i>Value</i>	<i>Std.Error</i>	<i>t-value</i>	<i>p-value</i>
(Intercept)	-45.66299	10.136028	-4.505019	0.0028
Year	0.02270	0.005046	4.498659	0.0028
Year:EJRA	-0.00002	0.000014	-1.137263	0.2929

There is strong evidence of steady annual growth ($p = 0.0028$) but no evidence of a change post 2011 ($p = 0.2929$). There is multiplicative growth of $\exp(0.02270) = 1.0230$ or a compound annual growth of 2.3% which is unchanged from 2006 - 2015..

Overall summary for the four grades

With the exception of RS IV, all grades showed significant evidence of annual growth in the Female/Male ratio and the annual growth rates did not change after 2011 in any grade.

For simultaneous inference from four such models, Bonferroni corrections should be applied to the p -values of the estimated coefficients of Year. Thus in the Statutory Professor grade, although the coefficient for $\log(\text{Female/Male})$ was higher than in any other grade at 0.03769, and gave a compound 3.84% growth per annum, with a Bonferroni-corrected p -value of 0.27 it is not close to being statistically significant. It should, however, be noted that, with comparatively low numbers of Statutory Professors, this result is hardly surprising.

However, the main result lies in there being no significant change detected in any of the growth rates post 2011. For Statutory Professors and Associate Professors the estimated change is zero to five decimal places and, although RS IV and Academic & Academic-related come out as slightly negative, the smallest uncorrected p -value is 0.2029 and there is no evidence for considering it to be anything other than zero.

Supplementary question

Given the available data, the question as stated is unanswerable. However, it is possible to compare the distributions of proportions across the different age groups for different years. In other words, the question “Has the distribution across age groups changed from 2006 to 2015?” is easily answered.

There are two possible interesting comparisons depending on whether or not the over age 67s are to be included. The figures for these are:

<i>Age group</i>	<i>Year 2006</i>	<i>Year 2015</i>	<i>Age group</i>	<i>Year 2006</i>	<i>Year 2015</i>
Under 30	15	14	Under 30	15	14
30 – 39	238	445	30 – 39	238	445
40 – 49	244	486	40 – 49	244	486
50 – 59	204	387	50 – 59	204	387
60 – 67	61	132	60 +	63	149

Excluding the over 67s a chi-squared test of independence of Age Group and Year has a p -value of 0.2999, so there is no evidence of dependence; in other words the relative proportions recorded in 2006 have not changed by 2015. This conclusion is not changed when the over 67s are included, the p -value being 0.1857. Whichever data are included, there is no evidence of any change in the distribution of relative proportions across the Age groups between 2006 to 2015.

In conclusion

There are two noticeable things in the reports sent to me by Professor Ewart which, even though not specified by him for comment, must be addressed.

- Practically every comparison in the report is made in terms of percentages or percentage changes and these are even plotted and compared when they are derived from different totals. This is worse than bad practice – it is incompetence. An illustration of how misleading this can be can be seen from the analyses above. A higher annual change was found in the Statutory Professor grade than in, say, the Associate Professor grade, yet the former, unlike the latter, was not statistically significant; as pointed out this was because the proportions were calculated from very different totals. If nothing else, it demonstrates that meaningful comparisons cannot be made in terms of percentages alone and that total numbers are essential if misleading conclusions are to be avoided.
- The only reference to statistical analysis I noticed anywhere showed a total and alarming lack of understanding of the subject. At the end of the paragraph below Figure 8 on the last page (page 20) of the *University of Oxford: Equality Report 2013/2014 Section B: Staff equality data* there appears the statement "...however, there was a large disparity (10%) in the Humanities division, which attained statistical significance" with a footnote "Pearson's Chi-squared test: 95% significance level". Clearly the author of this nonsensical statement does not understand the concept of Type 1 error in hypothesis testing. Furthermore, even if one were to take the charitable view that this is a typographical error and should read "less than 5%", it is clear that a Bonferroni-type correction for simultaneous inference has not been made or even considered to be relevant.

In conclusion, it has to be said that everything I have seen in the report sent to me by Professor Ewart displays a lack of statistical understanding and competence which has no place in any published report produced by a University.