PREDICTING RETIREMENT PATTERNS: PREDICTION FOR A MULTINOMIAL DISTRIBUTION WITH CONSTRAINED PARAMETER SPACE

by

James O. Berger and Ming-Hui Chen
Purdue University

Technical Report #92-44C

Department of Statistics
Purdue University

September 1992

*

Predicting Retirement Patterns : Prediction for a Multinomial Distribution with Constrained Parameter Space *

James O. Berger and Ming-Hui Chen

Department of Statistics, Purdue University, West Lafayette, IN 47907, U. S. A.

Abstract

A major problem facing American universities is the end of mandatory retirement in 1994. Currently, tenured professors must retire by age 70. There is concern that large numbers of professors will choose not to retire, leading to an expensive older faculty and preventing infusion of 'new blood'. We have undertaken a study of the problem at Purdue University, with the primary goal being to predict the number of faculty retirements in each of the next ten years.

Methodologically, the problem is one of inference for multinomial cell probabilities, based on independent data arising from various subgroups of cells. The cells correspond to age and reason for departure from the university (retirement or 'other'). The posterior distribution of these cell probabilities, together with current faculty demographics, can then be used to predict the future retirement numbers.

Prior information of several types exists, including orderings among the cell probabilities, retirement rates at other universities, and mortality rates. No data was available concerning retirement rates for professors older than 70, so inference within this important domain is highly dependent on the prior; a sensitivity study was thus carried out within this region.

Computations were done using a hybrid Markov Chain sampling scheme. This includes a "random direction" component that is particularly well suited for handling constrained parameter spaces.

Keywords: Bayesian posteriors, Gibbs sampler, Hit-and-Run sampler, Markov chain sampling.

*This research was supported by National Science Foundation grants DMS-8717799 and DMS-8923071 at Purdue University.

1 Introduction

In 1986, the Age Discrimination in Employment Act was amended by the United States Congress to prohibit mandatory retirement on the basis of age for all workers, except for a few exempt professions including faculty in higher education. This exemption will end on January 1, 1994, at which time no university can require a professor to retire; only seriously inadequate performance - not age - can be cause for dismissal.

Among the concerns that have been expressed at having professors over 70 (the current mandatory retirement age) are (i) concern that productivity or innovation may decline at such ages; (ii) the fact that older professors have higher salaries, so that the total cost of the faculty would increase; (iii) worry that universities would be unable to hire 'new blood' for a period of time. Solutions that have been proposed include 'buyouts' of older professors or increases in retirement benefits to encourage retirement. As a rough indication, estimated costs of such 'solutions' range from $5 million to $20 million per year at Purdue University.

The purpose of this study was to estimate the magnitude of the problem by predicting, for the next twelve years, the number of professors that would choose to work past 70 and predicting the number of upcoming retirements at Purdue . A variety of sources of information were available. Most important was data concerning the actual retirement and departure rates of professors at Purdue over the last ten years. Also available (and utilized) were mortality rates for professors (see Appendix A), and information from other universities. A survey of the Purdue faculty was also undertaken, but resulted in unusable data on the issue of likely retirement age. The reason the survey was a failure is of interest for those designing Bayesian elicitation schemes - see Appendix B for discussion.

The data can be modelled as a series of independent multinomial observations, though on possibly different cells. The key piece of prior information is information concerning monotonicity of cell probabilities. Thus the underlying statistical scenario is that of order-constrained multinomial inference. Bayesian analysis is particularly well-suited to this scenario, because it can easily handle monotonicity constraints. The prior information from mortality rates and other universities is incorporated through a hierarchical structure.

Posterior computations are done through a hybrid algorithm that combines features of Gibbs sampling, Hit and Run sampling, and a generalized Metropolis algorithm. This is described in

Section 4.

Section 5 develops the predictive model for this scenario, and gives the actual predictions of retirement rates and numbers. Because the predictions are, in part, based on projections past the range of the data (to ages over 70), a partial sensitivity study to assumptions in this domain was undertaken.

The conclusions of the study are discussed in Section 5.3. In brief, the conclusion is that the end of mandatory retirement poses no serious difficulties over the next ten years.

2 The Data and Modeling

2.1 The Data

Table 1 gives the data on retirements (R) and other net departures (D) of tenured faculty from Purdue University over the past ten years. Data prior to this time could not be easily utilized because of differences in the retirement system. Indeed, the numbers in parentheses in the table are unusable for this reason. (Until 1981-1982, retirement at age 65 was mandatory, except for a variety of special cases; the numbers in parentheses involve professors whose decisions were made under this earlier system.)

"Net departures" are those who left the faculty for reasons other than retirement, such as death or obtaining a different job, net of those who joined the tenured faculty in the given age group. (Rather than separately modeling this hiring process, we view it as a process of replacement; this causes a problem only when there are more "hires" than "other departures", something which only occurred twice.) Note that only tenured faculty are considered here. The reason is that non-tenured faculty have never been subject to mandatory retirement.

In understanding the data, it is crucial to realize that it is a collection of age cohorts. By an age cohort we mean a group of professors of the same age, imagining the group moving through time. Thus the 73 professors of age 51 in 81-82 are in the age cohort that becomes 68 professors of age 52 in 82-83, 67 professors of age 53 in 83-84, *etc.* . In Table 1, there are a total of 23 such age cohorts (not counting those corresponding to numbers in parentheses).

Table 1. Retirements (R) and other net Departures (D) out of (N) tenured faculty by age group and year.

	Year																													
Age	81–82			82–83			83–84			84–85			85–86			86–87			87–88			88–89			89–90			90–91		
Group	N	R	D	N	R	D	N	R	D	N	R	D	N	R	D	N	R	D	N	R	D	N	R	D	N	R	D	N	R	D
51	73		5	52		2	51	1		58			61			53			65			60			55		-5	38		
52	53			68		1	50			50			58	1		61			53			65			60			60		
53	59		3	53			67			50			50		1	57			61			53		1	65			60		-1
54	34			56			53			67		1	50			50			57	1		61	1		52		1	65		
55	38	1		34			56			53	1		66	2		50			50			56	1		60	1		51	2	
56	44			37			34			56			52	1		64	1		50			50			55			59	1	
57	43			44			37		1	34		1	56	1		51			63	2		50		2	50			55	1	
58	41			43			44			36			33			55	2		51	1		61			48			50	1	
59	40	1		41	1		43		2	44	1		36	1	1	33	3		53	1		50	1		61	2		48		2
60	60			39	2		40			41	2		43			34			30	1		52	1		49	2		59	6	
61	56	2		60	4		37			40		1	39			43	2		34	1		29			51	3		47	1	
62	39	2		54	3		56	3		37			39	2		39	3		41	2		33	2		29	1		48	1	
63	23	1		37		1	51	2	1	53	2		37	3		37			36	3		39	2		31	3		28	3	
64	25	2	1	22	4		36	3		48	4	1	51	3		34	1		37	3		33	3		37	4		28	2	
65	(26)	(12)		22	5		18	7		33	10		43	17		48	18		33	12		34	9		30	8		33	12	
66	(6)	(2)		(14)	(5)		17	1		11			23	8	1	26	3		30	7		21	4		25	6		22	4	
67	(6)	(2)		(4)			(9)			16	3		11	1		14	4		23	5		23	3		17	2		19	1	
68	(5)			(4)	(1)		(4)			(9)			13	3		10	1		10			18	3		20	1		15	5	
69	(3)	(1)		(5)	(2)		(3)			(4)			(9)	(3)		10	2		9	1		10	2		15			19	6	
70				(2)	(2)		(3)	(3)		(3)	(3)		(4)	(4)		(6)	(6)		8	8		8	8		8	8		15	15	

2.2 Modeling

The age cohorts will be considered independent. In a given cohort, the departures due to retirement and the net other departures will be modeled by a joint multinomial distribution, depending on unknown parameters p_i and r_i. Let

$$E_i = \{\text{departure from university at age} \geq 50 + i\}. \tag{2.1}$$

Then the relevant retirement probabilities are defined by

$$p_i = P\,(\text{retirement at age } 50 + i \mid E_1)\,;$$

the net non-retirement departure probabilities are given by

$$r_i = P\,(\text{net non-retirement departure at age } 50 + i \mid E_1)\,.$$

For convenience, also define

$$v_i = P(E_i|E_1) = \sum_{j=i}^{\infty}(p_j + r_j),$$

and note that, for $i = 1, 2, \cdots,$

$$\begin{aligned} p_i + r_i &= P\,(\text{net departure at age } 50 + i \mid E_1)\,, \\ \frac{p_i + r_i}{v_i} &= \text{net departure rate at age } 50 + i \text{ given } E_i. \end{aligned}$$

We actually will consider retirement and departure probabilities only up to age $(50 + k)$, where $k = 30$ or $k = 35$ years. Thus the "tail" will be summarized by

$$p^*_{k+1} = \sum_{j=k+1}^{\infty} p_j \quad \text{and} \quad r^*_{k+1} = \sum_{j=k+1}^{\infty} r_j.$$

For age cohort i (any labelling is okay), let $(50 + \underline{i})$ and $(50 + \bar{i})$ be the lower and upper age limits for the cohort in the data set. (Thus the age cohort discussed in Section 2.1 has $\underline{i} = 1$ and

$\bar{i} = 10$.) Define, for $j = \underline{i}$ to $\bar{i}$,

$$N_i \;=\; \text{number originally in age cohort } i; \tag{2.2}$$

$$R_{ij} \;=\; \text{number of retirements at age } 50 + j \text{ from age cohort } i; \tag{2.3}$$

$$D_{ij} \;=\; \text{net number of nonretirement departures at age } 50 + j \text{ from age cohort } i. \tag{2.4}$$

Then it is reasonable to assume that

$$(R_{i\underline{i}}, \cdots, R_{i\bar{i}}, D_{i\underline{i}}, \cdots, D_{i\bar{i}}) \sim \text{Multinomial}(N_i, \underset{\sim}{p}_i, \underset{\sim}{r}_i)\,, \tag{2.5}$$

where

$$\underset{\sim}{p}_i \;=\; \left(\frac{p_{\underline{i}}}{v_{\underline{i}}}, \frac{p_{\underline{i}+1}}{v_{\underline{i}}}, \cdots, \frac{p_{\bar{i}}}{v_{\underline{i}}}\right),$$

$$\underset{\sim}{r}_i \;=\; \left(\frac{r_{\underline{i}}}{v_{\underline{i}}}, \frac{r_{\underline{i}+1}}{v_{\underline{i}}}, \cdots, \frac{r_{\bar{i}}}{v_{\underline{i}}}\right);$$

in the notation we supress the last cell, which corresponds to retirement or departure after $\bar{i}$ and has cell probability $(v_{\bar{i}+1}/v_{\underline{i}})$. Note that the number of cells in each multinomial model is $2(\bar{i} - \underline{i}) + 3$, and that this varies from 3 to 21 cells for the various cohorts.

Assuming independence of the cohorts, the multinomial likelihoods for each cohort can simply be multiplied together which, ignoring multiplicative constants, results in the observed likelihood function

$$\begin{aligned} L^*(\underset{\sim}{p}, \underset{\sim}{r}, data) = \;& r_1^2\, r_2^1\, r_3^4\, r_4^2\, r_7^4\, r_9^5\, r_{11}^1\, r_{13}^2\, r_{14}^2\, r_{16}^1 \\ & \cdot p_1^1\, p_2^1\, p_4^2\, p_5^8\, p_6^3\, p_7^4\, p_8^4\, p_9^{11}\, p_{10}^{14}\, p_{11}^{13}\, p_{12}^{19}\, p_{13}^{19}\, p_{14}^{29}\, p_{15}^{98}\, p_{16}^{33}\, p_{17}^{19}\, p_{18}^{13}\, p_{19}^{11} \\ & \cdot v_2^{-15}\, v_3^1\, v_4^{27}\, v_5^{27}\, v_6^5\, v_7^{15}\, v_8^{13}\, v_9^9\, v_{10}^{-14}\, v_{11}^{-3}\, v_{12}^7\, v_{13}^{24}\, v_{15}^{26}\, v_{16}^{21}\, v_{17}^{18}\, v_{18}^{18}\, v_{19}^{10}\, v_{20}^{52}, \end{aligned} \tag{2.6}$$

where $\underset{\sim}{p} = (p_1, p_2, \cdots)$ and $\underset{\sim}{r} = (r_1, r_2, \cdots)$.

3 Prior Information

3.1 Mortality Rates

To incorporate known mortality rates into the analysis, it is necessary to reparameterize. In Section 2.2, we introduced r_i, which is the probability of a net non-retirement departure at age $(50+i)$ given E_1, i.e., given departure from the university at age greater than 50. Define

$$m_i = P\left(\text{death at age } 50+i \mid E_1\right),$$

$$t_i = P\left(\text{departure at age } 50+i \text{ for reasons other than death or retirement} \mid E_1\right).$$

Then we have

$$r_i = m_i + t_i, \text{ for } i = 1, 2, \cdots. \tag{3.1}$$

Denote the mortality rates for professors by

$$m_i^* = P\left(\text{death at age } 50+i \mid \text{lived to age } 50+i\right). \tag{3.2}$$

These were obtained from the *Teachers Insurance and Annuity Association* (see Appendix A), and can be used with the p_i and t_i to obtain a recursive formula for computing the r_i.

Proposition 3.1 *For* $i = 1,\ 2,\ \cdots,$

$$r_i = m_i^* \cdot v_i + t_i. \tag{3.3}$$

Also, in recursive computation, a convenient formula for v_i *is*

$$v_i = 1 - \sum_{j=1}^{i-1}(p_j + r_j).$$

Proof: Clearly

$$\begin{aligned} m_i &= P\left(\text{death at age } 50+i \mid E_1\right) \\ &= P\left(\text{death at age } 50+i \mid \text{lived to age } 50+i\right) \cdot P\left(E_i \mid E_1\right) \end{aligned}$$

$$= m_i^* \cdot v_i.$$

The conclusion follows from Equation (3.1). ∎

Note that the unknown parameters have now become $\underset{\sim}{p} = (p_1, p_2, \cdots, p_k)$ and $\underset{\sim}{t} = (t_1, t_2, \cdots, t_k)$, and that the likelihood function for $\underset{\sim}{p}$ and $\underset{\sim}{t}$ can be found from (2.6), using (3.3).

3.2 Constraints on Parameters

For many years, 65 was the standard retirement age in the United States. Because of this and the fact that full government retirement benefits begin at 65, it is still by far the most common age of retirement. No other ages possess unique features, so it is natural to assume that the p_i are increasing to age 65, and decreasing thereafter, i.e.,

$$p_1 \le p_2 \le \cdots \le p_{15} \ge p_{16} \ge \cdots \ge p_k. \tag{3.4}$$

It is less clear if the t_i can reasonably be constrained. Note that t_i is essentially just the net probability of leaving for a new job. We decided that $t_1, t_2, \cdots, t_{10}$ have no clear ordering, but that $t_{11}, t_{12}, \cdots, t_{20}$ are reasonably judged to be decreasing, with $t_{21}, t_{22}, \cdots$, being zero. Thus we assume that

$$t_1, t_2, \cdots, t_{10} \text{ are free}; \quad t_{11} \ge t_{12} \ge \cdots \ge t_{20}; \quad t_{21} = t_{22} = \cdots = 0. \tag{3.5}$$

3.3 Constraints in the Tail

For ages greater than 70, no data is available. Thus answers will be very sensitive to p_i in the tail, i.e., to $p_{20}, p_{21}, \cdots, p_k$. We will therefore also consider additional constraints

$$p_{i+1} \ge c\, p_i, \text{ for } i = 19, 20, \cdots, k, \tag{3.6}$$

and for $c = 0, 0.5, 0.75$. Motivation for this is given in Section 3.6.

3.4 Information from Other Universities

Define, conditional on E_1,

$$\begin{aligned}\rho_1 &= \text{net probability of departure in the 51-55 age group} \\ &= \sum_{i=1}^{5}(p_i + r_i), &(3.7) \\ \rho_2 &= \text{net probability of departure in the 56-60 age group} \\ &= \sum_{i=6}^{10}(p_i + r_i), &(3.8) \\ \rho_3 &= \text{net probability of departure in the 61+ age group} \\ &= \sum_{i=11}^{\infty}(p_i + r_i). &(3.9)\end{aligned}$$

Note that $\rho_3 = 1 - (\rho_1 + \rho_2)$.

Rees and Smith (1991) present information on net departures at other public universities. We are interested in using this information to augment the paucity of the Purdue data in the 51-60 age period. This information from Rees and Smith (1991) is summarized in Table 2; note that only information for five-year intervals was given.

Table 2: Net Departure Probabilities At Other Public Universities

	ρ_1	ρ_2	ρ_3
Estimate	.023	.078	.899
Standard Error	.0064	.0142	.015
Corrected Standard Error	.011	.019	.021

The estimates and standard errors in Table 2 refer to the net departure probabilities given in Rees and Smith (1991) for the five-year intervals averaged over the surveyed public universities. The estimates can be used as the prior means for the ρ_i, but the standard errors are not directly relevant, since they refer to the estimate of the public university average. The variance between universities is thought to be between $(0.01)^2$ and $(0.015)^2$. Adding this variance to that of the estimate provides a more accurate reflection of the accuracy for Purdue of the estimate; the corresponding standard deviations are given in the last row of Table 2.

To convert this to a prior distribution for $(\rho_1,\ \rho_2,\ \rho_3)$, we will match the estimates and corrected standard errors in Table 2 with those for a *Dirichlet* $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ distribution. Thus we want

$$(E[\rho_1], E[\rho_2], E[\rho_3]) = \left(\frac{\alpha_1}{\alpha_0}\ , \frac{\alpha_2}{\alpha_0}\ , \frac{\alpha_2}{\alpha_0}\ \right) \cong (.023,\ .078,\ .899),$$

$$(var(\rho_1),\ var(\rho_2),\ var(\rho_3)) \cong \left((.011)^2,\ (.019)^2,\ (.021)^2\right),$$

where

$$var(\rho_i) \ = \ \frac{\alpha_i}{\alpha_0} \cdot \left(1 - \frac{\alpha_i}{\alpha_0}\right) \cdot \frac{1}{(\alpha_0 + 1)}.$$

Approximate solution of these equations yields, as the prior for $(\rho_1,\ \rho_2,\ \rho_3)$, a

$$\textit{Dirichlet}\ (200, 4.6, 15.6, 179.8) \tag{3.10}$$

distribution. We also add the constraints

$$\rho_1 < \rho_2 < \rho_3, \tag{3.11}$$

to ensure consistency with (3.4).

3.5 The Prior Distribution

The prior distribution for $\underset{\sim}{p}$ and $\underset{\sim}{t}$ will be chosen to be

$$\pi(\underset{\sim}{p}, \underset{\sim}{t}) = \pi(\underset{\sim}{p}, \underset{\sim}{t} \mid \rho_1, \rho_2, \rho_3) \cdot \pi(\rho_1, \rho_2, \rho_3), \tag{3.12}$$

where $\pi(\rho_1, \rho_2, \rho_3)$ is the constrained *Dirichlet* distribution given in (3.10) with the constraint (3.11); and

$$\pi(\underset{\sim}{p}, \underset{\sim}{t} \mid \rho_1, \rho_2, \rho_3) \propto 1_{\Omega_{\underset{\sim}{\rho}}}(\underset{\sim}{p}, \underset{\sim}{t}), \tag{3.13}$$

which is the constant density on $\Omega_{\underset{\sim}{\rho}}$, where $\underset{\sim}{\rho} = (\rho_1, \rho_2, \rho_3)$ and

$$\Omega_{\underset{\sim}{\rho}} = \{\underset{\sim}{p},\ \underset{\sim}{t} :\ \text{expressions (3.4) through (3.9) are satisfied}\}. \tag{3.14}$$

The choice of a uniform density in (3.13) is an effort to be noninformative. Because $\Omega_{\underset{\sim}{\rho}}$ is so sharply constrained, different choices in (3.13) would not matter greatly, except in the "tail" (i.e., for $p_{20} \geq p_{21} \geq \cdots \geq p_k$) where there is no direct data. A uniform prior over this tail is actually somewhat dramatic; for instance, the prior means of the p_i decrease roughly quadratically. Since equal p_i would be the opposite extreme, we study sensitivity to the prior tail by considering, also, the constraint (3.6), for $c = 0$, 0.5, and 0.75. Recall also that we limit k to $k = 30$ or $k = 35$. We do not assume that the p_i are zero past theses points; rather, the presence of k is an effort to limit any undue influence of the uniform prior in the tail.

4 Posterior Computation

4.1 The Posterior and Posterior Expectation

Denote

$$\Omega_c = \{\underset{\sim}{p}, \underset{\sim}{t} : (\underset{\sim}{p}, \underset{\sim}{t}) \in \Omega_{\underset{\sim}{\rho}},\ 0 \leq \rho_1 \leq \rho_2 \leq \rho_3\}, \tag{4.1}$$

where $\Omega_{\underset{\sim}{\rho}}$ is defined in (3.14). The posterior distribution is

$$\pi(\underset{\sim}{p}, \underset{\sim}{t} \mid data) \ \propto\ \text{Likelihood} \times \text{Prior} \ =\ L(\underset{\sim}{p}, \underset{\sim}{t})\pi(\underset{\sim}{p}, \underset{\sim}{t}), \tag{4.2}$$

for $(\underset{\sim}{p}, \underset{\sim}{t}) \in \Omega_c$, where $L(\underset{\sim}{p}, \underset{\sim}{t}) = L^*(\underset{\sim}{p}, \underset{\sim}{r}, data)$ is given in (2.6); and the prior $\pi(\underset{\sim}{p}, \underset{\sim}{t})$ is given in (3.12).

Most Bayesian quantities of interest involve posterior expectation, such as

$$E^{\text{posterior}}[\psi(\underset{\sim}{p}, \underset{\sim}{t})] = \frac{\int_{\Omega_c} \psi(\underset{\sim}{p}, \underset{\sim}{t}) L(\underset{\sim}{p}, \underset{\sim}{t}) \pi(\underset{\sim}{p}, \underset{\sim}{t}) d\underset{\sim}{p} d\underset{\sim}{t}}{\int_{\Omega_c} L(\underset{\sim}{p}, \underset{\sim}{t}) \pi(\underset{\sim}{p}, \underset{\sim}{t}) d\underset{\sim}{p} d\underset{\sim}{t}}. \tag{4.3}$$

This requires high dimensional integration (up to 54 dimensions), and will be computed using a Monte-Carlo method as

$$E^{\text{posterior}}[\psi(\underset{\sim}{p}, \underset{\sim}{t})] \cong \frac{1}{M} \sum_{i=1}^{M} \psi(\underset{\sim}{p}^{(i)}, \underset{\sim}{t}^{(i)}), \tag{4.4}$$

where $\left\{(\underset{\sim}{p}^{(i)}, \underset{\sim}{t}^{(i)}),\ i = 1, 2, \cdots, M\right\}$ is a stream of "simulated" values from the posterior distribution $\pi(\underset{\sim}{p}, \underset{\sim}{t} \mid data)$ and M is suitably large. Section 4.2 describes the hybrid algorithm used to generate $\left\{(\underset{\sim}{p}^{(i)}, \underset{\sim}{t}^{(i)}),\ i = 1, 2, \cdots, M\right\}$ from $\pi(\underset{\sim}{p}, \underset{\sim}{t} \mid data)$.

4.2 The Hybrid Algorithm

4.2.1 Outline of the Algorithm

The hybrid algorithm has three major components: grouping, an approximate Gibbs step, and Metropolis-Hit and Run generation. The parameters of interest, $(p_1, p_2, \cdots, p_k, t_1, t_2, \cdots, t_k)$ for $k = 30$ or 35, are first grouped as follows:

$$\begin{aligned}
&\textit{Group } 1. \quad \underset{\sim}{G}_1 = (p_1, p_2, \cdots, p_{10}); \\
&\textit{Group } 2. \quad \underset{\sim}{G}_2 = (p_{11}, p_{12}, \cdots, p_{20}); \\
&\textit{Group } 3. \quad \underset{\sim}{G}_3 = (p_{21}, p_{22}, \cdots, p_k); \\
&\textit{Group } 4. \quad \underset{\sim}{G}_4 = (t_1, t_2, \cdots, t_{20}).
\end{aligned}$$

Reasons for this grouping are given in Section 4.2.2. The approximate Gibbs step of the algorithm is to iteratively sample over these four groups as follows:

Step 0. Choose a starting point $\left(\underset{\sim}{G}_1^{(0)}, \underset{\sim}{G}_2^{(0)}, \underset{\sim}{G}_3^{(0)}, \underset{\sim}{G}_4^{(0)}\right)$, and set $i = 0$.

Step 1. Generate $\underset{\sim}{G}_1^{(i+1)}$ (approximately) from the conditional posterior distribution

$$\left[\underset{\sim}{G}_1 \mid \underset{\sim}{G}_2^{(i)}, \underset{\sim}{G}_3^{(i)}, \underset{\sim}{G}_4^{(i)}\right].$$

Step j, j=2,3,4. Generate $\underset{\sim}{G}_j^{(i+1)}$ (approximately) from the conditional posterior distribution

$$\left[\underset{\sim}{G}_j \mid \left\{\underset{\sim}{G}_l^{(i+1)}, l < j\right\}, \left\{\underset{\sim}{G}_h^{(i)}, h > j\right\}\right].$$

Step 5. Set $i = i + 1$ and go to Step 1.

For each of Step 1 through Step 4, we approximately sample from the relevant conditional distribution using one-step of a Metropolis-Hit and Run sampler. For example, details of generating $\underset{\sim}{G}_1^{(i+1)}$ from one step of a generalized Metropolis algorithm for $\left[\underset{\sim}{G}_1 \mid \underset{\sim}{G}_2^{(i)}, \underset{\sim}{G}_3^{(i)}, \underset{\sim}{G}_4^{(i)}\right]$ are:

(i) Generate a candidate $\underset{\sim}{G}_1^*$ by

- Choosing a uniformly distributed direction $\underset{\sim}{d}$;

- Finding the line segment L through $\underset{\sim}{G}_1^{(i)}$ in the direction $\underset{\sim}{d}$ which lies in the constrained parameter space $\Omega_c^*(\underset{\sim}{G}_2^{(i)}, \underset{\sim}{G}_3^{(i)}, \underset{\sim}{G}_4^{(i)})$, which is the subspace of Ω_c in (4.1) given the components $\underset{\sim}{G}_2^{(i)}$, $\underset{\sim}{G}_3^{(i)}$, and $\underset{\sim}{G}_4^{(i)}$;
- Choosing $\underset{\sim}{G}_1^*$ uniformly on L.

(ii) Move to $\underset{\sim}{G}_1^*$ or stay at $\underset{\sim}{G}_1^{(i)}$ according to the general Metropolis algorithm for $\left[\underset{\sim}{G}_1 \mid \underset{\sim}{G}_2^{(i)}, \underset{\sim}{G}_3^{(i)}, \underset{\sim}{G}_4^{(i)}\right]$, i.e., move to $\underset{\sim}{G}_1^*$ with probability

$$\min\left\{1, \frac{g\left(\underset{\sim}{G}_1^* \mid \underset{\sim}{G}_2^{(i)}, \underset{\sim}{G}_3^{(i)}, \underset{\sim}{G}_4^{(i)}\right)}{g\left(\underset{\sim}{G}_1^{(i)} \mid \underset{\sim}{G}_2^{(i)}, \underset{\sim}{G}_3^{(i)}, \underset{\sim}{G}_4^{(i)}\right)}\right\},$$

where $g\left(\cdot \mid \underset{\sim}{G}_2^{(i)}, \underset{\sim}{G}_3^{(i)}, \underset{\sim}{G}_4^{(i)}\right)$ is the density function of $\left[\underset{\sim}{G}_1 \mid \underset{\sim}{G}_2^{(i)}, \underset{\sim}{G}_3^{(i)}, \underset{\sim}{G}_4^{(i)}\right]$.

Note that, for Gibbs step 3, we always move to $\underset{\sim}{G}_3^*$ since $\left[\underset{\sim}{G}_3 \mid \underset{\sim}{G}_1^{(i+1)}, \underset{\sim}{G}_2^{(i+1)}, \underset{\sim}{G}_4^{(i)}\right]$ is uniform.

4.2.2 Discussion and References

Random direction methods are typically well-suited to highly constrained parameter spaces. The constraints will typically result in a rather small line segment L, so that uniform random variate generation over L is reasonably efficient. Random direction methods are especially attractive when, as is the case here, the conditional distributions do not have standard form.

There are several reasons for the grouping of parameters. First, the random direction part of the algorithm requires rough orthogonality; otherwise iterations of the algorithm will move too slowly. The four groups that were created do have roughly orthogonal posteriors. The second reason for grouping is that computing the line segment L is fairly expensive, but with a cost that is fairly constant across dimensions. Grouping results in needing to compute only four line segments per iteration. A final reason for grouping is to keep the algorithm compatible with the structure of the prior (and posterior). The two relevant structural features are (ρ_1, ρ_2, ρ_3), which are preserved under the grouping, and the fact that the conditional posterior of $\underset{\sim}{G}_3$ is uniform (since the prior was uniform and there was no data in the tail).

General discussion of the Metropolis-Hastings algorithms can be found in Metropolis *et al.* (1953) and Hastings (1970). Extensive discussion and illustration of Gibbs sampling can be found in Geman and Geman (1984) and Gelfand and Smith (1990). Boneh and Golan (1979) and Smith

(1980, 1984) independently proposed Hit and Run samplers; and Belisle, Romeijn and Smith (1990), Schmeiser and Chen (1991), and Chen and Schmeiser (1992) proposed and discussed more-general versions of Hit and Run samplers. Discussions or instances of use of hybrid methods include Müller (1991), Tierney (1991), Tanner (1991), Geweke (1992), and Chen and Deely (1992). Gelfand, Smith and Lee (1992) discussed Bayesian inference for constrained parameter spaces using Gibbs sampling. They also had an example of a multinomial model with ordered parameters.

4.2.3 Computational Details

The hybrid algorithm was implemented in single precision Fortran-77 using the IMSL library. The computation time was about 4 hours for 1,000,000 iterations on a Sun Sparc-Station 1. (Note that this included accumulation of a $(k+1) \times (k+1)$ covariance-matrix, for $k = 30$ and 35.) Since the hybrid algorithm is a special version of Metropolis-Hastings, asymptotic convergence is guaranteed. The formal proof of convergence is very similar to that given in Chen and Deely (1992). Empirically, convergence seemed to be achieved at about 100,000 iterations.

We chose one single long run of the simulated Monte Carlo Markov chain to obtain the posterior means and standard errors for the parameters of interest. The Monte Carlo standard errors of these quantities were computed using non-overlapping batch means with batch size of 100,000. The Monte Carlo standard errors of the posterior means were from 2% to 10% of the posterior standard errors, which is acceptably small for our purposes. The non-overlapping/overlapping batch statistics for computing Monte Carlo standard errors can be found in Schmeiser (1982) and Schmeiser, Avramidis, and Hashem (1990).

4.3 Posterior Means and Standard Errors of the Parameters of Interest

The posterior means and standard errors of the retirement probabilities, p_i, net departure probabilities, $p_i + r_i$, and net departure rates $(p_i + r_i)/v_i$ for $c = 0$ and $k = 30, 35$ are given in Tables 3 and 4. The Bayesian estimates of the net departure probabilities and net departure rates for the different priors corresponding to $c = 0, 0.5, 0.75$ and $k = 30, 35$ are plotted in Figures 1 and 2. An examination of the tables and figures reveals considerable robustness with respect to c and k, except possibly for the "tails" (especially, the departure rates for age ≥ 70). Even here, however, the results are qualitatively similar, and would lead to the same policy conclusions. (The large differences in departure rates for age ≥ 80 make little difference at the policy level, because only a very few professors would be in this category.)

Table 3 : Posterior Means (Standard Errors) for Retirement and Net Departure Probabilities and Rates for $c = 0$ and $k = 30$ years

age $50+i$	Retirement Probability p_i	Net Departure Probability $p_i + r_i$	Net Departure Rate $(p_i + r_i)/v_i$
51	.0010 (.00061)	.0062 (.00235)	.0062 (.00235)
52	.0017 (.00079)	.0061 (.00156)	.0061 (.00158)
53	.0025 (.00103)	.0093 (.00260)	.0094 (.00264)
54	.0042 (.00139)	.0098 (.00230)	.0100 (.00235)
55	.0068 (.00156)	.0117 (.00208)	.0121 (.00214)
56	.0079 (.00169)	.0133 (.00220)	.0139 (.00232)
57	.0096 (.00196)	.0182 (.00379)	.0193 (.00404)
58	.0120 (.00251)	.0179 (.00294)	.0194 (.00321)
59	.0176 (.00329)	.0281 (.00501)	.0310 (.00551)
60	.0226 (.00361)	.0290 (.00387)	.0330 (.00444)
61	.0275 (.00413)	.0390 (.00483)	.0459 (.00575)
62	.0347 (.00495)	.0444 (.00518)	.0547 (.00650)
63	.0424 (.00610)	.0512 (.00619)	.0668 (.00829)
64	.0593 (.00926)	.0674 (.00928)	.0942 (.01303)
65	.2030 (.01746)	.2104 (.01747)	.3245 (.02569)
66	.0791 (.01146)	.0847 (.01150)	.1933 (.02448)
67	.0568 (.00785)	.0615 (.00790)	.1741 (.02131)
68	.0465 (.00618)	.0505 (.00624)	.1734 (.02117)
69	.0397 (.00562)	.0431 (.00566)	.1793 (.02464)
70	.0335 (.00568)	.0364 (.00571)	.1847 (.03206)
71	.0281 (.00538)	.0304 (.00538)	.1901 (.03930)
72	.0234 (.00476)	.0255 (.00476)	.1973 (.04512)
73	.0193 (.00433)	.0212 (.00434)	.2055 (.05311)
74	.0158 (.00393)	.0174 (.00396)	.2143 (.06225)
75	.0127 (.00359)	.0141 (.00362)	.2229 (.07377)
76	.0100 (.00324)	.0112 (.00328)	.2323 (.08992)
77	.0076 (.00290)	.0087 (.00293)	.2391 (.10977)
78	.0054 (.00252)	.0064 (.00255)	.2410 (.13623)
79	.0034 (.00206)	.0043 (.00211)	.2269 (.16525)
80	.0016 (.00149)	.0024 (.00157)	.1805 (.18610)
81+	.	.0203 (.01755)	1.000 (.00000)

Table 4: Posterior Means (Standard Errors) for Retirement and Net Departure Probabilities and Rates for $c = 0$ and $k = 35$ years

age $50+i$	Retirement Probability p_i	Net Departure Probability $p_i + r_i$	Net Departure Rate $(p_i + r_i)/v_i$
51	.0010 (.00060)	.0059 (.00196)	.0059 (.00196)
52	.0017 (.00079)	.0065 (.00209)	.0065 (.00210)
53	.0025 (.00103)	.0097 (.00312)	.0098 (.00316)
54	.0042 (.00140)	.0100 (.00276)	.0103 (.00283)
55	.0067 (.00156)	.0117 (.00207)	.0120 (.00214)
56	.0078 (.00169)	.0135 (.00260)	.0141 (.00273)
57	.0095 (.00196)	.0184 (.00382)	.0196 (.00407)
58	.0119 (.00245)	.0177 (.00281)	.0191 (.00308)
59	.0176 (.00323)	.0276 (.00446)	.0305 (.00493)
60	.0225 (.00363)	.0291 (.00400)	.0331 (.00457)
61	.0273 (.00416)	.0386 (.00480)	.0454 (.00571)
62	.0345 (.00491)	.0440 (.00511)	.0542 (.00643)
63	.0421 (.00608)	.0507 (.00615)	.0662 (.00824)
64	.0586 (.00924)	.0666 (.00925)	.0930 (.01301)
65	.2014 (.01733)	.2088 (.01734)	.3212 (.02516)
66	.0779 (.01153)	.0834 (.01158)	.1889 (.02444)
67	.0551 (.00787)	.0598 (.00792)	.1672 (.02109)
68	.0445 (.00629)	.0485 (.00634)	.1631 (.02112)
69	.0371 (.00559)	.0406 (.00563)	.1632 (.02342)
70	.0302 (.00523)	.0331 (.00527)	.1594 (.02716)
71	.0248 (.00440)	.0273 (.00443)	.1565 (.02804)
72	.0209 (.00376)	.0233 (.00379)	.1585 (.03036)
73	.0176 (.00315)	.0198 (.00321)	.1603 (.03008)
74	.0150 (.00281)	.0170 (.00289)	.1642 (.03230)
75	.0128 (.00253)	.0146 (.00264)	.1692 (.03486)
76	.0109 (.00234)	.0127 (.00245)	.1766 (.03924)
77	.0093 (.00215)	.0109 (.00228)	.1847 (.04374)
78	.0079 (.00201)	.0094 (.00214)	.1955 (.05201)
79	.0066 (.00186)	.0079 (.00199)	.2054 (.06013)
80	.0054 (.00168)	.0065 (.00182)	.2165 (.07123)
81	.0043 (.00155)	.0053 (.00169)	.2275 (.08636)
82	.0033 (.00138)	.0042 (.00151)	.2374 (.10636)
83	.0024 (.00118)	.0032 (.00132)	.2414 (.13103)
84	.0015 (.00095)	.0022 (.00111)	.2314 (.15757)
85	.0007 (.00066)	.0014 (.00087)	.1906 (.17469)
86+	·	.0101 (.00928)	1.000 (.00000)

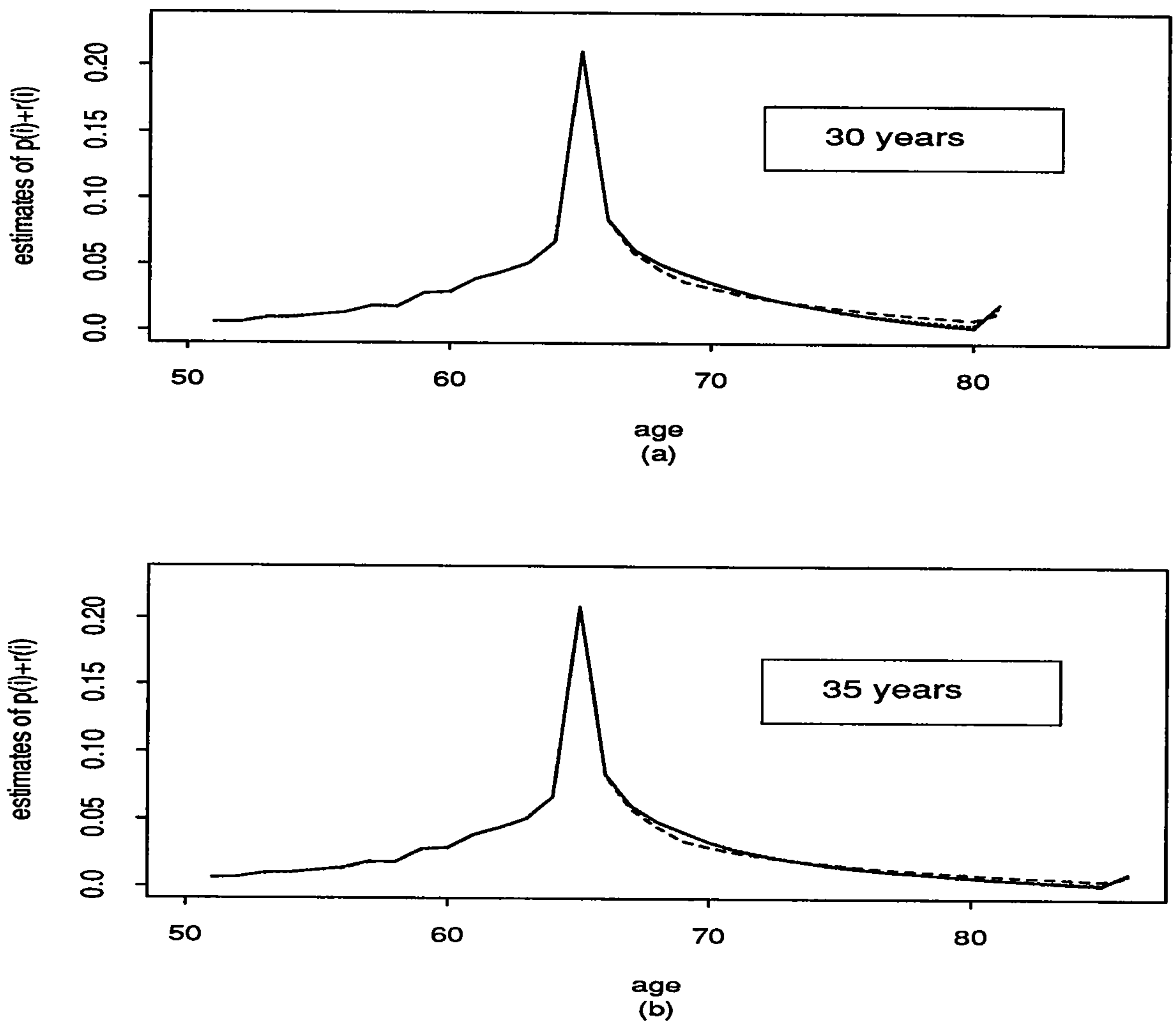

Figure 1: **The Bayesian Estimates of the Net Departure Probabilities for Three Different Priors. The solid curves are for $c = 0$; the dotted curves are for $c = 0.50$; the dashed curves are for $c = 0.75$.**

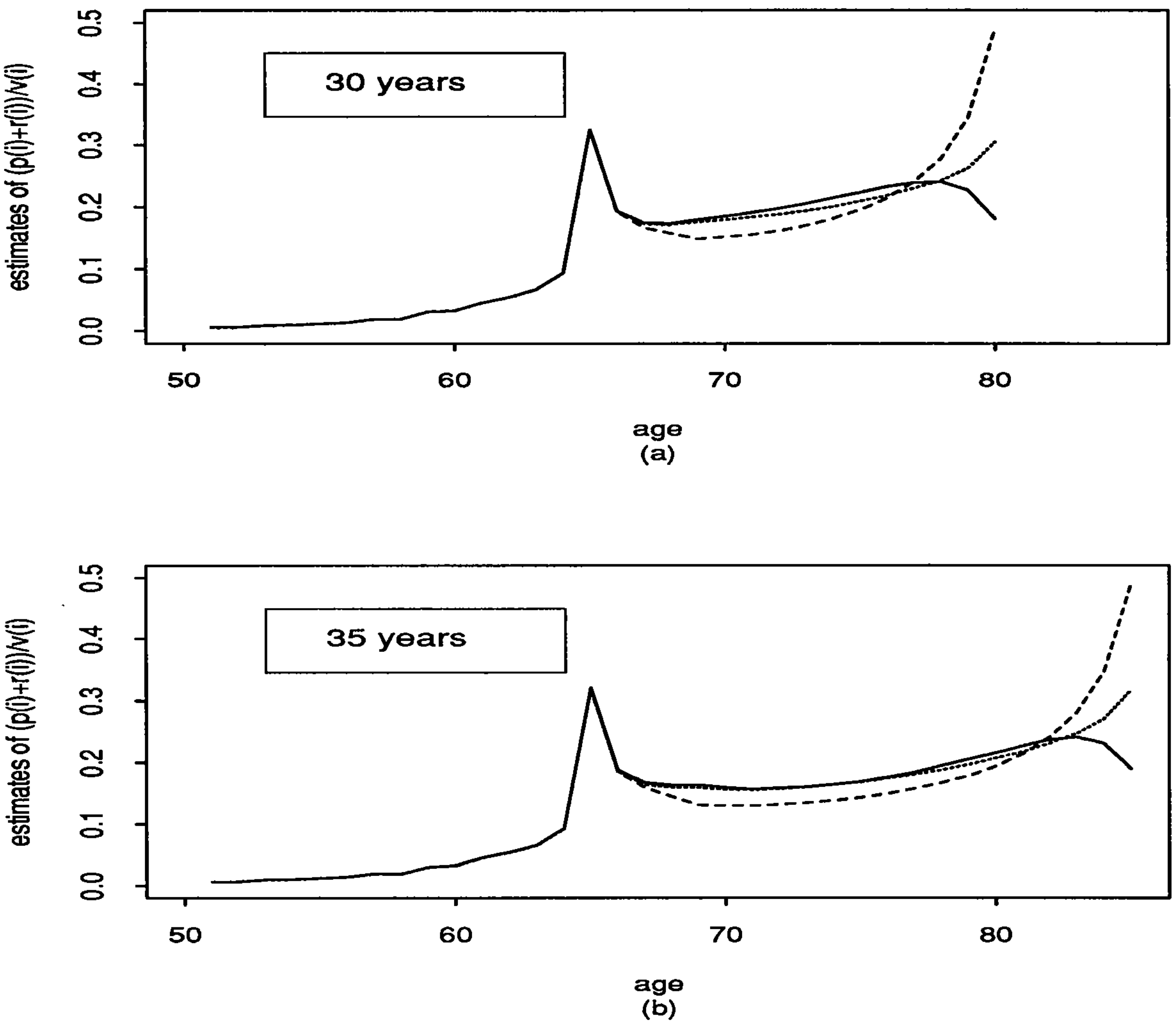

Figure 2: **The Bayesian Estimates of the Departure Rates for Three Different Priors. The solid curves are for $c = 0$; the dotted curves are for $c = 0.50$; the dashed curves are for $c = 0.75$**

Another quantity of interest is the mean extended duration (say, $\mu_{e.d.}$) of a tenured professor's stay at Purdue due to mandatory retirement. This is defined as

$$\mu_{e.d.} = \sum_{i=1}^{\infty} i \left(p_{20+i} + r_{20+i}\right). \tag{4.5}$$

The posterior means and standard errors for $\mu_{e.d.}$ are given in the following table, for $c = 0, 0.5, 0.75$ and $k = 30, 35$.

Table 5: Posterior Means (Standard Errors) for the Mean Extended Duration $\mu_{e.d.}$ at Purdue University

k (years)	$c = 0$	$c = 0.50$	$c = 0.75$
30	.7506 (.18475)	.7803 (.16835)	.8907 (.13948)
35	1.0022 (.19439)	1.0350 (.18514)	1.2391 (.17026)

On average, at most 1 year will be added to the length of a professor's stay at Purdue, which is a rather modest effect. Note that $\mu_{e.d.}$ is a summary of the effect of the "tail", and so Table 5 also shows the extent of robustness to the tail.

5 Prediction of Retirement Rates

5.1 The Predictive Model

The ultimate goal is to predict, over the next twelve years, the number of tenured faculty in various age categories and the net yearly number of departures of faculty age 51 or older. To this end, define

$$Y_{ij} = \text{the number of tenured faculty, from age cohort } 50+i \text{ in 1991-92, that remain at age } 50+j, \tag{5.1}$$

$$X_{ij} = Y_{ij} - Y_{i,j+1} = \text{the number of tenured faculty, from age cohort } 50+i, \text{ that depart at age } 50+j, \tag{5.2}$$

$$\underset{\sim}{X}_i = (X_{ii},\ X_{i,i+1},\ \cdots,\ X_{ik}). \tag{5.3}$$

Then, again suppressing the last cell "departure at age $\geq k+1$" in the notation,

$$\underset{\sim}{X}_i \mid \underset{\sim}{p},\ \underset{\sim}{t} \sim \text{Multinomial}\ (Y_{ii},\ \underset{\sim}{q}_i)\,, \tag{5.4}$$

where

$$\underset{\sim}{q}_i = (q_{ii},\ q_{i,i+1},\ \cdots,\ q_{ik})\,, \tag{5.5}$$

$$\begin{aligned} q_{ij} &= \text{the conditional probability that a tenured faculty member departs at age } 50+j, \\ &\quad \text{given that they are currently of age } 50+i \\ &= (p_j + r_j)/v_i. \end{aligned} \tag{5.6}$$

Noting, from (5.1) and (5.2), that

$$Y_{ij} = Y_{ii} - \sum_{l=i}^{j-1} X_{il}, \tag{5.7}$$

we can derive the following result for predicting the Y_{ij}.

Theorem 5.1

$$\hat{Y}_{ij} = E\left[Y_{ij}\right] = Y_{ii}\ \cdot\ E^{posterior}\left[v_j/v_i\right]; \tag{5.8}$$

and

$$Var(Y_{ij}) = Y_{ii}\ (Y_{ii}-1)\ E^{posterior}\left(1-\frac{v_j}{v_i}\right)^2 + \left(Y_{ii}-\hat{Y}_{ij}\right)\left[1-(Y_{ii}-\hat{Y}_{ij})\right]. \tag{5.9}$$

Proof: From (5.7) and (5.4),

$$\begin{aligned} \hat{Y}_{ij} &= E\left[Y_{ij}\right] = E^{\text{posterior}}\left(E\left[Y_{ii} - \sum_{l=i}^{j-1} X_{il} \mid \underset{\sim}{p},\ \underset{\sim}{t}\right]\right) \\ &= Y_{ii}\left(1 - E^{\text{posterior}}\left[\sum_{l=i}^{j-1} q_{il}\right]\right) = Y_{ii}\ \left(1 - E^{\text{posterior}}\left[\frac{v_i - v_j}{v_i}\right]\right), \end{aligned} \tag{5.10}$$

$$\begin{aligned} Var(Y_{ij}) &= E\left(Y_{ii} - \hat{Y}_{ij}\right)^2 \\ &= E\left(Y_{ii} - \sum_{l=i}^{j-1} X_{il} - \left[Y_{ii} - Y_{ii}\sum_{l=i}^{j-1} E^{\text{posterior}}(q_{il})\right]\right)^2 \end{aligned}$$

$$= E\left(\sum_{l=i}^{j-1}\left[X_{il} - Y_{ii}E^{\text{posterior}}(q_{il})\right]\right)^2$$

$$= E\left(\sum_{l=i}^{j-1}\left[\ (X_{il} - Y_{ii}q_{il}) + Y_{ii}\ \left(q_{il} - E^{\text{posterior}}(q_{il})\right)\right]\right)^2$$

$$= E\left(\sum_{l=i}^{j-1}[X_{il} - Y_{ii}q_{il}]\right)^2 + Y_{ii}^2 E^{\text{posterior}}\left(\sum_{l=i}^{j-1}\left[q_{il} - E^{\text{posterior}}(q_{il})\right]\right)^2. \quad (5.11)$$

Since

$$\sum_{l=i}^{j-1} X_{il} \mid \underset{\sim}{p},\ \underset{\sim}{t} \ \sim\ \text{Bin}\left(Y_{ii},\ \sum_{l=i}^{j-1} q_{il}\right),$$

$$E\left(\sum_{l=i}^{j-1}[X_{il} - Y_{ii}q_{il}]\right)^2 = E^{\text{posterior}}\left[Y_{ii}\ \left(\sum_{l=i}^{j-1} q_{il}\right)\left(1 - \sum_{l=i}^{j-1} q_{il}\right)\right]; \quad (5.12)$$

and

$$\sum_{l=i}^{j-1} q_{il} = \frac{v_i - v_j}{v_i}, \quad \hat{Y}_{ij} = Y_{ii}E^{\text{posterior}}\left[\sum_{l=i}^{j-1} q_{il}\right]. \quad (5.13)$$

Equation (5.9) follows by algebra. ■

5.2 Predictive Estimates

The Y_{ii} are the number of tenured professors at Purdue of age $(50+i)$ during 1991-92 (negative i are allowed). These numbers are given in Table 6.

Table 6: Tenured Purdue Faculty During 1991-92

Age	45	46	47	48	49	50	51	52	53	54	55	56	57
Number	60	65	50	63	58	66	59	38	60	61	65	49	58

Age	58	59	60	61	62	63	64	65	66	67	68	69	70
Number	54	49	46	53	46	47	25	26	21	18	18	10	13

Using these, and Monte Carlo estimates of $E^{\text{posterior}}[v_j/v_i]$ and $E^{\text{posterior}}[(1 - v_j/v_i)^2]$, Theorem 5.1 can be used to predict the $\hat{Y}_{ij}$ over the next twelve years. These predictions are given in Tables 7 and 8, by age category. The prediction for a given age category is found by summing the corresponding $\hat{Y}_{ij}$; because of the independence of the age cohorts, a reasonable approximation to the variance for a given age category was found by simply summing the variances of the corresponding $\hat{Y}_{ij}$.

Table 7: Predicted Percentages (Standard Errors) of Tenured Faculty in Various Age Categories, Assuming Constant Faculty Size

Year	Age Categories					
	0-50	**51-55**	**56-60**	**61-65**	**66-70**	**71+**
1991-92	49.66	17.46	15.79	12.15	4.94	0
1992-93	48.82 (.37)	17.41 (.08)	16.65 (.14)	12.65 (.21)	4.47 (.25)	0
1993-94	48.18 (.49)	17.16 (.11)	17.12 (.19)	13.35 (.30)	4.19 (.32)	0
1994-95	47.09 (.58)	17.31 (.11)	17.25 (.23)	13.26 (.34)	4.48 (.37)	.61 (.14)
1995-96	46.93 (.65)	17.99 (.13)	15.89 (.24)	13.50 (.37)	4.68 (.41)	1.01 (.20)
1996-97	46.00 (.70)	18.36 (.13)	16.31 (.26)	13.00 (.39)	5.03 (.44)	1.30 (.24)
1997-98	45.40 (.74)	18.00 (.13)	16.30 (.27)	13.73 (.41)	5.13 (.46)	1.44 (.27)
1998-99	45.50 (.77)	17.51 (.13)	16.07 (.27)	14.10 (.43)	5.31 (.48)	1.51 (.29)
1999-2000	44.22 (.79)	18.20 (.12)	16.24 (.28)	14.20 (.44)	5.32 (.48)	1.82 (.33)
2000-01	44.27 (.81)	18.42 (.13)	16.83 (.29)	13.01 (.44)	5.45 (.49)	2.02 (.36)
2001-02	44.03 (.83)	18.00 (.13)	17.17 (.29)	13.32 (.45)	5.22 (.48)	2.26 (.38)
2002-03	44.53 (.84)	17.32 (.13)	16.85 (.29)	13.40 (.45)	5.57 (.50)	2.33 (.39)
2003-04	45.02 (.85)	17.24 (.13)	16.37 (.29)	13.26 (.45)	5.70 (.51)	2.41 (.41)

Table 8: **Predicted Numbers (Standard Errors) of Tenured Faculty in Various Age Categories and Net Yearly Number of Departures of Faculty 51 or Older**

Year	Age Categories						Net # of ≥ 51 Departures
	0-50	**51-55**	**56-60**	**61-65**	**66-70**	**71+**	
1991-92	805	283	256	197	80	0	52.42
1992-93	791.42 (5.95)	282.23 (1.35)	269.85 (2.29)	205.04 (3.42)	72.46 (4.08)	0	47.55
1993-94	780.97 (7.90)	278.13 (1.72)	277.46 (3.10)	216.47 (4.81)	67.97 (5.17)	0	45.40
1994-95	763.37 (9.42)	280.61 (1.86)	279.58 (3.65)	214.92 (5.49)	72.55 (6.06)	9.97 (2.25)	47.38
1995-96	760.75 (10.51)	291.64 (2.11)	257.54 (3.90)	218.88 (6.03)	75.85 (6.64)	16.34 (3.21)	49.98
1996-97	745.74 (11.34)	297.61 (2.12)	264.35 (4.27)	210.73 (6.25)	81.59 (7.16)	20.98 (3.94)	50.14
1997-98	735.87 (11.95)	291.82 (2.07)	264.18 (4.38)	222.53 (6.67)	83.19 (7.42)	23.41 (4.45)	51.63
1998-99	737.51 (12.44)	283.79 (2.07)	260.52 (4.44)	228.56 (6.98)	86.15 (7.71)	24.47 (4.74)	53.28
1999-2000	716.79 (12.83)	295.00 (2.02)	263.28 (4.47)	230.21 (7.19)	86.18 (7.76)	29.54 (5.36)	54.83
2000-01	717.62 (13.15)	298.61 (2.12)	272.77 (4.72)	210.86 (7.09)	88.27 (7.90)	32.87 (5.78)	54.23
2001-02	713.77 (13.39)	291.81 (2.07)	278.38 (4.76)	215.93 (7.32)	84.57 (7.78)	36.54 (6.19)	56.91
2002-03	721.76 (13.61)	280.83 (2.07)	273.22 (4.67)	217.24 (7.29)	90.21 (8.06)	37.74 (6.40)	56.96
2003-04	729.74 (13.76)	279.51 (2.14)	265.41 (4.65)	214.89 (7.24)	92.36 (8.20)	39.09 (6.60)	56.58

These Tables were computed under the $c = 0$ and $k = 30$ prior. The "0 - 50" age category was estimated by simply assuming a constant tenured faculty size (at the 1991-92 size of 1621). The last column of Table 8 gives the predicted net number of yearly departures from Purdue of tenured faculty over age 50. These were computed simply by summing the predicted $\hat{X}_{ij}$ (see Equation (5.2)).

5.3 Conclusions

The subject - matter conclusion is that the end of mandatory retirement will pose few problems at Purdue. The percentage of tenured faculty over 70 will rise to only 2.41% (±.41%) by 2003-04. The actual number of such faculty will be only about 40, which cannot justify adopting any of the actions for encouraging retirement (costing $5 million to $20 million).

There is not even a problem concerning infusion of 'new blood', since the net number of departures from Purdue of tenured faculty over age 50 will be roughly constant, in the 45-55 range, over the next 12 years. This at first seems surprising, but is a consequence of overall faculty demographics. A large 'bump' of faculty is approaching retirement age, and this bump will tend to counteract the end of mandatory retirement in terms of number of departures.

The major methodological conclusion here is simply that Bayesian methods were able to handle a highly complex problem, involving constrained parameters, multiple sources of information, and prediction. Any of these complications, by itself, would have made classical analysis difficult.

Appendix A: Mortality Rates

Table 8: Mortality Rates from Teachers Insurance and Annuity Association

age $50+i$	mortality rates $m_i^* \times 100$	age $50+i$	mortality rates $m_i^* \times 100$
51	0.26	69	1.17
52	0.28	70	1.29
53	0.31	71	1.42
54	0.34	72	1.57
55	0.37	73	1.73
56	0.40	74	1.91
57	0.43	75	2.11
58	0.46	76	2.34
59	0.50	77	2.59
60	0.54	78	2.87
61	0.57	79	3.18
62	0.62	80	3.53
63	0.67	81	3.92
64	0.73	82	4.35
65	0.80	83	4.82
66	0.87	84	5.35
67	0.96	85	5.93
68	1.06	86	6.56

Appendix B: The Survey Question and Errors in Answers

The third part of the *Faculty Survey on Retirement Issues* at Purdue University, that was sent to a stratified sample of 680 faculty, read as follows:

3. **Retirement plans:** Use the codes given to fill in the grid shown below. Fill in <u>**all**</u> boxes.

0 = no chance; 1 = slight chance; 2 = probable; 3 = certainty

	55-59	60-64	65-69	70+
Age of expected retirement				
Age you would consider beginning a phased-in voluntary partial retirement plan with full benefits				

The intent here was to elicit personal probabilistic information about age of expected retirement, rather than simply a best guess. A four point scale was used, this being judged to be about as complex as feasible in a questionnaire. Of course, we anticipated answers such as 0; 1; 2; 1 for the first four boxes, or maybe 0; 0; 3; 0 if someone was certain. These would then be normalized to obtain a probability vector.

Unfortunately, most respondents did not seem to understand the question. We received responses such as 0; 2; 3; 3, which clearly indicated confusion. There were a fair number of usable 0; 0; 3; 0 type responses, but most responses were confused. (The usable responses did give results that were very consistent with our previous analysis.) Together with other bad experiences we have had in elicitation, this leads us to recommend that probabilistic information be elicited only if the elicitor can have personal contact with the statistician, preferably in a training session.

References

Belisle, Claude J.P., Romeijn, H.E. and Smith, R.L. (1990) Hit-and-Run Algorithms for Generating Multivariate Distributions, Technical Report 90-18, The University of Michigan, Department of Industrial and Operations Engineering.

Boneh, A. and Golan A. (1979) Constraints' Redundancy and Feasible Region Boundedness by Random Feasible Point Generator (RFPG), Third European Congress on Operations Research, EURO III, Amsterdam (April 9-11).

Chen, M.-H. and Deely, J. (1992) Application of a New Gibbs Hit-and-Run Sampler to a Constrained Linear Multiple Regression Problem, Technical Report 92-21, Purdue University, Center for Statistical Decision Sciences and Department of Statistics.

Chen, M.-H. and Schmeiser, B.W. (1992) Performance of the Gibbs, Hit-and-Run, and Metropolis Samplers, Technical Report 92-18, Purdue University, Department of Statistics.

Gelfand, A. E. and Smith, A.F.M. (1990) Sampling Based Approaches to Calculating Marginal Densities, *Journal of American Statistical Association, 85*, pp. 398-409.

Gelfand, A.E., Smith, A.F.M. and Lee, T.M. (1990) Bayesian Analysis of Constrained Parameter and Truncated Data Problems Using Gibbs Sampling, *Journal of American Statistical Association, 87*, pp. 523-532.

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*, pp. 721-741.

Geweke, J. (1992) Priors for Macroeconomic Time Series and Their Application, To appear in the *Proceedings of the Conference on Bayes Methods and Unit Roots*, a special issue of *Econometric Theory.*

Hastings, W.K. (1970) Monte Carlo Sampling Methods Using Markov Chains and Their Applications, *Biometrika, 57*, pp. 97 -109.

Müller, P. (1991) A Generic Approach to Posterior Integration and Gibbs Sampling, Technical Report 91-09, Purdue University, Department of Statistics.

Rees, A. and Smith, S.P. (1991) Faculty Retirement in the Arts and Sciences, Princeton Univ. Press, Princeton.

Schmeiser, B.W. (1982) Batch Size Effects in the Analysis of Simulation Output, *Operations Research 30*, pp. 556-568.

Schmeiser, B.W., Avramidis, Thanos N., and Hashem, S. (1990) Overlapping Batch Statistics, In *Proceedings of the 1990 Winter Simulation Conference*, pp. 395-398.

Schmeiser, B.W. and Chen, M.-H. (1991) On Hit-and-Run Monte Carlo Sampling for Evaluating Multidimensional Integrals, revision of Technical Report 91-39, Purdue University, Department of Statistics.

Smith, R. L. (1980) A Monte Carlo Procedures for Generating Random Feasible Solutions to Mathematical Programs, A Bulletin of the ORSA/TIMS Joint National Meeting, Washington, D.C., 101.

Smith, R. L. (1984) Efficient Monte Carlo Procedures for Generating Points Uniformly Distributed over Bounded Regions, *Operations Research, 32, 6*, pp. 1297-1308.

Tanner, M.A. (1991) Tools for Statistical Inference: Observed Data and Data Augmentation Methods, *Lecture Notes in Statistics 67*, Springer Verlag, NY.

Tierney, L. (1991) Markov Chains for Exploring Posterior Distributions, Technical Report No. 560, University of Minnesota, School of Statistics.