

BAYESIAN ESTIMATION OF MANUFACTURING EFFECTS IN A FUEL ECONOMY MODEL

R. W. ANDREWS

School of Business Administration, University of Michigan, Ann Arbor, MI 48109-1234, USA

J. O. BERGER

Statistics Department, Purdue University, West Lafayette, IN 47907, USA

AND

M. H. SMITH

Mathematics Department, University of Canterbury, Christchurch, New Zealand

SUMMARY

The analysis of fuel economy data results in estimates of the technology utilization by manufacturer and vehicle line. The analysis employs a hierarchical Bayesian regression model with random components representing vehicle lines and manufacturers. The model includes predictor variables which describe vehicle features, such as type of transmission, and vehicle line specific measurements, such as compression ratio. Non-informative priors with novel modifications are used and the Bayes estimates are obtained by use of Gibbs sampling. The results show there is substantial variability among manufacturers in efficiently utilizing technology for fuel economy.

1. INTRODUCTION

With the ever-growing importance of energy conservation and environmental protection the fuel efficiency of passenger cars and trucks must be constantly monitored. In the United States there is a Corporate Average Fuel Economy (CAFE) requirement of 27·5 miles per gallon (mpg). The requirement is a sales-weighted harmonic average for all vehicles sold by a manufacturer within a year. If this requirement is not met, the automobile manufacturer is required to pay a penalty on every vehicle sold of US\$5 times the number of tenths of a mile per gallon below the requirement. Therefore, the incentive for producing the most fuel-efficient vehicles possible is driven by energy conservation, environmental awareness, customer expectations, and a monetary penalty.

With the growing awareness of energy and environmental considerations the US legislature has been studying and, in some instances, supporting an increase in the minimum CAFE standards, from 27·5 mpg to as much as 40·0 mpg. Some proponents of this increase state that it can be achieved with established technologies. As there is considerable controversy over this issue, we utilized a large database to investigate the potential gain in fuel economy due to established technologies. That analysis is described in Andrews *et al.* (1993). It combined the data analysis with a subjective engineering assessment scheme to predict the increase in fuel efficiency that would be possible by the years 1995 and 2001.

For the data analysis part of that project, it was only necessary to obtain regression estimates of the effects of the studied technologies. A wealth of other information can be extracted from the data, however, particularly residual fuel-efficiency effects of specific manufacturers and even specific vehicle lines. These effects measure the degree to which a given manufacturer or a given vehicle line has achieved fuel economy for its given level of technology utilization. Thus, a high residual manufacturer mean indicates that the manufacturer either does an above-average job of implementing the technologies for fuel economy or calibrates and optimizes the vehicle for fuel economy in other ways.

The data analysis utilizes a hierarchical Bayesian model with a number of interesting features. These include: (1) a regression (fixed effects) component; (2) multilevel random effects with a correlation structure; (3) constraints on the parameters; and (4) unbalanced design and unequal random effect variances.

Effective analysis of a model of this complexity virtually requires Bayesian methods; alternative approaches do not produce reliable accuracy estimates. For the Bayesian analysis we utilized non-informative priors on fixed effects and hyperparameters. For the fixed effects, however, a *shrinkage* non-informative prior was employed.

Computation was done using the Gibbs sampling simulation scheme, with estimates (posterior means) and standard errors (square root of posterior variances) being produced for all 326 parameters. The *shrinkage* prior and the constraints on the parameters required some novel modifications.

In Section 2 the data used in this study are described. The statistical model relating fuel economy to characteristics of the vehicle, including the manufacturer, is presented in Section 3, followed by the prior distributions in Section 4. The method of solution, Gibbs sampling, is given in Section 5 and the results are discussed in the final section.

2. DATA

The CAFE legislation requires that all passenger vehicles sold in the United States be tested for fuel economy on a city and a highway driving cycle. In this study, the fuel economy test data for automobiles and light trucks for the model years 1988, 1989, and 1990 was the starting point. Additional variables for each configuration, not included in the test results, were needed in order to properly account for fuel economy in the statistical model. These additional variables were cam activation and coastdown time as defined in the Appendix, which lists and defines all the variables.

The original data consisted of 2264 automobile and 1509 truck records for a total of 3773. Not all the records were usable in the analysis and 1019 automobile and 561 truck configurations were eliminated, resulting in 2193 records in the final data set. The final data set covered approximately 80% of the vehicles sold in the United States. Data deficiencies and model constraints were the two broad reasons for eliminating records.

The main data deficiency was that a number of records were discovered to be simply the same record carried into the following model year. In other cases the same test results were used for a second configuration when that configuration differed only in a minor way. Elimination of such duplicates reduced the data set by 656 records. Because of the non-availability of the cam activation and/or the coastdown time variables, an additional 534 configurations had to be eliminated.

On the modelling side, preliminary regression runs revealed a number of configurations with very large negative residuals. Often these were very high-performance vehicles and as such were

eliminated as being optimized for performance at the expense of fuel economy; unfortunately, the data did not include performance measures. Others which were of extreme configuration in terms of the model variables were also eliminated. Examples of eliminated vehicles are Rolls-Royce, Lamborghini, Yugo, and all police vehicles. All vehicles with 12-cylinder engines, rotary engines, and diesel engines were also eliminated. All these vehicles (124 configurations), which have a very small market share, could be described as exotic and their elimination simplified and reduced stress on the model.

Turbo- and supercharged vehicles were removed from the data set because they are not considered to be general-use technologies for improving fuel efficiency. Only a very small number of vehicles in the data set had five-speed automatic or continuously variable transmissions. Because estimation of the fuel economy benefit would be very unreliable in these vehicles, they were also eliminated.

The final data set used in the statistical model consisted of 2193 configurations for the 24 variables described in the Appendix. Note that the fuel economy value used is the combined city and highway measurements, computed by

$$MPG = \left(\frac{0.55}{CITY} + \frac{0.45}{HWAY} \right)^{-1} \quad (1)$$

where *CITY* is the miles per gallon measured by the city driving test and *HWAY* is the miles per gallon measured by the highway driving test. The weighted harmonic average attempts to capture a driving pattern that contains 55% of the driving distance in the city and 45% on the highway. This harmonic average is the quantity that is addressed in policy matters, such as legislation.

The lack of any variable in the data set related to tyre-rolling resistance and drive-train friction was a serious shortcoming and it was decided to augment the data with coastdown time, which is a test track measurement of how long it takes a vehicle, not under power, to coast from 55 miles per hour to 45 miles per hour. A variable which measures tyre rolling resistance and drive train friction is tyre and chassis loss (*TCL*) which is computed from coastdown time (*CDT*), dynamometer setting (*DPA*), and equivalent test weight (*ETW*) by

$$TCL = \left(\frac{1}{CDT} - \frac{DPA}{0.06172 \times ETW} \right)^{-1} \quad (2)$$

3. MODEL

The regression model for the *i*th record in the data set is

$$\log Y_i = \beta_0 + \mathbf{X}_i \vec{\beta} + C_i + \varepsilon_i \quad (3)$$

where

$Y_i = MPG$;

$\vec{\beta} = (\beta_1, \dots, \beta_p)^t$ is a $p \times 1$ vector of parameters, and β_0 is the constant term;

\mathbf{X}_i is the $1 \times p$ vector of covariates;

C_i is the random effect for the vehicle-line l_i of the *i*th record;

ε_i is the random error.

The ε_i are independent $N(0, \sigma^2)$ random variables. The prior and hyperprior structure on the regression parameters, the vehicle-line random effects, and the error variance are described in

the next section. The covariates are: $\mathbf{X}_i = (X_{i1}, \dots, X_{ip})$, where $p = 23$. They are of two types: X_{i1}, \dots, X_{i8} , the logarithms of the eight characteristic variables, *ETW*, *TCL*, *DPA*, *N/V*, *AXLE*, *CID*, *COMP* and *HP/CID*; X_{i9}, \dots, X_{i23} , the indicators of the presence of various technologies and/or transmission types.

There are eight continuous variables (logarithm transformed) available as fuel economy predictor variables and 15 indicator variables are used to indicate the presence of various technologies and manual/automatic transmissions. The corresponding β_i in the model can be thought of as the fuel economy effects from the characteristic variables and the technologies and transmission types. The vehicle-line random effects provides interesting information about vehicle lines and manufacturers. Some experimentation with interaction terms revealed that, with the limited available data, interaction terms would not provide a significantly better estimated model.

4. PRIOR DISTRIBUTION

It is reasonable to model the C_{il} for manufacturer k (note that numerous records in the data set can arise from the same vehicle line, l) as

$$C_{il} \stackrel{\text{i.i.d.}}{\sim} N(\mu_k, V_k), \quad l = N_{k-1} + 1, \dots, N_k$$

with μ_k and V_k representing overall manufacturer residual mean and variance and $(N_k - N_{k-1})$ being the number of vehicle lines in the k th manufacturer's fleet. We also model the μ_k and V_k , hierarchically. To allow for possible correlation among firms (due, for example, to competition) we set

$$\tilde{\boldsymbol{\mu}} = (\mu_1, \mu_2, \dots, \mu_M) \sim N_M(\mathbf{0}, \lambda \mathbf{I} + \gamma \mathbf{1}) \quad (4)$$

where \mathbf{I} is the $M \times M$ identity matrix and $\mathbf{1}$ is the $M \times M$ matrix of ones. There is less reason to suppose correlation among the V_k , so we simply assume

$$V_k \stackrel{\text{i.i.d.}}{\sim} IG(\alpha, \tau), \quad k = 1, \dots, M$$

where M is the total number of manufacturers and $IG(\alpha, \tau)$ is the inverse gamma distribution with density

$$f(v | \alpha, \tau) = \frac{\tau^\alpha}{\Gamma(\alpha) v^{(\alpha+1)}} \exp\left\{-\frac{\tau}{v}\right\} \quad \text{for } v > 0$$

The choice of these distributions is for computational convenience, but because of the hierarchical nature of the model there is a reasonable degree of built-in robustness.

The prior used for the hyperparameters $(\sigma^2, \lambda, \tau, \gamma)$ and the constant β_0 (which could, alternatively, be thought of as a hyperparameter, namely the mean of the μ_k) is

$$\pi(\beta_0, \sigma^2, \lambda, \tau, \gamma) = 1/(\sigma^2 \gamma^{3/4})$$

reflecting the usual non-informative prior for σ^2 , and being constant in β_0 , λ , and τ . Although λ , τ , and γ are scale parameters, they occur at higher levels of the hierarchical model and hence cannot be assigned priors proportional to their inverses. Any 'small' power greater than -1 is acceptable; the choices $0, 0, -\frac{3}{4}$ for λ, τ , and γ , respectively, prove convenient later. The remaining hyperparameter, α , was found to have only a slight effect on the answers, and it sufficed to assign α a ten-point discrete prior, giving equal mass of $\frac{1}{10}$ to each of $\{0.75, 1.0, 1.5, 2.0, 2.25, 2.5, 2.75, 3.0, 3.5, 4.0\}$.

The prior distribution for the regression parameters, $\vec{\beta}$, is the constrained non-informative prior

$$\pi(\vec{\beta}) = |\vec{\beta}|^{-(p-1)} 1_{\Omega}(\vec{\beta}) \quad (5)$$

Here 1_{Ω} is the indicator function on the set:

$$\Omega = \{\vec{\beta}: \beta_{10} > 0, \beta_{15} > 0, \beta_{18} > 0\} \quad (6)$$

The prior $\pi(\vec{\beta})$ has considerable justification. When used to estimate $\vec{\beta}$ in the absence of random effects, $\pi(\vec{\beta})$ results in a standard shrinkage estimator for $\vec{\beta}$, an estimator that can be justified from many perspectives (robust Bayes, minimaxity, adaptive ridge regression). The presence of the indicator on Ω in equation (5) is because it is known from engineering principles, that β_{10} , β_{15} , and β_{18} are positive. These coefficients correspond to the technologies *TBI*, *ACTR*, and *ELEC*, respectively. The reason these three technologies can be singled out for such treatment is that they are the only ones whose presence has no effect on the characteristic variables, and hence which are not directly confounded with the characteristic variables.

At first, accommodation of $\pi(\vec{\beta})$ within Gibbs sampling appears to be difficult, because the resulting posterior distribution of $\vec{\beta}$, given the other parameters, is not a nice distribution. This problem can be easily circumvented, however, by introducing an auxiliary parameter $\xi > 0$, and using the prior distribution

$$\begin{aligned} \pi(\vec{\beta} | \xi) &= N_p(\mathbf{0}, \xi^{-1} \mathbf{I}_p) 1_{\Omega}(\vec{\beta}) \\ \pi(\xi) &= \xi^{-3/2} I_{(0, \infty)}(\xi) \end{aligned} \quad (7)$$

The use of equation (7) corresponds to the desired marginal prior, $\pi(\vec{\beta})$, because

$$\int_0^{\infty} \xi^{-3/2} \frac{1}{(2\pi\xi^{-1})^{p/2}} \exp\left\{-\frac{\xi}{2} |\vec{\beta}|^2\right\} d\xi \propto 1/|\vec{\beta}|^{(p-1)}$$

Furthermore, the conditional posteriors for $\vec{\beta}$ and for ξ , given all other parameters, are nice, so that Gibbs sampling can be easily employed.

A final relevant observation concerning $\pi(\vec{\beta})$ is that the marginal prior of some k -dimensional subvector of $\vec{\beta}$, call it $\vec{\beta}_I$ with $\vec{\beta}_r$ denoting the remaining co-ordinates, is

$$\begin{aligned} \pi_2(\vec{\beta}_I) &= \int |\vec{\beta}|^{-(p-1)} 1_{\Omega}(\vec{\beta}) d\vec{\beta}_r \\ &\propto \int \int_0^{\infty} \xi^{-3/2} \frac{1}{(2\pi\xi^{-1})^{p/2}} \exp\left\{-\frac{\xi}{2} |\vec{\beta}|^2\right\} 1_{\Omega}(\vec{\beta}) d\xi d\vec{\beta}_r \\ &\propto \int_0^{\infty} \xi^{-3/2} \frac{1}{(2\pi\xi^{-1})^{k/2}} \exp\left\{-\frac{\xi}{2} |\vec{\beta}_I|^2\right\} d\xi 1_{\Omega^*}(\vec{\beta}_I) \\ &\propto |\vec{\beta}_I|^{-(k-1)} 1_{\Omega^*}(\vec{\beta}_I) \end{aligned}$$

where Ω^* is the appropriate projection of Ω . Hence if one is seeking inferences for a subvector of $\vec{\beta}$, the observation stream from the Gibbs sampler for the full $\vec{\beta}$ can also be used for any subvector, with assurance that the answer will be consistent with the answer that would have been obtained from a Gibbs sampler with the subvector and use of $\pi(\vec{\beta})$ in that lower dimension.

5. GIBBS SAMPLING

The purpose of this section is to describe the algorithm used to construct the posterior distribution of the parameters of interest. A FORTRAN program was developed to execute a Gibbs sampler Monte Carlo algorithm. See Geman and Geman (1984), Gelfand and Smith (1990), Geyer (1991), and Tierney (1991) for descriptions of the background and theoretical properties of the Gibbs sampler methodology.

Initialization

In order to start the Gibbs sampling iterative algorithm, initial values must be set for the parameters. These initial values were established by using OLS (ordinary least squares) estimation of our model without the random effects terms, i.e. the model

$$\log Y_i = \beta_0 + \mathbf{X}_i \vec{\beta} + \varepsilon_i, \quad i = 1, \dots, n$$

The results of this OLS analysis yielded significant coefficients for all but one of the predictor variables. The achieved r^2 was 0.957 and the standard error of regression was 0.054.

The initial value for the vector of coefficients is the OLS estimate, denoted $\vec{\beta}^{(0)}$. For the initial value of σ^2 , we use the sample variance of the OLS residuals. For the initial values of the vehicle-line effects, C_i , we use the average OLS residuals over that car line:

$$C_i^{(0)} = \frac{\text{Sum of the residuals within car-line } l}{\text{Number of tests within car-line } l}$$

For μ_k and V_k the initial values are:

$$\mu_k^{(0)} = \text{Sample mean of the } C_l^{(0)}\text{'s, within manufacturer } k$$

$$V_k^{(0)} = \text{Sample variance of the } C_l^{(0)}\text{'s, within manufacturer } k$$

If there is only one car line within a particular manufacturer, then $V_k^{(0)}$ was set equal to the sample variance of all $C_l^{(0)}$'s. For $\lambda^{(0)}$ we used the sample variance of all the $\mu_k^{(0)}$'s and for $\tau^{(0)}$ we used the sample variance of all the $V_k^{(0)}$'s. The initial value of α was $\alpha^{(0)} = 1$.

Various other starting values were used to see if the Gibbs sampler is sensitive to the initial settings. Little sensitivity was found.

The Full Conditionals

The conditional posterior distributions used in the Gibbs sampler will be listed in the same order that they are simulated in the algorithm. The symbol $\pi(\cdot | \cdot)$ will be used generically to label these densities, with the parameters to the left of the bar indicating those parameters for which this is a density, and the variables to the right denoting the other variables (parameters or data) upon which the conditional distribution actually depends. (Note that these are all full conditionals, i.e. the conditioning is on all variables not listed on the left. However, we conventionally list on the right of the bar only those variables that are relevant.)

For notational convenience, define

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} C_{1n} \\ \vdots \\ C_{ln} \end{pmatrix}, \quad \mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

$$\tilde{\boldsymbol{\mu}} = (\mu_1, \dots, \mu_M), \quad \mathbf{V} = (V_1, \dots, V_M),$$

and assume that the data have been ordered by manufacturer and vehicle line with M_k being the first index in the data for a vehicle of manufacturer k , and L_j being the first index for vehicle line j . Let $M_0 = L_0 = 0$.

- (i) $\pi(\sigma^2 | \mathbf{Y}, \mathbf{X}, \beta_0, \vec{\beta}, \mathbf{C}) = IG(\frac{1}{2}(n-2), \frac{1}{2}\Sigma_{i=1}^n (Y_i - \beta_0 - \mathbf{X}_i \vec{\beta} - C_{in})^2)$.
(ii) $\pi(\xi | \vec{\beta}) = G(\frac{1}{2}(p-1), \frac{1}{2}|\vec{\beta}|^2)$. Here $G(\nu, \rho)$ refers to the Gamma distribution with density

$$f(\nu | \nu, \rho) = \frac{1}{\Gamma(\nu)} \rho^\nu \nu^{\nu-1} \exp\{-\nu\rho\}$$

- (iii) To avoid extreme inefficiency in the Gibbs sampler due to high correlation between β_0 and $\vec{\beta}$, we generate them jointly as follows: define

$$\vec{\beta} = \begin{pmatrix} \beta_0 \\ \vec{\beta} \end{pmatrix}, \quad \tilde{\mathbf{X}} = (\mathbf{1} \ \mathbf{X}), \quad \Sigma^* = \begin{pmatrix} 0 & 0 \\ 0 & \xi \mathbf{I}_{p \times p} \end{pmatrix} + \frac{1}{\sigma^2} \tilde{\mathbf{X}}' \tilde{\mathbf{X}}^{-1}$$

Then, defining Ω as in equation (6),

$$\pi(\vec{\beta} | \sigma^2, \xi, \mathbf{Y}, \mathbf{X}, \mathbf{C}) = \mathcal{N}\left(\frac{1}{\sigma^2} \Sigma^* \tilde{\mathbf{X}}' (\mathbf{Y} - \mathbf{C}), \Sigma^*\right) 1_\Omega(\vec{\beta})$$

Note that computing Σ^* can be organized efficiently by computing, in the initialization of the program, the orthogonal \mathbf{O} such that $\mathbf{O}' \tilde{\mathbf{X}}' \tilde{\mathbf{X}} \mathbf{O} = \text{diag}\{d_1, \dots, d_{24}\}$, and observing that

$$\Sigma^* = \mathbf{O} \mathbf{D}^* \mathbf{O}' + \frac{\xi}{(1 - \xi \mathbf{O}_{(1)} \mathbf{D}^* \mathbf{O}'_{(1)})} \mathbf{O} \mathbf{D}^* \mathbf{O}'_{(1)} \mathbf{O}_{(1)} \mathbf{D}^* \mathbf{O}'$$

where $\mathbf{O}_{(1)}$ is the first row of \mathbf{O} and

$$\mathbf{D}^* = \text{diag}\left\{\left(\frac{d_1}{\sigma^2} + \xi\right)^{-1}, \dots, \left(\frac{d_{24}}{\sigma^2} + \xi\right)^{-1}\right\}$$

This avoids matrix inversions in the iterative portion of the program. The easy way to account for $1_\Omega(\vec{\beta})$ is merely to repeatedly generate $\vec{\beta}$ until $\vec{\beta} \in \Omega$ is obtained, but this turned out to be quite expensive. Hence we used a truncated normal generator for the positive β_i , based on the inverse c.d.f. method.

- (iv) For each $k = 1, \dots, M$ and $N_{k-1} + 1 \leq l \leq N_k$,

$$\pi(C_l | \sigma^2, \mathbf{Y}, \mathbf{X}, \beta_0, \vec{\beta}, \mu_k, V_k)$$

$$= \mathcal{N}\left(\frac{\left[\sigma^2 \mu_k + V_k \sum_{i=L_{l-1}+1}^{L_l} (Y_i - X_i \vec{\beta} - \beta_0)\right]}{[(L_l - L_{l-1})V_k + \sigma^2]}, \frac{V_k \sigma^2}{[(L_l - L_{l-1})V_k + \sigma^2]}\right)$$

- (v) A convenient way to deal with the prior dependence among the μ_k in equation (4) is to represent μ_k via $\mu_k = \eta_k + \delta$, where the η_k are i.i.d $N(0, \lambda)$ and δ is $N(0, \gamma)$ (for all k). The conditionals of the μ_k , δ , λ , and γ that are convenient to use here (we do not condition δ on μ_k) are, for each $k = 1, \dots, M$,

$$\pi(\delta | \lambda, \mathbf{V}, \mathbf{C}, \gamma) = N\left(\frac{\gamma \sum_{k=1}^M \bar{C}_{(k)} / (\lambda + V_k^*)}{1 + \gamma \sum_{k=1}^M 1 / (\lambda + V_k^*)}, \frac{\gamma}{1 + \gamma \sum_{k=1}^M 1 / (\lambda + V_k^*)}\right)$$

$$\pi(\mu_k | \lambda, \mathbf{V}, \mathbf{C}, \delta) = N\left(\frac{\lambda \bar{C}_{(k)} + \delta V_k^*}{\lambda + V_k^*}, \frac{\lambda V_k^*}{\lambda + V_k^*}\right)$$

$$\pi(\lambda | \tilde{\mu}, \delta) = IG\left(\frac{1}{2}(M-2), \frac{1}{2} \sum_{k=1}^M (\mu_k - \delta)^2\right)$$

$$\pi(\gamma | \delta) = IG\left(\frac{1}{4}, \frac{1}{2} \delta^2\right)$$

where

$$\bar{C}_{(k)} = \frac{1}{(M_k - M_{k-1})} \sum_{l=M_{k-1}+1}^{M_k} C_l$$

$$V_k^* = \frac{V_k}{(M_k - M_{k-1})}$$

The rather unusual choice of $\gamma^{-3/4}$ as a non-informative prior for γ was made to ensure that $\pi(\gamma | \delta)$ is proper, as well as the overall posterior.

- (vi) $\pi(V_k | \alpha, \tau, \mathbf{C}, \mu_k) = IG\left(\alpha + \frac{1}{2}(M_k - M_{k-1}), \tau + \frac{1}{2} \sum_{l=M_{k-1}+1}^{M_k} (C_l - \mu_k)^2\right)$
- (vii) $\pi(\tau | \alpha, \mathbf{V}) = G\left(\alpha M + 1, \sum_{k=1}^M 1/V_k\right)$
- (viii) $\pi(\alpha | \tau, \mathbf{V}) \propto [\Gamma(\alpha)]^{-M} \tau^{\alpha M} (\prod_{k=1}^M V_k)^{-\alpha}$; recall we discretized α , with $\alpha \in \{0.75, 1.0, 1.5, 2.0, 2.25, 2.5, 2.75, 3.0, 3.5, 4.0\}$, so $\pi(\alpha | \tau, \mathbf{V})$ needs to be calculated only for these ten values of α , with α then being generated from the resultant discrete density.

Note: The gamma and inverse gamma random variables above were generated using the IMSL (version 10.0) routine DRNGAM, and the normal random variables using the IMSL routines RNNOF and DRNMVN. (See IMSL, 1989.)

Iteration and Burn-in

The Gibbs sampler begins with $\theta^{(0)}$, the initial vector of all parameters. One then passes through each of the nine conditional distributions listed above, generating new realizations of the parameters and resulting in a new vector $\theta^{(1)}$. For each conditional distribution, the values of the parameters used to the right of the bar are the *latest* generated values. This leads to a sequence $\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \dots$, of iterations of the parameters.

We rather arbitrarily chose the first 100 iterations of the Gibbs sampler to be 'burn-in', in

order hopefully to reach the stationary distribution of the Gibbs sampler. Probably, far fewer than 100 would suffice, but the cost of 100 iterations is negligible.

The Gibbs sampler was then run for an additional m iterations past burn-in. To investigate stationarity, numerous plots were made and numerous running averages were computed. None indicated any cause for alarm.

Estimation and Monte Carlo Variance

To estimate $E[g(\theta)]$, where $g(\theta)$ is some function of the parameters θ and E denotes posterior expectation, we use

$$\hat{E}[g] = \overline{g(\theta)} = \frac{1}{m} \sum_{i=1}^m g(\theta^{(i)})$$

The Monte Carlo variance of $\hat{E}[g]$ is (cf. Geyer, 1991)

$$\frac{1}{m} \sigma_g^2 = \frac{1}{m} \sum_{t=-\infty}^{\infty} \gamma_t \quad (8)$$

where $\gamma_t = \gamma_{-t} = \text{Covariance}(g(\theta^{(j)}), g(\theta^{(j+t)}))$, for the stationary distribution of the Gibbs sampler. The covariance of lag t can be estimated by

$$\hat{\gamma}_t = \hat{\gamma}_{-t} = \frac{1}{m} \sum_{j=1}^{m-t} (g(\theta^{(j)}) - \overline{g(\theta)})(g(\theta^{(j+t)}) - \overline{g(\theta)})$$

Because the estimator of $\hat{\gamma}_t$ is unstable for large values of t , one cannot directly plug this into equation (8). Instead, it is standard to use a down-weighted version, such as

$$\frac{1}{m} \hat{\sigma}_g^2 = \frac{1}{m} \left[\hat{\gamma}_0 + 2 \sum_{t>0} w(t) \hat{\gamma}_t \right]$$

We chose the weights (as recommended by Geyer, 1991) to be

$$w(t) = \begin{cases} 0 & \text{if } t \geq T \\ \cos(t/T) & \text{if } 1 \leq t \leq T \end{cases}$$

where T is the first value of t for which $\hat{\gamma}_t \leq 0$. The estimated standard deviations, $\sqrt{\hat{\sigma}_{\beta_i}^2}$, based on a run of size $m = 500$ were found for $i = 0, 1, \dots, 23$.

The Monte Carlo sample size, m , for the final run was determined by setting the $m^{-1} \hat{\sigma}_{\beta_i}^2$ equal to the desired accuracy, and solving for m . The accuracy we desired was 1/100 of the posterior variance of the β_i , so that the Monte Carlo variance would be negligible compared with the statistical variance. This procedure indicated that a sample size of $m \cong 1000$ was needed. We decided to use $m = 5000$ (after the 'burn-in' of 100 iterations), to produce the final numbers. This turned out to be a fortunate choice, as the $\hat{\sigma}_{\beta_i}^2$, determined above turned out to be as much as four times too small; by choosing $m = 5000$ we nevertheless achieved our goal of negligible Monte Carlo variance.

6. RESULTS

We will first discuss features of the posterior distribution of the β 's. Table I gives the estimated mean and standard deviation of the 24 β 's. For any particular coefficient, the mean and standard deviation are estimated by the sample average, $\hat{\beta}_i$, and the usual sample standard deviation, $\sqrt{s_{\beta_i}^2}$, of the $m = 5000$ replications of the Gibbs Sampler. The estimated means are

all compatible with engineering knowledge. Note that the standard deviations of these estimates are moderate.

The mean and standard deviation of the parameter, σ^2 , and the hyperparameters, ξ , λ , τ , α and δ are also given in Table I. From that table, the mean for σ^2 , 0.00146, is approximately half the OLS error variance, indicating that the vehicle-line random effects do pick up considerable residual variance. Recall that ξ was simply an artifactual parameter introduced for simplicity in analysis, so its moments have no interesting interpretation. The hyperparameters α , τ , and δ also are of little interpretational interest.

The variance of the manufacturer means, has posterior mean 0.00232, indicating that the standard error of the manufacturer effect is 4.8%. This shows that there is a substantial variability among manufacturers in their ability (or desire) to efficiently utilize technology for fuel economy. This is seen by looking at actual estimated manufacturing effects, the μ_k , given in Table II for 37 manufacturers. Note that these are residual means. The random effect C_l can be thought of as the residual fuel economy of vehicle line l due to factors other than the

Table I. Mean and standard deviation of parameters from posterior distribution (5000 replications)

Parameter	Mean	Std. dev.
Constant (β_0)	9.5607	0.1720
ETW (β_1)	-0.4527	0.0204
TCL (β_2)	0.0660	0.0062
DPA (β_3)	-0.1812	0.0116
N V (β_4)	-0.2911	0.0166
AXLE (β_5)	-0.0363	0.0158
CID (β_6)	-0.4108	0.0107
COMP (β_7)	0.2714	0.0316
HP CID (β_8)	-0.1578	0.0125
MPI (β_9)	0.0276	0.0056
TBI (β_{10})	0.0407	0.0057
OH2V (β_{11})	0.0098	0.0047
OH3V (β_{12})	0.0274	0.0070
OH4O (β_{13})	0.0263	0.0054
ACTD (β_{14})	0.0244	0.0081
ACTR (β_{15})	0.0260	0.0044
AUTO4 (β_{16})	-0.0323	0.0060
LKUP (β_{17})	0.0280	0.0042
ELEC (β_{18})	0.0024	0.0024
FWD (β_{19})	-0.0177	0.0073
MAN4 (β_{20})	0.0760	0.0066
MAN5 (β_{21})	0.0486	0.0052
CRP3 (β_{22})	0.0627	0.0060
CRP4 (β_{23})	0.0620	0.0117
σ^2	0.00146	0.000055
ξ	35.287	10.851
λ	0.00221	0.000747
τ	0.00285	0.000957
α	3.110	0.6948
δ	0.00235	0.00733

$\{X_i\}$. The mean, μ_k , of the C_l over all vehicles of manufacturer k can then be thought of as the residual fuel economy of the manufacturer due to other factors.

The correlation between the manufacturer means, as modelled by equation (4), is calculated by using the output stream from Gibbs sampling. The mean of the posterior correlation is 0.271 and the standard deviation is 0.368. The tenth, fiftieth, and ninetieth percentiles for the correlation are 0.000046, 0.049, and 0.974, respectively. This shows that there is a very large variance associated with the correlation and that most of the distribution is skewed towards zero.

It is tempting to ascribe these manufacturer effects to varying 'abilities' in utilization of technology to achieve fuel efficiency. However, a manufacturer could, instead, have been

Table II. Mean and standard deviation of μ (ordered on mean)

Manufacturer	Mean	Std. dev.
Suzuki	0.1193	0.0312
Honda	0.0624	0.0149
Toyota	0.0491	0.0134
Isuzu	0.0475	0.0322
JCI	0.0461	0.0315
Mercedes-Benz	0.0437	0.0202
Hyundai	0.0345	0.0301
Daihatsu (Truck)	0.0265	0.0346
Nissan (Truck)	0.0250	0.0180
Ford	0.0249	0.0114
Nissan	0.0247	0.0103
Mazda	0.0238	0.0172
Ford (Truck)	0.0199	0.0116
Suzuki (Truck)	0.0191	0.0278
General Motors (Truck)	0.0180	0.0116
Mitsubishi	0.0126	0.0124
General Motors	0.0092	0.0104
Mazda (Truck)	0.0081	0.0233
Nummi	0.0040	0.0281
Fuji	0.0022	0.0112
Alfa	0.0010	0.0342
Toyota (Truck)	0.0010	0.0221
Isuzu (Truck)	-0.0060	0.0198
AMC (Truck)	-0.0079	0.0179
Chrysler	-0.0114	0.0131
Fuji (Truck)	-0.0134	0.0167
Land Rover (Truck)	-0.0155	0.0354
AMC	-0.0161	0.0291
Peugeot	-0.0225	0.0174
Shelby (Truck)	-0.0256	0.0354
Mitsubishi (Truck)	-0.0284	0.0148
BMW	-0.0331	0.0174
VW (Truck)	-0.0504	0.0323
Chrysler (Truck)	-0.0529	0.0153
VW	-0.0596	0.0157
Saab	-0.0634	0.0250
Audi	-0.0701	0.0132

attempting to optimize with respect to performance, ride, or handling, not fuel efficiency. Since we do not have such performance variables in our database, interpretation of the cause of manufacturer effects requires care. Also, note that the standard errors of many of these means are comparatively large.

The manufacturer effects do, in any case, measure the residual manufacturer fuel efficiency. Thus Suzuki, for which the posterior mean of μ is 0.1193, achieves very high fuel efficiency relative to the technologies it uses. Of the three US firms, General Motors and Ford both have positive values for the mean of μ but the value for Chrysler is slightly negative; this holds also for the truck side. This runs contrary to the usual thinking that Chrysler does well on fuel economy because their fleet-wide average is usually higher than the other two; for example, in 1989 the CAFE ratings were: Chrysler: 28.0; GM: 27.2; and Ford: 26.6. However, when conditioned on vehicle characteristics and technology utilization, Chrysler has a somewhat worse manufacturer fuel economy. (This is statistically clear only for Chrysler Truck, because of the large standard errors.)

It is also worth investigating the residual fuel economy at the vehicle-line level. The 37 manufacturers had a total of 223 vehicle lines or models in the data set. The automobile side of General Motors had the most vehicle lines with 23. Table III gives a partial list of these vehicle lines and for each the Gibbs generated posterior mean and standard deviation of C_i are reported, where C_i is the vehicle-line random effect in the model.

Many of the 223 C_i 's have a large standard deviation, as demonstrated by the Caprice Wagon from General Motors with a standard deviation of 0.0302. Note that the Honda Civic CRX HF has one of the largest values of the mean of C_i out of the 223 models. This vehicle was intentionally manufactured for fuel economy, and in fact, the HF stands for high fuel.

The four vehicle lines from Suzuki all have exceptionally high values for C_i and therefore demonstrate that Suzuki, as a manufacturer, is concentrating on fuel economy. Also, the four vehicle lines from Mercedes-Benz all have reasonably large values for C_i , which indicates that it is not just the small vehicles that register large residual fuel economy. The value of 0.0699 for Jaguar is interesting because it is a vehicle that is perceived to be designed for performance, but yet on this residual fuel economy measurement it has a relatively large positive value. The smallest value of C_i for any automobile is given by the -0.0937 for the BMW M3; and likewise, the smallest for the trucks is -0.1143 for the Chrysler W250.

Table III. Mean and standard deviation of C_i (partial list)

Vehicle line	Manufacturer	Mean	Std. dev.
Metro XFI	Suzuki	0.2061	0.0283
Civic CRX HF	Honda	0.1425	0.0207
Metro	Suzuki	0.1337	0.0200
Metro LS <i>i</i>	Suzuki	0.1324	0.0186
Sprint	Suzuki	0.1232	0.0219
Caprice Wagon	General Motors	0.0742	0.0302
Jaguar	JCI	0.0699	0.0183
560.X	Mercedes-Benz	0.0600	0.0232
300.X	Mercedes-Benz	0.0489	0.0143
500.X	Mercedes-Benz	0.0425	0.0272
420.X	Mercedes-Benz	0.0416	0.0273
M3	BMW	-0.0937	0.0338
W250	Chrysler	-0.1143	0.0221

It is anticipated that the importance of fuel economy measurement will increase in the the immediate future and procedures that analyse the test data will similarly increase in prominence. The models used in these analyses should account for the random effects due to manufacturing and vehicle lines, as described in this paper.

APPENDIX: DATA SET VARIABLES

<i>MPG</i>	The combined city and highway miles per gallon, computed from the city and highway test cycle values by equation (1).
<i>ETW</i>	The equivalent test weight, which is the curb weight plus 300 lb for load. It is recorded in increments of 125 lb.
<i>TLC</i>	The tyre and chassis loss, which measures tyre rolling resistance and drive train friction. It is computed from equation (2).
<i>DPA</i>	The dynamometer setting, which is used on the fuel economy test equipment to compensate for aerodynamic drag losses.
<i>NV</i>	The ratio of engine revolutions to vehicle velocity in top gear.
<i>AXLE</i>	The final drive ratio on front-wheel vehicles and the axle ratio on rear-wheel vehicles.
<i>CID</i>	The cubic inch displacement of the engine.
<i>COMP</i>	The engine compression ratio.
<i>HPCID</i>	The engine horsepower divided by the cubic inch displacement, which is referred to as specific horsepower.
<i>IMPI</i>	An indicator of multi-point fuel injection, in which atomized fuel sprays are generated, one for each cylinder.
<i>ITBI</i>	An indicator for throttle-body injection, in which a metered single atomized spray is supplied to the engine.
<i>IOH2V</i>	An indicator for an overhead cam engine, with two valves per cylinder.
<i>IOH3V</i>	An indicator for an overhead cam engine, with three valves per cylinder.
<i>IOH4V</i>	An indicator for an overhead cam engine, with four valves per cylinder.
<i>IACTR</i>	An indicator for a roller cam follower as the mechanism for valve activation.

<i>IACFD</i>	An indicator for a direct acting cam follower as the mechanism for valve activation.
<i>IAUTO4</i>	An indicator for an automatic transmission with four forward gear ratios.
<i>ILKUP</i>	An indicator for a mechanical system that locks the transmission torque converter.
<i>IELEC</i>	An indicator for a transmission which has an electronic control for selecting the appropriate gear.
<i>IFWD</i>	An indicator for a vehicle with front-wheel drive.
<i>IMAN4</i>	An indicator for a manual transmission with four forward gears.
<i>IMAN5</i>	An indicator for a manual transmission with five forward gears.
<i>ICRP3</i>	An indicator for a transmission with a creeper gear and three other forward gears.
<i>ICRP4</i>	An indicator for a transmission with a creeper gear and four other forward gears.

ACKNOWLEDGEMENT

J. O. Berger acknowledges support under National Science Foundation Grant DMS-8923071.

REFERENCES

- Andrews, R. W., J. O. Berger and M. H. Smith (1993), 'Bayesian estimation of fuel economy potential due to technology improvements', in C. Gatsonis *et al.* (eds), *Case Studies In Bayesian Statistics*, Springer-Verlag, New York.
- Berger, J. O. (1985), *Bayesian Analysis and Statistical Decision Theory*, 2nd edn, Springer-Verlag, New York.
- Gelfand, A. E. and A. F. M. Smith (1990), 'Sampling-based approaches to calculating marginal densities', *Journal of the American Statistical Association*, **85**, 398–409.
- Geman, S. and D. Geman (1984), 'Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images', *IEEE Trans. Pattern Anal. Machine Intell.*, **6**, 721–741.
- Geyer, C. (1991), 'Markov chain Monte Carlo maximum likelihood. Technical Report', School of Statistics, University of Minnesota, Minneapolis.
- IMSL (1989), *STATLIBRARY: FORTRAN Subroutines for Statistical Analysis*, IMSL, Houston, Texas.
- Tierney, L. (1991), 'Markov chains for exploring posterior distribution', Technical Report No. 560, School of Statistics, University of Minnesota, Minneapolis.