Bayesian Analysis (2006) **1**, Number 3, pp. 547–550

Rejoinder

William J. Browne*, and David Draper†

We are grateful to Gelman, Kass and Natarajan, and Lambert for their thoughtful comments (and indeed for the original research that they summarize in their papers), and we offer the following remarks by way of rejoinder.

- Many of the results presented in our article were obtained more than a few years ago (based, as they were, on part of the work in Browne (1998)) and are only now seeing the light of publication largely due to, shall we say, the vagaries of non-Bayesian refereeing. We focused on the $\Gamma^{-1}(\epsilon, \epsilon)$ prior for random-effects variances in some of our work because—under the influence of the `WinBUGS` package and the examples distributed with it—this was very much the most common prior in use in hierarchical/multilevel modeling in the mid to late 1990s. Lambert expresses the opinion that this is still true today, although it appears to us that the pendulum is shifting away from this prior, for reasons like those mentioned by Gelman. (To be fair to the `WinBUGS` development group, in many of the examples distributed with release 1.4.1 they currently offer analyses with both $\Gamma(0.001, 0.001)$ priors on random-effects precision parameters τ and Uniform priors on the corresponding standard deviation parameters $\sigma = \tau^{-1/2}$, although they send a distinctly mixed message by building in default values of 0.001 for each of the shape and scale parameters whenever a parameter is given a Gamma distribution in the `DoodleBUGS` part of the package.)

  It is interesting to see that in 2006 there is still no consensus on a general-purpose choice of diffuse prior for this situation, although the work summarized in both the Gelman and Kass-Natarajan contributions to this discussion may go some distance toward achieving this goal. We have found ourselves recently gravitating toward Uniform priors on random-effects standard deviations, which accord with one of Gelman's suggestions, although instead of using Uniform$(0, \infty)$ (or Uniform$(0, A)$ for huge A) we prefer Uniform$(0, c)$ where c is chosen just large enough not to truncate the marginal likelihood for σ (and, in an interesting resurrection of the sometimes appropriately maligned Gamma prior, c can often be chosen well by making a preliminary fitting with a $\Gamma^{-1}(0.001, 0.001)$ prior on σ^2 and looking at the marginal posterior for σ). It is also interesting that $\Gamma^{-1}(\epsilon, \epsilon)$ priors were originally chosen for computational convenience (through their conditional conjugacy), and the half t family mentioned by Gelman again has surfaced due to computational benefits, this time arising from model expansion. One of us (Browne (2004)) has also seen these benefits in a more complex random effects model, reinforcing Gelman's comments on efficiency of MCMC chains.

*Division of Statistics, School of Mathematical Sciences, University of Nottingham, UK http://www.maths.nottingham.ac.uk/personal/pmzwjb/

†Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA, http://www.ams.ucsc.edu/~draper

 ba0003