

SUPPLEMENT TO “A BAYESIAN HIERARCHICAL MODEL FOR INFERRING PLAYER STRATEGY TYPES IN A NUMBER GUESSING GAME”

BY P. RICHARD HAHN, INDRANIL GOWSWAMI AND CARL F. MELA

University of Chicago and Duke University

This supplement contains computational details and a sensitivity analysis corresponding to the hierarchical spline model presented in “A Bayesian hierarchical model for inferring player strategy types in a number guessing game”.

1. Computational details. Posterior sampling is performed via a Metropolis-within-Gibbs approach. We sequentially sample parameters, given the current value of all other parameters, according to the Metropolis-Hastings acceptance probabilities. A challenge with this approach is to devise suitable proposal densities. We utilize a random-walk approach. As many of the model parameters are restricted to the unit interval, we introduce a latent variable representation using the “wrapping function”:

$$g(x) = x - \lfloor x \rfloor + \mathbb{1}(x < 0),$$

where $\lfloor x \rfloor$ denotes the integer part of x . This function maps numbers in the unit interval to themselves, while numbers outside the unit interval get mapped back to their fractional part, in the case of positive numbers, or one minus their fractional part, in the case of negative numbers. So 0.8 gets mapped to itself; 1.3 gets mapped to 0.3; -2.4 gets mapped to 0.6. For a parameter, such as μ_0 , restricted to the unit interval, we conduct a random walk over parameters $\tilde{\mu}$ on the whole real line and define $\mu \equiv g(\tilde{\mu})$. The $\tilde{\mu}$ parameters are unidentified, but this is irrelevant as we report inferences and define priors on the original, identified scale. (Note that η_i is restricted to $[0.3, 0.7]$ rather than $[0, 1]$, so $g(\cdot)$ must be modified accordingly by a simple rescaling.)

For the ϕ_i parameters, which can be exactly zero with positive probability, we also introduce a binary latent variable z_i and define $\phi_i = z_i g(\tilde{\phi}_i)$. We conduct the random walk on the real line with $\tilde{\phi}$ and transform to ϕ_i for likelihood evaluations.

1. For each i , sample $(\eta_i, \phi_i, \nu_i, \mu_i \mid -)$ according to a Metropolis ratio using likelihood and prior given in (2) and (3), (4) and (5) (of the

main text) respectively. Proposal density evaluations can be avoided by using a symmetric random walk over the elements of $\tilde{\eta}_i, \tilde{\phi}_i, \tilde{\nu}_i$ and $\tilde{\mu}_i^0$, centered at the current values.

2. For each i , sample $(z_i | -)$ using a straightforward application of Bayes rule.
3. For each i , sample $(s_i | -)$ using a straightforward application of Bayes rule.
4. Sample $(\mathbf{w} | -)$ from a Dirichlet distribution with parameter $\alpha^* = \alpha + \kappa$, where κ records counts of how many observations are currently assigned to each level of s .
5. Sample $(\rho | -)$ as a Beta(γ, β) random variable. Let $n_1 = \sum_i z_i$ and $n_0 = \sum_i (1 - z_i)$. Then $\gamma = 3 + n_0$ and $\beta = 1 + n_1$.
6. Sample $(q_0 | -)$ as a Beta(γ, β) random variable. Let n_{q_0} be the number of players for whom both $\phi_i = 0$ and $\nu_i < \mu_i^0 \eta_i$. Then $\gamma = 3 + n_{q_0}$ and $\beta = 1 + n_0 - n_{q_0}$.
7. Sample $(q_1 | -)$ as a Beta(γ, β) random variable. Let n_{q_1} be the number of players for whom both $\phi_i \neq 0$ and $\nu_i < \mu_i^0 \eta_i$. Then $\gamma = 3 + n_{q_1}$ and $\beta = 1 + n_0 - n_{q_1}$.

2. Model diagnostics. This section presents various model diagnostics: measures of Markov chain Monte Carlo convergence, a sensitivity analysis to determine if our conclusions are impacted by our specific choice of games, $p \in \{0.3, 0.4, 0.5, 0.6, 0.7, 1\}$, and an assessment of our three-knot spline model in terms of approximating true k -step strategies.

2.1. Convergence diagnostics. Figures 1 and 2 show trace plots for key parameters from the MCMC sampler. By this rough metric, mixing looks adequate. Subject-specific samples are shown, arbitrarily, for participant #100. Although hard to discern in the plot for ϕ , 75% of the samples drawn are $\phi = 0$. Figure 3 shows a Geweke-Brooks diagnostic plot for ρ, q_0 and q_1 ; for an explanation of this plot, see Plummer et al. [2006].

2.2. Design sensitivity analysis. Although our study represents an advance on the current literature in terms of collecting multiple responses from the same subject across different values of p without feedback, it is only natural to wonder if our particular experimental design of $p \in \{0.3, 0.4, 0.5, 0.6, 0.7, 1\}$ strongly influenced our results. Here we address this question in two parts. First, we conduct a sensitivity analysis, where we re-run our analysis, but omitting the data from individual values of p , one by one. Figure 4 shows how our posterior conclusions vary across these different designs. Essentially the only relevant held-out design point was $p = 0.3$,

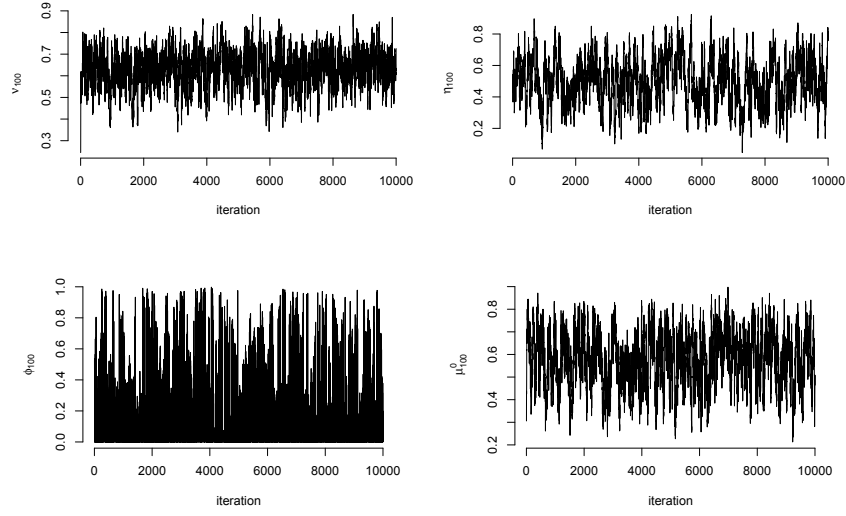


FIG 1. Trace plots of model parameters ν , η , ϕ , and μ for subject #100 (for illustration purposes).

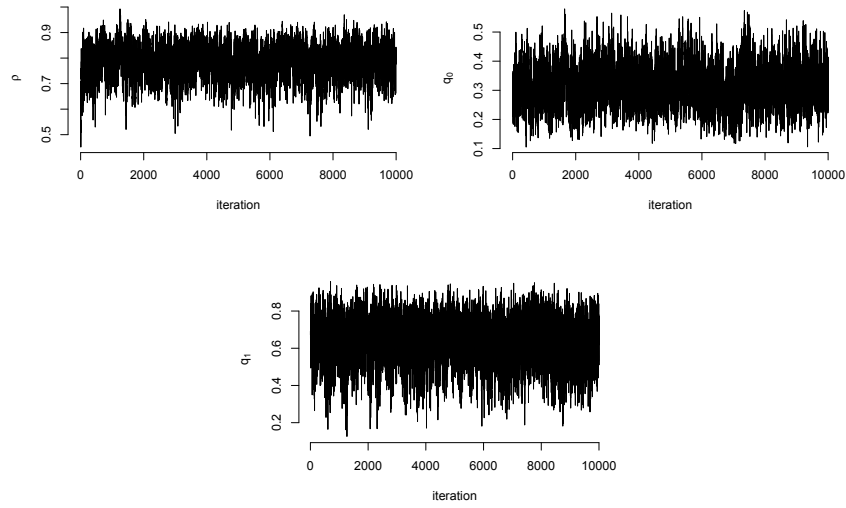


FIG 2. Trace plots for shared parameters ρ , q_0 and q_1 .

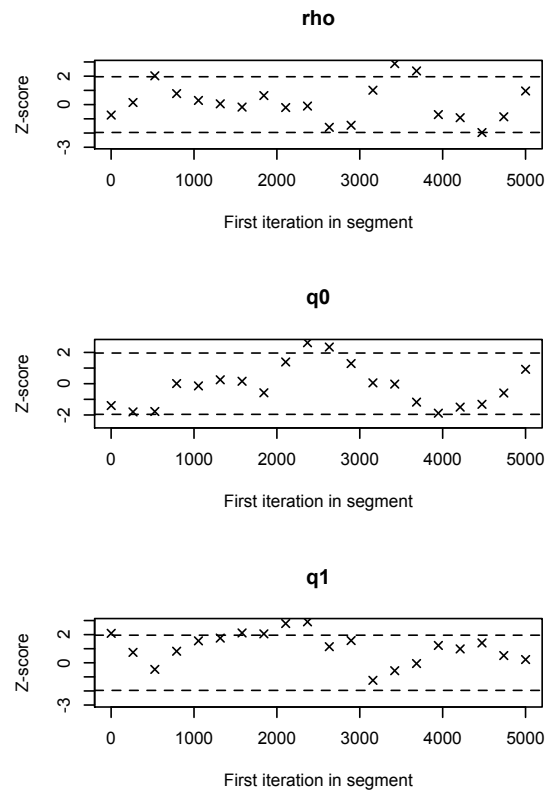


FIG 3. Geweke-Brooks diagnostic plot for ρ , q_0 and q_1 ; see Plummer et al. [2006] for a detailed explanation of how to read this plot.

which yielded higher posterior probability that any given subject playing a strategy through the origin ($\phi = 0$), leading to higher posterior estimates of ρ . Interestingly, this increased estimate of ρ was counteracted by a decrease in the probability of playing a concave strategy, yielding an essentially unchanged posterior over the probability of k -step compliance ρq_0 .

Even so, as $p = 0.3$ was an important design point, one might wonder how our results would differ had we included $p = 0$ in our study. In fact, our pilot data did include $p = 0$ as one of the design points. Based on this pilot data, we found a surprisingly high proportion of subjects who did not play 0 at $p = 0$. In our follow-up data collection, we opted not to include $p = 0$ in an effort to give every benefit of the doubt to the k -step hypothesis. In fact, when our model is applied to the pilot data, posteriors over the key quantity ρq_0 are notably lower, with posterior mean of 10% (compared to the 25% reported above).

2.3. Three-knot spline approximation and residual analysis. In light of the fact that k -step strategies are potentially high-order monotone polynomials on $[0, 1]$, it is natural to wonder how good or bad our three-knot spline approximation is. To quantify the goodness of this approximation formally is beyond the scope of this paper, but here we present numerical evidence that the approximation is adequate. To visualize our approximation, Figure 5 shows 6 example draws of bona fide k -step strategies and the corresponding best-fitting (in terms of squared error) three-knot spline approximation. While there is obvious disagreement between the two curves, the magnitude of the mis-fit is small relative to the curvature of the underlying strategy. Specifically, we see that errors due to this approximation may exhibit negative auto-correlation, for example, as shown in the sixth panel of Figure 5, on either side of $p = 0.8$.

Because our game response data is restricted to the unit interval (after scaling by dividing by 100), a straightforward residual analysis looking at the simple difference from of the observed data from the mean strategy curve is inappropriate. Instead, we evaluate the appropriateness of our beta error model by examining the normalized residual defined as

$$r_{i,j} = \Phi^{-1}\{F_{i,j}(y_{i,j})\}$$

where $\Phi^{-1}(\cdot)$ is the inverse-CDF of a standard normal distribution and $F_{i,j}(\cdot)$ is the cumulative distribution function of the beta distribution corresponding to subject $i \in \{1, \dots, 106\}$ at design point $j \in \{0.3, 0.4, 0.5, 0.6, 0.7, 1\}$. In fact, we have $M = 15,000$ posterior samples of each $r_{i,j}$. For visualization, we consider the posterior mean residuals $\bar{r}_{i,j} = M^{-1} \sum_{l=1}^M r_{i,j}^l$, where $r_{i,j}^l$ denotes the l th Monte Carlo sample.

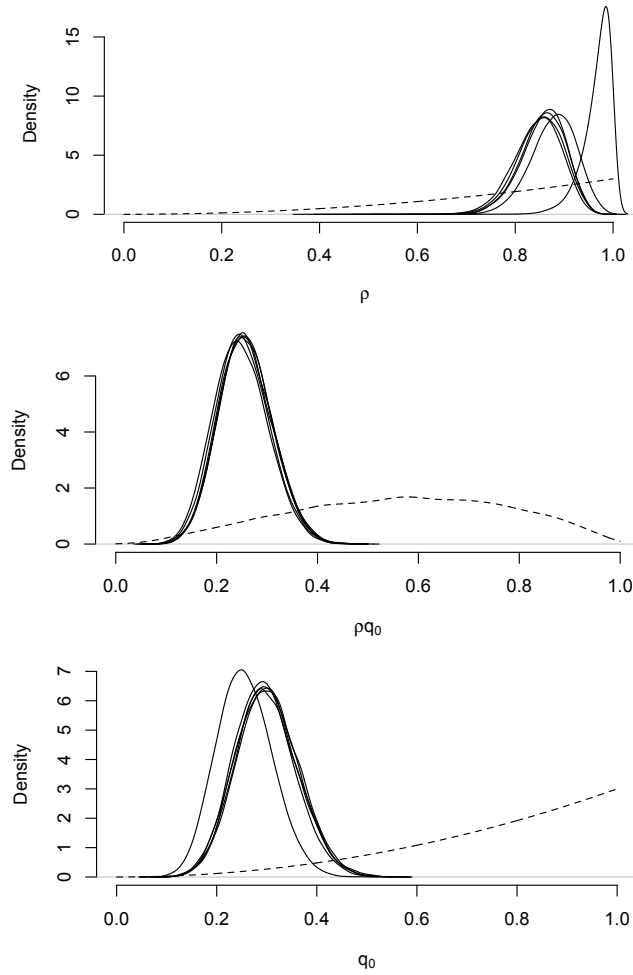


FIG 4. Posterior densities from analyses with a single design point $p \in \{0.3, 0.4, 0.5, 0.6, 0.7, 1\}$ omitted at a time. The dashed line shows the prior density. Only the $p = 0.3$ design point had a material impact, leading to a posterior for ρ that is shifted to the right and a posterior for q_0 that is shifted to the left. Interestingly, these changes offset when considering ρq_0 , the probability of k -step compliance, which appears relatively unchanged.

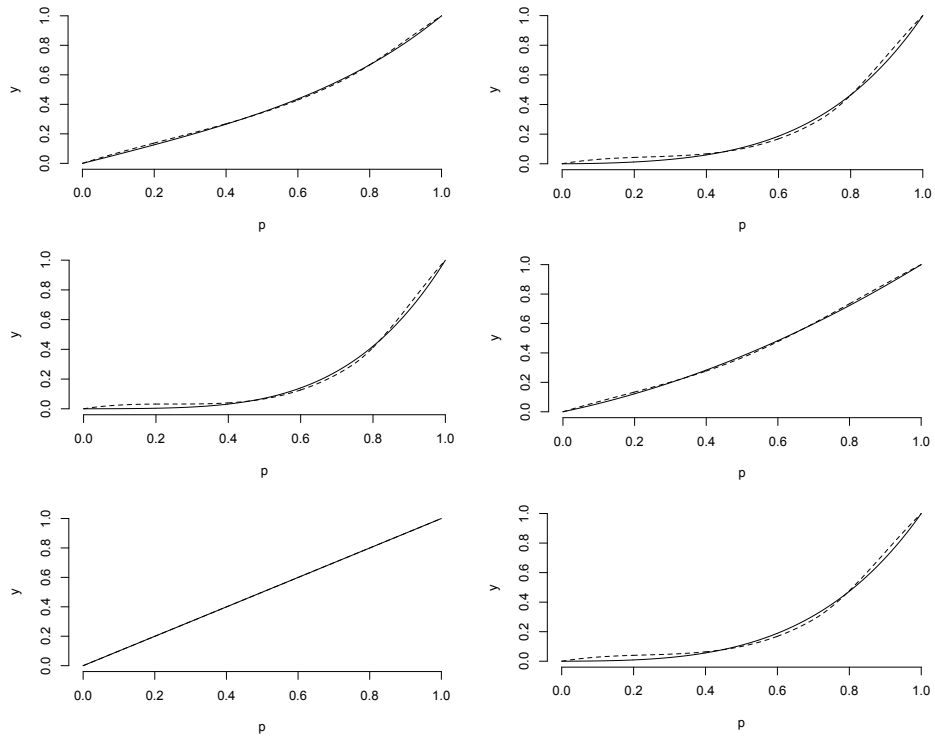


FIG 5. *The three-step spline approximation (dashed) versus true k -step compliant strategies (solid). All approximations are k -step compatible and deviations are visually minor.*

To explore this approximation-induced autocorrelation, we examine the correlation matrix of $(\bar{r}_{i,j}, \bar{r}_{i,j'})$ for all pairs (j, j') ; see Table 2.3. To compute this correlation, we consider only players with mean precision parameter $\bar{s}_i > 5$; this rules out those players (24 out of 106) whose game play exhibits essentially no errors, most of whom played perfectly linear strategies. After excluding these ultra-precise players, we indeed see evidence of approximation-induced autocorrelations, consistent with the approximation mis-fit shown in Figure 5, panels two, three and six especially.

	$p = 0.3$	$p = 0.4$	$p = 0.5$	$p = 0.6$	$p = 0.7$	$p = 1$
$p = 0.3$	1.00					
$p = 0.4$	-0.08	1.00				
$p = 0.5$	-0.44	-0.03	1.00			
$p = 0.6$	-0.11	-0.20	0.17	1.00		
$p = 0.7$	-0.22	0.01	-0.19	0.21	1.00	
$p = 1$	0.42	0.15	-0.12	-0.43	-0.76	1.00

TABLE 1

Correlations between posterior mean normalized residuals, computed across participants at various values of the game parameter p . Statistically significant correlations are shown in bold. This pattern of correlations suggests approximation-induced dependence between errors, consistent with regions of spline mis-fit shown in Figure 5.

Future work may wish to refine our model by explicitly allowing such approximation-induced autocorrelation, whereas we assume independent errors across values of p . An alternative approach would be to use additional knots in the spline model, however, this approach would complicate the characterization of k -step compliance. Despite the presence of approximation-induced autocorrelation in our residuals, we note that the best-fitting approximation is always concave increasing, suggesting that our primary question of interest would not be misguided by the spline approximation.

References.

- M. Plummer, N. Best, K. Cowles, and K. Vines. Coda: Convergence diagnosis and output analysis for mcmc. *R News*, 6(1):7–11, 2006. URL <http://CRAN.R-project.org/doc/Rnews/>.