BAYESIAN ESTIMATION OF FUEL ECONOMY POTENTIAL DUE TO TECHNOLOGY IMPROVEMENTS

by

Richard W. Andrews#

University of Michigan

and

James O. Berger#

Purdue University

and

Murray H. Smith#

University of Canterbury

#The authors acknowledge the financial support of the Ford Motor Company in preparing this report. Also, parts of the statistical methodology were developed under National Science Foundation Grants DMS-8702620, DMS-8717799, and DMS-8923071. We gratefully acknowledge the following individuals who assisted in the engineering assessment phase of this project: N.T. Barnes, B. Beeson, R.C. Belaire, W.C. Follmer, R.J. Genik, R.C. Heathfield, T.E. Kenney, D.L. Kulp, J.G. LaFond, J.D. Medrick, N.M. Novelli, G. Pietron, N.D. Postma, T.S. Shih, T.A. Smart, R.J. Tabaczynski, W.E. Taylor, J.J. Tripi, and especially the Ford project directors, J.J. Greenbaum and T.M. Watson. Thanks also are due a referee and the editors for helpful comments.

$\triangle$

1. INTRODUCTION

1.1 Background and Overview

In 1975 the Energy Policy and Conservation Act was enacted into law. As part of this legislation, methods were determined for setting a minimum corporate average fuel economy (CAFE) standard. This legislation separates each manufacturer's production into imports and domestics for both Automobiles and Light Trucks. Each combination must meet the CAFE standard. As an example, the standard is now 27.5 miles per gallon for Automobiles. The corporate average fuel economy is calculated by taking a weighted harmonic average, for each manufacturer, of the vehicles sold. If a manufacturer does not meet the minimum standards, they are required to pay a fine. Almost all the manufacturers have taken steps to increase the fuel economy of their fleets of vehicles. For example, the overall size of automobiles was reduced. In addition, technology improvements, such as replacing the old push-rod type of valve activators with modern overhead cam arrangements, have aided in increasing the miles produced per gallon of gas. These improvements and size reductions were instituted on some vehicles but not all.

Presently, the U.S. Congress is considering legislation which will increase the minimum CAFE value the manufacturers are required to meet or exceed. Some of the increases being considered are stated as percent increases over a specified base year. For example, one bill specifies that the CAFE minimum requirement for Automobiles be increased by 20% from the base year of 1987 before the model year of 1995. This means that if a manufacturer has a weighted average of 28.0 miles per gallon in 1987, they would be required to have a weighted average of 33.6 MPG or greater by 1995. These increases in CAFE are being proposed with the anticipation that they can be achieved primarily by the manufacturers instituting technology improvements in vehicles yet to be produced. The overall goal of our study was to determine whether or not these standards could be met through technology improvement.

There are numerous technology improvements that can be considered in an attempt to increase the CAFE value for a manufacturer. The amount of fuel efficiency gain from each technology improvement (TI) can be estimated. These individual estimates of fuel economy due to each TI can be aggregated across all technologies to give an estimate of the amount of gain in MPG feasible by employing these TI's.

The current legislation before Congress is based on *non-statistical* estimates of the available fuel efficiency gain from introduction of TI's. The primary source of these estimates is a series of studies done by a consulting firm, Energy and Environmental Analysis Incorporated (EEA), which utilized engineering studies, expert opinion, and one-off comparisons of vehicles – with and without certain technologies – to arrive at estimates.

There are numerous reasons why a statistical analysis of the problem is highly desirable. These include the following:

(i) There is substantial disagreement among engineers as to the gains available from TI's, with the manufacturers' engineers typically claiming that the estimates obtained by EEA are too large. Statistical analysis could settle the issues.

(ii) Studying TI's in isolation is dangerous, because numerous TI's will ultimately be combined on the same vehicle, and synergistic effects amongst the TI's could be large. It is virtually impossible to analyze the effect of TI's in groups without statistical analysis.

(iii) Even isolating the effect of a TI is difficult, because it is very rare to simply add a single TI to a vehicle. Usually, TI's are introduced as part of a major vehicle remodeling. Thus in trying to assess the benefits of individual TI's by non-statistical means, there is considerable danger of double counting or of assigning a benefit to a specific TI that is actually due to other changes.

(iv) Measurements of fuel efficiency have a substantial variability. The main measurement process that is relevant to the issue is that of the Environmental Protection Agency (EPA). The result of this measurement is the MPG stated on new vehicle sales stickers. The standard deviation of this process, for a given vehicle, is roughly 3%. In addition, there is vehicle-to-vehicle variation, for identically configured vehicles of a given model and year, of approximately 3%. Finally, if one attempts to compare "matched" vehicles of different makes (as is often done in the analysis of EEA), the model-to-model variation can be enormous. Dealing effectively with these sources of variation can best be done via statistics.

Fortunately, there exists a large data base concerning fuel efficiency, namely the EPA data for the years 1988, 1989, and 1990, which provides the fuel efficiencies of all vehicles sold in the United States for those years. This data base is fully described in Section 3.1; the key feature that makes it usable for our analysis is its inclusion of numerous descriptive variables of each of the vehicles tested, including presence of TI's. Statistical analysis of this data, and its utilization to determine the effects of TI's, was the statistical goal of our study.

The data analysis itself is discussed in Section 3 of the paper. Its key features are employment of a loglinear model, with hierarchically modelled random effects. Gibbs sampling was employed to carry out the computations.

In attempting to utilize the data analysis to determine the effects of TI's, it was realized that a major problem existed: adding a TI would typically change the vehicle performance in ways other than just fuel efficiency, invalidating direct estimation of fuel efficiency gain. Our solution was to incorporate auxiliary engineering assessments of the effects of TI's on other performance features. This is briefly described in Section 1.2, with Section 4 presenting the details.

Section 2 presents the actual results, showing that the CAFE standards being proposed by Congress indeed appear to be unattainable, without severe reduction in vehicle size or performance. Section 5 describes how these results were obtained, i.e., how the results of the data analysis were combined with the auxiliary engineering assessments of performance. Several interesting methodological problems were encountered here.

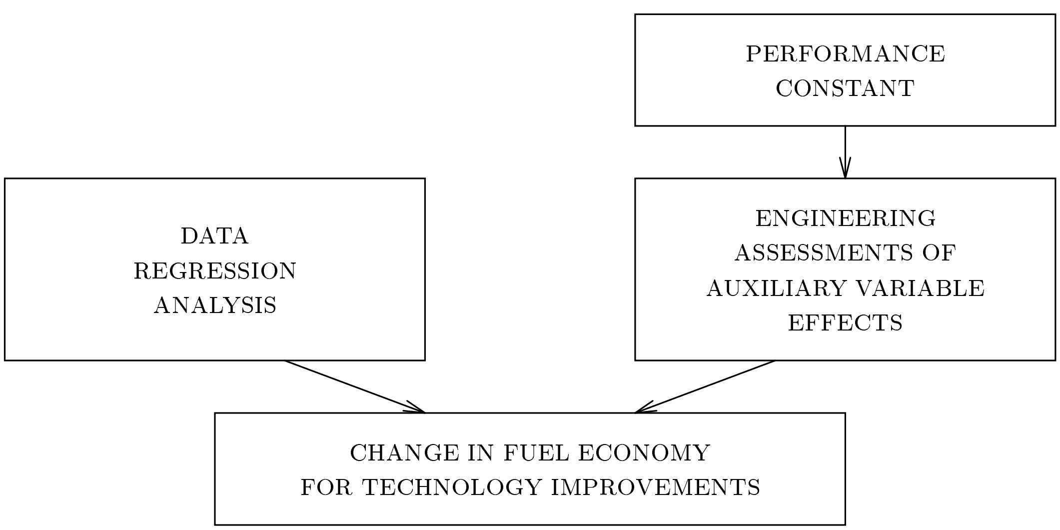
Sections 1.3 and 6 describe the roles that Bayesian analysis played in the study. The study could not have been done without substantial use of Bayesian reasoning.

1.2 The Two Pronged Estimation System

As mentioned in the previous section, our system for estimating the effect of a TI contains two major components: the statistical analysis of the EPA test data, and auxiliary engineering assessments of the effects of TI's on performance. The system is best described by the schematic given in Figure 1.1. In order to describe this system, using the schematic as a guide, the data analysis will be described first. Then the assessment procedure will be described and demonstrated. Finally, the method of combining the assessments with the data analysis to produce the estimates of fuel economy change will be introduced. In this section only a generic discussion of the system will be given, with an example to demonstrate the conceptual framework. The actual data, estimated coefficients, assessed engineering changes to auxiliary variables, and all numerical details are given in Sections 3, 4, and 5.

Figure 1.1

FUEL ECONOMY ANALYSIS SYSTEM



The Data Regression Analysis

The statistical model utilized in analyzing the data is the (random effects) regression model for a vehicle

$$\log\ \text{MPG}\ = \beta_0 + \sum X_i \beta_i + C_\ell + \varepsilon. \tag{1.1}$$

The dependent variable is the log of the MPG of a vehicle, while the independent variables, X_i, are of two types. The first type is that of (logs) of vehicle characteristic variables, such as weight and engine size. The second type is that of indicator variables representing the presence or absence of TI's, such as four speed automatic transmission or overhead cam engines. Here C_ℓ is a "vehicle line effect," and represents a random effect specific to a given vehicle line (e.g., Ford Taurus). The ε are assumed to be i.i.d. $N(0, \sigma^2)$, with σ^2 unknown. A complete description of this model, and justification thereof, is given in Section 3.2.

The C_ℓ are further modeled, using a hierarchical prior distribution, as i.i.d. normal random variables arising from a given manufacturer, with the overall manufacturer mean and variance also being modeled as i.i.d. normal random variables arising from the population of all manufacturers. (Because of the distinct natures of light trucks and automobiles, the Truck and Automobile divisions of a manufacturer are considered to be different manufacturers.) This is fully described in Section 3.3, where full specification of the prior is given.

Because of the hierarchical modeling and the unbalanced nature of the model (differing numbers of vehicles in each vehicle line), closed form statistical analysis is not possible. We utilized the Gibbs sampling Monte Carlo method of computing desired features of the posterior distribution of the unknown coefficients of the model. This is described fully in Sections 3.4 and 3.5.

Engineering Assessments

Soon after the project began, we realized that we were facing a major problem due to the issue of performance control. To see the problem, consider addition of the TI "4-speed automatic transmission" to a vehicle. In terms of the model (1.1), this would involve a change from 0 to 1 of the indicator variable, X_i, for that TI. However, addition of the TI would also result in other changes in the vehicle. Weight would change slightly, and the axle ratio and the N/V ratio (see Appendix B for definition) would typically change. Hence the corresponding variables in the model (1.1) would also need to be altered to determine the effect of adding the TI.

The scenario under which we are operating is the scenario of constant performance; the goal is to determine the effect of adding TI's, assuming that the performance of the vehicle is to remain constant. Thus, if a TI significantly alters performance, it is necessary to alter the other vehicle characteristics so as to maintain constant performance. Unfortunately, the data lacked enough measures of performance to enable this performance control to be done as part of the data analysis. It was thus necessary to subject each TI to an engineering assessment procedure.

This was done by assembling a group of automotive engineers and, for each TI, asking them the question "If this technology improvement is added to a vehicle, in order to keep performance approximately constant, what other characteristics of that vehicle will have to change and, on the average vehicle, by how much?" For Trucks, the "how much" of this question was requested as a three point distribution, by asking the engineers to state a lowest value, a highest value, and a most likely value. For Automobiles, a simpler two point scheme had been utilized. Correlations in the assessments were also requested. Details of this assessment scheme, and the problems encountered, are discussed in Section 4.

Constant performance for light trucks is defined to mean that the following four measures remain approximately the same: (1) start–up acceleration, (2) maximum gradability in 1^{st} gear, (3) general size (shadow volume) and (4) 3% grade speed in top gear (partial throttle). As an example of this subjective assessment procedure, consider the responses received for the TI of a 4-Speed Automatic transmission. The base for this TI is a 3-Speed Automatic transmission, meaning that the goal is to measure the change in fuel efficiency in going from 3-Speed Automatic to 4-Speed Automatic. In order to hold performance approximately constant when making this change, the engineering changes given in Table 1.1 would also have to be made.

For Automobiles, constant performance is defined to mean that the following three measures would remain the same: (1) start-up acceleration, (2) size of vehicle, and (3) gradability at 65 miles per hour. The performance control assessments for Automobiles thus typically differ from those for Truck.

Table 1.1

ENGINEERING CHANGES TO
ACCOMPANY AUTOMATIC 4–SPEED
TECHNOLOGY IMPROVEMENT IN LIGHT TRUCK

Auxiliary Variable	**Low Change**	**Likely Change**	**High Change**
Weight	20#	30#	40#
Axle (FDR)	+5%	+10%	+15%
N/V Ratio	-30%	-20%	-15%

Section 4 gives the results of the assessment procedure for all TI's considered in this report. It should be stressed that these assessments are assessments on characteristics that are related to the TI in such a way that performance is held constant. In that sense they are best described as indirect assessments. They are not direct assessments on the impact of the TI on fuel economy. At no time was the question asked, "If this TI is added, what will be the impact on fuel economy?" Our approach allows the data to answer that question.

Combining the Two Prongs

If a TI is added, there is a resulting change in the model (1.1); the indicator for that TI changes from 0 to 1, and vehicle characteristics may change (according to the engineering assessments) to maintain constant performance. Letting Δ generically denote "change" we are thus interested in estimating

$$\Delta \log \text{ MPG } = \sum (\Delta X_i)\beta_i + \Delta\varepsilon. \tag{1.2}$$

In estimating such quantities, note that we will have available the posterior distribution of the β_i. To proceed, we converted the engineering assessments concerning the ΔX_i into probability distributions. For Truck and the 3 point assessments, we utilized split normal distributions for the ΔX_i. For Automobiles and the 2 point assessments, we utilized normal distributions.

The problem thus becomes that of estimating (1.2), and associated measures of accuracy, when the β_i and the ΔX_i have specified probability distributions. Interestingly, simply using the posterior mean for such quantities is not necessarily wise, as will be discussed in Section 5. Indeed the estimation problem encountered is of substantial independent theoretical interest.

The issue facing Congress is how to set CAFE levels for the future. Of interest to Congress is thus, not the effects of individual TI's, but the aggregated total effect of all TI's. To assess this, assumptions must be made concerning the feasible adoption rates of TI's in vehicle fleets over the coming years. In our analysis, we simply utilize the "penetration rates" for TI's that have been proposed by EEA. These are given in Section 2, along with a discussion of the actual results of the analysis. Of interest here is that this aggregate analysis is simply another analysis of the form (1.2), where now the ΔX_i are the aggregate changes in the variables when averaged over all TI's weighted by their predicted penetration rates.

1.3 Roles for Bayesian Analysis

Bayesian analysis was important for three aspects of the analysis. First, the data analysis was done using a hierarchical Bayesian model. Use of a random effects model here is important for reasons discussed in Section 3.1, and the hierarchical Bayesian version thereof is an improvement over ordinary random effects models. Furthermore, because of a lack of balance in the data, and the fact that some of the parameters are known to lie in restricted parameter spaces, Bayesian analysis is virtually the only way to feasibly compute estimates and standard errors. (There is not enough data to place much faith in asymptotically motivated alternatives.)

Note that the answers we obtained from the Bayesian data analysis were substantially different from the answers obtained for ordinary least squares regression (see Appendix C). Also, the random effects model provides very interesting additional information. For instance, the C_ℓ can be interpreted as the residual fuel efficiency of a given vehicle-line, after accounting for vehicle characteristics and TI's. This is a measure of how well fuel efficiency has been optimized within the given vehicle-line. Similarly, the overall manufacturer means have a similar residual fuel efficiency interpretation for each manufacturer.

The second place in which Bayesian analysis was utilized was in the engineering assessments for performance control. Indeed, this was entirely a subjective Bayesian assessment procedure, with no data being utilized. It should be emphasized that analysis of the problem would have been impossible without this subjective Bayesian prong. (There had been earlier attempts to treat the problem from a pure regression viewpoint, but they had failed because of a lack of control for performance.)

The final place in which Bayesian analysis had a role was in the estimation of the quantities in (1.2). The difficulties in estimation of such quantities is, arguably, best addressed through Bayesian methods.

2. RESULTS

2.1 Fuel Efficiency Gains

In this section we present the results of our analysis, in the format we have used for presentation to non-statisticians. In applications, it is crucial to constantly keep in mind the ultimately desired goal, and so we begin with this desired "end product."

There is considerable interest in both the overall potential fuel economy gain through incorporation of TI's and the individual potential of each TI. Table 2.1 presents, for Automobiles and Light Trucks, the estimated fuel economy gain and its standard error for each TI. For example, the fuel economy gain for adding Lock-Up to a Truck is estimated to be 3.75%, with a standard error of 0.48%. Figure 2.1 provides a graphical presentation of the estimated fuel economy gains for the TI's common to Truck and Automobile, comparing the values of each. The intervals depict one standard error on each side of the point estimate. Note that there is substantial uncertainty about the fuel economy gains of many of the TI's. This uncertainty arose both because of the limits of the data and because certain of the engineering assessments relating to constant performance had a substantial degree of uncertainty.

To find the aggregate effect of all TI's, it is clearly necessary to determine the feasible increase in implementation of TI's, within the vehicle fleets. Tables 2.2 and 2.3 present such "penetration rates" of TI's. Legislation being considered typically gives 1987 as the "base" year, with CAFE standards to be set for years 1995 and 2001 in terms of percent increase from this base. Penetration rates are thus given in the tables for these years. For instance, Front Wheel Drive was used in 74% of the Automobiles in 1987, and could be implemented in 86% of the Automobiles by 1995 and 91% or 99% of the Automobiles by 2001. The two figures given for 2001 are projections under two different scenarios: the *Product Plan* scenario is supposedly an estimate of what manufacturers could readily do, while the *Max Technology* scenario presumes incorporation of TI's to the greatest extent possible, with no consideration of cost or efficiency.

The numbers in Tables 2.2 and 2.3 were produced by EEA, and have typically been used by all participants in the area. We followed this "tradition," even to the extent of presenting the Max Technology figures as our "totals" for 2001, since all others were doing so. In retrospect this was a strategic error, because when "final" numbers enter the public domain the assumptions under which they were developed are forgotten. Thus the Max Technology totals are reported as our best estimates of potential fuel economy gain by 2001, even though it is patently ludicrous to suppose that $10,000 of new technology will be placed on all vehicles.

Table 2.1		
PERCENT GAIN IN FUEL ECONOMY FOR INDIVIDUAL TECHNOLOGY IMPROVEMENTS		
TECHNOLOGY IMPROVEMENT	**ESTIMATE OF % GAIN* IN AUTOMOBILES**	**ESTIMATE OF % GAIN* IN LIGHT TRUCKS**
ENGINE		
Throttle–Body Fuel Injection for Carburetor	4.19% (0.57%)	4.19% (0.57%)
Multi–Point Fuel Injection for Carburetor	4.04% (0.78%)	5.08% (0.84%)
2–Valve Overhead Cam for Push–Rod	1.03% (0.65%)	2.04% (0.96%)
3–Valve Overhead Cam for Push–Rod		2.90% (0.99%)
4-Valve Overhead Cam for Push–Rod	4.28% (1.33%)	4.40% (1.70%)
Roller Cam Followers for Sliding Cam Followers	2.64% (0.45%)	2.64% (0.45%)
TRANSMISSION		
Electronic Transmission Control	0.24% (0.23%)	0.24% (0.23%)
Torque Converter Lock–Up from Open Torque Converter	2.83% (0.44%)	3.75% (0.48%)
4–Speed Automatic from 3–Speed Automatic	2.05% (2.23%)	3.20% (1.93%)
OTHER		
Weight Reduction (Materials, 2.5%)		1.90% (0.19%)
Weight Reduction (Materials, 5.0%)	3.96% (0.38%)	3.88% (0.39%)
Weight Reduction (Materials, 10.0%)	8.29% (0.83%)	
Aero Drag Reduction (5.0%)	1.34% (0.12%)	1.53% (0.17%)
Aero Drag Reduction (10.0%)	2.74% (0.24%)	3.14% (0.35%)
Aero Drag Reduction (20.0%)	5.75% (0.52%)	
Tire Resistance Reduction (5%)	0.38% (0.03%)	0.38% (0.03%)
Front Wheel Drive**	0.17% (0.79%)	

*Standard errors are in parentheses.
**If FWDII assessment is used (see Table 4.1) the figures would be 3.87% (1.04%).

Table 2.2				
PENETRATION RATES IN AUTOMOBILES FOR TECHNOLOGY IMPROVEMENTS				
TECHNOLOGY IMPROVEMENT	**1987**	**ProdPlan 1995**	**ProdPlan 2001**	**MaxTech 2001**
ENGINE				
Throttle–Body Fuel Injection	28%	40%	0%	0%
Multi–Point Fuel Injection	48%	60%	100%	100%
2–Valve Overhead Cam	24%	39%	39%	0%
4–Valve Overhead Cam	0%	30%	60%	100%
Roller Cam Followers	55%	95%	95%	95%
TRANSMISSION				
Electronic Transmission Control	0%	80%	80%	80%
Torque Converter Lock–Up	60%	80%	80%	80%
4–Speed Automatic	40%	80%	80%	80%
OTHER				
Weight Red.(5%)	0%	0%	80%	0%
Weight Red.(10%)	0%	0%	0%	80%
Drag Red.(10% over 1987)	20%	85%	85%	85%
Drag Red.(20% over 1987)	0%	15%	15%	15%
Drag Red.(5% over 1995)	0%	0%	80%	0%
Drag Red.(10% over 1995)	0%	0%	0%	80%
Tire Resistance Red. (5% over 1987)	0%	100%	100%	100%
Tire Resistance Red. (5% over 1995)	0%	0%	20%	100%
Front Wheel Drive	74%	86%	91%	99%

Table 2.3				
PENETRATION RATES IN LIGHT TRUCKS FOR TECHNOLOGY IMPROVEMENTS				
TECHNOLOGY IMPROVEMENT	**1987**	**ProdPlan 1995**	**ProdPlan 2001**	**MaxTech 2001**
ENGINE				
Throttle–Body Fuel Injection	40%	30%	0%	0%
Multi–Point Fuel Injection	40%	70%	100%	100%
2–Valve Overhead Cam	10%	20%	30%	25%
3–Valve Overhead Cam	0%	0%	15%	40%
4–Valve Overhead Cam	0%	0%	10%	20%
Roller Cam Followers	50%	100%	100%	100%
TRANSMISSION				
Electronic Transmission Control	0%	65%	75%	75%
Torque Converter Lock–Up	60%	75%	75%	75%
4–Speed Automatic	45%	65%	75%	75%
OTHER				
Weight Red.(2.5% over 1987)	0%	20%	20%	20%
Weight Red.(2.5% over 1995)	0%	0%	90%	10%
Weight Red.(5.0% over 1995)	0%	0%	10%	90%
Drag Red.(5.0% over 1987)	0%	50%	50%	50%
Drag Red.(10.0% over 1987)	0%	50%	50%	50%
Drag Red.(5.0% over 1995)	0%	0%	50%	10%
Drag Red.(10.0% over 1995)	0%	0%	40%	90%
Tire Resistance Red.	0%	100%	100%	100%

Even Product Plan figures are questionable, because they do not completely account for retooling; it is extremely expensive and takes several years to retool an assembly line for incorporation of many of the TI's. In an earlier report (Berger, Smith, and Andrews, 1991) we considered penetration rates more compatible with actual manufacturer constraints.

Tables 2.4 and 2.5 give the aggregated estimated fuel economy increases, over all the TI's we have considered, for the periods 1987-1995, 1995-2001, and 1987-2001, in the rows labeled "Regression Modelled Technologies." The standard errors of these estimates are given in parentheses. Note that the standard errors are quite small.

Listed in the footnotes of Tables 2.4 and 2.5 are a number of TI's that are relevant to the fuel economy question, but which could not be incorporated into our analysis (primarily because they were not present in the data base). We had little to say about these technologies, but felt compelled to include them in our final analysis so that our final numbers would be relevant to the legislative debate. The estimates we used for the effects of these other technologies was a weighted average of the various estimates that existed in the literature, with the weights determined by the extent of the agreement between the literature source and our estimates for the regression modelled technologies.

To place the grand totals in context, current legislation is proposing CAFE increases of 20% by 1995 and 40% by 2001. The total estimated fuel economy gains from Tables 2.4 and 2.5 for these periods are 10.59% and 28.76% (Automobile) and 8.33% and 21.66% (Truck), under the Max Technology scenario, and 10.59% and 19.76% (Automobile) and 8.33% and 16.09% (Truck) for the more realistic Product Plan scenario. Hence our conclusion is that the proposed CAFE increases cannot be met with current technology, without drastically decreasing the performance of the fleet. Such major changes to vehicles would have profound effects on safety, economics of the industry, and consumers.

2.2 Reactions to the Statistical Analysis

The issue of setting new CAFE standards is extremely volatile, with high stakes involved on both sides. There has thus been extensive discussion and criticism of all reports on the matter; this is especially true of our analyses (Berger, Smith, and Andrews, 1990, 1991), because they have been the only completely open studies with all assumptions public and with repeatable methodology.

The criticisms have been interesting, because of what they reveal about the failure to appreciate the value of statistics. We list, here, a few of these criticisms, with the obvious rejoinders.

Criticism 1: The effects of individual TI's are quite uncertain, and therefore the aggregated fuel economy estimate for legislation is also unreliable.

Rejoinder: Of course, averaging tends to reduce error, and the accumulation over TI effects is, in part, an averaging process, so that the aggregate totals have quite small standard errors. This basic statistical principle is not well understood. Even worse, the mere fact that a standard error is given seems to *reduce* the credibility of the number in the eyes of many. The notion that every stated estimate should be accompanied by an indication of its accuracy is not well accepted.

Criticism 2: If subsets of the data are analyzed, the answers can change dramatically.

Rejoinder: The data are, in no sense, homogeneous. Different subsets, such as Asian Automobiles or European Automobiles, tend to have very distinctive characteristics, and so do not typically span the space of independent variables (the Xs). Also, any such subset has a probably-too-limited sample size. So wide variation in analyses of subsets of the data is not contradictory. Note that one of the reasons for introducing a random effects model was to partly account for the lack of homogeneity.

Criticism 3: The statistical analysis measures the "average" effect of a TI, not the "best possible" effect, which is what legislation should address.

Rejoinder: The notion that one should legislate for a "best possible" technology implementation is troubling, but something of a side issue. The more telling rejoinder is to question the possibility of determining the "best possible" effect. For instance, if one looks at 10 vehicles in which a TI has been implemented, the maximum of the 10 measured TI effects will be substantial, due to the comparatively large measurement error, even if the TI has no effect whatsoever. In practice the situation is even worse, because TI's are virtually never implemented in isolation, and a "best implementation" might be simply due to a positive synergy with another TI or vehicle characteristic.

Criticism 4: The subjective engineering assessments of performance control greatly reduce the objectivity of the analysis.

Rejoinder: This is true, at least in terms of the popular conception of statistical objectivity (which we could see no value in criticizing). We stressed, therefore, the objectivity of the *method*, in that anyone with differing assessments is free to compute "their" answer. Indeed, there has been a lively debate between the interested parties concerning these assessments.

Criticism 5: There have been a number of "nonsense" criticisms. One such was the "observation" that our model was "biased" against fuel efficient vehicles, because the MPG predicted by the model for each vehicle tended to be lower than the actual vehicle MPG for vehicles with high fuel efficiency, and tended to be higher for vehicles with low fuel efficiency. (Failure to recognize a form of regression-to-the-mean would have been amusing, had it not been a criticism produced by a "statistical expert" hired by DOE to provide a critique of our analysis.)

Another set of criticisms arose out of complete misunderstandings of the meanings of correlation, synergy, interaction, additivity, double counting, multicolinearity, ... Because of continually facing this, we have included Appendix D in every report, as a partial non-technical explanation.

3. DATA ANALYSIS

3.1 The Data Set

The data used in the data analysis portion of the method is the EPA fuel economy and emission test data with some additional data obtained from automotive industry publications. The EPA data for automobiles and light duty trucks for the Model Years 1988, 1989 and 1990 is used.

Each record in the EPA data is one configuration of a particular make and model of vehicle. Some of the variables included in a record are:

- manufacturer and model name
- city and highway fuel economy (separate figures are reported)
- emissions data
- engine description variables:

  displacement, maximum horsepower, compression ratio, number of cylinders and valves per cylinder, etc.
- transmission and drive train variables:

  automatic/manual transmission, number of gears, axle ratio, top gear N/V, front/rear/4-wheel drive, etc.
- dynamometer settings:

  equivalent test weight, dynamic power adjustment.

The 1990 Model Year records included cam activation data and, since roller cam activation has a fuel economy benefit, the cam activation for 1988 and 1989 configurations were added to the original EPA data. Reduction in tire rolling resistance also has a fuel economy benefit. In order to estimate this benefit and for completeness of the model (see Section 3.2), the target coast down times for as many vehicle configurations as possible were obtained and added to the data.

The original data comprise 2264 automobile and 1509 light duty truck records for a total of 3773. Not all the records are usable in the analysis and 1019 automobile and 561 truck configurations were eliminated, resulting in 2193 records in the final data set. Data deficiencies and model constraints were the two broad reasons for eliminating records.

A number of records were discovered to simply be 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.

Coastdown times and cam activation data were not available for 534 configurations which were thus not retained.

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. Others which were of extreme configuration in terms of the model variables were also eliminated. Examples of eliminated vehicles are Rolls Royce, Lambourgini and Yugo. All vehicles with 12 cylinder engines, rotary engines and diesel engines were also eliminated. Police vehicles have large numbers of accessories, reduced fuel economy, and are not included in CAFE calculations and so were removed from the data set. All these vehicles (124 configurations), which have a tiny market share, could be described as exotic and their elimination improved and simplified the model.

Turbo and super-charged vehicles have extreme configurations (very low compression ratios and very high maximum horsepower) and are clearly performance oriented. A very small number of vehicles have 5-speed automatic or continuously variable transmissions (only 6 variants of the Subaru Justy in the latter). Because estimation of the fuel economy benefit would be very unreliable in these vehicles they were also eliminated.

The final data set comprising $n = 2193$ records still covers about 80% of the vehicles sold in the U.S.

3.2 The Model

Models relating fuel consumption over the EPA city and highway test cycles to vehicle characteristics have appeared in the engineering literature (see, for example, Sovran and Bohn, 1981, and An and Ross, 1991). These tend to be non-linear in some of the parameters and are models for individual vehicles. Because we are interested in an average proportional change both in MPG and in vehicle characteristics, the logarithmic model has proven to be highly satisfactory. In early comparative studies, using OLS estimation, the logarithmic model gave a fit superior to that of any of the other models considered.

The regression model chosen is, for the i^{th} record in the data set,

$$\log Y_i = \beta_0 + \boldsymbol{X}_i \boldsymbol{\beta} + C_{\ell_i} + \varepsilon_i, \tag{3.1}$$

where

Y_i is the combined MPG (over the City and Highway cycles);

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

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

C_{ℓ_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 Section 3.3. The covariates, definitions of which can be found in Appendix B, are

$$\boldsymbol{X}_i = (X_{i1}, \ldots, X_{ip}),$$

where $p = 23$. They are of two types:

$X_{i1}, \ldots, X_{i8},$ the logarithms of the 8 auxiliary variables, ETW, TCL, DPA, N/V, AXLE, CID, COMP and HP/CID;

$X_{i9}, \ldots, X_{i23},$ the indicators of the presence of various technologies and/or transmission types.

The regression model is used primarily to determine the sensitivity of MPG to changes in the vehicle characteristics when certain TI's are added. It is important to have as complete a set of predictor variables of fuel economy as possible. In our analysis we were constrained to use the variables recorded in the EPA data set. The lack of any variable related to tire rolling resistance and drive train friction was a serious shortcoming and it was decided to augment the EPA data with coastdown times. A variable which measures tire rolling resistance and drive train friction is tire and chassis loss (TCL) which is defined from the coastdown time, DPA, and ETW of the vehicle (Appendix B). The addition of this variable improves the model and enables estimates of fuel economy gain from reduction in tire rolling resistance to be made.

Thus there were 8 continuous variables (logarithm transformed) available as fuel economy predictor variables. The term auxiliary variables is used to describe these variables in relation to the engineering assessments.

A large number of indicator variables (15 in total) are used to indicate the presence of various technology improvements and manual/automatic transmissions. The corresponding β_i in the model can be thought of as representing the residual fuel economy effects from the TI's that are not explained by the auxiliary variables.

At first, we used models which did not include vehicle-line random effects, but these exhibited quite large (of the order of 0.4) serial correlations of the residuals. The introduction of the vehicle-line random effects overcame this problem and improved the model, as well as providing interesting information about vehicle-lines and manufacturers.

Interaction terms were not included in the model. See Appendix D for a discussion of the reasons that interactions were not used.

In earlier analyses of the data using OLS, separate models for city and highway MPG were used. As was expected, the city and highway coefficients of many of the variables were quite different. The city and highway estimates of fuel economy gain were then combined to give the overall gain for a technology improvement. In practice, the overall estimate obtained this way differed very little (usually in the third significant figure) from the estimate obtained using the coefficients of combined MPG as the dependent variable. Using combined fuel economy as the dependent variable also greatly simplified the error calculations for our estimates, since a complicated correlation structure would be necessary for a full Bayesian model of both city and highway MPG. Note that fuel economy legislation addresses only the combined MPG.

3.3 The Prior Distribution

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

$$C_\ell \overset{\text{i.i.d.}}{\sim} N(\mu_k, V_k), \qquad \ell = N_{k-1}+1, \ldots, N_k, \tag{3.2}$$

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, as

$$\begin{aligned} \mu_k &\overset{\text{i.i.d.}}{\sim} N(0, \lambda), \qquad k = 1, \ldots, M; \\ V_k &\overset{\text{i.i.d.}}{\sim} IG(\alpha, \tau), \qquad k = 1, \ldots, M, \end{aligned} \tag{3.3}$$

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

$$f(v|\alpha, \tau) = \frac{\tau^\alpha}{\Gamma(\alpha) v^{(\alpha+1)}} \exp\left\{-\frac{\tau}{v}\right\} \qquad \text{for } v > 0. \tag{3.4}$$

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

$$\pi(\beta_0, \sigma^2, \lambda, \tau) = 1/\sigma^2,$$

reflecting the usual noninformative prior for σ^2, and being constant in β_0, λ and τ; although λ and τ are scale parameters, they occur at higher levels of the hierarchical model and hence cannot be assigned $1/\lambda$ and $1/\tau$ as priors. 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 $\{0.75, 1.0, 1.5, 2.0, 2.25, 2.5, 2.75, 3.0, 3.5, 4.0\}$.

Two different prior distributions for the regression parameters, $\boldsymbol{\beta}$, were considered, namely

$$\pi_1(\boldsymbol{\beta}) = 1_\Omega(\boldsymbol{\beta}) \quad \text{and} \quad \pi_2(\boldsymbol{\beta}) = |\boldsymbol{\beta}|^{-(p-1)} 1_\Omega(\boldsymbol{\beta}); \tag{3.5}$$

here 1_Ω is the indicator function on the set

$$\Omega = \{\boldsymbol{\beta}\colon\ \beta_{10} > 0,\ \beta_{15} > 0,\ \beta_{18} > 0\}. \tag{3.6}$$

The decision was made at the beginning of the study to utilize a noninformative prior for $\boldsymbol{\beta}$. The reason was so that the analysis of the data could be sold as objective, since considerable controversy surrounds direct engineering assessments of fuel efficiency. The usual noninformative prior would be the constant $\pi_1^*(\boldsymbol{\beta}) = 1$, but $\pi_2^*(\boldsymbol{\beta}) = |\boldsymbol{\beta}|^{-(p-1)}$ has considerable justification. When used to estimate $\boldsymbol{\beta}$ in the absence of random effects, π_2^* results in a standard shrinkage estimator for $\boldsymbol{\beta}$, an estimator that can be justified from many perspectives (robust Bayes, minimaxity, adaptive ridge regression, ...). More importantly, use of $\pi_1^*(\boldsymbol{\beta}) = 1$ is suspect when it comes to estimation of variances of quantities such as (1.2), whereas π_2^* is quite satisfactory. These issues are more fully discussed in Section 5.4.

The presence of the indicator on Ω in (3.5) is because it is known, from engineering principles, that β_{10}, β_{15}, and β_{18} are positive. These coefficients correspond to the TI's TBI, ACTR, and ELEC, respectively. The reason these three TI's can be singled out for such treatment is that they are the only TI's whose presence has no effect on the auxiliary variables, and hence which are not directly confounded with the auxiliary variables.

At first sight, accommodation of π_2 within Gibbs sampling appears to be difficult, because the resulting posterior distribution of $\boldsymbol{\beta}$, given the other parameters, is not a nice distribution. This problem can be easily circumvented, however, by introducing an auxiliary parameter $\xi > 0$, and choosing the prior distributions

$$\begin{aligned} \pi(\boldsymbol{\beta}|\xi) &= N_p(\mathbf{0}, \xi^{-1}\boldsymbol{I}_p) I_\Omega(\boldsymbol{\beta}), \\ \pi_1(\xi) &= 1_{\{0\}}(\xi), \quad \pi_2(\xi) = \xi^{-3/2} I_{(0,\infty)}(\xi); \end{aligned} \tag{3.7}$$

use of $\pi_1(\xi)$ (a point mass at zero) yields the same posterior distribution as does use of $\pi_1(\boldsymbol{\beta}) = 1_\Omega(\boldsymbol{\beta})$, while use of $\pi_2(\xi)$ corresponds to the desired marginal prior, $\pi_2(\boldsymbol{\beta})$, because

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

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

A final relevant observation concerning $\pi_2(\boldsymbol{\beta})$ is that the marginal prior of some k-dimensional subvector of $\boldsymbol{\beta}$, call it $\boldsymbol{\beta}_I$ with $\boldsymbol{\beta}_{I^c}$ denoting the remaining coordinates, is

$$\begin{aligned}
\pi_2(\boldsymbol{\beta}_I) &= \int |\boldsymbol{\beta}|^{-(p-1)} 1_\Omega(\boldsymbol{\beta}) d\boldsymbol{\beta}_{I^c} \\
&\propto \int \int_0^\infty \xi^{-3/2} \frac{1}{(2\pi\xi^{-1})^{p/2}} \exp\left\{-\frac{\xi}{2}|\boldsymbol{\beta}|^2\right\} 1_\Omega(\boldsymbol{\beta}) d\xi d\boldsymbol{\beta}_{I^c} \\
&\propto \int_0^\infty \xi^{-3/2} \frac{1}{(2\pi\xi^{-1})^{k/2}} \exp\left\{-\frac{\xi}{2}|\boldsymbol{\beta}_I|^2\right\} d\xi 1_{\Omega^*}(\boldsymbol{\beta}_I) \\
&\propto |\boldsymbol{\beta}|^{-(k-1)} 1_{\Omega^*}(\boldsymbol{\beta}_I)
\end{aligned} \tag{3.8}$$

where Ω^* is the appropriate projection of Ω. Hence if one is seeking inferences for a subvector of $\boldsymbol{\beta}$ (as will be the case in our problem when a specific TI is being considered), the observation stream from the Gibbs sampler for the full $\boldsymbol{\beta}$ 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 π_2 in that lower dimension.

3.4 Computation

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 description 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

$$\log Y_i = \beta_0 + \boldsymbol{X}_i\boldsymbol{\beta} + \epsilon_i, \quad i = 1, \ldots, n.$$

The results of this OLS analysis are also of interest for comparative purposes, and are given in Appendix C.

The initial value for the vector of coefficients is the OLS estimate, to be denoted $\boldsymbol{\beta}^{(0)}$. For each $i, i = 1, \ldots, n$, the residuals from the OLS estimation are $e_i = Y_i - \beta_0 - \boldsymbol{X}_i\boldsymbol{\beta}^{(0)}$. For the initial value of σ^2, we use the sample variance of the residuals,

$$\sigma^{2(0)} = \frac{n \sum_{i=1}^{n} e_i^2 - (\sum_{i=1}^{n} e_i)^2}{n(n-1)}.$$

For the initial values of the vehicle-line effects, C_ℓ, we use the average OLS residuals over that car-line,

$$C_\ell^{(0)} = \frac{\text{Sum of } e_i\text{'s within car-line } i}{\text{\# of tests within car-line } i}.$$

For μ_k and V_k the initial values are:

$$\begin{aligned} \mu_k^{(0)} &= \text{ Sample mean of the } C_\ell^{(0)}\text{'s, within manufacturer } k. \\ V_k^{(0)} &= \text{ Sample variance of the } C_\ell^{(0)}\text{'s, within manufacturer } k. \end{aligned}$$

If there is only one car-line within a particular manufacturer, then $V_k^{(0)}$ was set equal to the sample variance of all $C_\ell^{(0)}$'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 of the bar indicating 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. We, however, conventionally list on the right of the bar only those variables that are relevant.)

For notational convenience, define

$$\boldsymbol{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \ \boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_n \end{pmatrix}, \ \boldsymbol{C} = \begin{pmatrix} C_{\ell_1} \\ \vdots \\ C_{\ell_n} \end{pmatrix}, \boldsymbol{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_M), \quad \boldsymbol{V} = (V_1, \ldots, V_m), \tag{3.9}$$

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) $\pi(\sigma^2|\boldsymbol{Y},\boldsymbol{X},\beta_0,\boldsymbol{\beta},\boldsymbol{C}) = IG(\frac{1}{2}(n-2),\ \frac{1}{2}\sum_{i=1}^{n}(Y_i - \beta_0 - \boldsymbol{X}_i\boldsymbol{\beta} - C_{\ell_i})^2)$.

(ii) $\pi(\xi|\boldsymbol{\beta}) = G(\frac{1}{2}(p-1),\ \frac{1}{2}|\boldsymbol{\beta}|^2)$, when the prior $\pi_2(\boldsymbol{\beta})$ is used; when $\pi_1(\boldsymbol{\beta})$ is used, ξ is set equal to zero. Here $G(\nu,\rho)$ refers to the Gamma distribution with density

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

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

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

Then, defining Ω as in (3.6),

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

Note that computing $\boldsymbol{\Sigma}^*$ can be organized efficiently by computing, in the initialization of the program, the orthogonal $\boldsymbol{O}$ such that $\boldsymbol{O}^t\tilde{\boldsymbol{X}}^t\tilde{\boldsymbol{X}}\boldsymbol{O} = \textbf{diag}\,\{d_1,\ldots,d_{24}\}$, and observing that

$$\boldsymbol{\Sigma}^* = \boldsymbol{O}^t\boldsymbol{D}^*\boldsymbol{O} + \frac{\xi}{(1-\xi\boldsymbol{O}_{(1)}\boldsymbol{D}^*\boldsymbol{O}_{(1)}^t)}\boldsymbol{O}^t\boldsymbol{D}^*\boldsymbol{O}_{(1)}^t\boldsymbol{O}_{(1)}\boldsymbol{D}^*\boldsymbol{O},$$

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

$$\boldsymbol{D}^* = \textbf{diag}\left\{\left(\frac{d_1}{\sigma^2}+\xi\right)^{-1},\ldots,\left(\frac{d_{24}}{\sigma^2}+\xi\right)^{-1}\right\}.$$

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

(iv) For each $k = 1,\ldots,M$ and $N_{k-1}+1 \le \ell \le N_k$,

$$\pi(C_\ell|\sigma^2,\boldsymbol{Y},\boldsymbol{X},\beta_0,\boldsymbol{\beta},\mu_k,V_k) = N\left(\frac{[\sigma^2\mu_k + V_k\sum_{i=L_{\ell-1}+1}^{L_\ell}(Y_i - X_i\boldsymbol{\beta} - \beta_0)]}{[(L_\ell - L_{\ell-1})V_k + \sigma^2]},\ \frac{V_k\sigma^2}{[(L_\ell - L_{\ell-1})V_k + \sigma^2]}\right).$$

(v) For each $k = 1,\ldots,M$,

$$\pi(\mu_k|\lambda,V_k,\boldsymbol{C}) = N\left(\frac{\lambda}{[(M_k - M_{k-1})\lambda + V_k]}\sum_{\ell=M_{k-1}+1}^{M_k}C_\ell,\ \frac{\lambda V_k}{[(M_k - M_{k-1})\lambda + V_k]}\right).$$

(vi) $\pi(V_k|\alpha,\tau,\boldsymbol{C},\mu_k) = IG(\alpha + \frac{1}{2}(M_k - M_{k-1}),\ \tau + \frac{1}{2}\sum_{\ell=M_{k-1}+1}^{M_k}(C_\ell - \mu_k)^2)$.

(vii) $\pi(\lambda|\boldsymbol{\mu}) = IG(\frac{1}{2}(M-2),\ \frac{1}{2}|\boldsymbol{\mu}|^2)$.

(viii) $\pi(\tau|\alpha,\boldsymbol{V}) = G(\alpha M + 1,\ \sum_{k=1}^{M} 1/V_k)$.

(ix) $\pi(\alpha|\tau,\boldsymbol{V}) \propto [\Gamma(\alpha)]^{-M}\tau^{\alpha M}\left(\prod_{k=1}^{M} V_k\right)^{-\alpha}$; 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 $\pi(\alpha|\tau,\boldsymbol{V})$ needs to be calculated only for these ten values of α, with α then being generated from the resultant discrete density. (For efficiency, compute the $[\Gamma(\alpha)]^{-M}$ at program initialization).

Note: The gamma and inverse gamma random variables above were generated using the IMSL (version 9.1) routine GGAMR, and the normal random variables using the IMSL routines GGNQF and GGNSM.

Iteration and Burn-In

The Gibbs sampler begins with $\boldsymbol{\theta}^{(0)}$, the initial vector of all parameters. One then passes through each of the nine conditional distributions listed in the previous section, generating new realizations of the parameters and resulting in a new vector $\boldsymbol{\theta}^{(1)}$. For each conditional distribution, the values of the parameters used to the right of the bar are the *latest* generated values. Thus $\sigma^{2(1)}$ is generated from (i), using $\boldsymbol{Y}, \boldsymbol{X}, \beta_0^{(0)}, \boldsymbol{\beta}^{(0)}$, and $\boldsymbol{C}^{(0)}$. Then $\xi^{(1)}$ is generated from (ii) using $\boldsymbol{\beta}^{(0)}$. Then $\tilde{\boldsymbol{\beta}}^{(1)}$ is generated from (iii) using $\boldsymbol{Y}, \boldsymbol{C}^{(0)}, \xi^{(1)}$, and $\sigma^{2(1)}$, etc. After finishing with $\boldsymbol{\theta}^{(1)}$, one starts over again with conditional (i), repeating the process. This leads to a sequence $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \boldsymbol{\theta}^{(3)}, \ldots$, of iterations of the parameters.

Figure 3.1

BATCH MEANS OF SIZE 100 FOR 2000 ITERATIONS OF THE GIBBS SAMPLER

We rather arbitrarily chose the first 100 iterations of the Gibbs sampler to be "burn-in," in order to hopefully 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. For illustration, Figure 3.1 is a graph of batch means of size 100 for the coefficient β_7 (COMP), based on a run of size $m = 2000$.

Estimation and Monte Carlo Variance

To estimate $E[g(\boldsymbol{\theta})]$, where $g(\boldsymbol{\theta})$ is some function of the parameters $\boldsymbol{\theta}$ and E stands for posterior expectation, we use

$$\hat{E}[g] = \overline{g(\boldsymbol{\theta})} = \frac{1}{m}\sum_{i=1}^{m} g(\boldsymbol{\theta}^{(i)}). \tag{3.10}$$

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

$$\frac{1}{m}\sigma_g^2 = \frac{1}{m}\sum_{t=-\infty}^{\infty} \gamma_t, \tag{3.11}$$

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

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

Because the estimator of $\hat{\gamma}_t$ is unstable for large values of t, one cannot directly plug this into (3.11). Instead, it is standard to use a down-weighted version, such as

$$\frac{1}{m}\hat{\sigma}_g^2 = \frac{1}{m}[\hat{\gamma}_0 + 2\sum_{t>0} w(t)\hat{\gamma}_t]. \tag{3.12}$$

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 $\hat{\gamma}_t \leq 0$. Table 3.1 provides, for the β_i, $i = 0, 1, \ldots, 23$, the estimated standard deviations, $\sqrt{\hat{\sigma}^2_{\beta_i}}$, based on a run of size $m = 500$.

The Monte-Carlo sample size, m, for the final run was determined by setting the $m^{-1}\hat{\sigma}^2_{\beta_i}$ 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. Comparing the $\sqrt{\hat{\sigma}^2_{\beta_i}}$ in Table 3.1 with the posterior standard deviations in Table 3.2, one can see that $m \cong 1000$ is needed. (Of course, the comparison was actually done for an earlier version of Table 3.2 based on the initial $m = 500$ Gibbs sampling run.) 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 $\hat{\sigma}^2_{\beta_i}$ determined above turned out to be as much as 4 times too small; by choosing $m = 5000$ we nevertheless achieved our goal of negligible Monte-Carlo variance.

Table 3.1

ESTIMATED MONTE-CARLO STANDARD DEVIATIONS $\times\sqrt{m}$
OF THE ESTIMATED POSTERIOR MEANS FOR THE β_i
($m = 500$ replications)

VARIABLE (β_i)	**M-C STD. DEV.** $\sqrt{\hat{\sigma}^2_{\beta_i}}$	**VARIABLE** (β_i)	**M-C STD. DEV.** $\sqrt{\hat{\sigma}^2_{\beta_i}}$
Constant (β_0)	0.467	OH3V (β_{12})	0.013
ETW (β_1)	0.036	OH4V (β_{13})	0.006
TCL (β_2)	0.011	ACTD (β_{14})	0.031
DPA (β_3)	0.032	ACTR (β_{15})	0.011
N/V (β_4)	0.033	AUTO4 (β_{16})	0.013
AXLE (β_5)	0.027	LKUP (β_{17})	0.004
CID (β_6)	0.011	ELEC (β_{18})	0.003
COMP(β_7)	0.031	FWD (β_{19})	0.069
HP/CID (β_8)	0.033	MAN4 (β_{20})	0.008
MPI (β_9)	0.015	MAN5 (β_{21})	0.008
TBI (β_{10})	0.022	CRP3 (β_{22})	0.006
OH2V (β_{11})	0.011	CRP4 (β_{23})	0.012

3.5 The Posterior Distribution

Means, variances, and correlations of parameters and hyperparameters

We will first discuss features of the posterior distribution of the β's. Tables 3.2 and 3.3 give the estimated mean and standard deviation of the 24 β's for the π_1-based posterior and π_2-based posterior, respectively (see (3.5)). For any particular coefficient, the mean and standard deviation are estimated by the sample average, $\overline{\beta}_i$, and the usual sample standard deviation, $\sqrt{s^2_{\beta_i}}$, 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 two priors seem to give very similar answers.

Table 3.2		
MEAN AND STANDARD DEVIATION OF PARAMETERS FROM POSTERIOR DISTRIBUTION (5000 replications, π_1-based posterior)		
PARAMETER	**MEAN**	**STD. DEV.**
Constant (β_0)	9.5814	0.1643
ETW (β_1)	-0.4554	0.0202
TCL (β_2)	0.0660	0.0062
DPA (β_3)	-0.1812	0.0118
N/V (β_4)	-0.2947	0.0165
AXLE (β_5)	-0.0350	0.0156
CID (β_6)	-0.4113	0.0108
COMP(β_7)	0.2811	0.0321
HP/CID (β_8)	-0.1600	0.0124
MPI (β_9)	0.0284	0.0055
TBI (β_{10})	0.0416	0.0056
OH2V (β_{11})	0.0095	0.0048
OH3V (β_{12})	0.0270	0.0067
OH4V (β_{13})	0.0262	0.0055
ACTD (β_{14})	0.0247	0.0079
ACTR (β_{15})	0.0264	0.0045
AUTO4 (β_{16})	-0.0335	0.0060
LKUP (β_{17})	0.0281	0.0042
ELEC (β_{18})	0.0024	0.0024
FWD (β_{19})	-0.0182	0.0075
MAN4 (β_{20})	0.0753	0.0066
MAN5 (β_{21})	0.0479	0.0052
CRP3 (β_{22})	0.0629	0.0060
CRP4 (β_{23})	0.0616	0.0118
σ^2	0.00146	0.000055
ξ		
λ	0.00223	0.000722
τ	0.00304	0.000973
α	3.278	0.6710

Table 3.3		
MEAN AND STANDARD DEVIATION OF PARAMETERS FROM POSTERIOR DISTRIBUTION (5000 replications, π_2-based posterior)		
PARAMETER	**MEAN**	**STD. DEV.**
Constant (β_0)	9.5473	0.1782
ETW (β_1)	−0.4502	0.0206
TCL (β_2)	0.0662	0.0062
DPA (β_3)	−0.1812	0.0119
N/V (β_4)	−0.2910	0.0166
AXLE (β_5)	−0.0372	0.0157
CID (β_6)	−0.4114	0.0107
COMP(β_7)	0.2722	0.0323
HP/CID (β_8)	−0.1585	0.0121
MPI (β_9)	0.0278	0.0054
TBI (β_{10})	0.0411	0.0055
OH2V (β_{11})	0.0097	0.0047
OH3V (β_{12})	0.0273	0.0066
OH4V (β_{13})	0.0263	0.0055
ACTD (β_{14})	0.0237	0.0077
ACTR (β_{15})	0.0260	0.0044
AUTO4 (β_{16})	−0.0323	0.0060
LKUP (β_{17})	0.0279	0.0042
ELEC (β_{18})	0.0024	0.0023
FWD (β_{19})	−0.0172	0.0074
MAN4 (β_{20})	0.0760	0.0066
MAN5 (β_{21})	0.0487	0.0053
CRP3 (β_{22})	0.0627	0.0060
CRP4 (β_{23})	0.0621	0.0118
σ^2	0.00146	0.000054
ξ	35.407	10.741
λ	0.00222	0.000712
τ	0.00279	0.000922
α	3.102	0.6744

Table 3.4

CORRELATION MATRIX FOR BETAS

	[β_{23}]	[β_{22}]	[β_{21}]	[β_{20}]	[β_{19}]	[β_{18}]	[β_{17}]	[β_{16}]	[β_{15}]	[β_{14}]	[β_{13}]	[β_{12}]	
β_0	1.000	[.1204]	[.3047]	[.2331]	[-.0027]	[-.0012]	[.1534]	[.2211]	[-.0120]	[-.0249]	[-.0087]	[.0159]	[β_{23}]
β_1	-.7386	1.000	[.1891]	[.1567]	[-.0110]	[-.0347]	[.2266]	[.0130]	[-.0271]	[.0071]	[.0192]	[-.0045]	[β_{22}]
β_2	-.2770	.1907	1.000	[.5191]	[.0162]	[.0533]	[.4699]	[.6497]	[-.0330]	[.0144]	[-.0096]	[-.0008]	[β_{21}]
β_3	.2220	-.4032	-.1977	1.000	[.0295]	[.0282]	[.3513]	[.2756]	[-.0230]	[-.0111]	[-.0208]	[.0300]	[β_{20}]
β_4	-.4401	.0591	-.0113	-.0260	1.000	[-.0189]	[-.0362]	[.0426]	[-.0756]	[.0122]	[.0231]	[.0614]	[β_{19}]
β_5	.2372	-.1909	.0879	-.1536	-.5622	1.000	[.0288]	[-.0551]	[-.0270]	[.0103]	[.0353]	[.0339]	[β_{18}]
β_6	.0247	-.5267	.0039	.0780	.2612	.1804	1.000	[-.2511]	[-.0114]	[.0195]	[.0038]	[-.0208]	[β_{17}]
β_7	-.4217	-.0261	-.0346	.0877	.0231	-.0190	.1011	1.000	[-.0223]	[.0145]	[-.0072]	[.0105]	[β_{16}]
β_8	.1325	-.1053	.0811	.1455	-.1840	.0040	.1938	-.0774	1.000	[.1635]	[-.0060]	[.0524]	[β_{15}]
β_9	.0951	-.1465	.0647	.0695	.0191	-.0118	-.0448	.0049	-.3473	1.000	[-.3589]	[.0070]	[β_{14}]
β_{10}	.0683	-.1239	.0384	.0276	.0018	-.0222	.1000	-.0361	-.1456	.7517	1.000	[.4566]	[β_{13}]
β_{11}	-.1372	.0781	-.0623	-.0248	-.0490	.0140	.1897	.0423	.0413	-.0146	.0658	1.000	[β_{12}]
β_{12}	-.1515	.1098	-.0109	-.0447	.0001	-.0394	.0460	.0752	-.0977	.1169	.1244	.5325	
β_{13}	-.0488	.0614	-.0392	-.0417	-.0195	-.0148	-.0104	.0086	-.1767	-.0404	-.0687	.4274	
β_{14}	.1081	-.0402	.0936	.0185	.0212	.0121	-.0102	-.2637	-.2273	.1385	.1648	.0465	
β_{15}	.0299	.0418	.0871	.0010	.0197	.0337	-.1614	-.1173	.0090	-.0115	-.1688	-.0377	
β_{16}	-.3222	.0472	-.0594	.0446	.7762	-.5768	.1325	.0404	-.1592	.0050	.0293	-.0262	
β_{17}	-.0404	.0156	.0374	-.0345	.0939	-.0214	.0382	-.0645	.0149	-.0505	-.0347	.0130	
β_{18}	-.0417	.0151	.0238	.0022	.0618	.0229	.0476	-.0322	-.0267	-.0030	.0092	.0243	
β_{19}	-.1979	.0657	.1189	.2161	-.0111	.0334	.1993	.0143	.0269	-.0854	-.0335	.0674	
β_{20}	-.2252	.1277	-.0460	.0207	.2921	-.2350	.0734	.0017	.0578	-.0633	-.0339	.0395	
β_{21}	-.3461	.1223	-.0527	.0057	.7082	-.5449	.1074	-.0079	-.1401	-.0280	.0023	-.0125	
β_{22}	.1151	-.0833	-.0581	.0553	-.0828	-.0116	-.0099	-.0198	.0271	-.0358	-.0102	.0173	
β_{23}	-.1151	-.0040	.0125	-.0239	.2544	-.1928	.1049	.0417	.0338	-.0190	-.0113	.0062	
	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	

NOTE: Bracketed correlations are between bracketed betas.

The mean and standard deviation of the parameter, σ^2, and the hyperparameters, ξ, λ, τ, and α are also given in Tables 3.2 and 3.3. From Table 3.3, the mean for σ^2, 0.001455, is half the OLS error variance of .002886 (see Appendix C), indicating that the vehicle-line random effects do pick up considerable residual variance. The hyperparameter λ, which is the variance of the manufacturer means, has posterior mean (from Table 3.3) 0.00222, indicating that the standard error of the manufacturer effect is 4.7%. Recall that ξ was simply an artifactual parameter introduced for simplicity in analysis, so its moments have no interesting interpretation. The hyperparameters α and τ also are of little interpretational interest.

Of considerable interest is the posterior distribution of the manufacturer effects, the μ_k. The random effect C_ℓ can be thought of as the residual fuel economy of vehicle-line ℓ due to factors other than the $\{X_i\}$. 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.

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 attempting to optimize with respect to performance, ride, or handling, not fuel efficiency. Since we do not have such performance variables in our data-base, interpretation of the cause of manufacturer effects requires care. Also, the standard errors of many of these means happen to be comparatively large. Because of these ambiguities, the rather sensitive nature of the numbers, and the fact that they are tangential to the purpose of the study, we will discuss manufacturer effects in a different forum.

Because our fuel economy analysis is based on the β's, it is important to estimate, and account for, correlation among the β's. Using the Gibbs sampler, the posterior covariance matrix (24×24) was estimated and utilized in the subsequent analysis (see Section 5). Table 3.4 gives the correlation matrix for the π_2-based posterior.

Marginal posterior distributions for β's

In Section 5, we will utilize a normal approximation to the posterior distribution of $\boldsymbol{\beta}$. To judge whether or not this approximation is reasonable, it is useful to plot marginal posterior densities for coordinates of $\boldsymbol{\beta}$. To do this we follow the usual technique (cf. Gelfand and Smith, 1990) of evaluating the marginal posterior density for $\boldsymbol{\beta}_I$ (some subset of the β's) at a grid of points $\{\boldsymbol{\beta}_{I,j}, j = 1, \ldots, J\}$ via

$$\pi(\boldsymbol{\beta}_{I,j}|\boldsymbol{Y}, \boldsymbol{X}) = \frac{1}{m}\sum_{i=1}^{m} \pi(\boldsymbol{\beta}_{I,j}|\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{C}^{(i)}),$$

where the superscripts in the variables to the right of the bar refer to the values of those variables at the i^{th} iteration of the Gibbs sampler. The posteriors on the r.h.s. are all easily determined from

(iii) in Section 3.4.

As an example, consider the univariate parameter, β_{17}, which is the coefficient of LKUP. The range of 500 generated values for β_{17} was from .0138 to .0399; so, $J = 101$ values were chosen equidistant over the interval $[.01, .04]$. The resulting estimated posterior marginal distribution of β_{17}, for the π_2-based posterior, is then plotted in Figure 3.2. As can be seen from Figure 3.2, the posterior assumption of normality is not contraindicated.

The bivariate posterior marginal distributions of (β_1, β_{17}) and (β_5, β_{16}) are plotted in Figures 3.3 and 3.4, respectively. The (β_1, β_{17}) pair, which are the coefficients of ETW and LKUP were chosen to demonstrate the case in which there is almost no correlation. The computed correlation is 0.0052. In contrast, the (β_5, β_{16}) pair, which are the coefficients of AXLE and AUTO4 were chosen because of their high negative correlation, -0.5768. The plots accurately demonstrate these correlations and further substantiate the assumption of normality of the posterior β's.

Figure 3.2

MARGINAL POSTERIOR DENSITY FOR β_{17}, THE COEFFICIENT OF LKUP.

Figure 3.3

BIVARIATE POSTERIOR DENSITY AND CONTOURS FOR β_1 (ETW) AND β_{17} (LKUP).

Figure 3.4

BIVARIATE POSTERIOR DENSITY AND CONTOURS FOR β_5 (AXLE) AND β_{16} (AUTO4).

4. SUBJECTIVE ENGINEERING ASSESSMENTS

In Section 1.2, the engineering assessment prong of our approach to fuel economy estimation was introduced. The engineering assessments were needed, not as direct estimates of fuel economy, but in order to ascertain the adjustments that would be needed in a vehicle, when a technology improvement is added, in order to maintain constant vehicle size and performance. Adjustments were allowed in any of the eight vehicle characteristic variables, ETW, TCL, DPA, N/V, AXLE, CID, COMP, and HP/CID. (See Appendix B for definitions.) The details of the assessment scheme are discussed in this section.

4.1 Modified Delphi Approach (Automobiles)

Specific Assessments Requested

For each technology improvement we requested that the engineers making the assessment consider what vehicle characteristics will be changed by the addition of this TI, given that performance and size would remain approximately constant. In addition, to stating which of the 8 vehicle characteristics are affected, they were asked to also judge by how much the characteristics would have to change in order to maintain size and performance.

The how much question requires that the assessors be considering the proper unit for assessment.

We explained to the engineers that they should consider a typical or average vehicle on which this TI may be implemented. We did not ask the engineers to consider the whole spectrum of vehicles which may obtain this TI. On this typical or average vehicle we requested that each affected characteristic be assessed with two numbers, representing the high change and the low change. The high and low changes were to be the 95^{th} and 5^{th} percentiles of the distribution of change for a characteristic due to that TI.

The changes were generally expressed as percentages, although the engineers preferred to express small weight changes in pounds and the change to compression ratios as a difference of ratios. For these two variables, assessments were converted to proportions by dividing by fleet averages (weight: 3000 pounds for automobiles and 4000 pounds for trucks; compression ratio: 9.1 for automobiles and 8.9 for trucks).

Correlation between assessed changes in the logarithms of the characteristics was ascertained, in effect, as zero or highly correlated. Certain characteristics would typically be assessed to change more or less proportionally, and, for such, a correlation of between 0.7 and 1.0 (in absolute value) would be assumed. Otherwise, the correlation would be assumed to be 0. This limited assessment of correlation was, in part, due to our lack of emphasis and, in part, due to the limited time and interest of the engineers.

General Assessment Procedure

The DELPHI METHOD (Dalkey, 1972) is a procedure for assessing group judgments and refining these judgments so as to obtain an accurate statement of the opinion of the whole group. The defining components of the DELPHI METHOD are:

- Independent initial assessments – each individual is given the opportunity to express in writing his/her initial assessment with respect to the judgment being rendered.
- All assessments are revealed – the initial and subsequent assessments are revealed but the the individuals attached to those assessments remain anonymous.
- Reasoning attached to assessments – the reasoning that led to each assessment is noted with that assessment but no reference to the individual assessor is allowed.
- Summary statistics – a summary of the assessments, using mean, median, range, and other simple measures is presented to all assessors.
- Assessments are iterated – each individual is given the opportunity to adjust his/her assessment and anonymously present these new assessments.
- Replication – there are replications of the procedure until some summary of the responses is considered a reflection of the group opinion.

Our procedure for obtaining the engineering assessments for each TI was similar to the DELPHI METHOD and had many of the same components but we were not able to keep the identity of the assessor separate from his/her assessment. The two reasons we were not able to do this were (1) the individuals were reluctant to give one–on–one or private assessments, especially if they had to sign–off on their judgments and (2) the engineers were so busy that they did not have time to prepare assessments in isolation. Therefore, we convened two major meetings at which the automobile assessments were obtained. The assembled engineers at these two meetings represented all the major vehicle components such as engine, power train, and transmission. In total we had 17 FORD engineers present at the first meeting and 12 of those 17 present at the second meeting.

As each TI was considered, the experts representing that area were asked to make the assessments and provide reasons for their values. Then all others present were allowed to present their judgments and the reasons for their opinions. This procedure, of course, violated one of the major tenants of the DELPHI METHOD, in that, direct confrontation was allowed. However, the meetings were controlled in a free give and take manner that made the resulting assessments a

Table 4.1
ASSESSED CHANGES IN VARIABLES DUE TO TI's (Automobiles)
(ENGINE)
Sliding to Roller Cam Followers Indicator Variable: **IACTR** changes from 0 to 1. No changes to auxiliary variables.
Carburetor to Throttle–Body Fuel Injection Indicator Variable: **ITBI** changes from 0 to 1. No changes to auxiliary variables.
Carburetor to Multi–Point Fuel Injection Indicator Variable: **IMPI** changes from 0 to 1. Auxiliary Variable: **CID** LOW: −8% HIGH: −4% Auxiliary Variable: **HP/CID** LOW: +9% HIGH: +18% Auxiliary Variable: **COMP** LOW: +0.2 HIGH: +0.4 Correlation (Δ **LCID**, Δ **LHPCID**) = −0.7 Corr (Δ **LCID**, Δ **LCOMP**) = −0.9
2–Valve Pushrod to 2–Valve Overhead Cam Indicator Variable: **IOH2V** changes from 0 to 1. Auxiliary Variable: **CID** LOW: −14% HIGH: −8% Auxiliary Variable: **HP/CID** LOW: +17% HIGH: +30% Auxiliary Variable: **AXLE** LOW: +3% HIGH: +6% Auxiliary Variable: **N/V** LOW: +3% HIGH: +6% Corr(Δ**LCID**, Δ**LHPCID**) = −0.9 Corr(Δ**LCID**, Δ**LAXLE**) = −0.8 Corr(Δ**LCID**, Δ**LN/V**) = −0.8 Corr(Δ**LHPCID**, Δ**LAXLE**) = +0.8 Corr(Δ**LHPCID**, Δ**LN/V**) = +0.8 Corr(Δ**LAXLE**, Δ**LN/V**) = +1.0
2–Valve Pushrod to 4–Valve Overhead Cam Indicator Variable: **IOH4V** changes from 0 to 1. Auxiliary Variable: **CID** LOW: −30% HIGH: −20% Auxiliary Variable: **HP/CID** LOW: +45% HIGH: +75% Auxiliary Variable: **AXLE** LOW: +8% HIGH: +15% Auxiliary Variable: **N/V** LOW: +8% HIGH: +15% Auxiliary Variable: **Weight** LOW: 0# HIGH: +40# Auxiliary Variable: **COMP** LOW: 0.0 HIGH: +0.6 Corr(Δ**LCID**, Δ**LHPCID**) = −0.9 Corr(Δ**LCID**, Δ**LAXLE**) = −0.8 Corr(Δ**LCID**, Δ**LN/V**) = −0.8 Corr(Δ**LHPCID**, Δ**LAXLE**) = +0.8 Corr(Δ**LHPCID**, Δ**LN/V**) = +0.8 Corr(Δ**LAXLE**, Δ**LN/V**) = +1.0
(TRANSMISSION)
Electronic Control Indicator Variable: **IELEC** changes from 0 to 1. No changes to auxiliary variables.

Table 4.1 (Continued)
Open Torque Converter to Lock–Up Torque Converter Indicator Variable: **ILKUP** changes from 0 to 1. No changes to auxiliary variables.
3–Speed Automatic to 4–Speed Automatic Indicator Variable: **IAUTO4** changes from 0 to 1. Auxiliary Variable: **CID** LOW: −10% HIGH: −5% Auxiliary Variable: **AXLE** LOW: +2% HIGH: +5% Auxiliary Variable: **N/V** LOW: −15% HIGH: 0% Auxiliary Variable: **Weight** LOW: +10# HIGH: +20# Correlation(Δ**LCID**, Δ**LAXLE**) = −1.0 Corr(Δ**LCID**, Δ**LN/V**) = +0.7 Corr(Δ**LAXLE**, Δ**LN/V**) = −0.8
(OTHER)
Rear Wheel Drive to Front Wheel Drive I Indicator Variable: **IFWD** changes from 0 to 1. Auxiliary Variable: **Weight** LOW: −80# HIGH: −50# Auxiliary Variable: **CID** LOW: −2.7% HIGH: −1.7% Correlation(Δ**LETW**, Δ**LCID**) = +0.9
Rear Wheel Drive to Front Wheel Drive II Indicator Variable: **IFWD** changes from 0 to 1. Auxiliary Variable: **Weight** LOW: −220# HIGH: −180# Auxiliary Variable: **CID** LOW: −7.3% HIGH: −4% Correlation(Δ**LETW**, Δ**LCID**) = +0.9
Weight Reduction (Materials) 5% Auxiliary Variable: **Weight** LOW: −5% HIGH: −5% Auxiliary Variable: **CID** LOW: −5% HIGH: −2.5%
Weight Reduction (Materials) 10% Auxiliary Variable: **Weight** LOW: −10% HIGH: −10% Auxiliary Variable: **CID** LOW: −10% HIGH: −5%
Aerodynamic Drag Reduction (5%) Auxiliary Variable: **DPA** LOW: −4% HIGH: −4% Auxiliary Variable: **N/V** LOW: −2.5% HIGH: −1.5%
Aerodynamic Drag Reduction (10%) Auxiliary Variable: **DPA** LOW: −8% HIGH: −8% Auxiliary Variable: **N/V** LOW: −5% HIGH: −3%
Aerodynamic Drag Reduction (20%) Auxiliary Variable: **DPA** LOW: −16% HIGH: −16% Auxiliary Variable: **N/V** LOW: −10% HIGH: −6%
Tire Rolling Resistance Reduction (5%) Auxiliary Variable: **TCL** LOW: +5.3% HIGH: +5.3% Auxiliary Variable: **DPA** LOW: −0.3% HIGH: −0.1%

product of those opinions that were justified by some engineering principle, previous study, or observed data. Some of the requested assessments were not decided at the first meeting because addition data or information was needed. In general, there was not an attitude at either meeting that any one individual should have a greater influence. The larger influence was directed to those opinions that were supported by reason or data.

At the conclusion of the first meeting, the FORD project director summarized the engineering assessments and the reasons for the assessments and provided this information to all participants. All participants were requested to comment on the assessments and to return a signed statement if they agreed. There was very little feedback and only 2 of the individuals returned signed statements. The engineers seemed to be reluctant to make a signed statement concerning specific engineering assessments. However, in a meeting environment they were very willing to provide feedback and opinion. Therefore, a second meeting was held to provide feedback and further agreement. The output of this meeting was an agreed upon set of assessments.

Only FORD engineers participated in these two original assessment meetings. However, the assessments were subsequently sent to the other two U.S. automobile manufac- t turers; they provided feedback on some of the assessments and some of this feedback was incorporated. In addition, as soon as a draft report of our project was prepared it was sent to a wide range of engineering practitioners, government employees interested in fuel economy, and engineering consultants to the government. We received numerous comments concerning the assessments and, in consultation with the FORD project engineer, we adjusted some of the assessments in light of the new information. These assessments continue to be dynamic in that, as additional relevant engineering information becomes available, the assessments will be adjusted if appropriate. This type of global feedback is necessary when the model has global implications, such as for fuel economy legislation.

The assessments that resulted from this process are given in Table 4.1. Explanation and justification for the individual assessments are given in Appendix E.

4.2 Attempted Formal Approach (Trucks)

The truck project on fuel economy was started approximately nine months after the start of the automobile project. It was decided that the assessment stage could be formalized by using forms instead of a plenary meeting. An assessment form was developed in which every two facing pages represented a different TI. Appendix F is an example of the two pages for the 2–Valve Overhead Cam TI. An instruction sheet accompanied the form.

On the first sheet, the individual engineer was requested to provide his/her assessments on the

Table 4.2
ASSESSED CHANGES IN VARIABLES DUE TO TI's (Truck)
(ENGINE)
Carburetor to Throttle–Body Fuel Injection Indicator Variable: **ITBI** changes from 0 to 1. No changes to auxiliary variables.
Carburetor to Multi–Point Fuel Injection Indicator Variable: **IMPI** changes from 0 to 1. Auxiliary Variable: **Weight** LOW = +3#, LIKELY = +5#, HIGH = +6# Auxiliary Variable: **HP/CID** LOW = +6%, LIKELY = +8%, HIGH = +10% Auxiliary Variable: **COMP** LOW = +0.2, LIKELY = +0.3, HIGH = +0.4 Auxiliary Variable: **CID** LOW = −8%, LIKELY = −6%, HIGH = −4% Correlation(**ΔLHPCID**, **ΔLCOMP**) = +0.8 Corr (**ΔLHPCID**, **ΔLCID**) = −0.8 Corr (**ΔLCOMP**, **ΔLCID**) = −0.8
2–Valve Pushrod to 2–Valve Overhead Cam Indicator Variable: **IOH2V** changes from 0 to 1. Auxiliary Variable: **HP/CID** LOW = +6%, LIKELY = +13%, HIGH = +19% Auxiliary Variable: **CID** LOW = −15%, LIKELY = −10%, HIGH = −5% Auxiliary Variable: **AXLE** LOW = +3%, LIKELY = +4.5%, HIGH = +6% Auxiliary Variable: **N/V** LOW = +3%, LIKELY = +4.5%, HIGH = +6% Corr (**ΔLHPCID**, **ΔLCID**) = −0.9 Corr (**ΔLHPCID**, **ΔLAXLE**) = +0.8 Corr (**ΔLHPCID**, **ΔLN/V**) = +0.8 Corr (**ΔLCID**, **ΔLAXLE**) = −0.8 Corr (**ΔLCID**, **ΔLN/V**) = −0.8 Corr (**ΔLAXLE**, **ΔLN/V**) = +1.0
2–Valve Pushrod to 3–Valve Overhead Cam Indicator Variable: **IOH3V** changes from 0 to 1. Auxiliary Variable: **HP/CID** LOW = +10%, LIKELY = +20%, HIGH = +30% Auxiliary Variable: **CID** LOW = −20%, LIKELY = −15%, HIGH = −10% Auxiliary Variable: **AXLE** LOW = +9%, LIKELY = +12%, HIGH = +15% Auxiliary Variable: **N/V** LOW = +9%, LIKELY = +12%, HIGH = +15% Corr (**ΔLHPCID**, **ΔLCID**) = −0.9 Corr (**ΔLHPCID**, **ΔLAXLE**) = +0.8 Corr (**ΔLHPCID**, **ΔLN/V**) = +0.8 Corr (**ΔLCID**, **ΔLAXLE**) = −0.8 Corr (**ΔLCID**, **ΔLN/V**) = −0.8 Corr (**ΔLAXLE**, **ΔLN/V**) = +1.0
2–Valve Pushrod to 4–Valve Overhead Cam Indicator Variable: **IOH4V** changes from 0 to 1. Auxiliary Variable: **HP/CID** LOW = +13%, LIKELY = +25%, HIGH = +37% Auxiliary Variable: **CID** LOW = −25%, LIKELY = −20%, HIGH = −15% Auxiliary Variable: **AXLE** LOW = +11%, LIKELY = +15%, HIGH = +20% Auxiliary Variable: **N/V** LOW = +11%, LIKELY = +15%, HIGH = +20% Auxiliary Variable: **COMP** LOW = 0.0, LIKELY = +0.3, HIGH = +0.5 Auxiliary Variable: **Weight** LOW = 0#, LIKELY = +20#, HIGH = +40# Corr (**ΔLHPCID**, **ΔLCID**) = −0.9 Corr (**ΔLHPCID**, **ΔLAXLE**) = +0.8 Corr (**ΔLHPCID**, **ΔLN/V**) = +0.8 Corr (**ΔLHPCID**, **ΔLCOMP**) = −0.8 Corr (**ΔLCID**, **ΔLAXLE**) = −0.8 Corr (**ΔLCID**, **ΔLN/V**) = −0.8 Corr (**ΔLCID**, **ΔLCOMP**) = −0.7 Corr (**ΔLAXLE**, **ΔLN/V**) = +1.0 Corr (**ΔLAXLE**, **ΔLCOMP**) = +0.7 Corr (**ΔLN/V**, **ΔLCOMP**) = +0.7

Table 4.2 (Continued)
Sliding to Roller Cam Followers Indicator Variable: **IACTR** changes from 0 to 1. No changes to auxiliary variables.
(TRANSMISSION)
Electronic Control Indicator Variable: **IELEC** changes from 0 to 1. No changes to auxiliary variables.
Open Torque Converter to Lock–Up Torque Converter Indicator Variable: **ILKUP** changes from 0 to 1. Auxiliary Variable: **AXLE** LOW = −4%, LIKELY = −3%, HIGH =−2%
3–Speed Automatic to 4–Speed Automatic Indicator Variable: **IAUTO4** changes from 0 to 1. Auxiliary Variable: **Weight** LOW = +20#, LIKELY = +30#, HIGH = +40# Auxiliary Variable: **AXLE** LOW = +5%, LIKELY = +10%, HIGH = +15% Auxiliary Variable: **N/V** LOW = −30%, LIKELY = −20%, HIGH = −15% Correlation(Δ**LAXLE**, Δ**LN/V**) = −0.9
(OTHER)
Weight Reduction (2.5%) Auxiliary Variable: **Weight** LOW = −2.5%, LIKELY = −2.5%, HIGH = −2.5% Auxiliary Variable: **CID** LOW = −2.5%, LIKELY = −1.675%, HIGH = −1.25%
Weight Reduction (5%) Auxiliary Variable: **Weight** LOW = −5%, LIKELY = −5%, HIGH = −5% Auxiliary Variable: **CID** LOW = −5%, LIKELY = −3.35%, HIGH = −2.5%
Aerodynamic Drag Reduction (5%) Auxiliary Variable: **DPA** LOW = −4%, LIKELY = −4%, HIGH = −4% Auxiliary Variable: **N/V** LOW = −3.5%, LIKELY = −2.5%, HIGH = −2%
Aerodynamic Drag Reduction (10%) Auxiliary Variable: **DPA** LOW = −8%, LIKELY = −8%, HIGH = −8% Auxiliary Variable: **N/V** LOW = −7%, LIKELY = −5%, HIGH = −4%
Tire Rolling Resistance Reduction (5%) Auxiliary Variable: **TCL** LOW = +5.3%, LIKELY = +5.3%, HIGH = +5.3% Auxiliary Variable: **DPA** LOW = −0.3%, LIKELY = −0.2%, HIGH = −0.1%

8 auxiliary variables. Instead of a two point assessment, as was used for automobiles, we requested a three point assessment, asking for *Lower Bound*, *Most Likely*, and *Upper Bound*. By requiring an additional point in the assessment we were able to allow for asymmetry. The request for the additional point did not cause problems in the assessment procedure.

On the second sheet, it was requested that correlation between the characteristic changes be provided, on a four point scale. This request, even though accompanied by a detailed explanation, proved to be too complex, at least when based on just written instructions.

Overall, this formal procedure of eliciting assessments was unsuccessful. Out of 17 forms distributed, only 3 were returned and one of those obviously did not understand the directions. As can be seen from Appendix F, there were liberal places for individual comments, and these sections were used by all three respondants; their comments were more valuable than their specific assessments. It is our belief that this formal method is still preferable to an initial assessment group meeting; however, a training session would be necessary to teach proper use of the form. It is not clear that, in a busy industrial setting, such deliberate and time–consuming activities will be supported. Furthermore, the engineers seemed reluctant to give their opinions in writing, but are very willing to express themselves fully in a meeting.

After this failed attempt to obtain assessments with a form, we reverted back to the Modified Delphi procedure used for automobiles. Several of the individuals that participated in the automobiles assessments also participated in the truck assessments; and furthermore many of the same shops or unit groups of the corporation were represented. Therefore, for some of the truck assessments, the difficult reasoning and discussion had already been accomplished. Also, since the truck assessments came after the automobile assessments, the assessors had longer to think about this process of expressing their opinion on these engineering questions and, often, either found supporting data or executed company simulation programs to improve their judgments. If these subsequent sources of information were also relevant to the automobile assessments, then the automobile assessments were adjusted accordingly.

The assessments that resulted from this process are given in Table 4.2. Explanation and justification for the individual assessments are given in Appendix E.

4.3 Conversion to a Prior Distribution

Conversion of the uncertainty about the mean change in the auxiliary variables to a prior distribution was carried out by first considering the upper and lower assessment points to be a three standard deviation range, in both 2 and 3-point assessments. (Recall that the engineers were told to provide a 90% range; converting this to 3 standard deviations is slightly conservative.) The "likely point" of a 3-point assessment was assumed to be the median of the prior. For a given TI, any uncertainty about changes in the auxiliary variables can be highly correlated and a correlation matrix was assessed for each TI from general discussions with the engineers. There is, of course, no uncertainty about any change to indicator variables.

Because of the possible asymmetry, we used a multivariate split normal distribution for the logarithm of the proportionate changes in the auxiliary variables, when modelling the three-point assessments. (Indeed, when transformed logarithmically, *none* of the nonconstant three point assessments are symmetric.) Consider a technology improvement and let $M_i, i = 1, 2, \ldots, p$, denote the mean proportionate change in the i^{th} variable; these are the quantities that we asked the engineers to assess. (Here we exclude the constant since no change occurs to it when a TI is added.) Using a 3-point assessment scheme, let $(\ell_i, m_i, u_i), i = 1, 2, \ldots, 8$, denote the assessment of M_i for the auxiliary variables, with $\boldsymbol{R}$ being the assessed correlation matrix. Let $m_i,\ i = 9, 10, \ldots, p$, denote the changes (if any) in the indicator variables. Define $\boldsymbol{X}^* = (X_1^*, X_2^*, \ldots, X_p^*)$ by:

$$X_i^* = \begin{cases} \log(1+M_i), & i = 1, 2, \ldots, 8, \\ m_i, & i = 9, 10, \ldots, p, \end{cases} \tag{4.1}$$

and write $\boldsymbol{X}^* = (\boldsymbol{X}_A^*, \boldsymbol{X}_I^*)$, where $\boldsymbol{X}_I^*$ is fixed and $\boldsymbol{X}_A^*$ is assigned the multivariate split normal prior distribution with parameters – median $\boldsymbol{\nu}$, left and right scale parameters $\boldsymbol{\sigma}^-, \boldsymbol{\sigma}^+$ and correlation matrix $\boldsymbol{R}$. Formally,

$$\begin{aligned} \boldsymbol{\nu} &= (\nu_1, \nu_2, \ldots, \nu_8) \quad \text{where} \quad \nu_i = \log(1+m_i), \\ \boldsymbol{\sigma}^- &= (\sigma_1^-, \sigma_2^-, \ldots, \sigma_8^-) \quad \text{where} \quad \sigma_i^- = (\log(1+m_i) - \log(1+\ell_i))/1.5, \\ \boldsymbol{\sigma}^+ &= (\sigma_1^+, \sigma_2^+, \ldots, \sigma_8^+) \quad \text{where} \quad \sigma_i^+ = (\log(1+u_i) - \log(1+m_i))/1.5, \end{aligned} \tag{4.2}$$

the 1.5 divisor coming from the fact that the 90% range was treated as three standard deviations.

Let $\mathbb{R}^+ = (0, \infty)$ and $\mathbb{R}^- = (-\infty, 0)$ and define

$$\mathbb{R}^{\boldsymbol{\alpha}} = \prod_{i=1}^{8} \mathbb{R}^{\alpha_i},$$

which is an 8-dimensional orthant indexed by $\boldsymbol{\alpha}$, any 8-vector of $+$'s and $-$'s. The 8-dimensional split normal prior distribution for $\boldsymbol{X}_A^* - \boldsymbol{\nu}$ is then given on each orthant, $\mathbb{R}^{\boldsymbol{\alpha}}$, by

$$\boldsymbol{X}_A^* - \boldsymbol{\nu} \sim N_8(\boldsymbol{0}, \boldsymbol{D}^{\boldsymbol{\alpha}} \boldsymbol{R} \boldsymbol{D}^{\boldsymbol{\alpha}}), \tag{4.3}$$

where

$$\boldsymbol{D}^{\boldsymbol{\alpha}} = \textbf{diag}\ \{\boldsymbol{\sigma}^{\boldsymbol{\alpha}}\} \quad \text{and} \quad \boldsymbol{\sigma}^{\boldsymbol{\alpha}} = (\sigma_1^{\alpha_1}, \sigma_2^{\alpha_2}, \ldots, \sigma_8^{\alpha_8}). \tag{4.4}$$

$\boldsymbol{R}$ is not the actual correlation matrix of $\boldsymbol{X}_A^*$, which will instead be some complicated function of all the parameters, but the terminology is retained as it represents the correlation structure within each orthant. It is easy to check that all marginal distributions of any dimension are multivariate split normal and that, in particular, X_i^* is split normal with median ν_i and left and right standard deviations σ_i^- and σ_i^+, respectively.

For the 2-point assessments $(\ell_i, u_i), i = 1, 2, \ldots, 8$, the prior distribution is

$$\boldsymbol{X}_A^* \sim N_8(\boldsymbol{\eta}_A, \boldsymbol{DRD}), \tag{4.5}$$

where $\boldsymbol{R}$ is the assessed correlation matrix, the components of $\boldsymbol{\eta}_A$ are

$$\eta_i = (\log(1 + \ell_i) + \log(1 + u_i))/2,$$

and $\boldsymbol{D} = \ \mathbf{diag}\ \{\boldsymbol{\sigma}\}$, where the components of $\boldsymbol{\sigma}$ are

$$\sigma_i = (\log(1 + u_i) - \log(1 + \ell_i))/3.$$

5. COMBINING DATA AND PERFORMANCE CONTROL

5.1 Estimation of Technology Improvement Benefits

The model described in Section 3.2 leads to a simple form for the mean proportionate change in miles per gallon for a TI. Recall that the model is, for a vehicle "0",

$$\log Y_0 = \beta_0 + \boldsymbol{X}_0\boldsymbol{\beta} + C_{\ell_0},$$

and hence the logarithm of the proportionate change in fuel economy for adding a TI is $\boldsymbol{X}^*\boldsymbol{\beta}$, where $\boldsymbol{X}^*$ is the mean change in the independent variables (log transformed for the 8 auxiliary variables), as described in Section 4.3. Thus we take the proportionate fuel economy gain from adding the TI to be

$$\gamma = \exp(\boldsymbol{X}^*\boldsymbol{\beta}), \tag{5.1}$$

with the net fuel economy benefit being $\gamma - 1$.

To estimate γ under uncertainty concerning $\boldsymbol{X}^*$ and $\boldsymbol{\beta}$, two estimates were considered. The first is

$$\hat{\gamma}_1 = E_1[\exp\{-\boldsymbol{X}^*\boldsymbol{\beta}\}]/E_1[\exp\{-2\boldsymbol{X}^*\boldsymbol{\beta}\}], \tag{5.2}$$

where E_1 stands for expectation with respect to the (independent) prior distribution of $\boldsymbol{X}^*$ and the π_1-based posterior distribution of $\boldsymbol{\beta}$ (i.e., the posterior assuming $\pi_1(\boldsymbol{\beta})$ to be the prior). The second estimate is

$$\hat{\gamma}_2 = E_2[\exp\{\boldsymbol{X}^*\boldsymbol{\beta}\}], \tag{5.3}$$

where E_2 stands for expectation with respect to the prior distribution of $\boldsymbol{X}^*$ and the π_2-based posterior distribution of $\boldsymbol{\beta}$. The motivation for these estimates is indicated in Section 5.4.

The above expectations cannot be evaluated in closed form. They can, however, be easily evaluated in a Monte-Carlo fashion, using simulated $\boldsymbol{X}^*$ from the split normal distribution and the values of $\boldsymbol{\beta}$ from the Gibbs sampler. One of our goals, however, is to produce a method of estimation that can be given in a simple closed form, so that continuing refinements of the $\boldsymbol{X}^*$ assessments (or alternative assessments by others) can easily be incorporated.

The first idea we tried was to use the linear approximation

$$\gamma - 1 = \exp\{\boldsymbol{X}^*\boldsymbol{\beta}\} - 1 \simeq \boldsymbol{X}^*\boldsymbol{\beta}, \tag{5.4}$$

which is reasonable because of the typically small magnitudes of the X_i^* and β_i. Indeed, this approximation was used in our earlier reports. We have found, however, that the approximation is somewhat inadequate in light of the greater resolving power of our current model.

Our second attempt was to estimate the exact γ, but using normal approximations to the distributions of $\boldsymbol{X}^*$ and $\boldsymbol{\beta}$. This turned out to be extremely successful, giving answers virtually indistinguishable from the Monte-Carlo estimates, and hence is the method we adopt.

Normal Approximations to Posteriors

In Section 3.5, the posterior mean and posterior covariance matrix for $\boldsymbol{\beta}$ were determined, under either π_1 or π_2. Denoting the mean by $\hat{\boldsymbol{\beta}}$ and the covariance matrix by $\boldsymbol{\Sigma}$, the normal approximation to the posterior distribution of $\boldsymbol{\beta}$ is simply

$$\boldsymbol{\beta} \sim N_p(\hat{\boldsymbol{\beta}}, \boldsymbol{\Sigma}). \tag{5.5}$$

Likewise, we will approximate the split-normal distribution of $\boldsymbol{X}^*$ (for the 3-point truck assessments) by

$$\boldsymbol{X}^* \sim N_p(\boldsymbol{\eta}, \boldsymbol{\Lambda}), \tag{5.6}$$

where $\boldsymbol{\eta}$ and $\boldsymbol{\Lambda}$ are given below.

To determine $\boldsymbol{\eta}$, write $\boldsymbol{\eta} = (\boldsymbol{\eta}_A, \boldsymbol{\eta}_I)$, corresponding to the means for the auxiliary and indicator variables. The components of $\boldsymbol{\eta}_I$ are all zero, except for the indicator of the TI (if any) being considered; the corresponding η_i is then 1. For the i^{th} characteristic variable in $\boldsymbol{\eta}_A$, the mean of the split-normal distribution in Section 4.3 can be computed to be

$$\eta_i = \nu_i + \frac{1}{\sqrt{2\pi}}(\sigma_i^+ - \sigma_i^-), i = 1, 2, \ldots, 8.$$

The covariance matrix, $\boldsymbol{\Lambda}$, is clearly all zeroes, except for the 8×8 upper left corner corresponding to $\boldsymbol{X}_A^*$. For this upper left corner, we use

$$\boldsymbol{\Lambda}_A = \boldsymbol{DRD},$$

where $\boldsymbol{R}$ is the assessed correlation matrix and $\boldsymbol{D} = \ \mathbf{diag}\ \{\sigma_1, \ldots, \sigma_p\}$, with

$$\sigma_i = \left[\frac{1}{2}((\sigma_i^+)^2 + (\sigma_i^-)^2) + \frac{1}{(2\pi)}(\sigma_i^+ - \sigma_i^-)^2\right]^{1/2}.$$

The σ_i^2 are the variances of the split-normal distribution in Section 4.3.

For the 2-point automobile assessments no approximation is necessary as the prior distribution for $\boldsymbol{X}_A^*$ is modelled by the multivariate normal. The mean vector and covariance matrix are given in (4.5).

It might appear to be superfluous to have used a split normal assessment scheme, and then use a normal approximation. From the viewpoint of accurate elicitation, however, we would argue for this approach; elicitors have considerable difficulty directly assessing means and variances in asymmetric situations.

The Closed-Form Estimates

The normal approximations yield closed-form expressions because of the following lemma.

Lemma. Suppose $\boldsymbol{X}^*$ and $\boldsymbol{\beta}$ are independent multivariate $N_p(\boldsymbol{\eta}, \boldsymbol{\Lambda})$ and $N_p(\hat{\boldsymbol{\beta}}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is non-singular. Then, provided

$$c^2 < \text{ (largest eigenvalue of } \boldsymbol{\Sigma}\boldsymbol{\Lambda})^{-1}, \tag{5.7}$$

$$E[\exp\{c\boldsymbol{X}^*\boldsymbol{\beta}\}] = |I - c^2\boldsymbol{\Sigma}\boldsymbol{\Lambda}|^{-1/2} \exp\{\frac{1}{2}\boldsymbol{\ell}^t\boldsymbol{T}^{-1}\boldsymbol{\ell} - \frac{1}{2}\hat{\boldsymbol{\beta}}^t\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\beta}}\}, \tag{5.8}$$

where

$$\boldsymbol{T}^{-1} = \boldsymbol{\Sigma}^{-1} - c^2\boldsymbol{\Lambda}, \quad \text{and} \quad \boldsymbol{\ell} = \boldsymbol{T}(\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\beta}} + c\boldsymbol{\eta}^t).$$

Proof: First take the expectation over $\boldsymbol{X}^*$ with $\boldsymbol{\beta}$ fixed, obtaining

$$E^{\boldsymbol{X}^*}[\exp\{c\boldsymbol{X}^*\boldsymbol{\beta}\}] = \exp\{c\boldsymbol{\eta}\boldsymbol{\beta} + \frac{1}{2}c^2\boldsymbol{\beta}^t\boldsymbol{\Lambda}\boldsymbol{\beta}\}.$$

Thus

$$\begin{aligned} E[\exp\{c\boldsymbol{X}^*\boldsymbol{\beta}\}] &= (2\pi)^{-p/2}|\boldsymbol{\Sigma}|^{-1/2}\int_{\mathbb{R}^p}\exp\{c\boldsymbol{\eta}\boldsymbol{\beta} + \frac{1}{2}c^2\boldsymbol{\beta}^t\boldsymbol{\Lambda}\boldsymbol{\beta} - \frac{1}{2}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^t\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})\}d_{\boldsymbol{\beta}} \\ &= |\boldsymbol{T}|^{1/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\left\{\frac{1}{2}\boldsymbol{\ell}^t\boldsymbol{T}^{-1}\boldsymbol{\ell} - \frac{1}{2}\hat{\boldsymbol{\beta}}^t\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\beta}}\right\} \times \int_{\mathbb{R}^p} f(\boldsymbol{\beta};\boldsymbol{\ell},\boldsymbol{T})d\boldsymbol{\beta}, \end{aligned}$$

where $f(\boldsymbol{\beta};\boldsymbol{\ell},\boldsymbol{T})$ denotes the density of the $N_p(\boldsymbol{\ell},\boldsymbol{T})$ distribution. The integral of f is one provided $\boldsymbol{T}$ is positive definite. This is the case if $\boldsymbol{T}\boldsymbol{\Sigma}^{-1} = (\boldsymbol{I} - c^2\boldsymbol{\Sigma}\boldsymbol{\Lambda})^{-1}$ is positive definite, which follows if c^2 is less than the reciprocal of the largest eigenvalue of $\boldsymbol{\Sigma}\boldsymbol{\Lambda}$. □

Lemma. An alternate expression for (5.8) is

$$\begin{aligned} E[\exp\{c\boldsymbol{X}^*\boldsymbol{\beta}\}] &= |I - c^2\boldsymbol{\Sigma}_{AA}\boldsymbol{\Lambda}_A|^{-1/2} \\ &\times \exp\{c\boldsymbol{\eta}\hat{\boldsymbol{\beta}} + \frac{1}{2}c^2\boldsymbol{\eta}\boldsymbol{\Sigma}\boldsymbol{\eta}^t + \frac{1}{2}c^2\boldsymbol{b}_A^t\boldsymbol{\Lambda}_A(\boldsymbol{I} - c^2\boldsymbol{\Sigma}_{AA}\boldsymbol{\Lambda}_A)^{-1}\boldsymbol{b}_A\}, \end{aligned} \tag{5.9}$$

where $\boldsymbol{b}_A = \hat{\boldsymbol{\beta}}_A + c[\boldsymbol{\Sigma}_{AA}|\boldsymbol{\Sigma}_{AI}]\boldsymbol{\eta}^t$, and

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{AA} & \boldsymbol{\Sigma}_{AI} \\ \boldsymbol{\Sigma}_{AI}^t & \boldsymbol{\Sigma}_{II} \end{pmatrix}.$$

Proof. Follows directly from partitioning all matrices. □

Table 5.1		
PERCENT GAIN IN FUEL ECONOMY FOR INDIVIDUAL TECHNOLOGY IMPROVEMENTS (π_1-based posterior distribution)		
TECHNOLOGY IMPROVEMENT	**ESTIMATE OF % GAIN* IN AUTOMOBILES**	**ESTIMATE OF % GAIN* IN LIGHT TRUCKS**
ENGINE		
Throttle–Body Fuel Injection for Carburetor	4.24% (0.58%)	4.24% (0.58%)
Multi–Point Fuel Injection for Carburetor	4.10% (0.78%)	5.16% (0.85%)
2–Valve Overhead Cam for Push–Rod	0.96% (0.65%)	1.97% (0.96%)
3–Valve Overhead Cam for Push–Rod		2.83% (0.98%)
4-Valve Overhead Cam for Push–Rod	4.18% (1.32%)	4.30% (1.71%)
Roller Cam Followers for Sliding Cam Followers	2.68% (0.46%)	2.68% (0.46%)
TRANSMISSION		
Electronic Transmission Control	0.24% (0.24%)	0.24% (0.24%)
Torque Converter Lock–Up from Open Torque Converter	2.84% (0.43%)	3.77% (0.48%)
4–Speed Automatic from 3–Speed Automatic	1.87% (2.25%)	3.12% (1.96%)
OTHER		
Weight Reduction (Materials, 2.5%)		1.91% (0.19%)
Weight Reduction (Materials, 5.0%)	3.99% (0.38%)	3.90% (0.39%)
Weight Reduction (Materials, 10.0%)	8.34% (0.83%)	
Aero Drag Reduction (5.0%)	1.34% (0.12%)	1.54% (0.17%)
Aero Drag Reduction (10.0%)	2.75% (0.24%)	3.15% (0.36%)
Aero Drag Reduction (20.0%)	5.78% (0.52%)	
Tire Resistance Reduction (5%)	0.38% (0.03%)	0.38 (0.03%)
Front Wheel Drive**	0.07% (0.80%)	

*Standard errors are in parentheses.
**If FWDII assessment is used (see Table 4.1) the figures would be 3.78% (1.05%).

Using (5.8) or (5.9), computation of $(\hat{\gamma}_1 - 1)$ and $(\hat{\gamma}_2 - 1)$ (assuming normality of $\boldsymbol{X}^*$ and $\boldsymbol{\beta}$) is easy; the results are given in Table 2.1 of Section 2 for $(\hat{\gamma}_2 - 1)$, and in Table 5.1, here, for $(\hat{\gamma}_1 - 1)$. Note that the sizes of the largest eigenvalues of $\boldsymbol{\Sigma\Lambda}$ which we observed were in the range of 10^{-5} to 10^{-4}, so that (5.7) was easily satisfied.

5.2 Aggregation and Indexing

The main aim of this study is to estimate the total potential fuel economy gain for the automobile and truck fleets, from the introduction of technology improvements, by the model years 1995 and 2001 relative to the base year 1987. As discussed in Section 2, in order to estimate the gain possible from technology it is necessary to know the market penetrations of the various technologies for the 3 key years.

How to combine penetration rates with estimated gains for individual technologies is a very complex question. Are the penetration rates independent of vehicle size? How are a mix of technologies distributed over the fleet? Different answers to these and other questions will give rise to different estimates.

The most straight-forward method is to multiply the gain in fuel economy by the change in penetration over the period in question for each TI. The sum of these products is then the estimated gain. This is a commonly used method, but it does not take account of the fact that gains tend to multiply up. It also ignores the fact that its implicit baseline is unimproved vehicles, while the correct baseline 1987 fleet already includes vehicles with many of the technology improvements. Extreme cases are: 74% of automobiles had front wheel drive and 60% of trucks had a locked-up Torque Converter by 1987. Any gains relative to 1987 will thus have to be scaled because the overall average fuel economy of the 1987 fleet has been boosted relative to unimproved vehicles.

To take this into account we use a Fuel Economy Index (FEI), similar to a Consumer Price Index, which has as its base (= 1000) the fuel economy of the 1987 fleet were it comprised only of unimproved vehicles. Using our estimates of individual gains the FEI for 1987 is about 1091.

To calculate the FEI for any year, the TI's are arranged into groups of mutually exclusive TI's together with their common base. These groups are {MPI, TBI, CARB (base)}, {OHC4V, OHC3V, OHC2V, PUSH (base)}, {AUTO4, AUTO3 (base)}, {FWD, RWD (base)}, and all other similar pairs consisting of the TI and its complement (as base). Any vehicle has one and only one of the TI's or base from each group. The group fuel economy factor contributing to the FEI for a given year is

$$f = (p_0 + \sum_k p_k/g_k)^{-1},$$

where p_k, g_k are the market penetration and estimate of γ, respectively, for the k^{th} TI, and $p_0 = 1 - \sum_k p_k$ is the market share of the group base. (The p_k's ($k \geq 1$) are given in Tables 2.2 and 2.3, and the g_k's in Table 2.1.) The sum is over all TI's in the group. The factor is a harmonic average since, in fuel economy calculation, it is always assumed that distances rather than gallons of gasoline are fixed. The estimated FEI for the year is then the product of the group factors. This assumes that the market shares of TI's from different groups are independent.

The estimated gain in fuel economy for year y^1 over base year y^0 is then

$$G = \widehat{\mathrm{FEI}}(y^1)/\widehat{\mathrm{FEI}}(y^0) - 1.$$

In all computations below we use the π_2-based posterior, which was also used in preparing Tables 2.4 and 2.5.

In practice, using arithmetic rather than harmonic averages in the group factors makes very little difference to G. For Trucks and 1987-1995, arithmetic averaging yields G of 7.21% as opposed to 7.26% for harmonic averaging. Even for 1987 to 2001, max technology, the difference is small, 17.60% as against 17.64%. Such minor differences suggest that direct estimation of the FEI's and the gains over 1987 might be appropriate, using the method described in Section 5.1.

Let $\boldsymbol{X}_k^*$ denote the vector representing the mean change in vehicle characteristics when the k^{th} TI is added. Consider a fixed year y and let

$$\boldsymbol{p} = (p_1, p_2, \ldots, p_t)$$

denote the penetrations for year y of each of the t TI's. Then

$$\boldsymbol{X}^*(\boldsymbol{p}) = \sum_{k=1}^{t} p_k \boldsymbol{X}_k^*, \tag{5.10}$$

represents the penetration weighted average of the mean changes vectors.. Using the model, it can then be argued that

$$\mathrm{FEI}\ (y) = \exp\{\boldsymbol{X}^*(\boldsymbol{p})\boldsymbol{\beta}\}.$$

Furthermore, the gain in fuel economy for year y^1 over year y^0, penetrations $\boldsymbol{p}^1$ and $\boldsymbol{p}^0$ respectively, is

$$\begin{aligned}\mathrm{FEI}\ (y^1)/\mathrm{FEI}\ (y^0) - 1 &= \exp\{(\boldsymbol{X}^*(\boldsymbol{p}^1) - \boldsymbol{X}^*(\boldsymbol{p}^0))\boldsymbol{\beta}\} - 1, \\ &= \exp\{\boldsymbol{X}^*(\boldsymbol{p}^1 - \boldsymbol{p}^0)\boldsymbol{\beta}\} - 1.\end{aligned}$$

As in Section 5.1, this can be estimated by $(\hat{\gamma}_1 - 1)$ or $(\hat{\gamma}_2 - 1)$, using the explicit formulas in (5.8) and (5.9) if the normality approximation is made. Note that the appropriate mean vector,

$\boldsymbol{\eta}(\boldsymbol{p}^1 - \boldsymbol{p}^0)$, for $\boldsymbol{X}^*(\boldsymbol{p}^1 - \boldsymbol{p}^0)$ is

$$\boldsymbol{\eta}(\boldsymbol{p}^1 - \boldsymbol{p}^0) = \sum_{k=1}^{t} (p_k^1 - p_k^0)\boldsymbol{\eta}_k, \tag{5.11}$$

where $\boldsymbol{\eta}_k$ is the mean of $\boldsymbol{X}_k^*$. To obtain the approximate covariance matrix, $\boldsymbol{\Lambda}(\boldsymbol{p}^1 - \boldsymbol{p}^0)$, for $\boldsymbol{X}^*(\boldsymbol{p}^1 - \boldsymbol{p}^0)$, we assume that the $\boldsymbol{X}_k^*$ are independent and hence

$$\boldsymbol{\Lambda}(\boldsymbol{p}^1 - \boldsymbol{p}^0) = \sum_{k=1}^{t} (p_k^1 - p_k^0)^2 \boldsymbol{\Lambda}_k, \tag{5.12}$$

where $\boldsymbol{\Lambda}_k$ is the approximate covariance matrix of $\boldsymbol{X}_k^*$ as defined in Section 5.1.

The estimates given in Section 2.1 are those obtained using $(\hat{\gamma}_2 - 1)$ with (5.11) and (5.12). Comparison of these estimates for Trucks and the π_2-based posterior, with those obtained using the harmonic and arithmetic factoring methods, is given in Tables 5.2 and 5.3. All 3 methods give very similar estimates. The direct method has the advantage that the variance can be easily calculated; see the next section. The results for Automobiles and for $(\hat{\gamma}_1 - 1)$ were very similar and are not reported here.

Table 5.2 Fuel Economy Index – Trucks

Method	1987	1995	Product Plan 2001	Max Technology 2001
Direct	1091.4	1170.4	1239.3	1283.6
Harmonic	1090.8	1170.0	1238.7	1283.2
Arithmetic	1092.0	1170.8	1239.8	1284.2

Table 5.3 Fuel Economy Gains – Trucks

Method	1987 to 1995	1987 to 2001 Product Plan	1987 to 2001 Max Technology
Direct	7.23%	13.54%	17.61%
Harmonic	7.26%	13.56%	17.64%
Arithmetic	7.21%	13.54%	17.60%

5.3 Variance Computation

The estimated variances of $(\hat{\gamma}_1 - 1)$ and $(\hat{\gamma}_2 - 1)$ are, respectively,

$$\hat{V}_1 = \frac{\psi(-2)}{\psi(-4)} - \frac{2\psi(-1)\psi(-3)}{\psi(-2)\psi(-4)} + \left(\frac{\psi(-1)}{\psi(-2)}\right)^2, \tag{5.13}$$

$$\hat{V}_2 = \psi(2) - (\psi(1))^2, \tag{5.14}$$

where $\psi(c) = E[\exp\{c\boldsymbol{X}^*\boldsymbol{\beta}\}]$, the expectation being under the π_1 or π_2 posterior, respectively. Note that $\hat{V}_2$ is simply the posterior variance w.r.t. the prior based on $\pi_2(\boldsymbol{\beta})$. Motivation for $\hat{V}_1$ is given in the next section.

Note that $\hat{V}_1$ and $\hat{V}_2$ can be explicitly computed, using (5.8) or (5.9), if the normality approximations for $\boldsymbol{X}^*$ and $\boldsymbol{\beta}$ are used. Table 2.1 gives (in parentheses) the resulting standard deviations ($\sqrt{\hat{V}_2}$) for the ($\hat{\gamma}_2 - 1$) estimates of fuel economy gain for each TI. The $\sqrt{\hat{V}_1}$ are given (in parentheses) in Table 5.1.

Tables 2.4 and 2.5 give (in parentheses) the standard deviations of the various accumulated fuel economy gains, under the π_2-based posterior.

5.4 Motivation for the Estimators

The motivation for estimating $\gamma = \exp\{\boldsymbol{X}^*\boldsymbol{\beta}\}$ by $\hat{\gamma}_1$, with variance $\hat{V}_1$, arises from decision theory. Suppose, first, that $\boldsymbol{X}^*$ is fixed and that the simple fixed-effects regression model is used, so that $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}^{\mathrm{OLS}} \sim N_p(\boldsymbol{\beta}, \boldsymbol{\Sigma})$. It can then be shown that estimation of γ is a scale parameter problem, and $\hat{\gamma}_1$ is the optimal "invariant" estimator, from a frequentist perspective. (A Bayesian "reference" prior version of the argument can also be given.) Likewise, $\hat{V}_1$ is the optimal "invariant" estimator of the accuracy of $\hat{\gamma}_1$.

For $\boldsymbol{X}^*$ random, partial theoretical results and extensive numerical studies indicate that $\hat{\gamma}_1$ remains a near optimal estimate. Details of these arguments can be found in Berger and Robert (1991).

The motivation for use of $\hat{\gamma}_2$ follows from consideration of the linear approximation (5.4). For a given TI, the natural estimate of $\boldsymbol{X}^*\boldsymbol{\beta}$ is

$$E[\boldsymbol{X}^*\boldsymbol{\beta}] = \boldsymbol{\eta}\hat{\boldsymbol{\beta}}, \tag{5.15}$$

where E stands for posterior expectation and $\hat{\boldsymbol{\beta}} = E[\boldsymbol{\beta}]$. The variance of this estimate is

$$\begin{aligned} E[\boldsymbol{X}^*\boldsymbol{\beta} - \boldsymbol{\eta}\hat{\boldsymbol{\beta}}]^2 &= E[\boldsymbol{\beta}^t\boldsymbol{\Lambda}\boldsymbol{\beta}] + \boldsymbol{\eta}\boldsymbol{\Sigma}\boldsymbol{\eta}^t \\ &= \hat{\boldsymbol{\beta}}^t\boldsymbol{\Lambda}\boldsymbol{\beta} + \boldsymbol{\eta}\boldsymbol{\Sigma}\boldsymbol{\eta}^t + tr(\boldsymbol{\Sigma}\boldsymbol{\Lambda}), \end{aligned} \tag{5.16}$$

using the notation from Section 5.1.

There are two reasons for *not* using the usual noninformative prior, $\pi(\boldsymbol{\beta}) = 1$, in these computations. The first is that (5.15) will be done for a variety of $\boldsymbol{\eta}$ (corresponding to different TI's) and simultaneous estimation of $\boldsymbol{\eta}\boldsymbol{\beta}$, for many $\boldsymbol{\eta}$, suggests that a shrinkage estimator of $\boldsymbol{\beta}$ might be appropriate. Use of $\pi_2(\boldsymbol{\beta})$ (see (3.5)) leads to a standard shrinkage estimator (Baranchik, 1964).

The more compelling motivation for using $\pi_2(\boldsymbol{\beta})$ is that $\pi(\boldsymbol{\beta}) = 1$ is notorious for yielding very poor estimates of quadratic forms, such as $\boldsymbol{\beta}^t \boldsymbol{\Lambda} \boldsymbol{\beta}$ in (5.16). This was first observed by Charles Stein; see Berger (1985) for discussion. Again, $\pi_2(\boldsymbol{\beta})$ can be shown to overcome the problem, from either a frequentist or a reference prior perspective. See Robert and Berger (1991) for details.

Because of the fact that $(\gamma - 1) \cong \boldsymbol{X}^* \boldsymbol{\beta}$, a prior that works well for $\boldsymbol{X}^* \boldsymbol{\beta}$, would likely be fine for $(\gamma - 1)$ also; this motivated our utilization of $\pi_2(\boldsymbol{\beta})$ and $\hat{\gamma}_2$.

The difference between the two approaches discussed above can usefully be thought of as follows. In the first approach, we desire to use $\pi(\boldsymbol{\beta}) = 1$, for simplicity, but then we view the problem from a decision-theoretic perspective and find a weighted quadratic loss for which the ensuing Bayes estimate is "optimal" from a frequentist perspective. (The weight is irrelevant from a frequentist perspective, but is needed to "adapt" the prior so as to yield the optimal invariant estimator.) In the second approach, we decide to utilize the posterior mean and variance as our estimates, but must then choose a (noninformative) prior carefully, so as to avoid known pitfalls of casual use of $\pi(\boldsymbol{\beta}) = 1$. Interestingly, the two approaches yield essentially the same answers here.

Indeed, the issues discussed above turn out to be mostly irrelevant for this particular data set; the answers are quite insensitive to the choice of π_1 or π_2 or the choice of estimator. We, of course, could not ascertain that this would be true, apriori.

6. CONCLUSIONS

The subject-matter conclusions have been summarized in Section 2 and in Section 3.5. Here we summarize general observations relevant to statistical practice.

Data and Modelling

(i) Even though this was supposedly an excellent data base, a major effort was needed to improve it to the point of being usable.

(ii) Utilization of hierarchical random effects seemed to improve the results, as well as Bayesian incorporation of known constraints on parameters.

(iii) Choice of the model can depend significantly on the *goal* of the analysis. Thus our choice of a loglinear model was partly based on the goal of estimating proportional changes due to incorporation of TI's, since "log proportional change" becomes additive. The choice of which dependent variables to include in the model was partly shaped by the realization that subjective assessments concerning these variables would be necessary (for performance control), so that the variables would need to be assessable by automotive engineers. Interaction terms were, in part, not utilized for this reason.

Subjective Assessments

Note that we were relative novices in the implementation of assessment procedures and made numerous mistakes, in part because we had not anticipated – in advance of the study – the need for subjective assessments. The following observations should be viewed in that light, as well as the fact that they are based on a "sample" of size one.

(i) When dealing with busy professionals, count on much less of their time than you feel is necessary. Thus keep the number of quantities to be assessed very small.

(ii) Begin with a training session if possible, explaining exactly which assessments it is hoped to obtain and why. Make sure that the assessors understand what is meant by asked-for quantities, such as mean, variance, correlation, and quantiles. Also, clearly convey that there are no "right" or "wrong" answers.

(iii) Obtaining preliminary individual assessments is probably helpful, as a means of having the assessors focus on the problem, and as a starting point for the Delphi analysis. Note, however, that there is considerable reluctance by engineers to go "on the record".

(iv) Groups sessions are the key to successful assessment. Different assessors possess differing information, and the information sharing process is the most crucial aspect of group assessment. Formal methods of combining individual probability assessments cannot capture this process.

(v) Multiple group sessions and/or other opportunities for feedback are extremely valuable, and should be planned.

Statistical Analysis

(i) Gibbs sampling allowed the use of a complex hierarchical random effects model and easy incorporation of constraints on parameters. We can thus attest to its value in allowing freedom in model development. Of course, all the usual strengths of using Bayesian methodology were also present, such as the ability to easily produce estimates of accuracy.

(ii) Nonlinear multivariate estimation problems can be tricky. Use of standard noninformative priors (e.g., a constant) and the posterior mean and variance may well be suboptimal. For problems with a regression component, the use of $\pi(\boldsymbol{\beta}) = 1/|\boldsymbol{\beta}|^{(p-1)}$, which can easily be incorporated into Gibbs sampling, might be better on a routine basis.

REFERENCES

An, F. and Ross, M. (1991). Automobile efficiency at part-load: the opportunity for improvement. Report, Physics Department, University of Michigan, Ann Arbor.

Baranchik, A. J. (1964). Multiple regression and estimation of the mean of a multivariate normal distribution. Technical Report No. 51, Dept. of Statistics, Stanford Univ., Stanford.

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

Berger, J. and Robert, C. (1991). Estimation of noncentrality parameters and quadratic functions. In preparation.

Berger, J., Smith, M. H., and Andrews, R. W. (1990). A system for estimating fuel economy potential due to technology improvements (automobile). Technical Report.

Berger, J., Smith, M. H., and Andrews, R. W. (1991). A system for estimating fuel economy potential due to technology improvements (truck). Technical Report.

Dalkey, N. C. (1972). *Studies in the Quality of Life: Delphi and Decision-Making*. D.C. Heath and Co., Lexington, Mass.

Difiglio, C., Duleep, K. G., and Greene, D. L. (1990). Cost effectiveness of future fuel economy improvements. *The Energy Journal* **11**, 65–86.

Energy and Environmental Analysis, Inc. (1988). Light duty truck fuel economy, review and projections 1980–1995, DOE/OR/21400-HO, prepared for U.S. Department of Energy, Washington, D.C.

Energy and Environmental Analysis, Inc. (1991). An assessment of potential passenger car fuel economy objectives for 2010. Report prepared for Air and Energy Policy Division of the Environmental Protection Agency by EEA, Arlington, Virginia.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* **85**, 398–409.

Geman, S. and Geman, D. (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.

Poyer, D. A. and Santini, D. J. (1990). The potential of 1989 light duty vehicle engine and drivetrain technology to reduce CO_2 emissions by improving fuel efficiency. Report, Center for

Transportation Research, Argonne National Laboratory, Argonne, IL.
Robert, C. and Berger, J. (1991). Optimal loss functions, with application to estimation of lognormal moments. In preparation.
Sovran, G. and Bohn, M. S. (1981). Formulae for the tractive-energy requirements of vehicles driving the EPA schedules. *Society of Automotive Engineers* Paper #810184.
Tierney, L. (1991). Markov chains for exploring posterior distributions. Technical Report No. 560, School of Statistics, University of Minnesota, Minneapolis.

APPENDIX A

DEFINITIONS OF TECHNOLOGY IMPROVEMENTS

The definitions of the technology improvements covered by this report are as follows.

Throttle–Body Fuel Injection: This is an improvement using the carburetor as a base. The metered fuel injection system supplies a single atomized fuel spray to the engine.

Multi–Point Fuel Injection: This is a metered fuel injection system but it differs from the throttle–body system in that multiple atomized fuel sprays are generated, one for each cylinder.

2–Valve Overhead Cam: This is an engine improvement using the 2-valve push-rod assembly as the base.

3–Valve Overhead Cam: This is an engine improvement using the 2-valve push-rod assembly as the base.

4–Valve Overhead Cam: This is an engine improvement using the 2-valve push-rod assembly as the base.

Roller Cam Followers: This engine improvement uses a roller bearing assembly instead of a sliding bearing as the mechanism for valve activation.

Electronic Control: This TI has sensors which detect engine and transmission states and selects the appropriate gear using solenoid valves. It is a part hydraulic and part electronic system with the choice of gear and adjustments of pressure depending on the electronic system. The base for this TI is an all hydraulic system with shift control of and lock–up in gears 2, 3, and 4.

Torque Converter Lock–Up: This mechanical system locks the torque converter and prevents slippage. An open converter is regarded as the base.

4-Speed Automatic: This TI uses a 3-Speed transmission without lock-up as base.

Weight Reduction (2.5%, 5%, and 10%): These TI's covers weight reduction due to material substitution. These reductions will be accomplished using aluminum, plastic, and composite material designs.

Aerodynamic Drag Reduction (5%, 10%, and 20%): These TI's assume a reduction in C_D of 5%, 10%, and 20%, under various scenarios.

Tire Rolling Resistance Reduction (5%): It is projected that new construction techniques for tires can decrease the rolling resistance.

Front Wheel Drive: This is the changing of the drive system from the rear wheels to the front.

APPENDIX B

REGRESSION VARIABLES

The variables in regression are defined as follows:

LMPG = the natural log of the combined city and highway miles per gallon, computed from CITY and HWAY miles per gallon via the formula

$$\mathrm{LCOMB} = -\ell n(.55/\mathrm{CITY} + .45/\mathrm{HWAY}).$$

LETW = the natural log of equivalent test weight.

LTCL = the natural log of tire and chassis loss(TCL); where TCL is computed by the following formula:

$$\mathrm{TCL} = \left(\mathrm{CDT}^{-1} - \frac{\mathrm{DPA}}{.06172 \times \mathrm{ETW}}\right)^{-1} \tag{B1}$$

where

$$\begin{aligned} \mathrm{CDT} &= \text{coastdown time,} \\ \mathrm{ETW} &= \text{equivalent test weight, and} \\ \mathrm{DPA} &= \text{dynamometer setting.} \end{aligned}$$

LDPA = the natural log of the dynamometer setting.

LNV = the natural log of N/V, the ratio of engine RPM to vehicle velocity in top gear.

LAXLE = the natural log of AXLE, which is the final drive ratio on front wheel vehicles and the axle ratio on rear wheel drive vehicles.

LCID = the natural log of the cubic inch displacement of the engine.

LCOMP = the natural log of the engine compression ratio.

LHPCID = the natural log of the engine horsepower divided by cubic inch displacement, which is referred to as specific horsepower.

IMPI = an indicator variable for multi–point injection.

An indicator variable has a value of one if the characteristic is present and a value of zero otherwise. For example,

$$\text{IMPI } = \begin{cases} 1, & \text{if MPI present;} \\ 0, & \text{if MPI not present.} \end{cases}$$

ITBI = an indicator variable for throttle–body injection.

IOH2V = an indicator variable for an overhead cam with 2 valves per cylinder.

IOH3V = an indicator variable for an overhead cam with 3 valves per cylinder.

IOH4V = an indicator variable for an overhead cam with 4 valves per cylinder.

IACTR = an indicator variable for a roller cam follower.

IACTD = an indicator variable for direct acting cam followers.

IAUTO4 = an indicator variable for an automatic transmission with 4 forward gear ratios.

ILKUP = an indicator variable for a transmission with a torque converter lock–up.

IELEC = an indicator variable for transmission with an electronic control.

IFWD = an indicator variable for a vehicle with front wheel drive.

IMAN4 = an indicator variable for a manual transmission with 4 forward gears.

IMAN5 = an indicator variable for a manual transmission with 5 forward gears.

ICRP3 = an indicator variable for a vehicle with a creeper and 3 forward gears.

ICRP4 = an indicator variable for a vehicle with a creeper and 4 forward gears.

APPENDIX C

OLS REGRESSION OUTPUT

Dependent Variable: LMPG

Analysis of Variance

Source	DF	Sum Sqrs	Mean Sqr	F-Stat	Signif
Regression	23	140.37	6.1032	2115.0	0.
Error	2169	6.25	.00289		
Total	2192	146.63			

Mult R= .97842 R-SQR = .95731 SE = .053719

Estimation of Regression Coefficients

Variable	Parameter Estimate	Std Error	T-Stat	Signif
0. CONSTANT	10.090	.14584	69.182	0.
1. LETW	− .45049	.01786	−25.218	0.
2. LTCL	.07809	.00642	12.148	.0000
3. LDPA	− .18771	.00824	−22.765	0.
4. LNV	− .33890	.01789	−18.942	.0000
5. LAXLE	− .08300	.01577	−5.2604	.0000
6. LCID	− .46001	.01148	−40.066	0.
7. LCOMP	.24541	.02980	8.2346	.0000
8. LHPCID	− .13137	.01143	−11.491	.0000
9. IMPI	.03287	.00534	6.1455	.0000
10. ITBI	.03745	.00527	7.1059	.0000
11. IOH2V	− .01853	.00362	−5.1056	.0000
12. IOH3V	.03097	.00540	5.7277	.0000
13. IOH4V	.04647	.00562	8.2666	.0000
14. IACTD	− .02993	.00573	−5.2211	.0000
15. IACTR	.00141	.00314	.4504	.6524
16. IAUTO4	− .03738	.00662	−5.6401	.0000
17. ILKUP	.03043	.00485	6.2750	.0000
18. IELEC	.01812	.00464	3.9050	.0001
19. IFWD	− .02201	.00440	−5.0022	.0000
20. IMAN4	.07816	.00803	9.7288	.0000
21. IMAN5	.04489	.00619	7.2478	.0000
22. ICRP3	.06773	.00732	9.2464	.0000
23. ICRP4	.05096	.01508	3.3783	.0007

APPENDIX D

ADDITIVITY, DOUBLE COUNTING and CORRELATION ISSUES

Any system of fuel economy assessment potentially suffers from double counting, additivity, and correlation problems. It is worthwhile to illustrate these notions and discuss the extent to which our analysis avoids these problems.

Example of Double Counting: Suppose it is observed that front wheel drive vehicles seem to be, on the average, 200 pounds lighter (the numbers here are hypothetical values – they are only for purposes of illustration) than rear wheel drive vehicles of the same interior volume. Hence a 200 pound "packaging improvement" is assigned to the TI *front wheel drive* (FWD). Suppose it is also noticed that 4-valve overhead cam engines seem to be 100 pounds lighter than 2-valve overhead cam engines of the same power. Hence a 100 pound weight improvement is assigned to the TI *4-valve overhead cam engine* (4VOHC). But, the vehicles with 4-valve overhead cam engines are predominantly front wheel drive, so a large portion of the 200 pound improvement assigned to FWD could be due to the 100 pound weight loss for 4-valve overhead cam engines; the 100 pounds is being "double counted".

Once pointed out, any particular instance of double counting can often be eliminated in ad hoc assessment systems, but it can be virtually impossible to perceive all sources of double counting in complicated systems. One of the major advantages of global assessment systems, such as the regression approach adopted here, is that double counting cannot occur. By putting all TI's together in the regression model, actual mileage improvements seen in the data can only be "divided up" among the TI's.

Even more insidious than double counting is the additivity issue. Here is a hypothetical example.

Example of Additivity Bias: Technology improvements **A** and **B** can be added to an engine to improve fuel burn efficiency. There are several vehicles in which only **A** has been adopted, and the fuel efficiency gain for these vehicles averages 4%. There are also several vehicles in which only **B** has been adopted, and the fuel efficiency gain for these average 6%. This would tempt one to say that the effect of adopting both **A** and **B** is 10%. But, for the vehicles in which both **A** and **B** are adopted, the fuel efficiency gain turns out to average only 7%. The two TI's do not have additive effects.

The best way to allow for such nonadditivity is to develop models with *interaction terms*. For instance, the part of the model relating to **A** and **B** in the above example might be

$$(.04)IA + (.06)IB - (.03)IAB,$$

the last term being the indicator for the presence of both **A** and **B**. Unfortunately, we were unable to utilize interaction terms in our model, because of the subsequent planned utilization of the model. Interaction terms involving vehicle characteristic variables would have required subject assessments of interactive characteristic changes in the performance control prong of the project. And interaction terms solely involving the TI indicator variable would have required detailed knowledge of penetration rates for simultaneous introduction of TI's, when estimating total fuel economy gain. Also, some experimentation with interaction terms revealed that, with the limited available data, interaction terms would not provide significantly better predictions. Hence, we proceeded without interactions. Note, of course, that even a purely additive model overcomes the nonadditivity problem mentioned above, to a large extent, as the following indicates.

Example of Additivity Bias (continued): Suppose that, in the data base, 10% of the vehicles have **A** only, 10% have **B** only, and 30% have both **A** and **B**. Then the additive model will divide up the actual 7% fuel efficiency gain for the "both **A** and **B**" vehicles between the two TI's, resulting (when averaged with the **A** only and **B** only vehicles) in a model such as $(.03)IA + (.05)IB$. Use of this will give considerably better answers than the crude $(.04) + (.06) = (.10)$. Again the key point is that the effect of all TI's must be analyzed simultaneously, as is done in the regression model.

What would happen if one could develop and utilize a model with interaction terms? The likely effect would be to lower the predictions of possible fuel efficiency gains. This is because the coefficients on interaction terms will tend to be negative (engineering experience suggests negative synergies) while the future will call for more simultaneous use of TI's. Hence, inclusion of interactions such would likely only lower our estimated fuel efficiency gains.

The final notion relevant to the discussion is that of correlation. Many of the TI's are highly correlated in the data set; that is, they tend either to occur together (positive correlation) or not to occur together (negative correlation). The presence of such correlation makes all analyses more difficult; it becomes all–too–easy to erroneously assign the benefit of one TI to another (positively correlated) TI. Ad hoc analyses are particularly susceptible to this problem, since the effect of this correlation is difficult to determine.

In contrast, regression analysis precisely measures the effect of correlation. The primary effect of correlation is to increase the variance of the regression coefficients. Since we report these variances, the effect of correlation is being accounted for. It was the case that some of our estimates of fuel efficiency gains had sizable variances, much of which was probably due to the correlated data. Again, therefore, the issue of correlation points out the necessity to quantify accuracy formally; correlation is too important and too subtle an issue to deal with by ad hoc methods.

Note that it is sometimes argued that regression analysis is inappropriate here because the data typically contains many runs of a specific model, the vehicles having only different TI's (or configurations) in each run. This, it is argued, allows certain vehicles to receive much more "weight" than their due and to possibly distort the picture of the true market.

Such arguments are based on a misunderstanding of the role of the regression analysis as we use it. We use it only to find the effect of a TI on an individual vehicle, not to find its global effect on the population of vehicles. If the model is correct, the data need not occur in correct "market proportions" to provide accurate estimates of the regression coefficients. The actual sales–weighted vehicle population is (properly) brought into the analysis only at the final stage, based on penetration rates of the TI's.

Of course it is true that the model is not absolutely correct, in that there is no *absolute* effect of a TI; TI's will affect different vehicles differently. Hence our model is, in some sense, measuring the "average" effect of a TI, and the averaging is over the vehicles in the data set. While not completely representative of the sales-weighted market, the composition of the data set is reasonably similar to this market. Also, the use of random effect terms for vehicle-lines will reduce the vehicle effect on a TI coefficient. Hence we do not feel that there is a need for concern over the representativeness of the data base.

APPENDIX E

EXPLANATION OF ENGINEERING ASSESSMENTS

For TI Implementation in Automobiles: The following explanations refer to the assessments in Table 4.1.

Sliding to Roller Cam Followers: There are no changes in the auxiliary variables associated with the change from sliding to roller cam followers. Therefore, the resulting change will be reflected solely by the indicator variable for this TI.

Carburetor to Throttle–Body Fuel Injection: This TI will improve the metering, an effect that will be captured by the corresponding indicator variable; again, no changes to other characteristics have to be considered in order to keep performance constant.

Carburetor to Multi–Point Fuel Injection: The multi–point fuel injection system delivers higher horsepower, assessed to be in the range 9 to 18 percent, because of better mixing of air and fuel. Since the power output of the engine, per unit of displacement, would increase, it is reasonable to assume that displacement would be diminished in order to maintain constant performance. This displacement reduction was assessed to be in the range of 4 to 8 percent. Due to the better fuel and air mixing, the engine is not as *knock limited* and therefore there is an opportunity to increase the compression ratio. As indicated in the table, the increase in the compression ratio is in the range of 0.2 to 0.4.

2–Valve Pushrod to 2–Valve Overhead Cam: Because of greater power output, as indicated by the assessed increase in HP/CID of 17% to 30%, reduction in displacement would be possible, of between 8 and 14 percent. To maintain gradability, as defined by constant performance, the final drive ratio was assessed to increase by 3 to 6 percent. Therefore, the N/V ratio in top gear is also assessed to increase by the same amount. There is no assessed weight change since the overhead cam hardware offsets the reduced weight that accompanies the reduction in displacement.

2–Valve Pushrod to 4–Valve Overhead Cam: Because of greater power output, as indicated by the assessed increase in HP/CID of 45% to 75%, the range for displacement reduction was assessed to be between 20 and 30 percent. The assessed changes to final drive ratio and N/V are, again, to maintain gradability. The 4–valve technology increases the weight of a V–configured engine by approximately 20 to 40 pounds, but it does not add any weight to I4 type engines. Therefore, the assessed weight interval was the range of zero to forty pounds. There is some evidence that an increase in COMP is possible, in the indicated range.

Electronic Control: There are no changes in the auxiliary variables associated with the introduction of electronic gear shifting.

Open Torque Converter to Lock–Up Torque Converter: There are no changes in the auxiliary variables associated with the introduction of lock–up.

3–Speed Automatic to 4–Speed Automatic: The change from 3–speed to 4–speed automatic transmission results in several synergistic changes which affect four of the auxiliary variables. In order to maintain constant performance, the displacement, final drive ratio, and N/V ratio have to be appropriately adjusted. The changes are given by the reported intervals in the table. Furthermore, these changes would be made in a dependent manner. At one extreme one could anticipate, for example, a change in displacement of -5% associated with an N/V change of -15% and a final drive ratio change of 2%. At the other extreme the changes would be: (CID: -10%); (N/V: 0%) and (AXLE: 5%). This dependence is modelled by assuming correlations of 1 in these assessments. Finally, there is a minor weight increase due to additional parts.

Note that the engineers did not simply consider changes in characteristics that would result from adding a fourth gear to a transmission, since doing only that would degrade performance (specifically, gradability). The suggested ensemble of changes would more properly maintain constant performance.

Rear Wheel Drive to Front Wheel Drive: Due to packaging improvements in changing from rear wheel drive to front wheel drive there is an assessed weight reduction of between 50 and 80 pounds. To maintain constant power to weight ratio, it is assessed that the displacement will be reduced by the same percentage. Using 3000 pounds as a typical vehicle weight, the 50 to 80 pounds implies a 1.7 to 2.7 percent reduction in CID.

There is a major difference between the weight assessment here and the weight change due to front wheel drive that is determined by EEA. The discrepancy might be due to the fact that older, rear wheel drive vehicles have not yet been optimized from a weight viewpoint, in which case additional gains from weight improvement would indeed be possible, but the situation is not clear. There is apparently a significant weight differential in the raw data, but as we have repeatedly stressed one must carry out such comparisons only under strict performance control. Until such is done we view the issue as unresolved. Nevertheless, we record as Rear Wheel Drive to Front Wheel Drive II, the assessments proposed by EEA.

Weight Reduction (Materials): Weight reductions of 5% and 10% are anticipated under various scenarios. In order to maintain performance at approximately a constant value, there is a reduction in displacement in the indicated ranges, corresponding to the weight reductions.

Aerodynamic Drag Reduction: Aerodynamic drag reductions of 5%, 10%, and 20% are anticipated under various scenarios. (The 20% reduction corresponds to so-called luxury vehicle removal.) There is no indicator variable for aerodynamic drag and there is no specific auxiliary variable for aerodynamic drag. However, drag affects the DPA auxiliary variable and the effect on fuel economy of a reduction in aerodynamic drag can be estimated by the appropriate change in the DPA auxiliary variable. In addition there is a proportional reduction in N/V.

Tire Rolling Resistance Reduction: There is no indicator variable for tires and there is no specific auxiliary variable for tires. However, the effects due to reduction of 5% in tire resistance affects the TCL and DPA auxiliary variables in the assessed ranges.

For TI Implementation in Light Trucks: The following explanations refer to the assessments in Table 4.2. Only explanations that differ substantially from the Automobile explanations are given.

2–Valve Pushrod to 3–Valve Overhead Cam: The explanation for this assessment is similar to that for 2–Valve Overhead Cam, except the extra valve requires a slight increase in each of the assessed values.

Open Torque Converter to Lock–Up Torque Converter: Based on outputs from CVSP (a Ford proprietary computer program which generates fuel economy output by using input settings that simulate a specific vehicle configuration), the performance in gradeability is held approximately constant by decreasing the final drive ratio by 2 to 4 percent.

3–Speed Automatic to 4–Speed Automatic: Changing the transmission from 3–speed to 4–speed results in several synergistic changes which affect three of the auxiliary variables. In order to maintain constant performance, the final drive ratio and the N/V ratio have to be appropriately adjusted. The changes are given by the reported intervals in the table. Furthermore, these changes would be made in a dependent manner. At one extreme one could anticipate, for example, a change in N/V of -30% associated with a final drive ratio change of 15%. At the other extreme the changes would be: (N/V: -15%) and (AXLE: 5%). This dependence is modelled by assuming a correlation of 1 in these assessments. Finally, there is a minor weight increase due to additional parts.

This assessment presumes that the TI is essentially the addition of an overdrive gear; complete regearing and CID adjustment is not reasonable for trucks, as was done for Automobiles, because of the truck performance criteria.

Weight Reduction (Materials): Weight reductions of 2.5% and 5% are anticipated under various scenarios. In order to maintain performance at approximately a constant value, there is a reduction in displacement in the indicated ranges, corresponding to the weight reductions. Output from CVSP verified these assessments.

Aerodynamic Drag Reduction: Aerodynamic drag reductions of 5% and 10% are anticipated under various scenarios. The explanation is as for Automobiles. Again, output from CVSP substantiated the assessments.