Beyond the Model Specification Problem: Model and Parameter Averaging Using Bayesian Techniques

Henry L. Bryant and George C. Davis

Selected Paper
American Agricultural Economics Association Annual Meeting
Chicago, IL, August 5 – August 8, 2001

Abstract

The model specification problem is perhaps the Achilles heel of applied econometrics. Rather than test down to a single model as is usually done, we estimate 72 different demand systems and use Bayesian averaging procedures over all 72 systems to generate meta estimates of the parameters (e.g., elasticities) of interest.

*Appreciation is expressed to Henry Kinnucan and Gary Brester for providing some of the data used in this analysis, to John Stevens for assistance with the Bayesian software used in the analysis, and to Melissa Weichert for technical assistance.

Contact Information: George C. Davis, Texas A&M University, 302 Blocker Bldg. College Station, Tx 77843-2124, Ph: (979) 845-3788. E-mail: gdavis@tamu.edu.

Beyond the Model Specification Problem: Model and Parameter Averaging Using Bayesian Techniques

In its infancy, the promise of econometrics was that it would allow economists to uncover permanent laws with fixed coefficients and place economics on an equal footing with physics (Morgan). In its childhood, it became clear that this was an overly optimistic hope because of, what is now called, the model specification problem (Epstein). The *model specification problem* simply stated is the difficult problem of selecting the form of the equations to be estimated and the variables to go in the equations. Though suggestive, economic theory is usually silent on how to *fully* specify an empirical model in terms of the functional form and all the explanatory variables.

The most common approach to the model specification problem has been what Pagan labels the 'test-test-test' approach, which is a general catch all name for the well known 'general-to-specific' approach (Hendry) and the 'probability reduction' approach (Spanos 1999, McGuirk et al. AJAE 1995). In this approach the researcher begins with some general model and then starts testing the underlying assumptions of the model. If some of these assumptions are violated then the model is respecified and tested again to see if the violations have been eliminated. The modeler continues this process until a single 'statistically adequate' model is obtained. Though preferred to no model specification testing, there are three limitations of this approach. One, the functional form is usually assumed to be the same throughout the search process. Two, the single final model is conditional on the results of many tests so the overall type I error is hard to access. Three, if there were a particular parameter of interest one would like an idea of its robustness or fragility across different models.

An alternative approach that overcomes most of these classical limitations is a Bayesian model averaging approach. Within a Bayesian context, model uncertainty is conceptually easily handled as it becomes just another unknown in the prior and posterior density functions. Within this frame density functions can be formed over different models and consequently model results can be formally combined or averaged based on the model's density functions, hence the term Bayesian model averaging. Bayesian model averaging is not a new concept but until the last decade the computational components of Bayesian analysis have been rather taxing on its actual implementation. However, within the last few years Monte Carlo (MC) methods have been exploited to approximate the needed integrals in Bayesian analysis (Dorfman; Geweke 1989). With MC methods it is now easier to implement Bayesian approaches to standard statistical problems such as the model specification problem and Bayesian model averaging has recently received a great deal of attention in the econometrics and statistics literature (e.g., George and McCulloch; Geweke (1999); Fernandez, Ley, and Steel; Hoeting, Madigan, Raferty, and Volinsky; Moulton; Raferty, Madigan, and Volinsky; Raferty, Madigan and Hoeting).

In the Bayesian model averaging approach, multiple nested or non-nested models are estimated. The models and/or specific parameters are then averaged over models using a weighting scheme based on the posterior ratios. The advantages to this approach are that (i) no single model is selected as 'the true model'; (ii) because the weighting is based on posterior odds, the fits of the different models are taken into account; (iii) the averaged parameter estimate is robust to alternative types of model specification issues, especially functional form and the inclusions and exclusions of certain variables.

In this paper, we summarize the Bayesian approach to the model selection problem and averaging procedure. We then apply the techniques to the study of meat demand in the U.S. because it has been so intensely studied and hotly debated in the literature.

A Bayesian Approach to the Model Selection Problem

Consider the general estimation problem faced by the empirically oriented economist. An economic theory suggests some functional relationship between variables,

(1) $$y = f(x_1, x_2, ..., x_k)\ .$$

where y is the variable or phenomenon the theory seeks to explain, the k variables denoted by x are identified as being the determinants of y, and $f : \Re^k \rightarrow \Re$. Now most theories, if not all, lack the required specificity to estimate (1). The two major shortcomings are (i) the functional form for f is not specified beyond stating that it is within a class of functions with certain properties (e.g., signs on partial derivatives, restrictions on functions of the partial derivatives, etc.); (ii) the k variables are not uniquely identified beyond the statement that some are expected to be more important than others. Thus the applied economist is forced to select from a functional form to represent f and the variables to be included in the function. This is the model selection problem.

A model M_m can be formally defined as a pair$\{f_i\ , \boldsymbol{X}_j\}$, where $f_i \in F$, $\boldsymbol{X}_j \subset \boldsymbol{X}$, and F is a class of functional forms and $\boldsymbol{X}$ is a matrix of all possible explanatory variables. The m index refers to a unique pair since the same variables can be used in different functional forms and so $M = \{M_1, M_2, \ldots, M_s\}$ denotes the set of all possible models to be considered. Now given a particular model M_m and observed data D, following standard Bayesian procedures the posterior distribution for the parameter vector for that model ω_m can be written as

(2) $$p(\omega_m \mid D, M_m) \propto p(\omega_m \mid M_m)\, p(D \mid M_m, \omega_m)$$

where $p(\omega_m \mid M_m)$ is the conditional prior density for ω_m and $p(D \mid \omega_m, M_m)$ is the conditional density for the data, which is proportional to the likelihood function.[1] The marginal likelihood for model M_m is then defined as

(3) $$p(D \mid M_m) = \int p(\omega_m \mid M_m)\, p(D \mid \omega_m, M_m) d\omega_m .$$

Following Zellner (1971), the posterior odds ratio K_{rs} favoring a model m over a model r is

(4) $$K_{ms} = \frac{P(M_m)}{P(M_r)} \frac{p(D \mid M_m)}{p(D \mid M_r)}$$

where the first part of this expression is the prior odds ratio and the second part the Bayes factor. For s possible models, the posterior probability of a particular model m is

(5) $$P(M_m \mid D) = \frac{K_{ms}}{\sum_n^s K_{ns}} .$$

From the Bayesian perspective, if a *single* model had to be chosen then it would be the model with the largest posterior probability. Of course there is always a danger in choosing one out of many possible models but fortunately within the Bayesian framework there is enough information available to coherently and formally average over models.

Suppose there is some quantity of interest, say η, common to all models but that can differ across models which is a function of the parameters, i.e., $\eta_m = h(\omega_m)$. This quantity may be something as simple as an elasticity estimate or something more complicated such as a forecast. In the present setting, we concentrate on estimating elasticities. By the rules of probabilities, then the expected value of ε with a discrete set of models can be denoted as

$$(6) \qquad E(\eta \mid D, M) = \sum_{i=1}^{r} E[h_i(\varpi_i) \mid D, M_i] P(M_i \mid D)$$

$$= \sum_{i=1}^{r} \hat{\eta}_i P(M_i \mid D)$$

$$= \hat{\eta},$$

where $\hat{\eta}_i$ is the estimate of the expected value for η from model M_i and $\hat{\eta}$ is the overall estimate for η. Note then that the estimate of the η is nothing more than a weighted average of the individual estimates across models with the weights being the probability that a particular model is consistent with the data based on the posterior odds ratio. Furthermore, from the definition of a variance, an estimate of the variance of $\hat{\eta}$ is

$$(7) \quad V(\hat{\eta}) = \left(\sum_{i=1}^{r} (\hat{\eta}_i)^2 p(M_i \mid D) \right) - (\hat{\eta})^2 .$$

Because the elasticity estimate given in (6) is based on all models under consideration, it may be considered robust to the uncertainty about the underlying model. Consequently, an elasticity such as (6) may be considered a meta-elasticity.

[1] Throughout p(·) denotes a density function and P(·) a cummulative distribution function.

An Application to Meat Demand in the United States

Many recent studies have modeled retail demand for meat in the United States. The researchers conducting these studies have identified a myriad of factors that may be important determinants of meat demand. Some examples include branded advertising (Brester and Schroder 1995), information on the health impacts of cholesterol (Kinnucan *et al.* 1997), and increasing participation of women in the labor force (McGuirk et al. 1995). But the debates surrounding meat demand always come back to the same fundamental issue: the model specification problem. While there has been much discussion of the model specification problem in meat demand (e.g., Alston and Chalfant AJAE 1993; Davis AJAE 1997; McGuirk et. al JARE 1995; Kinnucan, et al. AJAE 1997) all of this analysis has been done within a classical framework. In the classical statistical framework, decisions regarding model specification are made on an "either-or" basis. After trying various possibilities and conducting various misspecification tests on each of these possible models, the researcher selects the single model believed to be most appropriate. Conducting multiple misspecification tests, however, compounds the nominal significance levels that were used in the individual tests. That is to say, the researcher who employs more than a very small number of such tests cannot be very confident in a concluding there is no model misspecification. This suggests that analysis of meat demand could benefit from a methodology that explicitly acknowledges that no single model can be confidently declared to be the "true" model. The Bayesian framework provides such a methodology.

Following the literature, we consider conditional demand systems consisting of the demand for beef, chicken, pork, and fish. As indicated, the model space is determined by the functional forms considered and the variables within each functional form. At this point, two of the most popular demand systems are considered: the Rotterdam (RDAM) and First differenced AIDS (FAIDS).

Neves(1994) has demonstrated how these demand systems are closely related. An important component in connecting these demand systems is the total differential of the expenditure share w_i

(8) $dw_i = w_i \, d \ln q_i + w_i \, d \ln p_i - w_i \, d \ln E$

with q_i, p_i, and E representing the per capita quantity and price on the ith good and E is the total expenditure on beef, chicken, pork, and fish. The Rotterdam model has the form

(9) $w_i \, d \ln q_i = \mu_i(d \ln E - d \ln P) + \sum_j \pi_{ij} \, d \ln p_j \qquad i = 1, 2, 3, 4,$

where μ_i is the constant marginal budget share for good i, π_{ij} is the price parameter, and P is the Divisia price index. Defining the parameters, $b_i = \mu_i - w_i$ and $\gamma_{ij} = \pi_{ij} - w_i w_j + w_i \delta_{ij}$, δ_{ij} being the Kronecker delta, the first difference version of the AIDS model using (8) and (9) is

(10) $w_i \, d \ln q_i = b_i(d \ln E - d \ln P) + \sum_j \gamma_{ij} \, d \ln p_j + w_i(d \ln E - d \ln p_i) \quad i = 1, 2, 3, 4.$

Expressed in these forms, the only difference between the RDAM model and the FAIDS model is the extra term $w_i(d \ln E - d \ln p_i)$, which will be called the defining term. From the two demand systems, price elasticities and expenditure elasticities can be calculated. Table 1 gives the elasticity formulas associated with each of the demand systems.

Data

The data used here are quarterly observations of all variables for 1976 through 1993. Per capita beef, pork, and poultry quantities and retail prices were obtained from Kinnucan *et* *al.*, with the original sources being Putman and Allshouse and the USDA's *Livestock and Poultry Situation and Outlook Report*. The fish quantity and price series are those used by Kinnucan. They were constructed using data from various sources. A cholesterol index intended to measure the impact of health information knowledge on demand is also considered, and also comes from Kinnucan, et al. Branded and generic advertising data were obtained from Brester and Schroeder, and following McGuirk, et al, a women's participation in the labor force, as percent of employment was obtained from the Bureau of Labor Statistics. See Kinnucan, et al., and Brester and Schroeder for discussions of the data.

Aside from the defining term, we define the full design matrix ***X*** to consist of 13 variables: four price variables, the total expenditure variable, a contemporaneous and lagged branded advertising variable for beef, pork, chicken, and fish, a contemporaneous and lagged generic advertising variable for beef and pork, a cholesterol information index variable, a women's participation in the labor force variable, and three quarterly dummy variables. All of these are expressed in log differential form except for the dummy variables. The prices, total expenditure, and dummy variables are taken as certain and included in all models. The other six variables are considered questionable and are allowed to be included and excluded. These variables are included or excluded in the following way. If advertising is included in a model, then all types of advertising are included (e.g., beef, pork, chicken, and fish branded advertising). Lagged advertising is not included without contemporaneous advertising. These restrictions lead to nine valid combinations of advertising variables for each combination of the two remaining variables, for a total of $36 = 9 \times 2^2$ possible model specifications for each of the two systems or $|M| = 72 = 36 \times 2$ possible models.

Priors and Computations

The estimation is based on the seemingly unrelated regression (SUR) framework of Zellner and the Bayesian treatment of the model can be found in Zellner (1971, chapter 8). Using standard notation, the jth demand system is expressed as

(11) $\mathbf{y} = \mathbf{X}_j\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j \qquad j = 1,2,...,144,$

(12) $\boldsymbol{\varepsilon}_j \mid \mathbf{X}_j \sim N_{3T}(\mathbf{0}, \boldsymbol{\Sigma}_j \otimes \boldsymbol{I})$

where $\mathbf{y}$ is $3T \times 1$, $\mathbf{X}_j$ is $3T \times k_j$, $\boldsymbol{\beta}_j$ is $k_j \times 1$, $\boldsymbol{\varepsilon}_j$ is $3T \times 1$, $\boldsymbol{\Sigma}_j$ is the 3×3 positive definite matrix and $\mathbf{I}$ is $T \times T$. As priors, we assume that

(13) $\boldsymbol{\beta}_j \sim N(\underline{\boldsymbol{\beta}}_j, \underline{\boldsymbol{H}}_j^{-1})$

(14) $\boldsymbol{H}_j \sim W(\underline{\mathbf{S}}^{-1}, \underline{v})$

where the underscore indicates the prior, $\underline{\boldsymbol{H}}_j^{-1}$ is the $k_j \times k_j$ precision matrix on the parameter vector prior, W refers to the Wishart distribution with mean $\underline{v}\, \underline{\mathbf{S}}^{-1}$ and degrees of freedom $\underline{v}$. This is the standard representation for the informative prior case in the multiple equation model and is usually referred to as the Normal-Wishart prior.

Priors

For all models we center our prior on $\boldsymbol{\beta}_j$ at zero, except for the parameter for the model defining term, which we center at one. For example, the prior on the parameters in equation (10) with only expenditures, prices, and the model defining term, would be (0, 0, 0, 0, 0, 1). The one in this prior on the mode defining term comes from the fact that we want the model defining term in every model with a one as its parameter and we will control this with a very small variance in the precision prior. For the precision matrix $\underline{\boldsymbol{H}}_j^{-1}$ we use a block diagonal form

(15) $$\underline{H}^{-1} = \begin{bmatrix} g(\boldsymbol{X}'_{j-1}\boldsymbol{X}_{j-1})^{-1} & 0 \\ 0 & c \end{bmatrix}$$

where the upper block is the Zellner g-prior for the precision matrix for all parameters except that parameter corresponding to the model defining term, which has a precision prior $\boldsymbol{c}$. For the Wishart prior in the present setting, the $\underline{\mathbf{S}}^{-1}$ will be a 3×3 and $\underline{v}$ the degrees of freedom parameter. In the present context, g, c, $\underline{\mathbf{S}}^{-1}$and $\underline{v}$ are hyper-parameters to be chosen. We do not have strong priors and do not want to impose any strong priors so we specify the priors to allow for large dispersion. Following the advice of Fernandez, Ley, and Steel we chose g = 1/ 3T ≈ .05, c = .00001 (a very small variance on the defining term parameter), $\underline{\mathbf{S}}$ = $diag$(.0001,.0001,.0001) and $\underline{v} = k_j$. These priors remain the same across models.

Computations

The Bayesian Analysis, Computation, and Communication (BACC) program developed by John Geweke is implemented (see Koop 1999 for a review). The BACC program uses Monte Carlo importance sampling techniques in generating the prior and posterior distributions. The present analysis is a straightforward application of the normal linear model in BACC. For the Monte Carlo integration, 1,010 samples are drawn for the prior and posterior, and the first 10 were removed. The Monte Carlo algorithm for the prior generates independent draws from the prior distributions. The algorithm for the posterior is a two-block Gibbs sampler, which is given in the BACC manual.[2]

Results

Figure 1 plots the logarithm of the marginal likelihood, described by equation (3), associated with each of the 72 models. The first 36 models are the Rotterdam specification, the last 36 are the First difference AIDS models. Within each set of 36, the first nine contain neither the cholesterol index nor women's labor force participation, the second nine contain the former but not the latter, the third nine contain both, and the last nine contain women's labor force participation but not the cholesterol index. Within each set of nine, the log marginal likelihood for the models generally declines as advertising variables are added. For example, model one contains no advertising variables at all, while model nine contains both contemporaneous and lagged observations of both branded and generic advertising. Thus the advertising variables are responsible for the saw tooth pattern observed in Figure 1.

Posterior probabilities, described by equation (5), were calculated for each of the 72 models. The posterior probability for model number one (this is the Rotterdam model containing none of the "optional" variables) was found to be effectively one, and effectively zero for all other models. Given these results, the meta-elasticities that we calculate are equivalent to those for model number one. We find the following compensated own-price elasticities, with standard errors given in parenthesis: –0.597 (0.133) for beef, -0.773 (0.085) for pork, and –0.169 (0.070) for poultry. Expenditure

[2] The BACC software and manuals are available free at http://www.econ.umn.edu/~bacc/bacc99/. The software is obtainable as a Gauss module and thus all of the Bayesian analysis is done in Gauss.

elasticities are 0.117 (0.111) for beef, 0.180 (0.122) for pork, and 0.169 (0.105) for poultry.

These extreme results are surprising and naturally raise suspicions about the priors and possible program. We have experimented with other models using a different data set and find that the results are not as extreme. However, in his review of BACC, Koop conducted an experiment where he got similar extreme results and McCausland in another setting got similar extreme results via a Monte Carlo experiment. We are in the process of exploring these issues.

Conclusions

The model specification problem is perhaps the Achilles heel of applied econometrics. In this paper we summarize the Bayesian Model Averaging approach to this problem, which incorporates model uncertainty directly into the analysis. In the Bayesian Model Averaging approach, a quantity of interest (e.g., an elasticity) is averaged over models based on the probability of each model occurring within the universe of models considered. We estimate 72 meat demand systems associated with the Rotterdam model and AIDS model and including various combinations of advertising variables, a cholesterol index and women's labor force participation. We find that the basic Rotterdam model including only price terms and the volume index, has a probability of almost one relative to the other 71 demand systems. Consequently the meta-elasticities obtained from averaging over the elasticity estimates from the different models weighted by the probability of the model are the same as though obtained from the basic Rotterdam model.

References

Alston, J. and J. Chalfant. "The Silence of the Lambdas: A Test of the Almost Ideal and Rotterdam models." *American Journal of Agricultural Economics.* 75(1993):304-313.

Brester, G. and T. Schroeder. "The impacts of brand and generic advertising on meat demand." *American Journal of Agricultural Economics*. 77(1995):969-979.

Davis, G. C. "The Logic of Testing Structural Change in Meat Demand: A Methodological Analysis and Appraisal." *American Journal of Agricultural Economics*. 79(1997):1186-92.

Dorfman, J. *Bayesian Economics Through Numerical Methods*: *A Guide to Econometrics and Decision Making with Prior Information*. New York: Springer-Verlag, 1997.

Epstein, R. (1987) *A History of Econometrics*. New York: North Holland.

Fernandez, C., L. Eduardo, and M. Steel. "Benchmark Priors for Bayesian Model Averaging." *Journal of Econometrics*. 100(2001):381-427.

George , E. I. and R. E. McCulloch. "Variable Selection Via Gibbs Sampling." *Journal of American Statistical Association*. Vol. 88. No. 423. (September 1993): 881-89.

Geweke, J. "Simulation Methods for Model Criticism and Robust Analysis." in *Bayesian Statistics 6*: 275-299. J.O. Berger, J.M. Bernado, A.P. David, and A.F.M. Smith (eds.), Oxford: Oxford University Press 1999.

Geweke, J. "Using Simulation Methods for Bayesian Econometric Modelling: Inference, Development, and Communication." *Econometric Reviews*. 18(1999): 1-74.

Hendry, D. *Dynamic Econometrics*. New York: Oxford University Press. 1995.

Hoeting, J.A., D. Madigan, A. Raferty, and C. Volinsky. "Bayesian Model Averaging: A Tutorial." *Statistical Science*. Vol. 14. No. 4. (1999): 230-242.

Kinnucan, H., Xiao, H., Hsia, C. and J. Jackson. "Effects of health information and generic advertising on U.S. meant demand." *American Journal of Agricultural Economics*. 79(1997):13-23.

Koop, G "Bayesian Analysis, Computation, and Communication Software." *J. Appl. Econometrics*. 14(1999):677-89.

McCausland, W. (1999), 'Using the BACC software for Bayesian inference', paper presented at the conference *Computing in Economics and Finance.*

McGuirk, A., P. Driscoll, J. Alwang, H. L. Huang. "System Misspecification Testing and Structural-Change in the Demand for Meats." *J. Agr. Res. Econ* 20(July, 1995):1-21.

Morgan, M. *The History of Econometrics Ideas*. Cambridge: Cambridge University Press. 1990.

Moulton, B. R. "A Bayesian Approach to Regression Selection and Estimation, With Application to a Price Index for Radio Services." *Journal of Econometrics*. 49 (1991):169-193.

Neves, P.D. "A Class of Differential Demand Systems," *Economics Letters*. 44(1994):83-86.

Pagan, A. "Three Econometric Methodologies: A Critical Appraisal." *Journal of Economic Surveys*. 1(1987):3-24.

Putman, J. and J. Allshouse. *Food Consumption, Price, and Expenditure, 1970-1990,* Washington, D.C.: U.S. Department of Agriculture, ERS Statistical Bulletin No. 840.

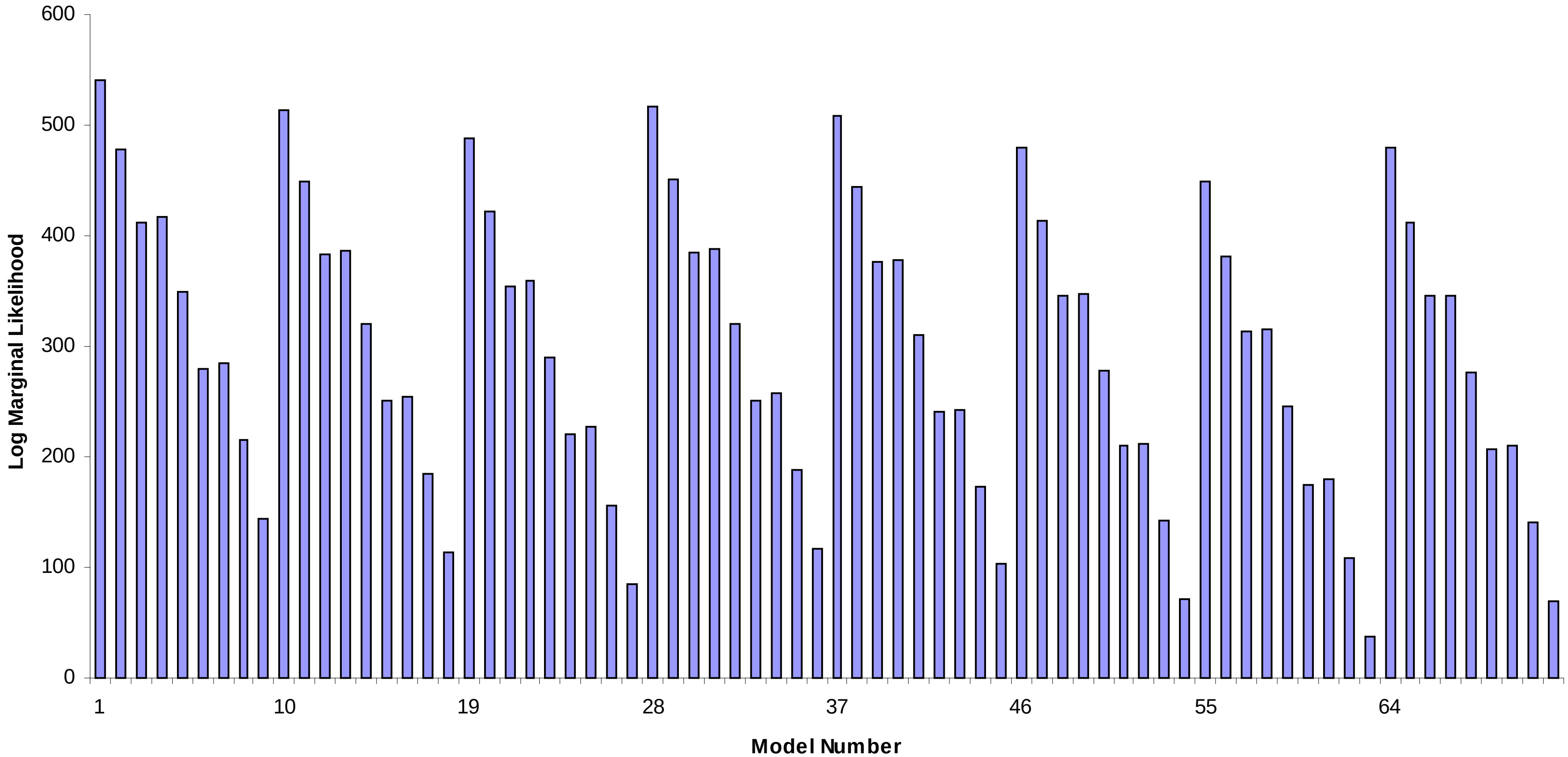
Raferty, A. E., D. Madigan, and C. Volinsky. "Accounting for Model Uncertainty in Survival Analysis Improves Predictive Performance," in *Bayesian Statistics* 5, J. M. Bernardo, J. O. Berger, A. P. David, and A. F. Smith, Eds. Oxford: Oxford University Press, 1996.

Raferty, A. E., D. Madigan, and J. Hoeting. "Bayesian Model Averaging for Linear Regression Models," *Journal of the American Statistical Association*. 92(1997):179-191.

Spanos, A. . *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data*. Cambridge University Press, Cambridge, 1999.

Zellner, A. *An Introduction to Bayesian Inference in Econometrics*. Wiley, New York, NY. 1971.

Figure 1: Log marginal likelihood estimates



This figure plots the log marginal likelihood values reported by the BACC software for each model. Models 1 thorugh 36 are Rotterdam models, models 37 through 72 are AIDS models.