Bayesian Modeling and Computations in Final Offer Arbitration

Tim Swartz *

Abstract

This paper develops a Bayesian model for the analysis of bidding behaviour in final offer arbitration. Posterior calculations are obtained using a Markov chain algorithm. An example is considered using salary data from major league baseball.

Keywords : latent variables, Markov chain Monte Carlo, major league baseball.

*T Swartz is Professor, Department of Statistics and Actuarial Science, Simon Fraser University, 8888 University Drive, Burnaby BC, Canada V5A1S6.

1. INTRODUCTION

In an attempt to curtail prolonged disputes and to avoid employee strikes, arbitration procedures are now commonplace in many labour contracts.

Final offer arbitration (FOA) is one such arbitration process whereby an employer offers a wage (or some other quantity) x, the employee independently requests z, and an arbitrator makes a decision by choosing either x or z (Stevens 1966). In conventional arbitration, the arbitrator is not constrained to choosing x or z. The motivation behind FOA over conventional arbitration is to reduce the "chilling" effect; i.e. knowing that either x or z will be chosen, negotiating teams ought to bargain more realistically and hence reduce the need for arbitration. Milner (1993) compares FOA with conventional arbitration on the basis of dispute deterrence. Ashenfelter et al (1992) provide laboratory evidence that FOA does not reduce the chilling effect.

In FOA theory, it is commonly assumed that the arbitrator behaves according to the parity concept (Farber 1980; Dworkin 1981). Under the parity concept, the arbitrator weighs the evidence and determines an undisclosed "fair" wage y. The arbitrator then chooses either x or z depending on which value is closer to y. Typically $x < y < z$.

Assuming parity, an active area of research concerns the behaviour of the arbitrator. For example, one approach (Fizel 1996; Dworkin 1981) is based on probit models where the arbitrator's decision is regressed against a set of covariates including $(z - y) - (y - x)$. However, a difficulty associated with such models is that the unknown y must be estimated and this estimation is unaccounted for in the regression. Faurot and McAllister (1992) also consider arbitrator behaviour where it is assumed that y is normally distributed, offers are optimal risk-neutral and arbitrators are exchangeable.

In this paper, we consider a related problem. We develop a model which focuses on the behaviour of the bidders; i.e. we are concerned with covariates that effect the bids x and z. One approach is to regress the "standardized gap" $(z - x)/x$ or some transformed value against a set of covariates as in Frederick et al (1996) and Frederick, Kaempfer and Wobbekind (1992). A difficulty with such an approach is that various information is intertwined in the standardized gap. For example, if large gaps are observed for female employees, it could be the case that employers discriminate against women. However, it could also be the case that females have a higher sense of self-worth or that they are less risk averse than males.

In Section 2, we propose a Bayesian model that investigates the behaviour of bidders. We define variables related to the standardized gap which help identify 3 quantities: (1) the mean amount by which the employer (employee) underestimates (overestimates) the fair wage, (2) the amount due to discrimination on the part of the employer and (3) the amount due to the misevaluation of self-worth on the part of the employee. These 3 quantities are all relative to the value y determined by the arbitrator. In addition, our model assumes the parity condition and we take into account the fact that y is unknown. In Section 3, the posterior calculations for this model are obtained using a Markov chain algorithm. Inference can thus be carried out for any function of the unknown parameters. In Section 4 we look at an example based on FOA data arising from major league baseball. In Section 5, we conclude with a short discussion.

2. A MODEL FOR FOA

For ease of presentation, we initially consider a single arbitration case and omit subscripts. Let x denote the employer's offer and z denote the employee's request. The "true" market value as determined by the arbitrator is a latent (i.e. unobserved) variable which we denote by y. The arbitrator now chooses either x or z. Under the parity concept (Farber 1980; Dworkin 1981), the arbitrator's decision is given by

$$d = \begin{cases} 1 & \text{if } x \text{ is chosen (i.e. } y - x < z - y) \\ 0 & \text{if } z \text{ is chosen (i.e. } y - x > z - y) \end{cases} . \tag{1}$$

The data therefore consist of x, z and d where typically $x < z$. We note that in major league baseball salary arbitration, should $x \geq z$, the FOA protocol is to inform the parties and abandon the arbitration hearing.

We propose a model where conditional on $(y, \alpha, \beta, \sigma_1, \sigma_2, w_1, w_2)$,

$$\begin{aligned} (y - x)/y &\sim \text{Normal}(w_1'\alpha, \sigma_1^2) \\ (z - y)/y &\sim \text{Normal}(w_2'\beta, \sigma_2^2) \end{aligned} \tag{2}$$

and the two distributions are independent. This model should be interpreted as an approximation, as it does not impose the condition $x < z$. In addition, the model is not necessarily behavioural as the employer and employee do not observe the fair wage y. Here $w_i : (r_i \times 1)$ is a covariate vector whose values may be thought to affect the bids in FOA. For example, components of w_i may include covariates such as the race, sex, age and occupation of the employee. The unknown parameters in (2) are $\alpha : (r_1 \times 1)$, $\beta : (r_2 \times 1)$, σ_1 and σ_2. In most applications, we would choose $w_1 = w_2$ as it is difficult to imagine criteria that would be important to one party but of absolutely no importance to the other in determining bids. However, for the sake of generality, we will maintain the distinction involving w_1 and w_2.

Our primary interest concerns the vector parameters α and β. Letting α_1 and β_1 denote the constant terms, consider the baseline case where the remaining $r_i - 1$ covariates in w_i are equal to zero. Then $y\alpha_1$ is the mean amount that the employer underestimates the arbitrator's fair wage y. Similarly, $y\beta_1$ is the mean amount that the employee overestimates the arbitrator's fair wage y. The quantities α_1 and β_1 may be tactical as suggested by the game theory literature or may simply represent what the employer and employee deem to be fair. We are also interested in α_i, $i = 1, \ldots, r_1$ and β_i, $i = 1, \ldots, r_2$. For example, if $\alpha_i > 0$, this implies that the employer discriminates according to the i-th covariate when compared against the baseline case. In other words, a smaller offer is made to employees for whom the i-th covariate in w exceeds zero. We may also be interested in quantities such as $w_1'\alpha - w_2'\beta$ for an employee with specified covariates w_1 and w_2. For if $w_1'\alpha > w_2'\beta$, then the employee would have a better chance than the employer of winning the FOA decision.

Introducing vector notation for the n independent cases and letting $[A \mid B]$ denote the conditional "density" of A given B, we obtain the posterior density

$$\begin{aligned}
[\alpha,\beta,\sigma_1,\sigma_2,\underline{y} \mid \underline{x},\underline{z},\underline{d},\underline{w}_1,\underline{w}_2] &\propto [\alpha,\beta,\sigma_1,\sigma_2,\underline{y},\underline{x},\underline{z},\underline{d} \mid \underline{w}_1,\underline{w}_2] \\
&= [\underline{d} \mid \alpha,\beta,\sigma_1,\sigma_2,\underline{y},\underline{x},\underline{z},\underline{w}_1,\underline{w}_2]\,[\alpha,\beta,\sigma_1,\sigma_2,\underline{y},\underline{x},\underline{z} \mid \underline{w}_1,\underline{w}_2] \\
&= [\underline{d} \mid \underline{y},\underline{x},\underline{z}]\,[\underline{z} \mid \alpha,\beta,\sigma_1,\sigma_2,\underline{y},\underline{x},\underline{w}_1,\underline{w}_2]\,[\underline{x} \mid \alpha,\beta,\sigma_1,\sigma_2,\underline{y},\underline{w}_1,\underline{w}_2]\,[\alpha,\beta,\sigma_1,\sigma_2,\underline{y} \mid \underline{w}_1,\underline{w}_2] \\
&= [\underline{d} \mid \underline{y},\underline{x},\underline{z}]\,[\underline{z} \mid \beta,\sigma_2,\underline{y},\underline{w}_2]\,[\underline{x} \mid \alpha,\sigma_1,\underline{y},\underline{w}_1]\,[\alpha,\beta,\sigma_1,\sigma_2]\,[\underline{y}] \qquad (3)
\end{aligned}$$

where we assume that $(\alpha,\beta,\sigma_1,\sigma_2,\underline{y})$ does not depend on $\underline{w}_1$ and $\underline{w}_2$ (relaxed in Section 4) and that the latent vector $\underline{y}$ is independent of $(\alpha,\beta,\sigma_1,\sigma_2)$. Here $[\underline{d} \mid \underline{y},\underline{x},\underline{z}]$ is a point mass according to the parity concept (1) and can be expressed as

$$\prod_{i=1}^{n} d_i I(y_i - x_i < z_i - y_i) + (1-d_i) I(y_i - x_i > z_i - y_i)$$

where I is the indicator function. The densities $[\underline{z} \mid \beta,\sigma_2,\underline{y}] \sim \prod_{i=1}^{n} \text{Normal}(y_i + w_{2i}'\beta y_i, \sigma_2^2 y_i^2)$ and $[\underline{x} \mid \alpha,\sigma_1,\underline{y}] \sim \prod_{i=1}^{n} \text{Normal}(y_i - w_{1i}'\alpha y_i, \sigma_1^2 y_i^2)$ follow immediately from (2). For $[\alpha,\beta,\sigma_1,\sigma_2]$, we use the improper prior density $1/(\sigma_1\sigma_2)$ which is a standard reference prior (Berger 1985). Finally, the modeling is complete by defining $[\underline{y}] \propto 1$. Although this is a standard reference prior, and is sensible given the application in Section 4, in other applications where good prior information is available, one might consider a subjective prior density for $[\underline{y}]$. We also remark that although it causes neither inferential nor computational difficulties, the posterior density

(3) is not defined when any $y_i = 0$, $i = 1, \ldots, n$. In the Appendix, we establish that the posterior is proper.

The posterior density (3) is a $(n + r_1 + r_2 + 2)$-dimensional function which fully describes the uncertainty in the parameters given the data. However in practice, our primary interest concerns marginal posterior characteristics (e.g. expectations with respect to α and β). As the calculation of these characteristics involves intractable high-dimensional integrals, we estimate the characteristics using a Markov chain Monte Carlo algorithm in Section 3.

We note that a simplification of the model can be obtained by imposing the restriction $\sigma = \sigma_1 = \sigma_2$. In this case, the posterior density becomes

$$[\alpha, \beta, \sigma, \underline{y} \mid \underline{x}, \underline{z}, \underline{d}] \;\propto\; [\underline{d} \mid \underline{y}, \underline{x}, \underline{z}]\; [\underline{z} \mid \beta, \sigma, \underline{y}, \underline{w}_2]\; [\underline{x} \mid \alpha, \sigma, \underline{y}, \underline{w}_1]\; [\alpha, \beta, \sigma]\; [\underline{y}] \tag{4}$$

where $[\alpha, \beta, \sigma]$ is given by the improper prior density $1/\sigma$. In practice, one might begin with model (3), and if the posterior analysis shows $\sigma_1 \approx \sigma_2$, then proceed with the more parsimonious model (4). We also note that model (3) can be viewed as a generalization of an ordinary regression model. For if we reduce the uncertainty in the latent variable $\underline{y}$ (i.e. $\underline{y}$ approaches a point mass at some known $\underline{y}_0$), then the posterior density converges to the density proportional to $[\underline{z} \mid \beta, \sigma_2, \underline{y}_0, \underline{w}_2]\; [\underline{x} \mid \alpha, \sigma_1, \underline{y}_0, \underline{w}_1]\; [\alpha, \beta, \sigma_1, \sigma_2]$.

In passing, we note that a classical likelihood analysis would involve only the first three factors of the posterior density (3). The maximization of the log-likelihood would most conveniently proceed by first maximizing with respect to $y = y(\alpha, \beta, \sigma_1, \sigma_2)$ for which an analytic expression is available. Substituting $y = \hat{y}(\alpha, \beta, \sigma_1, \sigma_2)$ into the log-likelihood, one could then numerically maximize with respect to $(\alpha, \beta, \sigma_1, \sigma_2)$. An inferential difficulty with maximum likelihood concerns the increasing dimensionality of the parameter space (i.e. $n + r_1 + r_2 + 2$) as the sample size grows and this argues against the use of standard asymptotic theory. In contrast, the Bayesian approach provides exact inferences from the posterior distribution and the Markov chain algorithm allows the investigation of any posterior characteristic of interest.

3. POSTERIOR CALCULATIONS VIA GIBBS SAMPLING

The Gibbs sampling algorithm (Geman and Geman 1984; Gelfand and Smith 1990) provides an iterative approach to simulation from a target distribution. In our problem, an implementation of Gibbs sampling proceeds by setting initial values $j = 1$ and $(\alpha^{(0)}, \beta^{(0)}, \sigma_1^{(0)}, \sigma_2^{(0)}, \underline{y}^{(0)})$. We then repeat the following steps $M + N$ times:

(i) generate $\alpha^{(j)} \sim [\alpha \mid \beta^{(j-1)}, \sigma_1^{(j-1)}, \sigma_2^{(j-1)}, \underline{y}^{(j-1)}, \underline{x}, \underline{z}, \underline{d}, \underline{w}_1, \underline{w}_2]$

(ii) generate $\beta^{(j)} \sim [\beta \mid \alpha^{(j)}, \sigma_1^{(j-1)}, \sigma_2^{(j-1)}, \underline{y}^{(j-1)}, \underline{x}, \underline{z}, \underline{d}, \underline{w}_1, \underline{w}_2]$

(iii) generate $\sigma_1^{(j)} \sim [\sigma_1 \mid \alpha^{(j)}, \beta^{(j)}, \sigma_2^{(j-1)}, \underline{y}^{(j-1)}, \underline{x}, \underline{z}, \underline{d}, \underline{w}_1, \underline{w}_2]$

(iv) generate $\sigma_2^{(j)} \sim [\sigma_2 \mid \alpha^{(j)}, \beta^{(j)}, \sigma_1^{(j)}, \underline{y}^{(j-1)}, \underline{x}, \underline{z}, \underline{d}, \underline{w}_1, \underline{w}_2]$

(v) generate $\underline{y}^{(j)} \sim [\underline{y} \mid \alpha^{(j)}, \beta^{(j)}, \sigma_1^{(j)}, \sigma_2^{(j)}, \underline{x}, \underline{z}, \underline{d}, \underline{w}_1, \underline{w}_2]$

(vi) set $j = j + 1$

The idea is to sample from the full conditional distributions in (i)-(v) until we are confident that the vector $\theta^{(j)} = (\alpha^{(j)}, \beta^{(j)}, \sigma_1^{(j)}, \sigma_2^{(j)}, \underline{y}^{(j)})$, $j = M+1, \ldots, N$ approximates a realization from the posterior distribution. Note that convergence to the posterior as $M \to \infty$ is a property of the Gibbs sampling algorithm. We then average the variates $\theta^{(M+1)}, \ldots, \theta^{(N)}$ to estimate various posterior characteristics.

To generate from the full conditional distribution of σ_1, we generate $v_1 \sim \text{Gamma}(n/2, k_1)$ where $k_1 = \sum_{i=1}^n (x_i - y_i + w_{1i}'\alpha y_i)^2/(2y_i^2)$ and then set $\sigma_1 = 1/\sqrt{v_1}$. Similarly, to generate from the full conditional distribution of σ_2, we generate $v_2 \sim \text{Gamma}(n/2, k_2)$ where $k_2 = \sum_{i=1}^n (z_i - y_i - w_{2i}'\beta y_i)^2/(2y_i^2)$ and then set $\sigma_2 = 1/\sqrt{v_2}$.

The derivation of the full conditional distribution for α requires some algebra. For example,

$$
\begin{aligned}
[\alpha \mid \cdot] &\propto \exp\left\{ \frac{-1}{2\sigma_1^2} \sum_{i=1}^n \left(\frac{x_i - y_i + w_{1i}'\alpha y_i}{y_i} \right)^2 \right\} \\
&\propto \exp\left\{ \frac{-1}{2\sigma_1^2} \left(\alpha' A_1 A_1' \alpha - 2\alpha' A_1 t_\alpha \right) \right\} \qquad (5)
\end{aligned}
$$

where $A_1 : (r_1 \times n) = (w_{11}, \ldots, w_{1n})$ and $t_\alpha' : (1 \times n) = ((y_1 - x_1)/y_1, \ldots, (y_n - x_n)/y_n)$. We recognize the form of (5) and therefore generate $\alpha \sim \text{Normal}_{r_1}((A_1A_1')^{-1}A_1 t_\alpha, \sigma_1^2 (A_1A_1')^{-1})$. Similarly, we generate $\beta \sim \text{Normal}_{r_2}((A_2A_2')^{-1}A_2 t_\beta, \sigma_2^2 (A_2A_2')^{-1})$ where $A_2 : (r_2 \times n) = (w_{21}, \ldots, w_{2n})$ and $t_\beta' : (1 \times n) = ((z_1 - y_1)/y_1, \ldots, (z_n - y_n)/y_n)$.

The remaining n full conditional distributions are non-standard as the density $[y_i \mid \cdot]$ is proportional to

$$g_i(y_i) = \begin{cases} \frac{1}{y_i^2} \exp\left\{ \frac{(x_i - y_i + w'_{1i}\alpha y_i)^2}{-2\sigma_1^2 y_i^2} + \frac{(z_i - y_i - w'_{2i}\beta y_i)^2}{-2\sigma_2^2 y_i^2} \right\} I\left(y_i < \frac{x_i + z_i}{2}\right) & d_i = 1 \\ \frac{1}{y_i^2} \exp\left\{ \frac{(x_i - y_i + w'_{1i}\alpha y_i)^2}{-2\sigma_1^2 y_i^2} + \frac{(z_i - y_i - w'_{2i}\beta y_i)^2}{-2\sigma_2^2 y_i^2} \right\} I\left(\frac{x_i + z_i}{2} < y_i\right) & d_i = 0 \end{cases} \tag{6}$$

where I denotes the indicator function. Rather than sample from (6) directly, we "imbed" a Metropolis step whereby we set $a_i = (x_i + z_i)/2$ and introduce the proposal densities

$$h_{i1}(y) = (3/a_i) \exp\{3(y - a_i)/a_i\} \qquad y < a_i$$

when $d_i = 1$ and

$$h_{i0}(y) = (3/a_i) \exp\{-3(y - a_i)/a_i\} \qquad y > a_i$$

when $d_i = 0$. The proposal densities are convenient as they allow efficient variate generation via inversion. The proposal densities also capture the essential features of the $g_i(y)$ and therefore yield high acceptance rates. Note that the proposal densities have exponentially decreasing tails extending from the estimate $a_i = (x_i + z_i)/2$ of the true market value y_i. The jth iteration of the Metropolis step (see Gilks, Richardson and Spiegelhalter 1996) then proceeds by generating $u \sim \text{Unif}(0, 1)$, generating $y_i^{(j)} \sim h_{d_i}(y)$ and setting $y_i^{(j)} = y_i^{(j-1)}$ if

$$u > \frac{g_i(y_i^{(j)}) h_{id_i}(y_i^{(j-1)})}{g_i(y_i^{(j-1)}) h_{id_i}(y_i^{(j)})}.$$

Therefore the Markov chain Monte Carlo algorithm is easily implemented and provides a methodology for studying the bidding behaviour in FOA. We note that minor simplifications arise in the algorithm when the posterior density (4) is considered.

4. ANALYSIS OF MAJOR LEAGUE BASEBALL SALARY DATA

We obtained salary data based on 161 cases of FOA in major league baseball during the period 1990 through 2001. The covariate $w = w_1 = w_2$ is 4-dimensional (i.e. $r_1 = r_2 = 4$) where the first coordinate is the constant term and the second coordinate is 1 (0) corresponding to a pitcher (non-pitcher). The third coordinate is a race variable given by 1 (0) for whites (non-whites) where the non-white setting refers to blacks and latinos. We remark that although race is not scientifically well-defined, there was complete agreement and no indecisiveness in two independent assignments of the race variable by viewing photos of the 161 baseball players. The fourth coordinate is an age variable given by age $-$ 22 where age is the player's age at the time of arbitration.

Table 1 provides estimates of the posterior means and standard deviations of various parameters of interest based on $N = 10^5$ iterations of the Markov chain algorithm. The calculations require approximately 10 minutes of computation on a SUN workstation. Since $\hat{\sigma_2} > \hat{\sigma}_1$, we conclude that players are more variable in their bids than owners. We also observe that there does not seem to be much of an effect due to position (pitcher/non-pitcher) as the bulk of the posterior distributions for α_2 and β_2 are centred near zero. Interestingly, there is no indication of race discrimination on the part of the owners (i.e. $\hat{\alpha}_3 \approx 0$) and there is mild evidence that white players are more risk averse than blacks and latinos (i.e. $\hat{\beta}_3 < 0$). The latter finding disagrees with the conclusions discussed in Fizel (1996). It is also of interest to note that $\hat{\alpha}_4 < 0$ and $\hat{\beta}_4 > 0$. This implies that owners discriminate against older players (i.e. offer less) whereas older players are less risk averse (i.e. request more). Perhaps owners de-value the limited future of older players whereas older players want to be primarily rewarded for past performance? It should be noted that one must be cautious in assigning a behavioural interpretation to the results as the employer and employee do not observe the fair wage y.

We now consider the analysis of a more complex model which allows us to investigate the sensitivity of the analysis with respect to the prior specification. Expanding on (3) with $w = w_1 = w_2$ and $r = r_1 = r_2$, we consider the posterior density

$$\begin{aligned}[\alpha, \beta, \sigma_1, \sigma_2, \lambda, \underline{y} \mid \underline{x}, \underline{z}, \underline{d}, \underline{w}] \quad \propto \quad & [\underline{d} \mid \underline{y}, \underline{x}, \underline{z}]\, [\underline{z} \mid \beta, \sigma_2, \underline{y}, \underline{w}] \\ & [\underline{x} \mid \alpha, \sigma_1, \underline{y}, \underline{w}]\, [\alpha, \beta, \sigma_1, \sigma_2]\, [\underline{y} \mid \lambda, \underline{w}]\, [\lambda]\,. \qquad (7)\end{aligned}$$

In motivating the new model, one might reason that if the gaps depend on the covariate w, then so might y. In (7), we assume that the y_i are conditionally independent, we introduce $\lambda : (r \times 1)$ and set

$$[y_i \mid \lambda, \underline{w}] \sim \text{Normal}(w_i'\lambda,\ (ka_0)^2)$$

where $a_0 = \$2,000,000$ represents a typical salary in major league baseball. Here, the prior mean of y_i can be interpreted as a base salary λ_1 which is then perturbed according to the non-constant covariates. The prior standard deviation of y_i can be interpreted as k typical salaries where k is a specified hyperparameter. We complete the model specification by assigning the standard reference prior $[\lambda] \propto 1$.

Under model (7), there are only two changes in the Markov chain computations. First, a multiplicative factor (i.e. $[y_i \mid \lambda, \underline{w}]$) is introduced to the full conditional densities (6). Second, an extra step is added to the Gibbs sampling algorithm corresponding to the generation of λ from its full conditional distribution. It is not difficult to show that $[\lambda \mid \cdot] \sim \text{Normal}(\gamma, V)$ where $V = (ka_0)^2(\sum_{i=1}^n w_i w_i')^{-1}$ and $\gamma = (\sum_{i=1}^n w_i w_i')^{-1}(\sum_{i=1}^n w_i y_i)$. The results of the major league baseball analysis based on $k = 1$ are given in Table 2. We observe close agreement with the results from the simpler model (i.e. Table 1) indicating that the FOA model is robust with respect to the prior specification of y. Furthermore, we note that λ_1 and λ_2 are the only important parameters in λ. That is, the arbitrator's fair wage is centered about $\$2,107,000$ and is adjusted downward for pitchers. On average then, amongst those who have gone through the arbitration process, pitchers do not command as high a salary as position players. Finally, as might be expected, when we choose k large (i.e. $k = 10$), the prior for y is quite flat and the subsequent analysis produces the same results as in Table 1 to three decimal places.

5. CONCLUDING REMARKS

In FOA, the data consist solely of the arbitrator's decision d, the employer's offer x and the employee's request z. With only such data, it is impossible to determine which of the 3 parties are acting in an unbiased fashion. For example, with all of the arbitration outcomes falling in the employer's favour, it could be the case that the arbitrator is unfair and/or the employer is submitting high offers and/or the employee is submitting high requests. Therefore, in studying FOA data, certain assumptions need to be made.

Most of the literature in FOA tends to focus on the behaviour of the arbitrator where, for example, y is estimated (Fizel 1996) or it is assumed that the employer and employee submit offers according to rational game theory considerations (Faurot and McAllister 1992).

In this paper, we focus on the behaviour of the employer and employee relative to the arbitrator. This is a convenient perspective as the assumptions are weak and it is often thought that the arbitrator is an unbiased but random decision maker (Ashenfelter 1987). In this case, the behaviour of the employer and employee can be interpreted in terms of departures from a position of "fairness". The major contribution of the paper is the development of a Bayesian model and the associated computations to investigate such departures.

ACKNOWLEDGMENTS

This work was partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada. The author thanks James Dworkin, David Faurot and Steve Gietschier for assistance in obtaining the data. The author also thanks the associate editor and a referee for helpful comments that lead to an improvement in the manuscript.

APPENDIX: PROPRIETY OF THE POSTERIOR

In light of the improper prior density $[\alpha, \beta, \sigma_1, \sigma_2][\underline{y}] \propto 1/(\sigma_1\sigma_2)$, we establish that the posterior (3) is proper by showing that the integral

$$\int\int\int\int\int \frac{1}{\sigma_1\sigma_2}\left[\prod_{i=1}^{n}\frac{1}{\sigma_1\sigma_2 y_i^2}\ g(\alpha,\beta,\sigma_1,\sigma_2,y_i)\ I_i\right]\ d\alpha\ d\beta\ d\sigma_1\ d\sigma_2\ d\underline{y} \tag{8}$$

is finite where

$$g(\alpha,\beta,\sigma_1,\sigma_2,y_i) = \exp\left\{\frac{(x_i - y_i + w_{1i}'\alpha y_i)^2}{-2\sigma_1^2 y_i^2} + \frac{(z_i - y_i - w_{2i}'\beta y_i)^2}{-2\sigma_2^2 y_i^2}\right\}$$

and

$$I_i = \begin{cases} I(y_i < (x_i + z_i)/2) & \text{if } d_i = 1 \\ I(y_i > (x_i + z_i)/2) & \text{if } d_i = 0 \end{cases}.$$

Integrating first with respect to α, we concentrate on the inner integral

$$\int \exp\left\{\frac{-1}{2\sigma_1^2}\sum_{i=1}^{n}\left(\frac{x_i - y_i + w_{1i}'\alpha y_i}{y_i}\right)^2\right\}\ d\alpha \quad = \quad \frac{\sigma_1^{r_1}}{[\det(A_1A_1')]^{1/2}}\exp\left\{\frac{-1}{2\sigma_1^2}\ t_\alpha' Q_1 t_\alpha\right\}$$

where we have used the notation from Section 3 and have set $Q_1 = (I - A_1(A_1A_1')^{-1}A_1)$ with A_1 assumed full rank. Similarly, letting $Q_2 = (I - A_2(A_2A_2')^{-1}A_2)$ with A_2 assumed full rank, we have

$$\int \exp\left\{\frac{-1}{2\sigma_2^2}\sum_{i=1}^{n}\left(\frac{z_i - y_i - w_{2i}'\beta y_i}{y_i}\right)^2\right\} d\beta \;=\; \frac{\sigma_2^{r_2}}{[\det(A_2A_2')]^{1/2}}\exp\left\{\frac{-1}{2\sigma_2^2}\; t_\beta' Q_2 t_\beta\right\}.$$

Returning to the original integral (8), we establish that the posterior is proper if

$$\int\left[\int \frac{1}{\sigma_1^{n-r_1+1}}\;\exp\left\{\frac{-1}{2\sigma_1^2}\; t_\alpha' Q_1 t_\alpha\right\}\; d\sigma_1\right]\cdot\left[\int \frac{1}{\sigma_2^{n-r_2+1}}\;\exp\left\{\frac{-1}{2\sigma_2^2}\; t_\beta' Q_2 t_\beta\right\}\; d\sigma_2\right]\frac{(\prod_{i=1}^n I_i)}{(\prod_{i=1}^n y_i^2)}\; d\underline{y}$$

is finite. Working on the two inner integrals and recalling the norming constant of the Inverse Gamma distribution, the posterior is proper if

$$\int (t_\alpha' Q_1 t_\alpha)^{-(n-r_1)/2}\; (t_\beta' Q_2 t_\beta)^{-(n-r_2)/2}\frac{(\prod_{i=1}^n I_i)}{(\prod_{i=1}^n y_i^2)}\; d\underline{y} \tag{9}$$

is finite. We first need to check the singularities of the integrand in (9). Since Q_1 and Q_2 are positive definite matrices, singularities only occur when any of the $y_i = 0$, $i = 1, \ldots, n$, and it is straightforward to show that the limit of the integrand is finite as we approach the singularities provided that $n > (r_1+r_2+4)/2$. Therefore it is only a matter of investigating the tail behaviour of (8), and since $t_\alpha' Q_1 t_\alpha$ and $t_\beta' Q_2 t_\beta$ are both bounded in the tails, integrability follows.

To establish the existence of moments for σ_1 and σ_2, the proof is modified by changing the norming constants in the Inverse Gamma distributions. To establish the existence of moments for α and β, we modify the proof using the known moments of the multivariate normal distribution.

REFERENCES

Ashenfelter, O. (1987). "A model of arbitrator behavior", *American Economic Review*, 77, 342-346.

Ashenfelter, O., Currie, J., Farber, H.S. and Spiegel, M. (1992). "An experimental comparison of dispute rates in alternative arbitration systems", *Econometrica*, 60, 1407-1433.

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, Springer Verlag, New York.

Dworkin, J.B. (1981). *Owners versus Players: Baseball and Collective Bargaining*, Boston: Auburn House.

Farber, H.S. (1980). "An analysis of final-offer arbitration", *Journal of Conflict Resolution*, 24, 683-705.

Faurot, D.J. and McAllister, S. (1992). "Salary arbitration and pre-arbitration negotiation in major league baseball", *Industrial and Labor Relations Review*, 45, 697-710.

Fizel, J. (1996). "Bias in salary arbitration: the case of major league baseball", *Applied Economics*, 28, 255-265.

Frederick, D.M., Kaempfer, W.H. and Wobbekind, R.L. (1992). "Salary arbitration as a market substitute", in *Diamonds are Forever: The Business of Baseball*, P.M. Sommers (eds.), The Brookings Institution, Washington, D.C., 29-49.

Frederick, D.M., Kaempfer, W.H., Ross, M.T. and Wobbekind, R.L. (1996). "Race, risk and repeated arbitration", in *Baseball Economics: Current Issues*, J. Fizel, E. Gustafson and L. Hadley (eds.), Greenwood Press, 129-141.

Gelfand, A.E. and Smith, A.F.M. (1990). "Sampling based approaches to calculating marginal densities", *Journal of the American Statistical Association*, 85, 398-409.

Geman, S. and Geman, D. (1984). "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (ed.) (1996), *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London.

Milner, S. (1993). "Dispute deterrence: evidence on final offer arbitration", in *New Perspectives on Industrial Disputes*, D. Metcalf and S. Milner (eds.), Routledge, 133-159.

Stevens, C.M. (1966). "Is compulsory arbitration compatible with bargaining", *Industrial Relations*, 5, 38-52.

Table 1: Estimates of posterior means and standard deviations based on model (3).

Parameter	Post Mean	Post Std Dev
α_1	.241	.028
α_2	-.011	.019
α_3	-.005	.020
α_4	-.007	.004
β_1	.198	.040
β_2	.004	.027
β_3	-.044	.029
β_4	.007	.005
σ_1	.101	.007
σ_2	.141	.012
$\sigma_1 - \sigma_2$	-.040	.013

Table 2: Estimates of posterior means and standard deviations based on model (7).

Parameter	Post Mean	Post Std Dev
α_1	.237	.028
α_2	-.013	.019
α_3	-.004	.020
α_4	-.007	.004
β_1	.203	.040
β_2	.007	.027
β_3	-.045	.028
β_4	.007	.006
σ_1	.103	.008
σ_2	.139	.012
$\sigma_1 - \sigma_2$	-.036	.013
λ_1	2107.608	863.256
λ_2	-836.505	351.106
λ_3	-236.098	382.086
λ_4	-65.895	119.514