Implications of fully modelling the National Student "Survey" (NSS)

Paul Hewson

June 23, 2011

Abstract

The NSS has been conducted annually in the UK since 2005. In common with other performance indicator systems there is concern that optimising results might take precedence over using the results to inform best practice. Moreover, the exact interpretation of the response is unclear - how does a tick in one box correspond to perception of a learning experience.

Nevertheless, this paper reports on Bayesian models fitted to the "survey" taking account of personal, institutional and subject specific characteristics as well as accounting for non-response bias. There are surprisingly clear patterns in terms of subject group responses to different questions; for example students who have been on placement rate their confidence more highly, students on subjects (such as mathematics) where objective marking is more readily attainable agreeing that their assessment was "fair".

Fuller results of the analysis are reported, and implications for enhancement are discussed.

1 Background

The UK origins of the National Student "Survey" (NSS) lie in a decision by the Higher Education Funding Council for England (HEFCE) to publish comparative data in order to reduce the cost of reviews [Williams and Brennan, 2003]. An early pilot took place on recent graduates in 2004 and the NSS went live in 2005 [Richardson et al., 2007] using an external data collection agency.

There are concerns that the philosophy behind the "survey" instrument is more about customer satisfaction than say the US National Survey of Student Engagement (relevant given current debates in HE, e.g. Ramsden [2008]). Timing is also important, as the pilot study suggested more positive responses were obtained after graduation. Although the validity of using league tables with performance indicators is contested [Bird et al., 2005] the NSS is used in this way and is a key marketing point. This paper briefly summarises some work which looks at other information that can be gleaned from the NSS.

One key problem is response rate, currently around 60%. Richardson et al. [2007] reported institutionally differential non-response. As the "survey" responses are matched to demographic data from the entire sample frame it has been possible to examine non-response directly. Non-response does not occur at random, for example in a model containing a number of other terms (ethnicity, institution, subject group) the 95% Credible Interval for the odds ratio of responding for a male versus a female is (0.60 0.78), indicating that (many other things being equal) males are less likely to respond to the "survey" than females. This makes careful modelling essential.

2 Modelling responses

In the interests of space we can only present results for three questions based on reducing the original five point response to a binary response [Fielding et al., 2010], i.e., we model whether a responding student agreed/strongly agreed with the statement or instead responded with neutral or disagreed/strongly disagreed. We have been fitting Bayesian hierarchical models for "agreement" by Markov chain Monte Carlo (MCMC) - a now standard approach [Gelman et al., 2003]. These full probability models provide information via so-called "posterior distributions" on the effect of individual level predictor variables (for example, all things being equal, females tend to be more likely to agree with the questions than males), the relative effects of institutions and subject groupings. Using a full probability model means that inference drawn from each posterior distribution takes into account both the missing data (if our model is correct) and the various explanatory variables and also enables posterior predictive distributions for the missing data and can form which can be used to provide imputed results for any subject / institution hopefully free of non-response bias.

Figure 1 summarises the posterior distribution for the context adjusted differential effect of subject groups to "agreement" with Question 22. In other

words, how a respondent in each subject group differed from respondents in other subject groups once we had accounted for overall agreement, individual level variables as well as institutional level random effects.

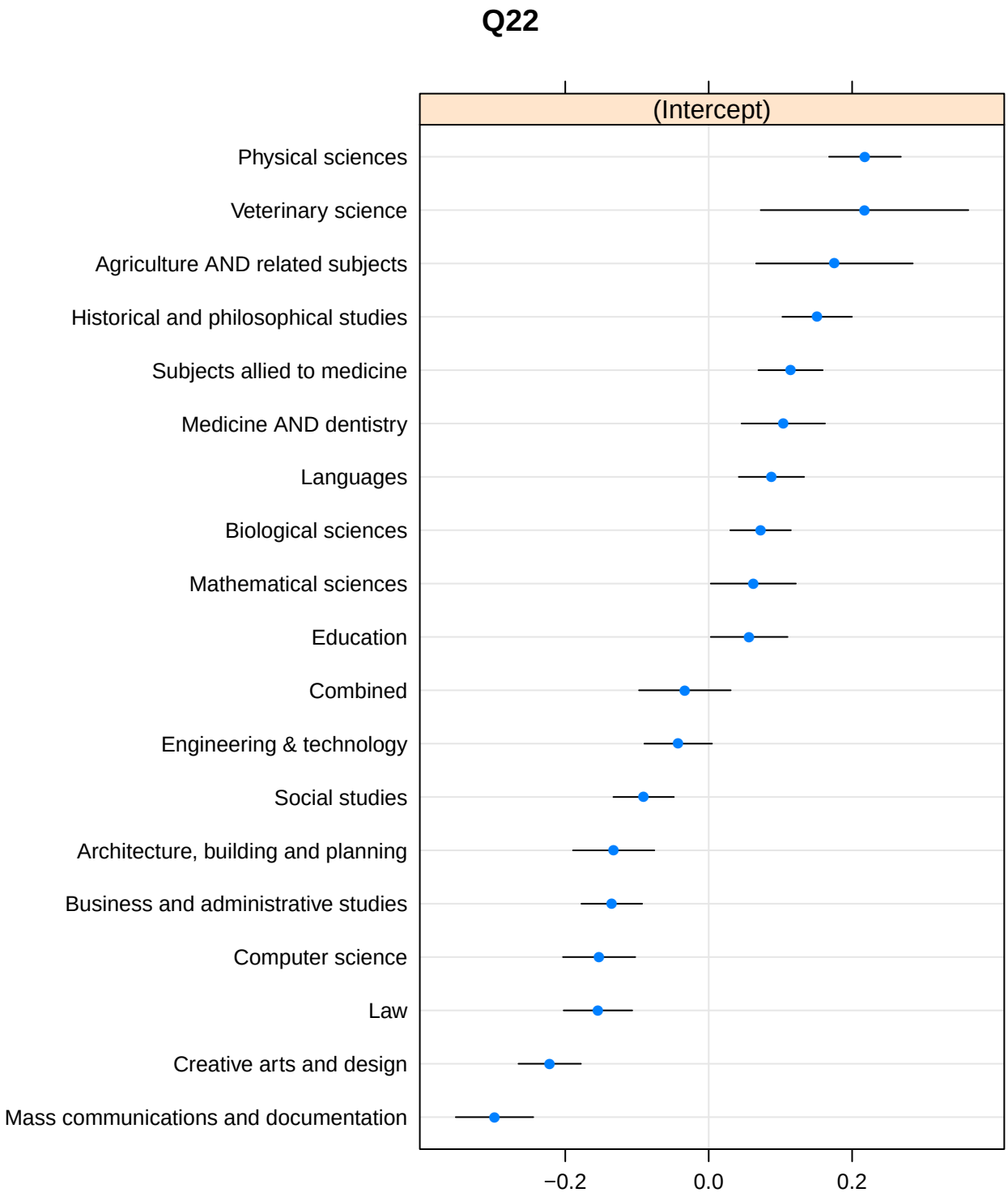

Figure 1: Random effects for subject group: Question 22 "Overall, I am satisfied with the quality of the course."

Figure 2 which gives the posterior distribution for the subject differential "agreement" with question 6. For example, Mathematics student respondents to the 2010 NSS reported favourably on the fairness of their assessments.

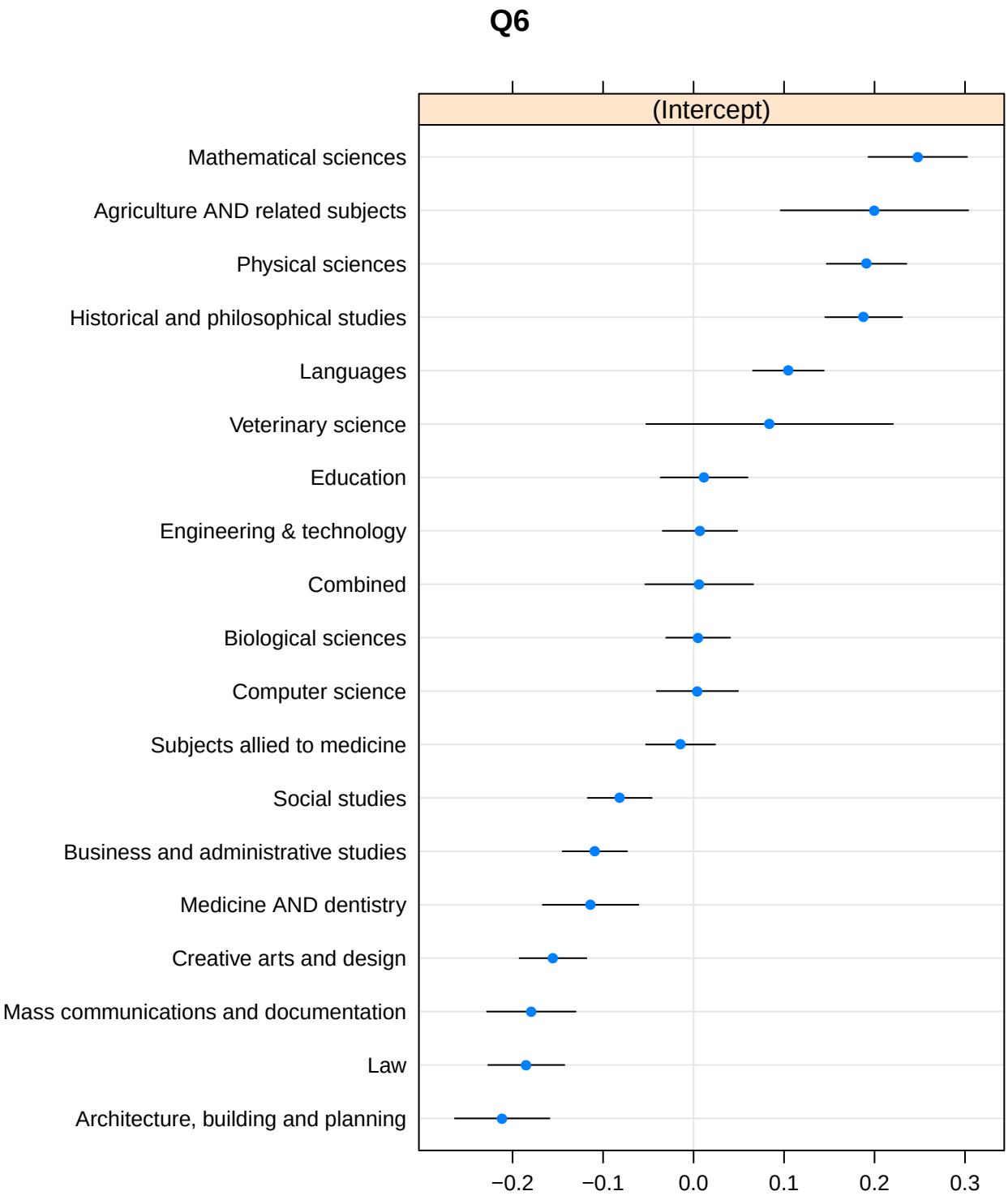

Figure 2: Random effects for subject group: Question 6 "Assessment arrangements and marking have been fair."

Finally, figure 3 summarises "differential effect" for Question 19. It is notable that all the subject groups that scored highly include practical placements.

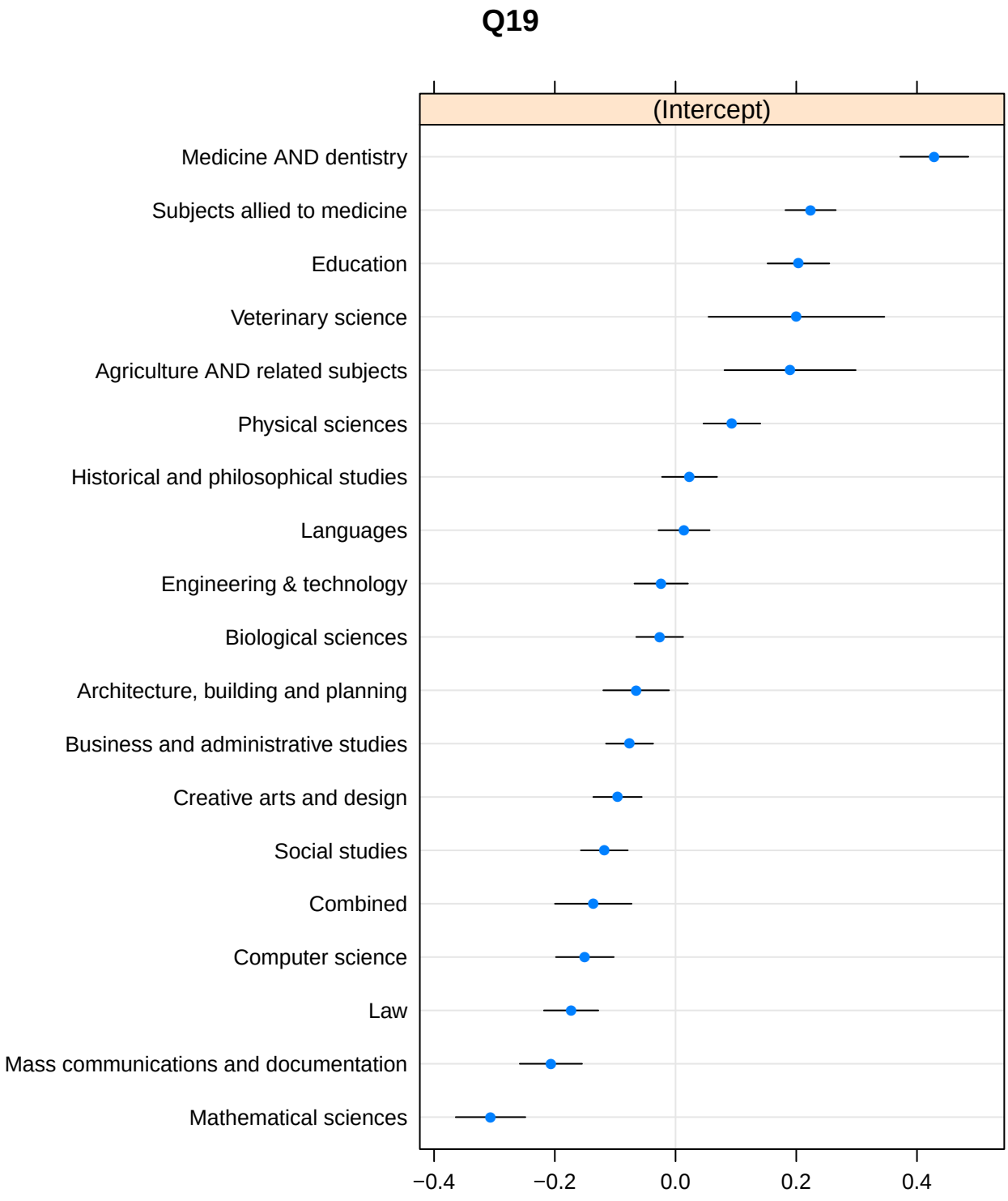

Figure 3: Random effects for subject group: Question 19 "The course has helped me present myself with confidence"

3 Conclusion

The missing data problem and the consequent need for modelling (and for the model assumptions to be correct) cannot be overstated. Neither can the need to understand student perceptions of the questions [Flint et al., 2009] as well as ticks in boxes.

The hint that students with placements rated their confidence more highly than others serves as a reminder that there may be instrinsic characterstics of subject groups that render comparison problematic. Nevertheless, the potential to use these results as a "can opener" (opening a can of worms) to suggest productive avenues for further investigation is clear.

References

S. M. Bird, C. Sir David, V. T. Farewell, G. Harvey, H. Tim, and S. Peter C. Performance indicators: good, bad, and ugly. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(1):1–27, 2005. ISSN 1467-985X. doi: 10.1111/j.1467-985X.2004.00333.x. URL http://dx.doi.org/10.1111/j.1467-985X.2004.00333.x.

A. Fielding, P. J. Dunleavy, and A. M. Langan. Interpreting context to the UKs national student (Satisfaction) survey data for science subjects. *Journal of Further and Higher Education*, 34(3):347, 2010. ISSN 0309-877X. doi: 10.1080/0309877X.2010.484054. URL http://www.informaworld.com/10.1080/0309877X.2010.484054.

A. Flint, A. Oxley, P. Helm, and S. Bradley. Preparing for success: one institution's aspirational and student focused response to the national student survey. *Teaching in Higher Education*, 14(6):607, 2009. ISSN 1356-2517. doi: 10.1080/13562510903315035. URL http://www.informaworld.com/10.1080/13562510903315035.

A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Chapman and Hall / CRC, Boca Raton, 2nd edition, 2003.

P. Ramsden. The future of higher education teaching and the student experience. Technical report, Higher Education Academy, York, 2008.

J. T. E. Richardson, J. B. Slater, and J. Wilson. The national student survey: development, findings and implications. *Studies in Higher Education*, 32 (5):557, 2007. ISSN 0307-5079. doi: 10.1080/03075070701573757. URL http://www.informaworld.com/10.1080/03075070701573757.

R. Williams and J. Brennan. Collecting and using student feedback on quality and standards of learning and teaching in higher education. http://oro.open.ac.uk/11876/, May 2003. URL http://oro.open.ac.uk/11876/.