ANALYSIS OF CENSORED DATA FROM FRACTIONATED EXPERIMENTS: A BAYESIAN APPROACH

M. Hamada and C.F.J. Wu

IIQP Research Report
RR-92-11

December 1992

ANALYSIS OF CENSORED DATA FROM FRACTIONATED EXPERIMENTS: A BAYESIAN APPROACH

M. Hamada and C.F.J. Wu

Department of Statistics and Actuarial Science
University of Waterloo
December 1992

ABSTRACT

Censored data arise naturally in industrial experiments whose observations are failure times or measurements with mixed continuous-categorical outcomes. When censored data are observed from a fractionated experiment, likelihood-based estimates quite often do not exist, especially when the opportunity for improvement is great. To circumvent this problem, we propose a Bayesian analysis strategy which has a straightforward implementation using the data augmentation and Monte Carlo EM algorithms. For non-regular designs with complex aliasing patterns, a modified analysis strategy is proposed which substantially reduces the computation needed for model selection. Non-Bayesian practitioners should have no qualms with the approach since the analysis results are insensitive to the specification of the prior. The proposed strategy is illustrated using data from three real industrial experiments.

Key Words: Complex aliasing; Conjugate and improper prior; Data augmentation; Fractional factorial design; Monte Carlo EM; Plackett-Burman design.

1. INTRODUCTION

Censored data are commonly observed in life testing in the industrial setting. While interest has traditionally been focused on reliability estimation rather than reliability *improvement*, Taguchi (1987) advocated the use of designed experiments to identify factor settings that can lead to increased lifetime or improved reliability. Although his proposed analysis method is faulty (Hamada, 1992), he deserves credit for promoting a pro-active approach to reliability engineering. Censored data also occur in other contexts such as measurements with mixed continuous-categorical outcomes, e.g., window open/not open (categorical) and window diameter (continuous) if window is open.

As argued in Hamada and Wu (1991), standard methods often do not work for censored data from highly fractionated experiments. They proposed a new analysis method and demonstrated its usefulness on some real data. The method has the limitation that it relies on the existence of the maximum likelihood estimates (MLEs) for the model that starts their algorithm. As we show in the next section, quite often the MLEs do not exist even for a simple model consisting of only main effects. Thus, the likelihood-based method fails in this simple setting. The main purpose of this paper is to propose a Bayesian approach that circumvents this problem. The resulting estimates are shown to be quite insensitive to the prior specification, which should assure industrial practitioners who may otherwise have reservations with its usage.

One novel feature of our proposal is that the Bayesian analysis strategy depends on the nature of the design. Before elaborating on this important point, we define two types of fractional factorial designs. We call a design *regular* if it can be defined by a group of defining relations; otherwise, it is called *nonregular*. An example of a regular design is given in Table 1. The experiment studied five factors in eight runs based on a 1/4-fraction of a full factorial design with defining relations D=AC and E=BC, i.e., column D is the product of columns A and C, and column E is the product of columns B and C. Here D denotes the main effect of factor D and AC the interaction between A and C. From these two relations, a third relation DE=AB can be generated. These three relations and the identity element form a group called the defining contrast subgroup (John, 1971). All the relations among the factorial effects can be inferred from the group. For example, assuming that all 3-factor and higher interactions are negligible, we can associate each of the seven degrees of freedom in the experiment with the seven strings of aliases: A=CD, B=CE, C=AD=BE, D=AC, E=BC, AB=DE, AE=BD. More details on regular designs can be found in John (1971) and Box, Hunter and Hunter (1978). Apart from the aliasing of effects, we can use these seven effects and the grand mean to form a saturated model for the data. Any other combination of effects can be viewed as a submodel. Therefore, standard Bayesian analysis methods can be applied to the saturated model. For censored data, the Bayesian computations can be easily implemented using the data augmentation algorithm and the Monte Carlo EM algorithm as described by Tanner and Wong (1987) and Wei and Tanner (1990a, b). In Section 3, we review these methods and describe an extension using conjugate priors. In Section 4 we illustrate the method on two reliability improvement experiments involving fluorescent lights and router bits. Unlike the method of Hamada and Wu (1991), the Bayesian approach for regular designs has the additional advantage of not requiring iterations in the model selection phase.

Most fractional factorial designs used in practice are of the regular type. Nonregular designs include the well-known Plackett-Burman (1946) designs and the 18-run design with seven 3-level factors and one 2-level factor. A large collection can be found in Dey (1985) and Wang and Wu (1991). The main reason for their practical use is run size economy. While they have traditionally been used only for screening factors (i.e., main effect estimation), Hamada and Wu (1992) recently showed that they can also be used to entertain and estimate interactions.

For nonregular designs, not every degree of freedom can be uniquely associated with a string of aliases. Therefore, unlike regular designs a saturated model that contains all other models as submodels cannot be found. Consider the 12-run Plackett-Burman design in Example 3. It is known that every 2-factor interaction is partially aliased with the main effects not present in the interaction with correlation $\pm 1/3$. As a result, there can be many saturated models by choosing k main effects and 11-k 2-factor interactions. Searching for a model by fitting all such possible models would be a formidable task. Instead in Section 5, we propose a method that streamlines the model search by adapting a strategy from Hamada and Wu (1991). For any initial model, apply Bayesian analysis methods to obtain estimates for the model parameters. Next, impute the censored observations based on these estimates. Then standard model selection techniques can be applied to the pseudo-complete data (i.e., the combined complete and imputed data) to obtain an updated model. The procedure is repeated until the selected model stops changing. It works well on the cast fatigue data in Example 3.

A common concern in the practical use of Bayesian methods is its reliance on the prior specification. This is especially acute if the prior distribution is not based on past data or prior experience. To alleviate this concern, we propose a sensitivity analysis of the model chosen by the methods of Sections 4 and 5. For a given model, we compute the Bayesian estimates for a range of prior specifications. If the prior specification has little effect on the estimates or their relative rankings, we call the result *robust*. The results are fairly robust for the three data sets considered here.

2. ESTIMABILITY PROBLEMS WITH THE MLE METHOD

Consider the simplest case of a one-factor experiment with two levels. If all the observations at the high level are censored while those at the low level are not, the likelihood function for any reasonable regression model increases to the maximum value as the regression parameter, i.e., the factor main effect, tends toward infinity. Therefore, the MLE for the main effect does not exist when, obviously, the factor exhibits a large effect and the opportunity for improvement is great. For a general class of regression models with vector covariates x, Silvapulle and Burridge (1986) and Hamada and Tse (1988) showed that the MLEs exist if and only if there does not exist a nonzero vector e such that (i) $e^T x_i = 0$ for x_i corresponding to the uncensored observations and (ii) $e^T x_i \geq 0$ for x_i corresponding to the censored observations. Implications of these conditions for factorial designs were given in Hamada and Tse (1992) and illustrated by the following example.

Example 1: Light Lifetime Experiment

Consider an experiment to improve the lifetime of fluorescent lights (Taguchi, 1987, page 930). An 8-run experiment with two replicates was performed to study five two-level factors denoted by A-E. The experiment, which used a 2^{5-2} fractional factorial design with defining relations D=AC and E=BC, was conducted over 20 days with inspections every two days. The design and lifetime data appear in Table 1, where "–" and "+" refer to the low and high factor levels, respectively. The lifetimes are given as intervals with (14,16) indicating that the light failed between the 14th and 16th day inspections and (20,∞) indicating that the light had not failed by the 20 day inspection, i.e., a censored observation. Besides the main effects, the experimenter also thought that the AB(=DE) interaction might be potentially important.

In addition to considering these six effects, estimation of the effect BD(=AE) can also be entertained. Because both replicates at run 5 are censored, according to the results in Hamada and Tse (1992), the maximum likelihood estimates for all eight effects (including an intercept) under the lognormal regression model (Lawless, 1982) do not exist. Consequently, a standard likelihood-based analysis of the full model is not feasible. Next, we consider how a Bayesian approach avoids this problem.

Table 1: Design and Lifetime Data for Light Experiment

	design					data	
run	A	B	C	D	E	(no. of days)	
1	+	+	+	+	+	(14,16)	(20,∞)
2	+	+	-	-	-	(18,20)	(20,∞)
3	+	-	+	+	-	(08,10)	(10,12)
4	+	-	-	-	+	(18,20)	(20,∞)
5	-	+	+	-	+	(20,∞)	(20,∞)
6	-	+	-	+	-	(12,14)	(20,∞)
7	-	-	+	-	-	(16,18)	(20,∞)
8	-	-	-	+	+	(12,14)	(14,16)

3. BAYESIAN COMPUTATIONS VIA DATA AUGMENTATION AND THE MONTE CARLO EM

Using the standard multiple linear regression model for the response y (or transformed response h(y)),

$$y_i = x_i{}^T\beta \ + \ \sigma\epsilon_i, \ i = 1, \ \ldots, \ n, \tag{1}$$

where the x_i are $k \times 1$ vectors and the ϵ_i are independent standard normal. For censored data, Bayesian computations are easily implemented using the methods proposed by Wei and Tanner (1990a, b). They showed how the posterior and posterior maximizer for β can be obtained using data augmentation (Tanner and Wong, 1987) and the Monte Carlo EM (MCEM), respectively, under an improper prior for β and σ. We consider the more general class of natural conjugate priors (Raiffa and Schlaifer, 1961) in order to study the sensitivity of the conclusions to the prior specification.

Data augmentation turns the calculation of the posterior for censored data into that for complete data by randomly imputing the censored data as follows (Wei and Tanner, 1990b). Randomly sample the current approximation of the (β, σ) posterior m times. For each sampled (β, σ), obtain a pseudo-complete data set by randomly imputing the censored data, i.e., by randomly sampling from a conditional normal distribution resulting from the actual lifetime exceeding the censoring time. The updated approximation of the posterior is then a mixture of the posteriors computed from the pseudo-complete data sets.

The natural conjugate prior (Raiffa and Schlaifer, 1961) is

$$p(\beta,\sigma) = \sigma^{-k} exp\{-(\beta-\beta_0)^T A_0(\beta-\beta_0)/2\sigma^2\} \times \sigma^{-(\nu_0+1)} exp(-\nu_0 s_0^2/2\sigma^2) \,. \qquad (2)$$

From the factorization (2), it can be shown that the marginal prior for σ^2 is the reciprocal of a Gamma distribution with parameters $(\nu_0/2,\ \nu_0 s_0^2/2)$ and the marginal prior of β given σ^2 is multivariate normal with mean β_0 and covariance matrix $\sigma^2 {A_0}^{-1}$. Thus, the factorization provides a straightforward way to sample from the prior for (β,σ^2) (and consequently, the posterior since it has the same form).

The data augmentation algorithm starts by treating the censored observations as actual observations (i.e., pseudo-complete data) to get an initial posterior. This posterior has the same form as (2) except that σ^2 has the reciprocal of a Gamma distribution with parameters $(\nu_1/2,\ \nu_1 s_1^2/2)$ and the marginal posterior of β given σ^2 is multivariate normal with mean $\tilde{\beta}$ and covariance matrix $\sigma^2 M^{-1}$, where $\nu_1 = n+\nu_0$, $M = X^T X + A_0$, $\tilde{\beta} = M^{-1}(X^T X\hat{\beta} + A_0\beta_0)$, $\nu_1 s_1^2 = \nu_0 s_0^2 + \nu s^2 + (\tilde{\beta}-\beta_0)^T A_0(\tilde{\beta}-\beta_0) + (\tilde{\beta}-\hat{\beta})^T X^T X(\tilde{\beta}-\hat{\beta})$, $X^T = (x_1^T,\ldots,x_n^T)$ with x_i given in (1), and $\hat{\beta}$ and νs^2 are the least squares estimate and residual sum of squares, respectively, based on the pseudo-complete data. See Abowd, Moulton and Zellner (1987) for details.

Next, the current approximation to the posterior of (β,σ) is randomly sampled m times using the results from the factorization above. For each sampled (β,σ), obtain a pseudo-complete data set by randomly imputing each censored datum, i.e., randomly sampling from a conditional normal distribution with density $\sigma^{-1}\phi(\{y-{x_i}^T\beta\}\sigma^{-1})[1-\Phi(\{y_i-{x_i}^T\beta\}\sigma^{-1})]^{-1}$ for $y > y_i$, where ϕ and Φ are the standard normal pdf and cdf. The updated approximation of the posterior is then a mixture of the posteriors computed from the pseudo-complete data sets. That is, a $\tilde{\beta}$ and $\nu_1 s_1^2$ are computed from each pseudo-complete data set as was done above in obtaining the initial posterior. The updated approximation of the posterior can then be randomly sampled by first randomly sampling from the integers 1 to m, say j, and then randomly sampling from the jth posterior (of the updated posterior approximation). The approximation of the posterior continues to be updated until convergence. We monitored the maximum relative and absolute componentwise difference of the first four moments and the 0.005, 0.025, 0.975 and 0.995 quantiles between successive approximations of the posterior and stopped when these criteria were sufficiently small. Note that at convergence, the final m $(\tilde{\beta},\ \nu_1 s_1^2)$ can be viewed as a sample from the posterior so that for sufficiently large m, a histogram (or smoothed histogram) and summary statistics such as the moments and quantiles can be computed to obtain estimates of these quantities.

Briefly, we discuss results for the non-informative prior whose density is $p(\beta,\sigma) = \sigma^{-1}$. The data augmentation algorithm is much the same as described above and is given in Wei and Tanner (1990b), although their non-informative prior was for σ^2 rather than σ. In fact, the non-informative prior can be obtained from the informative prior by setting A_0, β_0 and s_0^2 to zero and ν_0 to -k. Thus, a relatively diffuse prior can be obtained by a suitable choice of prior parameter values close to these particular values. In fact, when the MLEs do not exist as discussed in the previous section, the latter approach is necessary since the posterior mode does not exist under the non-informative prior. Therefore, the conjugate prior, albeit relatively diffuse, is used in the following.

For nonregular designs, initially we will only need the posterior maximizer rather than the entire posterior. Wei and Tanner (1990a) presented the MCEM algorithm for obtaining the posterior maximizer under the non-informative prior. Here we present the results for the informative prior given in (2). They showed that the maximizer could be obtained in a similar fashion as the entire posterior by randomly sampling the posterior based on the current maximizer.

The initial approximation to the maximizer is obtained by treating the censored data as actual observations so that $\tilde{\sigma^2} = (\nu_1 s_1^2)/(n+\nu_0+k+1)$ and $\tilde{\beta} = M^{-1}(X^T X\hat{\beta}+A_0\beta_0)$, where ν_1, $\nu_1 s_1^2$ and M are defined as above and $\hat{\beta}$ and νs^2 (used in calculating $\nu_1 s_1^2$) are the least squares estimates and residual sum of squares, respectively, based on the initial pseudo-complete data. Subsequent approximations to the maximizer are obtained by using the m pseudo-complete data sets denoted by $\{z_j\}$ (arising from the random imputations): $\tilde{\beta} = \sum_j \tilde{\beta}_j/m$, where $\tilde{\beta}_j = M^{-1}(X^TX\hat{\beta}_j + A_0\beta_0)$, $\tilde{\sigma^2} = \sum_j\{\nu_0 s_0^2 + \nu s_j^2 + (\tilde{\beta} - \tilde{\beta}_j)^T M(\tilde{\beta} - \tilde{\beta}_j)\}/m(\nu_1 + k + 1)$ and $\hat{\beta}_j$ and νs_j^2 are the least squares estimate of β and the residual sum of squares from z_j, respectively. The approximation of the posterior maximizer continues to be updated until convergence.

Thus far, the posterior and posterior maximizer were discussed when the data were either complete or right-censored. Left-censored and interval-censored are also easily handled by randomly sampling the appropriate conditional normal distribution to obtain the pseudo-complete data sets. In fact, left-, right- and interval-censored data can be represented as intervals (a,b), whose corresponding conditional normal distribution has the density $\sigma^{-1}\phi(\{y - x_i{}^T\beta\}\sigma^{-1})/[\Phi(\{b - x_i{}^T\beta\}\sigma^{-1}) - \Phi(\{a - x_i{}^T\beta\}\sigma^{-1})]$ for $a < y_i < b$. Note that b and (a+b)/2 can be used as the initial pseudo-complete data for left- and interval-censored data, respectively.

4. ILLUSTRATION OF BAYESIAN ANALYSIS AND ITS ROBUSTNESS TO PRIOR SPECIFICATION

Consider the 8-run design in Example 1. Of the seven available degrees of freedom, five are used for estimating the main effects (A-E). The two remaining degrees of freedom can be used to estimate the two pairs of aliased interactions, AB=DE and AE=BD, since none of the four interactions is aliased with any of the main effects. Therefore, we can entertain a linear model consisting of the seven effects A, B, C, D, E, AB and AE and an intercept term. Because of the two aliasing relations, AB and AE can be substituted by DE and BD in the model. Apart from the aliasing of the interactions, we can use this regression model in the Bayesian analysis.

Example 1 (Continued): Analysis of Light Lifetime Experiment

Next, we obtain the posterior for this full model in the log lifetime, i.e., the lognormal regression model. Initially, a relatively diffuse prior for (β, σ) was used whose parameters were: $\beta_0 = 0$ except that the intercept is set equal to 3, the average of the logged experimental data; $A_0 = 0.0001 I_8$, where I_8 is the 8×8 identity matrix (i.e., a prior variance for the intercept and factorial effects of $10000\sigma^2$); $\nu_0 = 1$ and $s_0^2 = 0.01$ ((0.035, 18.039) is the central 99% of the prior distribution for σ^2.). The data augmentation algorithm required 13 iterations whose sample sizes (m) were 4(100), 3(1000), 3(10000), 3(50000); i.e., m=100 for iterations 1 to 4, m=1000, for iterations 5 to 7, etc. Histograms of the marginal posteriors based on the last iteration for the seven effects of interest are displayed in Figure 1. The quantiles corresponding to the central 95% and 99% of the marginal posteriors are presented in Table 2. (The intercept is denoted by Int.) These results suggest that only main effects A, B, D and E are important, with D having the largest effect, and B being perhaps somewhat larger than A and E.

We also obtained the marginal posteriors under more peaked priors on the factorial effects. Instead of a prior variance of $10000\sigma^2$, prior variances of $100\sigma^2$ and σ^2 were used (with lower right 7×7 matrix of A_0 being $0.01 I_7$ and I_7, respectively). Table 2 reports the results for the more peaked prior with variance σ^2, the magnitude of the error variance. In spite of the ten thousand fold decrease in prior variance, there are hardly any differences in the extreme quantiles and therefore no differences in the conclusions; namely, the A, B, D and E main effects are important and in the same order of importance.

The next example is more complicated. The experimental design is obtained from a regular design with 32 runs and 13 two-level factors by grouping one set of three two-level factors into a four-level factor for column D and another set for column E as shown in Table 3. Because the effects are either orthogonal or fully aliased, a comprehensive model can be specified as demonstrated in Hamada and Wu (1991).

Table 2: Posterior Quantiles for Light Lifetime Experiment

	diffuse prior				peaked prior			
	quantile				quantile			
effect	.005	.025	.975	.995	.005	.025	.975	.995
Int	2.83	2.84	3.02	3.07	2.83	2.84	3.03	3.10
A	-0.24	-0.19	-0.02	-0.00	-0.21	-0.16	-0.01	0.01
B	0.07	0.09	0.26	0.31	0.06	0.08	0.24	0.29
C	-0.06	-0.05	0.12	0.17	-0.07	-0.05	0.10	0.15
AB	-0.15	-0.10	0.07	0.08	-0.13	-0.08	0.07	0.09
E	0.02	0.04	0.21	0.26	0.01	0.03	0.18	0.23
D	-0.40	-0.35	-0.18	-0.17	-0.38	-0.32	-0.16	-0.15
BD	-0.10	-0.05	0.11	0.13	-0.08	-0.03	0.12	0.14
σ	0.10	0.10	0.20	0.24	0.12	0.13	0.25	0.31

Example 2: Router Bit Lifetime Experiment

Phadke (1986) reported on an experiment to improve router bit life for a routing process that cuts 8x4 inch printed wiring boards from an 18x24 inch panel. When the router bit becomes dull, it produces boards with rough edges, thereby requiring an extra cleaning process. Also, frequently changing the router bits is expensive. Failure is determined by evidence of an excessive amount of dust, where router bit life is measured in (x100) inches of cut in the x-y plane. A 32-run design was used to study the nine factors denoted by **A**-**I** (number of levels in parentheses) which are suction (2), x-y feed (2), in-feed (2), bit type (4), spindle position (4), suction foot (2), stacking height (2), slot depth (2) and speed (2), respectively. The design and data appear in Table 3. The router bits were inspected every 1(x100) inches up to 17(x100) inches which resulted in left-censored (14 failed before the first inspection), interval-censored and right-censored (eight had not failed when the experiment was stopped) data. The experimenters were interested in the relative importance of the nine main effects and four two-factor interactions, BI, CI, GI, and BG. The experimental objective was to identify factors which improve router bit life.

Table 3: Design and Lifetime Data for Router Bit Experiment

	design									
run	A	B	C	D	E	F	G	H	I	data
1	1	1	1	1	1	1	1	1	1	(3,4)
2	1	1	1	2	2	2	2	1	1	(0,1)
3	1	1	1	3	4	1	2	2	1	(0,1)
4	1	1	1	4	3	2	1	2	1	(17,∞)
5	1	2	2	3	1	2	2	1	1	(0,1)
6	1	2	2	4	2	1	1	1	1	(2,3)
7	1	2	2	1	4	2	1	2	1	(0,1)
8	1	2	2	2	3	1	2	2	1	(0,1)
9	2	1	2	4	1	1	2	2	1	(17,∞)
10	2	1	2	3	2	2	1	2	1	(2,3)
11	2	1	2	2	4	1	1	1	1	(0,1)
12	2	1	2	1	3	2	2	1	1	(3,4)
13	2	2	1	2	1	2	1	2	1	(0,1)
14	2	2	1	1	2	1	2	2	1	(2,3)
15	2	2	1	4	4	2	2	1	1	(0,1)
16	2	2	1	3	3	1	1	1	1	(3,4)
17	1	1	1	1	1	1	1	1	2	(17,∞)
18	1	1	1	2	2	2	2	1	2	(0,1)
19	1	1	1	3	4	1	2	2	2	(0,1)
20	1	1	1	4	3	2	1	2	2	(17,∞)
21	1	2	2	3	1	2	2	1	2	(0,1)
22	1	2	2	4	2	1	1	1	2	(17,∞)
23	1	2	2	1	4	2	1	2	2	(14,15)
24	1	2	2	2	3	1	2	2	2	(0,1)
25	2	1	2	4	1	1	2	2	2	(17,∞)
26	2	1	2	3	2	2	1	2	2	(3,4)
27	2	1	2	2	4	1	1	1	2	(17,∞)
28	2	1	2	1	3	2	2	1	2	(3,4)
29	2	2	1	2	1	2	1	2	2	(0,1)
30	2	2	1	1	2	1	2	2	2	(3,4)
31	2	2	1	4	4	2	2	1	2	(0,1)
32	2	2	1	3	3	1	1	1	2	(17,∞)

By studying the design structure of this experiment, Hamada and Wu (1991) observed that the interactions AF, CH, AI, FI, and HI could also be entertained. In addition, because factor E was not thought to be important, interactions AH, BF and CG could be studied as well. The model with these 23 effects (intercept, seven two-level factor main effects, three two-level pseudofactors for the D main effect (denoted by D1, D2 and D3) and 12 interactions) could not be fit using a lognormal regression model because the maximum likelihood estimates do not exist. Note that Hamada and Wu (1991) initially entertained the smaller model with main effects and interactions (BI, CI, GI and BG) model whose maximum likelihood estimates did exist.

Next, we obtain the posterior for the 23-effect model assuming a lognormal regression model, i.e., the lifetimes are logged. The prior parameters are: $\beta_0 = 0$ except that the intercept is set equal to 1.5, the average of the logged experimental data; $A_0 = 0.0001 I_{23}$; $\nu_0 = 1$ and $s_0^2 = 0.01$. The data augmentation algorithm required 12 iterations whose sample sizes were 3(500), 4(1000), 2(5000), 3(10000). Histograms of the marginal posteriors are displayed in Figure 2. The quantiles corresponding to the central 95% and 99% of the marginal posteriors for the significant effects are presented in Table 4. The results for this relatively diffuse prior suggest that I, B, (D1,D3), BF, CG, F, G, GI, AF, which were also identified by Hamada and Wu (1991), are important. In addition to these effects, the Bayesian analysis identifies BI, CI, FI and HI as important.

The sensitivity of these results were investigated using more peaked priors on the factorial effects. The results for a prior variance of σ^2, a ten thousand fold decrease, (i.e., the lower right 22×22 matrix of A_0 equal to I_{22}) appear in Table 4. There are some minor differences with BI and CI, two of the smallest effects under the peaked prior, no longer being significant, although they remain significant for a less peaked prior with a variance of $10\sigma^2$. Note that the relative rankings of the factorial effects are affected much less than the quantile values.

Table 4: Posterior Quantiles for Router Bit Lifetime Experiment

	diffuse prior				peaked prior			
	quantile				quantile			
effect	.005	.025	.975	.995	.005	.025	.975	.995
int	0.81	0.87	1.15	1.20	0.47	0.60	1.17	1.25
I	0.45	0.48	0.70	0.73	0.32	0.38	0.84	0.94
B	-0.66	-0.61	-0.32	-0.28	-0.94	-0.83	-0.30	-0.23
D1	0.23	0.28	0.56	0.62	0.11	0.19	0.72	0.84
D3	-1.12	-1.05	-0.78	-0.74	-1.46	-1.32	-0.76	-0.70
BF	0.14	0.18	0.44	0.50	0.04	0.13	0.64	0.75
CG	-0.52	-0.46	-0.18	-0.14	-0.78	-0.67	-0.14	-0.06
F	-0.69	-0.64	-0.39	-0.34	-0.98	-0.86	-0.33	-0.27
BI	-0.40	-0.36	-0.10	-0.06	-0.52	-0.42	0.04	0.13
G	-0.95	-0.91	-0.64	-0.60	-1.31	-1.18	-0.62	-0.54
CI	-0.40	-0.37	-0.09	-0.03	-0.56	-0.47	0.02	0.12
GI	0.32	0.35	0.57	0.61	0.22	0.28	0.73	0.82
AF	0.31	0.34	0.61	0.68	0.22	0.29	0.81	0.94
FI	0.12	0.16	0.38	0.43	0.01	0.08	0.55	0.64
HI	0.06	0.09	0.29	0.33	-0.04	0.03	0.45	0.54
σ	0.03	0.04	0.16	0.19	0.35	0.37	0.58	0.64

5. ANALYSIS STRATEGY FOR DESIGNS WITH COMPLEX ALIASING

As was demonstrated in Examples 1 and 2, each degree of freedom in a regular design can be associated with a string of aliases, which in group-theoretic terms is a coset of the defining contrast subgroup. This fact allows us to find a comprehensive model for the design. For nonregular designs, this simplicity is lost because the partial aliasing of effects makes it impossible to associate each degree of freedom with a unique string of effects. Consider the 12-run Plackett-Burman design (see Table 5) for which each two-factor interaction is partially aliased with all main effects not present in the interaction. If six factors are significant, to use up the remaining five degrees of freedom, only five of the 15 two-factor interactions can be entertained. There are, however, $\binom{15}{5} = 3003$ models! Adopting the Bayesian analysis strategy in Section 3 would require the computation of posteriors for these 3003 models which is impractical. For more complex situations, the total number of models can be much larger. Recognizing this computational problem, Hamada and Wu (1991) proposed a method that drastically reduced the effort needed in model fitting and selection. The main idea is to impute the censored observations based on the parameter estimates obtained from the previous model fit, and then do model selection with the pseudo-complete data consisting of the uncensored data and the imputed data.

Formally, the method adapted to the Bayesian approach is as follows.

A.0 Initial Model Specification

A.1 Bayesian Model Fitting

A.2 Imputation

A.3 Model Selection

Repeat Steps A.1 through A.3 until model selection termination.

A.4 Inference about effects and recommendation on optimum factor levels for the final model from Steps A.0-A.3

B Robustness study of the final model from Steps A.0-A.3

Details for the procedure are as follows.

A.0 The experimenter chooses $\mu = X_0\beta_0$ (Model 0) which includes main effects and interactions thought to be potentially important. For highly fractionated designs, one may be restricted to a main-effects model.

A.1 Use the MCEM algorithm in Section 2 to obtain the posterior maximizer, $\tilde{\beta}$ and $\tilde{\sigma}$.

A.2 Impute the censored data by their conditional means (or medians):

$$E(h(y)|y \in (a,b)) = x_i\tilde{\beta}_i + \tilde{\sigma}(\phi(z_a) - \phi(z_b))/(\Phi(z_b) - \Phi(z_a)), \quad (3)$$

where $z_w = (h(y) - x_i{}^T\tilde{\beta}_i)/\tilde{\sigma}$. Since we are interested in identifying location effects, we use the conditional mean or median as a typical value.

A.3 Informally apply standard forward or stepwise selection procedures to the pseudo-complete data from Step A.2 to select a model. In addition to the initial model effects, entertain interactions between factors having large main effects in the initial model. Also, if there is one dominating factor in the initial model, consider its interactions with the other factors. See the example below and Hamada and Wu (1992) for more details. Stop when the current model selected is the same as the previous model, i.e., $X_i\tilde{\beta}_i = X_{i-1}\tilde{\beta}_{i-1}$. There is no need to apply Bayesian techniques here to the pseudo-complete data. If a selected model has a small number of effects, it may be helpful to compute its posterior to provide an additional assessment of the importance of its effects.

A.4 For the final model chosen in Step A.3, compute its posterior and make inference about each effect.

B Repeat Steps A.0-A.3 for a range of prior specifications. If the final model chosen in A.3 is nearly the same for several priors, the analysis result is declared to be robust. Otherwise, care should be taken in interpreting the choice of prior.

Note that the imputation in Step A.2 is different in nature from the imputation in the data augmentation algorithm. While both use imputation to convert an incomplete data problem into a complete data problem, the former allows model selection and fitting to be performed efficiently over a large number of candidate models; the latter allows Bayesian computations to be performed in a straightforward manner.

Some discussion of imputation is warranted for replicated experiments. If there are runs with two or more of the replicates censored, instead of imputing all censored values (for the same run) with the same value, other alternatives might be considered. Imputation can be done by randomly or systematically sampling the conditional distribution. As an example of the latter, the 25th, 50th and 75th percentiles can be used if there are three replicates.

Next, we illustrate this modified strategy by reanalyzing the data from an experiment which used a 12-run Plackett-Burman design.

Example 3: Weld Repaired Cast Fatigue Experiment

Hunter, Hodi and Eager (1982) used a 12 run Plackett-Burman design to study the effects of seven factors on fatigue life of weld repaired castings. The factors denoted by A-G are initial structure, bead size, pressure treat, heat treat, cooling rate, polish and final treat, respectively, and were assigned to the first seven columns of the design matrix given in Table 3, where "+" and "–" denoting the high and low factor levels. The experiment was stopped at 7.000; the logged lifetimes also appear in Table 3 which shows that one censored observation occurred at run 5. In the original analysis, heat treat (factor D) and polish

Table 5: Design and Lifetime Data for Cast Fatigue Experiment

	design											logged
run	A	B	C	D	E	F	G	8	9	10	11	data
1	+	+	−	+	+	+	−	−	−	+	−	6.058
2	+	−	+	+	+	−	−	−	+	−	+	4.733
3	−	+	+	+	−	−	−	+	−	+	+	4.625
4	+	+	+	−	−	−	+	−	+	+	−	5.899
5	+	+	−	−	−	+	−	+	+	−	+	$(7.000,\infty)$
6	+	−	−	−	+	−	+	+	−	+	+	5.752
7	−	−	−	+	−	+	+	−	+	+	+	5.682
8	−	−	+	−	+	+	−	+	+	+	−	6.607
9	−	+	−	+	+	−	+	+	+	−	−	5.818
10	+	−	+	+	−	+	+	+	−	−	−	5.917
11	−	+	+	−	+	+	+	−	−	−	+	5.863
12	−	−	−	−	−	−	−	−	−	−	−	4.809

(factor F) were identified as significant, although heat treat had a much smaller effect (with a p-value around .2).

The initial model (Model 0) considered was the main effects model. Note that the log lifetimes were analyzed by Hodi et al. (1982) so that the lognormal regression model was assumed. The prior parameters are: $\beta_0 = 0$ except that the intercept is set equal to 5, the average of the logged experimental data; $A_0 = 0.0001 I_8$; $\nu_0 = 1$ and $s_0^2 = 0.01$. The data augmentation algorithm required ten iterations whose sample sizes were 3(500), 3(1000), 4(10000). Histograms of the marginal posteriors are displayed in Figure 3 which suggest that F has the largest effect, then D followed by A, B and C, with the latter three being approximately the same.

Because of this design's complex aliasing patterns, interactions between (A, B, C, D, F) and also EF and FG were entertained following the strategy proposed in Hamada and Wu (1992). The censored observation (at run 5) was imputed using $(\tilde{\beta}, \tilde{\sigma^2})$, the maximizer of the Model 0 posterior. For the right-censored observation in this experiment, the imputation formula given in Step A.2 reduces to $x_i^T\tilde{\beta} + \tilde{\sigma}\phi(\{y_i - x_i^T\tilde{\beta}\}\tilde{\sigma}^{-1})/[1 - \Phi(\{y_i - x_i^T\tilde{\beta}\}\tilde{\sigma}^{-1})]$ which was 7.156.

Using the resulting pseudo-complete data set, a forward selection procedure was performed entertaining the (A, B, C, D, F) main effects, the interactions between them and the EF and FG interactions, which indicated that FG appears to be important. In fact, FG was the first effect selected and together with F have an R^2 of .87. Note that the model (F,D) identified in the original analysis has an R^2 of .58, whereas the model containing (F,D,A,B,C) has an R^2 of .73. Hence, the next model (Model 1) we considered contained all main effects plus the FG interaction. Figure 4 displays the marginal posteriors obtained from the data augmentation algorithm which required six iterations with sample sizes of 2(500), 2(1000), 2(10000). These results suggest that F and FG are the largest effects followed by much smaller ones for D, E and G. Note that now A, B and C appear negligible. A subsequent imputation using the Model 1 maximizer (7.042 at run 5) and forward selection procedure again identified FG as significant. A final application of data augmentation to obtain the posterior for model (D,E,F,G,FG), which required the same number of iterations and sample sizes as the previous model, produced the marginal posteriors whose histograms are displayed in Figure 5. The quantiles corresponding to the central 95% and 99% of the marginal posteriors are displayed in Table 6 for both the diffuse and peaked priors. For the peaked prior, a ten thousand decrease in the prior variance for the factorial effects was used (i.e., the lower right 5×5 matrix of A_0 equal to I_5). Note the insensitivity (or robustness) of the results to the prior specification. These results suggest that F and FG dominate with D, E, G being approximately equal but much smaller.

Table 6: Posterior Quantiles for Cast Fatigue Experiment

	diffuse prior				peaked prior			
	quantile				quantile			
effect	.005	.025	.975	.995	.005	.025	.975	.995
Int	5.73	5.73	5.76	5.77	5.73	5.73	5.78	5.80
D	-0.13	-0.13	-0.11	-0.11	-0.15	-0.14	-0.11	-0.11
E	-0.14	-0.12	-0.07	-0.07	-0.14	-0.11	-0.06	-0.06
F	0.46	0.46	0.49	0.50	0.42	0.42	0.46	0.48
G	0.05	0.06	0.09	0.09	0.02	0.04	0.08	0.08
FG	-0.50	-0.48	-0.45	-0.45	-0.48	-0.46	-0.41	-0.41
σ	0.15	0.15	0.21	0.23	0.23	0.23	0.31	0.35

ACKNOWLEDGEMENTS

This research was supported by General Motors of Canada Limited, the Manufacturing Research Corporation of Ontario, and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Abowd, J.M., Moulton, B.R. and Zellner, A. (1987), "The Bayesian Regression Analysis Package User's Manual," University of Chicago Graduate School of Business Technical Report.

Box, G., Hunter, W.G. and Hunter, J.S. (1978), *Statistics for Experimenters*, New York: John Wiley.

Dey, A. (1985), *Orthogonal Fractional Factorial Designs*, New York: Halsted Press.

Hamada, M. (1992), "An Explanation and Criticism of Minute Accumulating Analysis," *Journal of Quality Technology*, 24, 70-77.

Hamada, M. and Tse, S.K. (1988), "A Note on the Existence of Maximum Likelihood Estimates in Linear Regression Models Using Interval-censored Data," *Journal of the Royal Statistical Society*, Ser. B, 50, 293-296.

Hamada, M. and Tse, S.K. (1992), "On Estimability Problems in Industrial Experiments with Censored Data," *Statistica Sinica*, 2, 381-391.

Hamada, M. and Wu, C.F.J. (1991), "Analysis of Censored Data from Highly Fractionated Experiments," *Technometrics*, 33, 25-38.

Hamada, M. and Wu, C.F.J. (1992), "Analysis of Designed Experiments with Complex Aliasing," *Journal of Quality Technology*, 24, 130-137.

Hunter, G.B., Hodi, F.S. and Eager, T.W. (1982), "High-Cycle Fatigue of Weld Repaired Cast Ti-6Al-4V," *Metallurgical Transactions* 13A, 1589-1594.

John, P.W.M. (1971), *Statistical Design and Analysis of Experiments*, New York: Macmillan.

Lawless, J.F. (1982), *Statistical Models and Methods for Lifetime Data*, New York: John Wiley.

Phadke, M.S. (1986), "Design Optimization Case Studies," *AT&T Technical Journal*, 65, 51-68.

Plackett, R.L. and Burman, J.P. (1946), "The Design of Optimum Multifactorial Experiments," *Biometrika* 33, 305-325.

Raiffa, H. and Schlaifer, R. (1961), *Introduction to Statistical Decision Theory*, New York: McGraw-Hill.

Taguchi, G. (1987), *System of Experimental Design,* White Plains NY: Unipub/Kraus International Publications.

Tanner, M.A. and Wong, W.H. (1987), "The Calculation of Posterior Distributions by Data Augmentation" (with discussion), *Journal of the American Statistical Association*, 82, 528-540.

Wang, J.C. and Wu, C.F.J. (1991), "An Approach to the Construction of Asymmetrical Orthogonal Arrays," *Journal of the American Statistical Association*, 86, 450-456.

Wei, G.C.G. and Tanner, M.A. (1990a), "Posterior Computations for Censored Regression Data," *Journal of the American Statistical Association*, 85, 829-839.

Wei, G.C.G. and Tanner, M.A. (1990b), "A Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithms," *Journal of the American Statistical Association*, 85, 699-704.

Figure 1: Light Experiment Posteriors

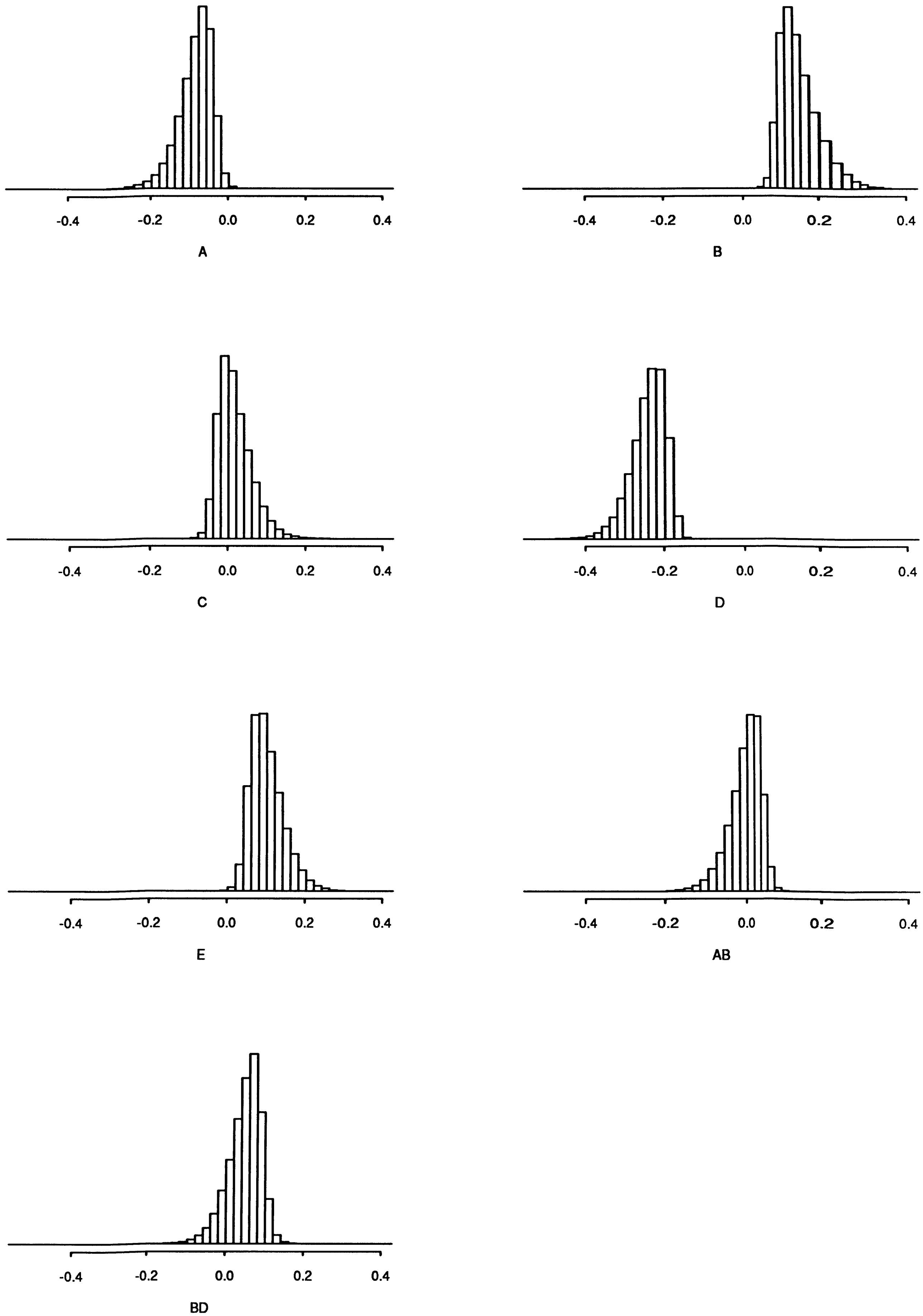

Figure 2: Router Experiment Posteriors

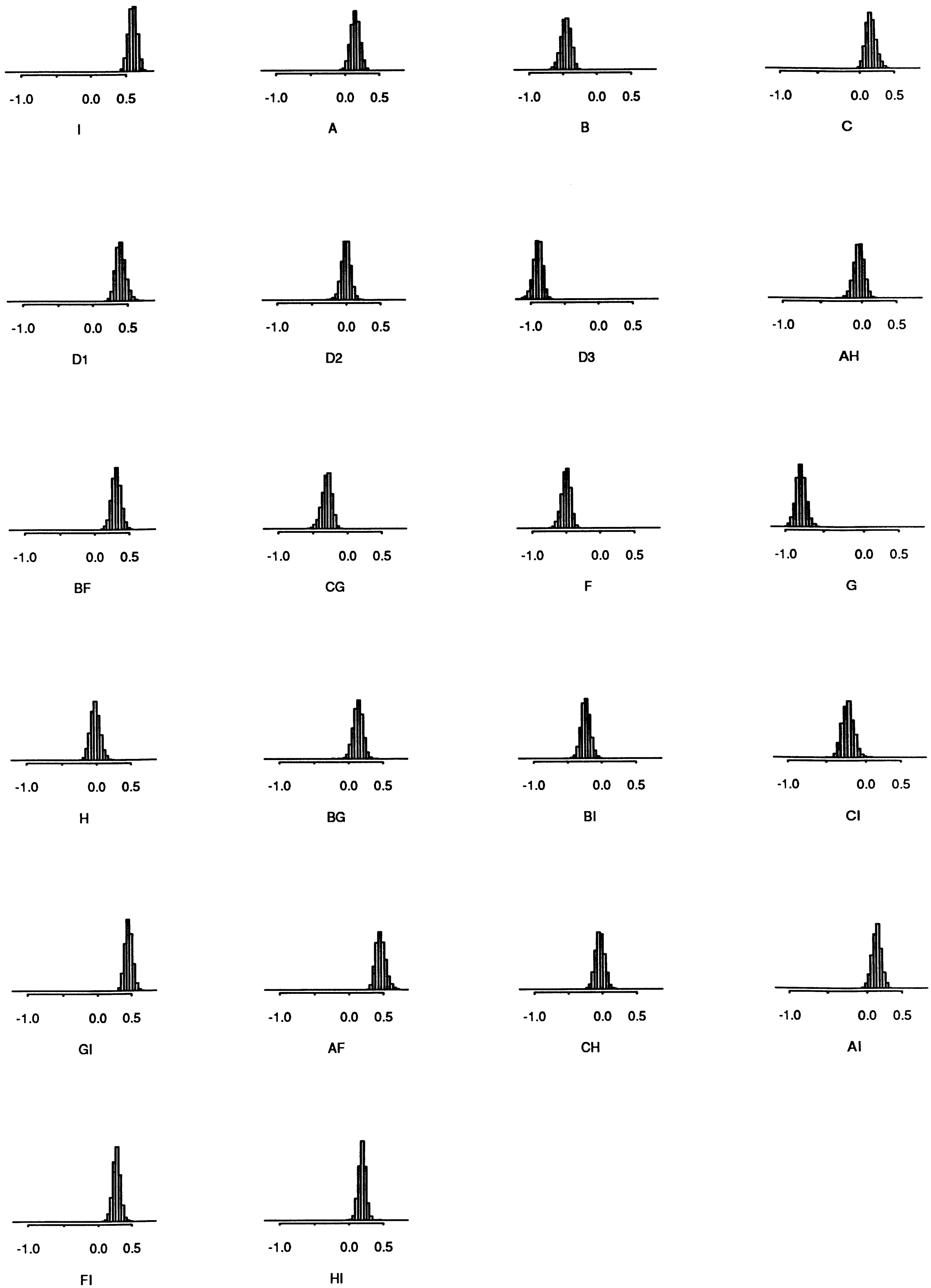

Figure 3: Cast Experiment Posteriors for Model 0

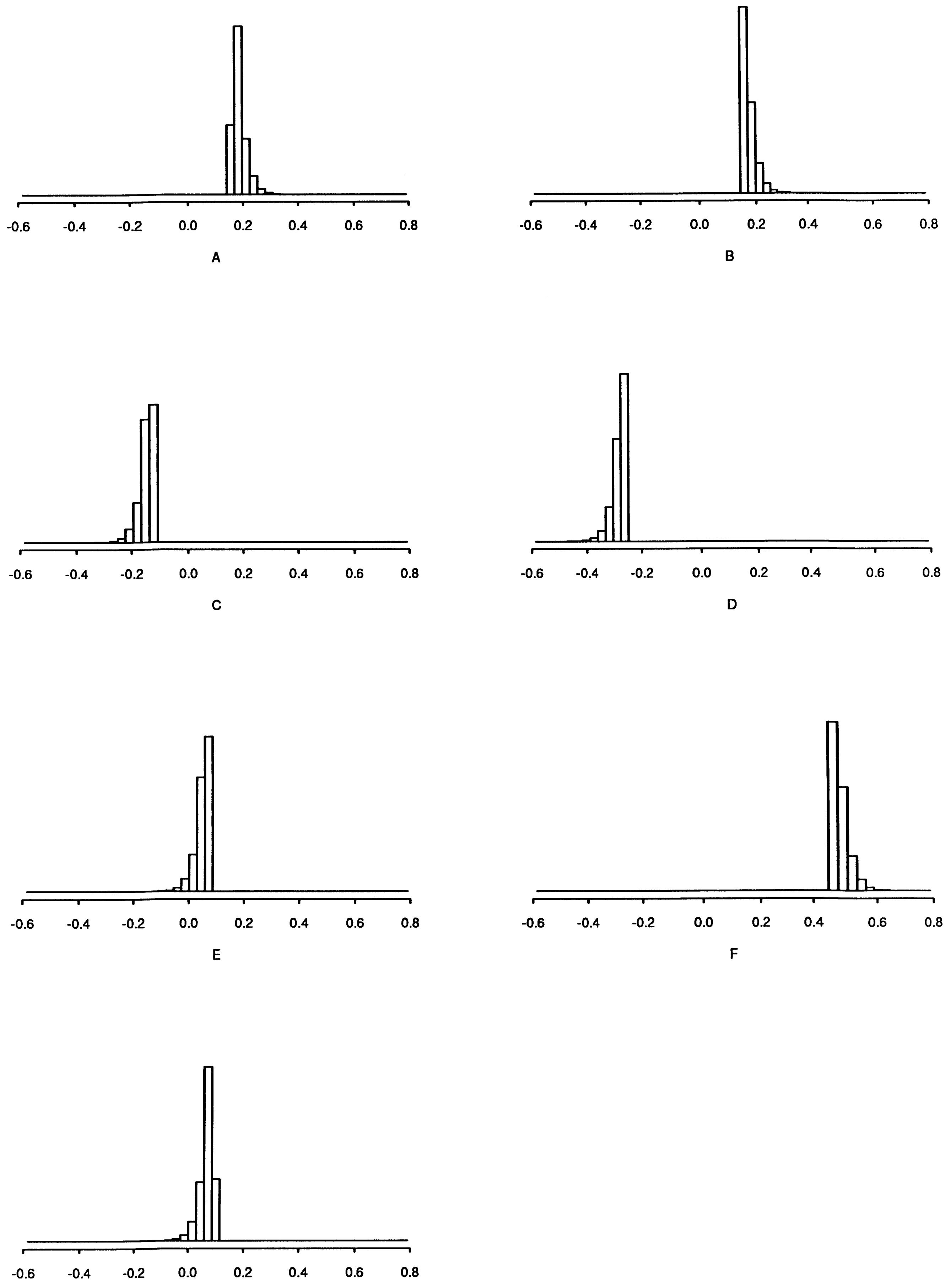

Figure 4: Cast Experiment Posteriors for Model 1

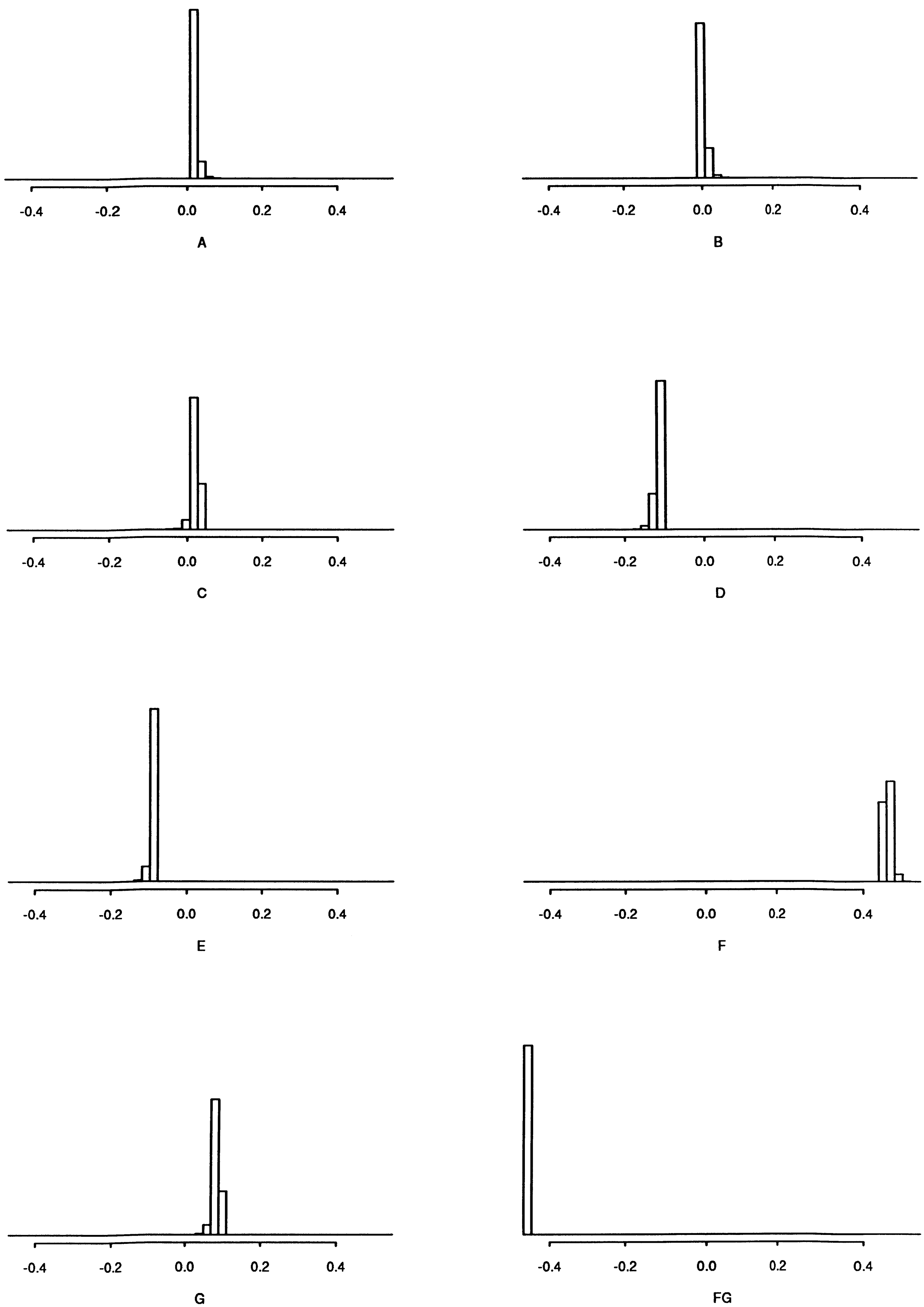

Figure 5: Cast Experiment Posteriors for Model 2