ECONOMETRICA

VOLUME 53 MARCH 1985 NUMBER 2

BAYESIAN ECONOMETRICS

BY ARNOLD ZELLNER[1]

The widespread use of prior information in formulating, estimating, and using econometric models is reviewed. Attempts to avoid the use of prior information by formulating multivariate statistical VAR and ARMA time series models for economic time series data have resulted in heavily over-parametrized models. A simple demand, supply, and entry model is presented to contrast models utilizing prior information provided by economic theory and other sources with multivariate statistical time series models. Formal Bayesian methods for incorporating prior information in econometric estimation, testing, and prediction are presented. A number of published applied Bayesian studies are cited in which Bayesian methods have proved to be effective. It is concluded that wise use of the Bayesian approach will produce improved econometric results.

1. INTRODUCTION

I AM GRATEFUL to have this opportunity to share some of my thoughts on Bayesian econometrics with you. Before doing this, I would like to say that I have great admiration and respect for the work of Irving Fisher and Henry Schultz. Henry Schultz, who spent many years at the University of Chicago, made many significant research contributions. Similarly, Irving Fisher's research has had a profound effect on economics and econometrics. While I could spend the entire lecture attempting to summarize their research, I shall just emphasize that both of them produced key results relating to relatively simple models that have endured over the years. For example, Fisher put forward the famous Fisher equation that relates the nominal interest rate to the anticipated real rate and the anticipated rate of inflation. Often, when I am asked, "Are there any laws in economics?" I point to the Fisher equation as an example. Schultz worked on the laws of supply and demand, relatively simple relationships that are additional examples of sophisticatedly simple laws in economics. I shall discuss the role of "simplicity" in model-building later in my lecture.

Since the early 1960's, Bayesian econometrics has developed rapidly with considerable interaction among econometricians and statisticians. After the publication of Bayes' essay in 1763, many statisticians and scientists including Laplace, Edgeworth, Jeffreys, de Finetti, Box, Savage, Lindley, Raiffa, Jaynes, Schlaifer, Pratt, Good, Geisser, DeGroot, Tiao, and others have made important contributions to the development of Bayesian statistics. As regards modern Bayesian econometrics,[2] it commenced in the early 1960's with papers by Jacques Drèze, Tom Rothenberg, Walter Fisher, Albert Ando, Gordon Kaufman, myself, and a few others. Since then, there has been an upswell of work in Bayesian econometrics

[1] This is a written version of my Fischer-Schultz Lecture, presented at the European Econometric Society Meetings, Pisa, Italy, August 29–September 2, 1983. My research was financed in part by the National Science Foundation and by income from the H. G. B. Alexander Endowment Fund, Graduate School of Business, University of Chicago.

[2] See Zellner [61] for a review of the development of modern Bayesian econometrics.

with a good deal of interaction among econometricians and statisticians.[3] Now we have many published Bayesian monographs and papers. Also, recent and current econometrics texts include varying amounts of Bayesian material and Bayesian computer programs are available. Then too, the Bayesian learning model has been utilized in many works in economic theory.[4]

During the last twenty years, I have found econometricians and economists to be very interested in Bayesian ideas and methods. Contrary to some theories of how science progresses, it was the case that many were quite interested in learning about Bayesian ideas and methods as long as they were given evidence that the new methods work well in practice. This pragmatic attitude on the part of the profession is very commendable. I always like to learn about new ideas and approaches but the bottom line is how well they work in practice. This pragmatic dictum gets a lot of weight in the thinking of many econometricians and economists.

In view of the importance of the pragmatic dictum, mentioned above, I shall not describe in detail the various axiom systems that have been put forward to justify the Bayesian approach. Harold Jeffreys [**26**], a leading natural scientist, has an axiom system in his book, *Theory of Probability*, that leads him to conclude that the only sound approach to inference in science is the Bayesian approach based on a "reasonable degree of belief" definition of probability. In addition, Savage's [**49**] axiom system, put forward in his *Foundations of Statistics*, implies that procedures for acting rationally in making decisions under uncertainty are Bayesian procedures. While these axiom systems are not perfect, they do exist and may be of interest to some of you. As an aside, Savage and Milton Friedman were very good friends and both admirers of Bruno de Finetti's work. Savage and Friedman co-authored two well known papers on choice under uncertainty.[5] These papers brought to the fore the importance of economic utility theory in statistics by emphasizing the central role of maximization of expected utility or equivalently minimization of expected loss in making choices under uncertainty, for example in choosing estimates, predictions, and models. Their work helped to introduce economists and statisticians to the intimate connection between utility theory and statistical theory. There is thus considerable axiomatic support for Bayesian inference and decision procedures that have been applied in analyzing many problems in econometrics, for example estimation, prediction, testing, control, and other decision problems with very good results.

[3] Since 1970, semi-annual meetings of the NBER-NSF Seminar on Bayesian Inference in Econometrics have been held at which Bayesian research papers have been discussed by econometricians and statisticians. Some of these papers have been published in Fienberg and Zellner [**13**] and Zellner [**62**].

[4] See, e.g., papers and references in Fienberg and Zellner [**13**] and Zellner [**62**]. Also see the book edited by M. Boyer and R. E. Kihlstrom, *Bayesian Models in Economic Theory*, published in the North-Holland Publishing Company's *Series on Bayesian Econometrics and Statistics*, 1983.

[5] These papers, "The Utility Analysis of Choices Involving Risk" and "The Expected-Utility Hypothesis and the Measurability of Utility" have been reprinted in *The Writings of Leonard Jimmie Savage*, American Statistical Association: Washington, D.C., 1981.

Rather than dwell on axiomatic matters, I shall consider more down-to-earth matters. Why are we interested in the Bayesian approach? My point of view is that in econometrics and in other areas of science a good deal of prior information is employed, usually informally. The Bayesian approach provides formal methods for handling background or prior information. Some of you may say that this is subjective and not objective. I believe that the words, "subjective" and "objective" are loaded words fraught with difficulty. I refer to the philosopher Bishop Berkeley who was a philosophical idealist. He claimed, rightly in my opinion, that it is impossible to prove that the "real" world exists. On the other hand, there are philosophical realists who believe, without proof, that the "real" world exists and that we as scientists can learn about its properties. Finally, there are solipsists who claim that only I exist and nothing else exists. Since it's impossible to settle these issues, who is "objective" and who is "subjective"? My position is that the best we can do at present is to build models and determine how well they perform in explaining past data, predicting as yet unobserved data, and analyzing policy problems. The point that I make to you, very strongly, is that non-Bayesians, as well as Bayesians, use a good deal of prior information in building and using models. Further, since the use of prior information is unavoidable, it should be used carefully and more formally than has been done in the past. Finally, it appears to me that most econometricians usually do not use a frequency concept of probability in model-building, estimation, testing, and prediction, but rather a degree of reasonable belief concept of probability.[6]

To illustrate some of these points, let us consider modeling of time series data as, for example, in demand and supply modeling or macroeconometric modeling. How do we generally approach the modeling problem? Some take the Jeffreys–Wrinch Simplicity Postulate, Ockham's Razor, and the Principle of Parsimony very seriously and formulate relatively simple models with the *belief* that they will probably work well. If there is a need to capture much detail, these workers would formulate a large, relatively simple model and investigate the extent to which it is adequate and accomplishes the objectives of a modeling project. Those who advocate sophisticated simplicity in model-building include Harold Jeffreys, Ronald Fisher, Milton Friedman, George Box, myself, and many others. Note that there are many examples of sophisticatedly simple models that work reasonably well, for example, Newton's laws, the gas laws, Einstein's laws, Schrödinger's equation, laws of supply and demand, the Fisher equation, Friedman's consumption model, the quantity theory of money, and so on. On the other hand, there are others who *believe* that complicated models will probably work well in practice. Witness the large number of macroeconometric models that have hundreds of nonlinear stochastic difference equations containing thousands of parameters. It is not clear whether such models have unique or many solutions. I have challenged many audiences to give me examples of large, complicated models that work well in any area of science and have not yet heard of any. A main point to be appreciated in this discussion is that *prior beliefs* about the

[6] This point has also been made with respect to physical scientists in Jeffreys [**26**].

simplicity/complexity issue have a vital impact on the way in which econometric models are formulated.

To pursue this last point further and more explicitly, let me review several types of linear time series models that are currently being studied and used. The first is the multiple time series model or multivariate autoregressive moving average (MVARMA) model put forward by Quenouille [**48**] and others. The MVARMA model for a mean-corrected $p \times 1$ vector of variables, z_t, is given by

$$H(L)z_t = F(L)e_t \qquad (t = 1, 2, \ldots, T), \tag{1}$$

where $H(L)$ and $F(L)$ are matrix lag operators given by

$$H(L) = H_0 + H_1 L + H_2 L^2 + \cdots + H_r L^r,$$
$$F(L) = I + F_1 L + F_2 L^2 + \cdots + F_q L^q,$$

with L the lag operator $(L^i z_t = z_{t-i})$ and e_t is a $p \times 1$ non-serially correlated error vector with zero mean and $p \times p$ covariance matrix Σ.

The issue is how successful will we probably be in using a MVARMA model given the types of samples that we often have in econometrics, say 20 years of quarterly data? One relevant consideration is the number of parameters in the model. This number, denoted by N is given by

$$N = (r+1)p^2 + qp^2 + p + p(p+1)/2,$$

that is, $(r+1)p^2$ parameters in the H_i matrices, qp^2 in the F_i matrices, p means, and $p(p+1)/2$ distinct parameters in the error covariance matrix. Clearly, the value of N, the number of parameters will depend on the values assigned to p, r, and q. Some may consider a model with 30 variables ($p = 30$) to be a small model but for illustration I shall assume that just six variables ($p = 6$) appear in the model, $r = 3$, and $q = 4$. With these assumed values for p, q, and r, there are 395 parameters in the model. Suppose that we have 20 years of quarterly data on the six variables, that is a total of 480 observations. The observation/parameter ratio is 480/395, not even two observations per parameter which is hardly satisfactory because with so few observations, parameter estimates and predictions will generally be imprecise. Also, if tests are performed to determine whether elements of the coefficient matrices are equal to zero, they probably will not be very powerful and will have an impact on subsequent inferences, the problem of pre-testing. These then are some of the consequences of entertaining a relatively complicated model. With nonlinear versions of the above model, there would be even greater difficulties. There is a need for restrictions or prior information, perhaps derived from economic theory, that can effectively reduce the number of parameters.

Some in economics and statistics have not employed MVARMA processes but have utilized vector autoregressions. Of course, if the matrix lag operator, $F(L)$ in the MVARMA model is invertible, the MVARMA process has an infinite autoregressive representation, $F^{-1}(L)H(L)z_t = e_t$, or

$$(I - G_1 L - G_2 L^2 - \cdots - G_n L^n - \cdots)z_t = e_t.$$

If we cut off the autoregressive operator at the term involving L^n, that is we consider an nth order vector autoregression to represent our data, note that each of the G matrices is $p \times p$ and we have n of them. Thus there are np^2 parameters in the G matrices, $p(p+1)/2$ in the error term covariance matrix, and p means. If again we assume that we have 6 variables and that $n = 10$, perhaps a reasonable value for quarterly, seasonally unadjusted data, then the system contains 387 parameters. If we have 20 years of quarterly data on the 6 variables, there are 480 observations and the observation/parameter ratio, 480/387 is abysmally low. Further, consider a single equation of the above vector autoregression, say that for z_{it}. There are $np = 10 \times 6 = 60$ coefficients in this relation. Fitting by least squares is like fitting a regression with 60 independent variables that are usually highly correlated. The precision of estimation and prediction is not going to be very high. Indeed Litterman [**35**] has fit such relations with U.S. quarterly macroeconomic data and found that estimates and predictions are not very precise, the latter relative to predictions yielded by Box-Jenkins models, univariate autoregressions, and several macroeconometric models. Thus an attempt to determine relations solely from the data, this time by use of unrestricted vector autoregressions is foiled by the appearance of an inordinate number of parameters in the system. Again there is a need for subject matter information, or prior information, that will help to reduce the number of free parameters.

Sometimes economic theory or prior information suggests that some of the variables in a MVARMA process are exogenously determined. If we partition the z_t vector as follows, $z_t' = (\boldsymbol{y}_t', \boldsymbol{x}_t')$, we can write the MVARMA model in the following form:

$$\begin{pmatrix} H_{11}(L) & H_{12}(L) \\ H_{21}(L) & H_{22}(L) \end{pmatrix} \begin{pmatrix} \boldsymbol{y}_t \\ \boldsymbol{x}_t \end{pmatrix} = \begin{pmatrix} F_{11}(L) & F_{12}(L) \\ F_{21}(L) & F_{22}(L) \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_{1t} \\ \boldsymbol{e}_{2t} \end{pmatrix}. \tag{2}$$

The assumption that the variables in $\boldsymbol{x}_t$ are exogenous implies, as pointed out in Zellner and Palm [**64**], that

$$H_{21}(L) \equiv 0, \quad F_{12}(L) \equiv 0, \quad \text{and} \quad F_{21}(L) \equiv 0.$$

Thus 3 matrix lag polynomials are identically zero, given that $\boldsymbol{x}_t$ is exogenous. With these restrictions imposed, the above system becomes

$$H_{11}(L)\boldsymbol{y}_t + H_{12}(L)\boldsymbol{x}_t = F_{11}(L)\boldsymbol{e}_{1t}, \tag{3}$$

$$H_{22}(L)\boldsymbol{x}_t = F_{22}(L)\boldsymbol{e}_{2t}. \tag{4}$$

The first set of relations in (3) is in the form of a linear dynamic structural econometric model with $\boldsymbol{y}_t$ endogenous and $\boldsymbol{x}_t$ exogenous. The second set of relations in (4) is a model for the random exogenous variables that is in the MVARMA form. The question has arisen as to whether a model for $\boldsymbol{x}_t$ is needed. For certain purposes it is not needed. For example in estimation of the structural equations, provided that they are identified, one can proceed conditional upon the observed values of the $\boldsymbol{x}_t$'s. However, if the system is to be used in forecasting, the future values of $\boldsymbol{x}_t$ are needed. Currently in forecasting with structural econometric models, it is often not clear how the future values of exogenous variables are obtained. In many cases workers *guess* the future values of the exogenous variables, a very important, informal use of prior information that usually has a vital impact on forecasts and forecast intervals. If the exogenous variables are generated by a MVARMA process, as in (4), it can be used to forecast their future values.

If in the above system, 2 of the 6 variables are assumed exogenous and if we use the same assumptions as above, namely $r=3$ and $q=4$, the total number of free parameters is 183. While this is an improvement relative to the unrestricted MVARMA process, the observation/parameter ratio, 480/183 is still quite low. However, we generally go further by assuming that particular variables do not appear in some of the equations of (3), assumptions that lead to zero restrictions which are important in achieving identification and representing our economic knowledge about the forms of structural relations. Such restrictions, as well as the assumption that some variables are exogenous, have come under vigorous attack in the recent literature as being "unrealistic." I do not believe that this is a reasonable criticism of these systems because in many applications these restrictions are reasonable and if there is some doubt about them, many of them can be tested.

To make this last point clearly, let's come down from these abstract algebraic systems to a very simple demand and supply system, the kind that Henry Schultz and many others have used. In this connection, I shall emphasize another area, that of model formulation in which prior information is very important. In formulating an initial variant of a model, say a demand and supply model, we all use much prior information in sorting through combinations of ideas. In this process, the role of unusual facts has been emphasized by Jacques Hadamard [**21**] in his book, *The Psychology of Invention in the Mathematical Field*. From a survey of his fellow mathematicians, he found that most of them made their major research breakthroughs by discovering unusual facts and providing explanations of them. In economics, econometrics and other sciences there are a number of examples of this process. Given that unusual facts are so important, in a recent paper (Zellner [**63**]) I suggested 8 ways to produce unusual facts, one of which is to look at the equations of current macroeconometric models. Almost every equation contains one or more unusual facts. I have had several students consider such unusual facts and investigate them using micro and industry data with very illuminating results—see Chau [**8**], Laub [**29**, **30**], Levedahl [**33**], Montmarquette [**38**, **39**], and Peck [**44**, **45**]. Another way to produce unusual facts is to push current theories to extremes. For example, with respect to the proportionality hypothesis of Friedman's [**16**] consumption theory, consider what happens as permanent income approaches zero. The ratio of permanent consumption to permanent income probably rises toward one as I suggested years ago (Zellner [**57**]) and recently work with low income Indian consumers has supported this view—see Bhalla [**4**]. Also, at very high permanent incomes, the proportion of permanent income consumed probably falls. Thus, while proportionality is a pretty good approximation over the mid-range of permanent incomes, at very low and very high levels of permanent incomes, it probably breaks down. Discovering such unusual facts and rationalizing them can lead to significant new theories of consumption and saving—see, e.g., Gersovitch [**18**].

Recently, with respect to theoretical and applied demand and supply studies, I noticed some very unusual facts. In Muth [**41**] and some other papers on rational expectations models of competitive industries, the number of firms is assumed to be a constant. On the other hand, in most applied econometric demand-supply analyses, the number of firms variable does not appear explicitly in estimated models. As is well known, Alfred Marshall and many economists have emphasized the important role of entry and exit of firms in producing long run equilibrium in a competitive industry. Also, simple aggregation considerations indicate that the number of firms in operation is a variable that should appear in a competitive industry's supply function. With these somewhat unusual facts in mind, William Veloce, a current student of mine, and I have formulated a three equation demand and supply model (Veloce and Zellner [**55**]) containing demand, supply, and net entry equations for the following endogenous variables:

$$y_{1t} = (1-L)\log Q_t, \qquad y_{2t} = (1-L)\log P_t, \qquad \text{and}$$
$$y_{3t} = (1-L)\log N_t,$$

where Q_t, P_t, and N_t, are quantity, real price, and the number of firms in operation, respectively in period t. The three equations of our demand, supply, and entry (DSE) model of a competitive industry are:

$$\text{Demand:}\quad y_{1t} = a_1 y_{2t} + \boldsymbol{b}_1'\boldsymbol{x}_{1t} + u_{1t}, \tag{5}$$

$$\text{Supply:}\quad y_{1t} = a_2 y_{2t} + y_{3t} + \boldsymbol{b}_2'\boldsymbol{x}_{2t} + u_{2t}, \tag{6}$$

$$\text{Net Entry:}\quad (1-cL)y_{3t} = \boldsymbol{b}_3'\boldsymbol{x}_{3t} + u_{3t}, \quad 0<|c|<1, \tag{7}$$

where $\boldsymbol{x}_{1t}$ and $\boldsymbol{x}_{2t}$ are vectors of exogenous variables appearing in the demand and supply equations, respectively, $\boldsymbol{x}_{3t}$ is a vector of predetermined variables appearing in the net entry equation, the a's, b's, and c are structural parameters, and u_{1t}, u_{2t}, and u_{3t} are zero-mean, serially uncorrelated disturbance terms.

From the net entry relation, we can solve for y_{3t}, that is

$$y_{3t} = (\boldsymbol{b}_3'\boldsymbol{x}_{3t} + u_{3t})/(1-cL) \tag{8}$$

and substitute this expression in the supply equation (6) to obtain the following "reduced" two equation system:

$$y_{1t} = a_1 y_{2t} + \boldsymbol{b}_1'\boldsymbol{x}_{1t} + u_{1t}, \tag{9}$$

$$(1-cL)y_{1t} = a_2(1-cL)y_{2t} + \boldsymbol{b}_3'\boldsymbol{x}_{3t} + \boldsymbol{b}_2'(1-cL)\boldsymbol{x}_{2t} + u_{3t} + (1-cL)u_{2t}. \tag{10}$$

This reduced two equation system is similar to those that have been estimated in applied econometric studies. Note that equation (9) is exactly the demand equation (5) of the three equation DSE system. However, the second equation (10) of the two equation reduced system is a mixture of the supply and net entry equations, (6) and (7) of the three equation DSE model. It contains lagged endogenous variables, autocorrelated disturbances, and predetermined variables and coefficients from the supply and net entry equations. In empirical work, lags and autocorrelated disturbances in an equation like (10) have been rationalized in terms of partial supply adjustment and/or adaptive expectations effects—see, e.g., Nerlove [**42**]. Here the lags and autocorrelation appear as a consequence of considering the three equation DSE model to be an appropriate model. In current work with annual data for a Canadian industry, Veloce has found that a three equation DSE model, like the one above, provides a good interpretation of the data.

Let us consider the number of variables in the three equation system, (5)-(7). There are three endogenous variables and the variables in three x vectors. If each of these vectors contains three exogenous variables, then there are 12 variables in the system. Consider modeling these 12 variables in a vector autoregressive approach. Each matrix in the autoregressive system is 12×12, that is contains 144 parameters. Further, the 12×12 disturbance covariance matrix contains 78 distinct elements. It is clear that with about 30 annual observations for each variable that a vector autoregressive approach, as well as an unrestricted MVARMA approach, will probably be very unsatisfactory. If we omit some of the relevant variables, we shall be analyzing marginal processes, like the two equation system (9)-(10), and it will be very difficult to interpret estimation results in behavioral terms. These then are some of the practical consequences of employing approaches that fail to take account of relevant, available economic information, that is prior information.

What is the economic information that has been employed to produce the restrictions on the three equation model above? First, for example, it is assumed that the number of firms variable does not appear in the demand equation, certainly a reasonable restriction. Also, variables such as the real wage rate and the price of capital services that might be included as exogenous variables in the supply equation probably do not belong in the demand equation. These restrictions and others like them have been embedded in the three equation model which contains about 20 parameters. If some of these restrictions seem doubtful, they can be tested. In these testing problems as well as in formulating the three equation model, it is apparent that much prior information is used. It is clearly important to use such prior information carefully and to have good methods for incorporating it flexibly in estimation, testing, forecasting, and policy analysis procedures.

To provide another example of a situation in which much prior information is used and has a vital impact on the results of analyses, consider the topic of seasonal adjustment. In much work on seasonal adjustment (see, e.g., Zellner [**59**]) an observed variable, y_t, is assumed to be represented as the sum of unobservable random trend-cycle, seasonal, and noise components, all assumed mutually uncorrelated. There does not seem to be any economic theory that suggests that the random trend-cycle component is uncorrelated with the random seasonal component. Yet this a priori assumption is employed to produce seasonally adjusted variables that are used in econometric analyses. See

Havenner and Swamy [**24**] for an analysis that casts doubt on the adequacy of this assumption.

The discussion presented above indicates that much prior information is needed and used in work with econometric models. Thus good procedures for incorporating prior information in analyses are needed. Fortunately, the Bayesian approach is one in which prior information, as little or as much as an investigator desires, can be flexibly and formally incorporated in estimation, prediction, testing, model selection, and policy analysis procedures. Let me briefly review these procedures and comment on some of their salient features.[7]

First, with respect to estimation, Bayes' Theorem yields the general result that a posterior probability density function (pdf) for the parameters of a model is proportional to a prior pdf times the likelihood function; that is,

$$\text{Posterior pdf} \propto (\text{Prior pdf}) \times (\text{Likelihood Function})$$

or

$$(11) \qquad p(\boldsymbol{\theta}|D) = cp(\boldsymbol{\theta}|I_0)p(\boldsymbol{y}|\boldsymbol{\theta}), \quad \boldsymbol{\theta} \subset \Theta, \quad \boldsymbol{y} \subset R_y,$$

where $\boldsymbol{\theta}$ is a vector of parameters, Θ the parameter space, $\boldsymbol{y}$ a vector of observations, R_y the sample space, I_0 the prior information, $D = (\boldsymbol{y}, I_0)$, c a normalizing constant given by $c^{-1} = \int_\Theta p(\boldsymbol{\theta}|I_0)p(\boldsymbol{y}|\boldsymbol{\theta})\, d\boldsymbol{\theta}$, $p(\boldsymbol{\theta}|I_0)$ the prior pdf, and $p(\boldsymbol{y}|\boldsymbol{\theta})$ the likelihood function. Thus for any estimation problem, given a prior pdf and a likelihood function, Bayes' Theorem provides an *exact*, *finite sample* posterior pdf for the parameters.[8]

The prior pdf, $p(\boldsymbol{\theta}|I_0)$, in (11) can be chosen to represent available prior information about possible values of $\boldsymbol{\theta}$, a formal, probabilistic representation. Note that the ranges of the elements of $\boldsymbol{\theta}$ can be restricted to particular intervals, say zero to one if this is suggested by economic theory or other considerations. Further, (11) can be viewed as a transformation of the prior pdf into a posterior pdf that incorporates both prior and sample information. In addition, assuming a pdf for $\boldsymbol{\theta}$ is closely related to assuming a pdf for parameters in non-Bayesian analyses of random parameter models such as random regression coefficient models, the structural form of the errors in the variables model, etc. The adequacy of the prior pdf for $\boldsymbol{\theta}$ can be assessed by use of various diagnostic checks including posterior odds, described below, for alternative priors, and most importantly by determining the quality of predictions yielded by a model incorporating the information in the prior pdf as was done in Litterman [**35**]. Of course, in structural econometric models, prior information must be introduced to identify structural parameters. On many occasions, Drèze [**11, 12**] and others have emphasized that this needed prior information can be flexibly and satisfactorily represented by use of a prior pdf.

[7] Some works that provide detailed presentations of Bayesian methods are Box and Tiao [**6**], DeGroot [**10**], Jeffreys [**26**], Leamer [**31**], Lindley [**34**], and Zellner [**58**].

[8] In Baxter [**3**], a posterior pdf for the mean rate of growth of the money supply is computed and employed in an imaginative way to provide a measure of the public's degree of confidence in a monetary reform.

Analytical or numerical integration procedures are available for analyzing properties of the posterior pdf in (11). For example, if $\boldsymbol{\theta}' = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')$, the marginal posterior pdf for $\boldsymbol{\theta}_1$ is

$$(12) \qquad p(\boldsymbol{\theta}_1|D) = \int_{\Theta_2} p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|D)\, d\boldsymbol{\theta}_2 = \int_{\Theta_2} p(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2) p(\boldsymbol{\theta}_2|D)\, d\boldsymbol{\theta}_2$$

where $\boldsymbol{\theta}_2 \subset \Theta_2$.[9] The integration in (12) is a very convenient means of getting rid of nuisance parameters, for example autocorrelation coefficients, disturbance variances and covariances, etc., a problem that is often difficult to solve in non-Bayesian approaches. The marginal posterior pdf for $\boldsymbol{\theta}_1$ in (12) can be used to make posterior probability statements about the elements of $\boldsymbol{\theta}_1$, to compute moments of the elements of $\boldsymbol{\theta}_1$, etc.

If a point estimate for $\boldsymbol{\theta}$ is desired, given that we have a convex loss function, $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, where $\hat{\boldsymbol{\theta}}$ denotes an estimate for $\boldsymbol{\theta}$, the optimal estimate is obtained by minimizing the posterior expectation of the loss function with respect to $\hat{\boldsymbol{\theta}}$, that is by solving the following problem:

$$(13) \qquad \min_{\hat{\boldsymbol{\theta}}} \int_{\Theta} L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) p(\boldsymbol{\theta}|D)\, d\boldsymbol{\theta}.$$

The solution to (13), say $\hat{\boldsymbol{\theta}}^*$ is the Bayesian estimate that is tailored to be optimal relative to the particular loss function that is deemed reasonable. For example, for the quadratic loss function, $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})' A (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})$, with A a given, positive definite symmetric matrix, $\hat{\boldsymbol{\theta}}^*$ = the posterior mean of $\boldsymbol{\theta}$, i.e., $\hat{\boldsymbol{\theta}}^* = \int_{\Theta} \boldsymbol{\theta} p(\boldsymbol{\theta}|D)\, d\boldsymbol{\theta}$. For an absolute error loss function, the median of the posterior pdf is optimal. If other loss functions are employed, the minimization problem in (13) can be solved analytically or numerically to obtain an optimal point estimate. Generally, Bayesian estimates have been shown to have very good sampling properties in both small and large samples. A sufficient condition for Bayesian estimators to be admissible is

$$(14) \qquad \text{Average Risk} = \int_{\Theta} r(\boldsymbol{\theta}) p(\boldsymbol{\theta}|I_0)\, d\boldsymbol{\theta} < \infty$$

where $r(\boldsymbol{\theta}) = \int_{R_y} L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) p(y|\boldsymbol{\theta})\, dy$, the risk function for $\hat{\boldsymbol{\theta}}$ relative to the loss function $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$. When (14) is satisfied, as it usually is when $p(\boldsymbol{\theta}|I_0)$ is a proper prior pdf, the Bayesian estimator, $\hat{\boldsymbol{\theta}}^*$, minimizes average risk given in (14). Thus Bayesian estimators have rather good sampling properties as shown analytically and in Monte Carlo studies.[10]

[9] The conditional posterior pdf, $p(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2, D)$ has been used in many analyses to explore the effects of various assumptions about $\boldsymbol{\theta}_2$'s value on inferences about $\boldsymbol{\theta}_1$. See, e.g., Box and Tiao [**6**] and Zellner [**58**] for examples.

[10] See, e.g., Chetty [**9**], Fomby and Guilkey [**15**], Gibbons [**19**], Griffiths and Dao [**20**], Lee, Judge, and Zellner [**32**], Park [**43**], Surekha and Griffiths [**50**], Thornber [**53**], and Zellner [**58**].

As regards prediction procedures, let $\boldsymbol{y}_f$ denote a vector of as yet unobserved variables with pdf, $p(\boldsymbol{y}_f|\boldsymbol{\theta})$. Since $\boldsymbol{\theta}$'s value is unknown, it is hard to use this pdf in practice. However, the joint pdf for $\boldsymbol{y}_f$ and $\boldsymbol{\theta}$ is $p(\boldsymbol{y}_f|\boldsymbol{\theta})\times p(\boldsymbol{\theta}|D)$, where $p(\boldsymbol{\theta}|D)$ is the posterior pdf for $\boldsymbol{\theta}$ given in (11). Thus the marginal or predictive pdf for $\boldsymbol{y}_f$, $p(\boldsymbol{y}_f|D)$, is readily obtained by integration, that is,

$$p(\boldsymbol{y}_f|D)=\int_{\Theta} p(\boldsymbol{y}_f|\boldsymbol{\theta})p(\boldsymbol{\theta}|D)\,d\boldsymbol{\theta} \tag{15}$$

which can be used to make probability statements about the elements of $\boldsymbol{y}_f$, to compute their moments, etc. Further, if $L(\boldsymbol{y}_f, \hat{\boldsymbol{y}}_f)$ is a convex loss function, where $\hat{\boldsymbol{y}}_f$ is a point prediction, an optimal value of $\hat{\boldsymbol{y}}_f$ can be obtained by minimizing expected loss with respect to $\hat{\boldsymbol{y}}_f$, that is by solving the following problem:

$$\min_{\hat{\boldsymbol{y}}_f} \int_{R_y} L(\boldsymbol{y}_f, \hat{\boldsymbol{y}}_f)p(\boldsymbol{y}_f|D)\,d\boldsymbol{y}_f. \tag{16}$$

The solution to (16), say $\hat{\boldsymbol{y}}_f^*$, is the Bayesian point prediction.[11] If, for example, the loss function is quadratic, then $\hat{\boldsymbol{y}}_f^*$ is the mean of the predictive pdf, $p(\boldsymbol{y}_f|D)$. As with Bayesian estimates, the Bayesian predictor has very good sampling properties in small and large samples.

With respect to large sample properties of Bayesian estimation procedures, it has been shown in the literature that the posterior pdf $p(\boldsymbol{\theta}|D)$ in (11) approaches a multivariate normal pdf as the sample size grows with mean equal to the maximum likelihood estimate, $\hat{\boldsymbol{\theta}}_{\text{MLE}}$, and covariance matrix equal to the inverse of the estimated Fisher information matrix, $\widehat{Inf}^{-1}$, that is $p(\boldsymbol{\theta}|D)\to N(\hat{\boldsymbol{\theta}}_{\text{MLE}}, \widehat{Inf}^{-1})$ as the sample size gets large. Heyde and Johnstone [**25**] and Hartigan [**23**] show that when the observations are i.i.d., the conditions needed for this result are identical to those needed for proving the asymptotic normality of maximum likelihood estimators. However, when the observations are stochastically dependent, as in time series analyses, Heyde and Johnstone [**25**] show that the conditions needed for the asymptotic normality of the posterior pdf are simpler and more robust than those needed to establish the asymptotic normality of the maximum likelihood estimator. While these results are interesting, it should be appreciated that Bayes' Theorem in (11) provides operational, exact finite sample results and thus there is no need in general to rely on approximate, asymptotic results that are often poor approximations in small sample cases.[12]

Bayesian posterior odds ratios are available for comparing and testing hypotheses and for solving model selection problems. For example, consider two mutually exclusive hypotheses, H_1 and H_2 and assume that the prior odds for these two hypotheses is Π_1/Π_2, where Π_i, $i=1,2$, is the prior probability associated with H_i. Then, from Bayes' Theorem, the posterior odds associated with H_1 versus

[11] See Varian [**54**] for a very interesting computation of $\hat{\boldsymbol{y}}_f^*$ for an asymmetric LINEX loss function. Also Litterman [**35**] computes $\hat{\boldsymbol{y}}_f^*$ in his work with Bayesian VAR models.

[12] Zellner and Rossi [**66**] provide analysis to illustrate this point with respect to logit models.

H_2, denoted by K_{12} is given by

$$(17) \qquad K_{12} = (\text{prior odds}) \times (\text{Bayes' factor})$$
$$= (\Pi_1/\Pi_2) p_1(\boldsymbol{y}|H_1)/p_2(\boldsymbol{y}|H_2)$$

where $\boldsymbol{y}$ is an observation vector and $p_i(\boldsymbol{y}|H_i)$ is the pdf for $\boldsymbol{y}$ given H_i, $i = 1, 2$. If H_1 and H_2 are both simple hypotheses, the Bayes' factor, $p_1(\boldsymbol{y}|H_1)/p_2(\boldsymbol{y}|H_2)$ is identical to the likelihood ratio. On the other hand, if H_1 and H_2 are composite hypotheses involving free parameters, $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively, then the Bayes' factor is given by

$$(18) \qquad p_1(\boldsymbol{y}|H_1)/p_2(\boldsymbol{y}|H_2)$$
$$= \int_{\Theta_1} p_1(\boldsymbol{y}|\boldsymbol{\theta}_1, H_1) p_1(\boldsymbol{\theta}_1|I_1)\, d\boldsymbol{\theta}_1 \Big/ \int_{\Theta_2} p_1(\boldsymbol{y}|\boldsymbol{\theta}_2, H_2) p_2(\boldsymbol{\theta}_2|I_2)\, d\boldsymbol{\theta}_2\,,$$

where $p_i(\boldsymbol{y}|\boldsymbol{\theta}_i, H_i)$ is the pdf for $\boldsymbol{y}$ given $\boldsymbol{\theta}_i$ and H_i and $p_1(\boldsymbol{\theta}_i|I_i)$ is the prior pdf for $\boldsymbol{\theta}_i \subset \Theta_i$ given prior information I_i, $i = 1, 2$. Since $p_i(\boldsymbol{y}|H_i)$ can be computed before observing $\boldsymbol{y}$, the Bayes' factor in (18) is in the form of a ratio of predictive pdfs. Bayes' factors have been computed for many hypotheses relating to means, variances, regression coefficients, coefficients of ARIMA models, etc.,[13] and in a number of cases related to sampling theory test statistics and P-values. Also, analyses of more than two hypotheses can be performed and asymptotic properties of posterior odds have been derived. While these asymptotic results are interesting, it should be noted that K_{12} in (17) is an exact finite sample measure of the relative degree of confidence associated with H_1 and H_2. Since non-Bayesian testing procedures do not involve the concept of the probability of a hypothesis, there is no non-Bayesian analogue of the posterior odds, K_{12}.

If the two hypotheses H_1 and H_2 are mutually exclusive and exhaustive, then $K_{12} = P/(1-P)$, where P is the posterior probability associated with H_1. Suppose, for example, that H_1 is $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, with $\boldsymbol{\theta}_0$ a given value of $\boldsymbol{\theta}$, and H_2 is $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. From (17), the value of $K_{12} = P/(1-P)$ can be computed and from it a value of P can be obtained. If our loss function is $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})' A (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})$, with A a given pds matrix, then expected loss is given by

$$EL(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = PL(\boldsymbol{\theta}_0, \hat{\boldsymbol{\theta}}) + (1-P) EL(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})|H_2$$

where E on the right-hand side denotes the expectation based on the posterior pdf for $\boldsymbol{\theta}$ under H_2. On inserting the quadratic loss function and minimizing expected loss with respect to $\hat{\boldsymbol{\theta}}$, the optimal point estimate, $\hat{\boldsymbol{\theta}}^*$ is

$$(19) \qquad \hat{\boldsymbol{\theta}}^* = P\boldsymbol{\theta}_0 + (1-P)\bar{\boldsymbol{\theta}}_2$$
$$= \boldsymbol{\theta}_0 + \frac{1}{1+K_{12}} (\bar{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_0)$$

[13] See e.g., Jeffreys [**26**], Geisel [**17**], Leamer [**31**], Monahan [**37**], and Zellner [**58**].

where $\bar{\boldsymbol{\theta}}_2$ is the posterior mean of $\boldsymbol{\theta}$ under H_2. $\hat{\boldsymbol{\theta}}^*$ is a Bayesian "pre-test" estimate.[14] The second line of (19) involves the "shrinkage" factor $1/(1+K_{12})$ and is in Stein-like form. The analysis can be easily modified to apply to other types of hypotheses and to combining predictions from alternative models.[15] That these procedures are relatively simple and permit flexible incorporation of prior information through the prior odds and prior pdfs for parameters is noteworthy.

With respect to control and other policy analysis problems, the Bayesian approach yields solutions that are operational, reflect uncertainty about parameters' values, and incorporate allowance for sequential learning. To illustrate, let $y_f = x\beta + u_f$ where y_f is a future value of a policy variable, x is a control variable, β is a regression coefficient, and u_f is a future error term with zero mean and variance σ^2. If the loss function is $L(y_f, y_*) = (y_f - y_*)^2$, where y_* is a given target value for y_f, expected loss is given by $E(x\beta + u_f - y_*)^2$, or

$$E(y_f - y_*)^2 = x^2 E\beta^2 - 2y_* x E\beta + Eu_f^2 + y_*^2$$

under the assumption that u_f and β are independent. The value of x that minimizes expected loss is

$$x^* = y_* E\beta / E\beta^2 = (y_*/E\beta)\frac{1}{1+\operatorname{var}\beta/(E\beta)^2} \tag{20}$$

where $E\beta$ and $\operatorname{var}\beta$ are the posterior mean and variance of β, respectively.[16] It is seen that the expression for x^* involves not only $E\beta$ but also $\operatorname{var}\beta$, the posterior variance of β, a measure of uncertainty about the value of β. In the literature, more elaborate one-period and multi-period control problems have been analyzed using Bayesian methods. In multiperiod problems, learning and design considerations enter. Comparisons of Bayesian and non-Bayesian control solutions have generally indicated that Bayesian solutions are better.[17]

Above, some Bayesian procedures and methods have been briefly reviewed. It is relevant to question whether these methods work well in applications. In this regard, it is important to realize that many, if not all, non-Bayesian results can be produced by Bayesian methods under special assumptions. For example, with a large sample assumption, it was mentioned earlier that a posterior mean that is optimal relative to quadratic loss is approximately equal to the maximum likelihood estimate. Also many non-Bayesian estimates such as SUR regression coefficient estimates, 2SLS and 3SLS estimates, Cochrane–Orcutt estimates, etc., have been shown to be means of *conditional* posterior pdfs, based on diffuse priors, that is conditional posterior pdfs in which nuisance parameters are set equal to sample estimates. In *large samples* such approximate results may be

[14] This estimate was presented in Zellner and Vandaele [**65**] and a similar approach was employed in Griffiths and Dao [**20**].

[15] Geisel [**17**] used this approach in an application to combine predictions from two non-nested models.

[16] See Fisher [**14**], Harkema [**22**], Prescott [**46**], and Zellner [**58**] for further analysis of this and other control problems.

[17] See, e.g., Harkema [**22**], Prescott [**46**], and Zellner [**58**].

satisfactory; however in small to moderate sized samples these conditional posterior pdfs are poor approximations to the relevant marginal posterior pdfs and thus such approximations lead to poor estimates and other inferences.[18]

Bayesian methods have been used successfully in a number of applied studies. I shall just mention a few examples. Litterman [**35**] has used Bayesian vector autoregressions in forecasting seven important U.S. quarterly macroeconomic variables including real GNP, the implicit deflator for GNP, the unemployment rate, etc. In his work, Litterman used and is using a cleverly formulated informative prior pdf for the many parameters of his seven variable VAR which effectively reduces the number of free parameters. His forecasts compare very favorably with those obtained from unrestricted VARs, univariate AR models, Box-Jenkins' models, and two large structural econometric models. Varian [**54**] has applied Bayesian regression techniques to the problem of real estate assessment employing an asymmetric LINEX loss function. Akaike and Ishiguro [**1**] and their colleagues have developed a Bayesian approach and computer program, BAYSEA, for seasonal adjustment. Comparative study of the output of BAYSEA with the results of other seasonal adjustment procedures indicate that the former results compare favorably with other results. Morris [**40**] has discussed and referenced many applications of empirical Bayes' procedures. Prescott [**46**], Bowman and Laporte [**5**], Harkema [**22**], and Zellner [**58**, Ch. 11] have compared Bayesian and non-Bayesian solutions to policy control problems with results indicating that the former are superior in general. Swamy [**51**] and Swamy and Mehta [**52**] have compared Bayesian and non-Bayesian solutions to the undersized sample problem with the finding that the former are as good or better than non-Bayesian solutions. Monahan [**37**] has developed and applied Bayesian procedures for Box-Jenkins' models that include very good model selection, estimation, and prediction methods. Wright [**56**] reports a very interesting application of Bayesian methods in solving an overhead cost allocation problem. Miller [**36**] reviews Bayesian applications in actuarial statistics. Many applications of Bayesian analysis are reported in Kanjii [**28**]. Fomby and Guilkey [**15**] and Griffiths and Dao [**20**] have shown in Monte Carlo experiments that Bayesian estimators have lower mean-squared error (MSE) than non-Bayesian estimators for a regression coefficient in a simple regression model with autocorrelated disturbances. Surekha's and Griffiths' [**50**] Monte Carlo results for estimation of regression coefficients when disturbance terms are heteroscedastic indicate that Bayesian estimators outperform leading non-Bayesian estimators according to a MSE criterion. Park's [**43**] Monte Carlo experimental results show that a Bayesian minimum expected loss (MELO) estimator performs somewhat better than OLS and 2SLS estimators according to several criteria including MSE and mean absolute deviation. Brown's [**7**], Bawa, Brown, and Klein's [**2**], and Jorion's [**27**] Bayesian procedures for portfolio choice perform better than certainty equivalence procedures.[19]

[18] For examples, see Box and Tiao [**6**], Leamer [**31**], and Zellner [**58**, **60**].

[19] For further references, see the many papers listed under the key words "Bayes" and "Bayesian" in annual issues of the ASA/IMS *Current Index to Statistics*, Washington, DC: American Statistical Association.

In summary, I have emphasized that substantial prior information is being used in econometric analyses, usually informally and sometimes incoherently.[20] The Bayesian approach permits investigators to use prior information formally in a flexible, coherent framework to solve estimation, testing, prediction, model selection, design, and policy analysis problems. Some studies have been cited to demonstrate that the Bayesian approach not only has good theoretical properties but has also been successfully applied in practice. It is to be expected that many more Bayesian applications will be reported in the future given that a growing number of Bayesian computer programs are becoming available—see, e.g., Press [**47**], and that many more econometricians are becoming familiar with Bayesian methods. Since wise use of Bayesian techniques can lead to better solutions to econometric problems, Bayesian methods can contribute importantly to the progress of econometric science.

University of Chicago

Manuscript received January, 1984; revision received May, 1984.

REFERENCES

[1] Akaike, H., and M. Ishiguro: "Comparative Study of the X-11 and BAYSEA Procedures of Seasonal Adjustment," in *Applied Time Series Analysis of Economic Data*, ed. by A. Zellner. U.S. Government Printing Office: Washington, DC., 1983 (presented at the ASA-CENSUS-NBER Conference on Applied Time Series Analysis of Economic Data, Washington, DC, 1981).
[2] Bawa, V. S., S. J. Brown, and R. W. Klein: *Estimation Risk and Optimal Portfolio Choice.* Amsterdam: North-Holland, 1979.
[3] Baxter, M.: "The Role of Expectations in Stabilization Policy," presented at the Econometric Society meeting, Pisa, Italy, 1983.
[4] Bhalla, S. S.: "Measurement Errors and the Permanent Income Hypothesis: Evidence from Rural India," *American Economic Review*, 69(1979), 295-307.
[5] Bowman, H. W., and A. M. Laporte: "Stochastic Optimization in Recursive Equation Systems with Random Parameters with an Application to Control of the Money Supply," in *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, ed. by S. E. Fienberg and A. Zellner. Amsterdam: North-Holland, 1975, pp. 427-440.
[6] Box, G. E. P., and G. C. Tiao: *Bayesian Inference in Statistical Analysis.* Reading, Massachusetts: Addison-Wesley, 1973.
[7] Brown, S. J.: "Optimal Portfolio Choice Under Uncertainty," Ph.D. Thesis, Graduate School of Business, U. of Chicago, 1976.
[8] Chau, L. C.: "Stock Adjustment, Liquidity, and Normal Income in the Determination of Household Expenditures on Durable Goods," Ph.D. Thesis, Department of Economics, University of Wisconsin at Madison, 1967.
[9] Chetty, V. K.: "*Bayesian Analysis of Some Simultaneous Equation Econometric Models*," Ph.D. Thesis, Department of Economics, University of Wisconsin at Madison, 1966.
[10] DeGroot, M. H.: *Optimal Statistical Decisions.* New York: McGraw Hill Book Co., 1970.
[11] Drèze, J. H.: "Econometrics and Decision Theory," *Econometrica*, 40(1972), 1-17.
[12] ——— "Bayesian Theory of Identification in Simultaneous Equation Models," in *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, ed. by S. E. Fienberg and A. Zellner. Amsterdam: North-Holland, 1975, pp. 159-174.
[13] Fienberg, S. E., and A. Zellner (Eds.): *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage.* Amsterdam: North-Holland, 1975.
[14] Fisher, W. D.: "Estimation in the Linear Decision Model," *International Economic Review*, 3(1962), 1-29.

[20] See Swamy [**51**] for some striking examples of incoherent use of prior information in some leading non-Bayesian "solutions" to the "undersized sample" problem.

[15] FOMBY, T. B., AND D. K. GUILKEY: "On Choosing the Optimal Level of Significance for the Durbin-Watson Test and the Bayesian Alternatives," *Journal of Econometrics*, 8(1978), 203-213.
[16] FRIEDMAN, M.: *A Theory of the Consumption Function*. Princeton, New Jersey: Princeton University Press, 1957.
[17] GEISEL, M. S.: "Bayesian Comparisons of Simple Macroeconomic Models," in *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, ed. by S. E. Fienberg and A. Zellner. Amsterdam: North-Holland, 1975, pp. 227-256.
[18] GERSOVITCH, M.: "Savings and Nutrition at Low Incomes," *Journal of Political Economy*, 91(1983), 841-855.
[19] GIBBONS, M. R.: "Estimating the Parameters of the Capital Asset Pricing Model (CAPM): A Minimum Expected Loss (MELO) Approach," Technical Report, Graduate School of Business, Stanford University, 1980.
[20] GRIFFITHS, W. E., AND D. DAO: "A Note on the Bayesian Estimator in an Autocorrelated Error Model," *Journal of Econometrics*, 12(1980), 389-392.
[21] HADAMARD, J.: *The Psychology of Invention in the Mathematical Field*. New York: Dover Publications, Inc., 1945.
[22] HARKEMA, R.: "An Analytical Comparison of Certainty Equivalence and Sequential Updating," *Journal of the American Statistical Association*, 70(1975), 348-350.
[23] HARTIGAN, J. A.: *Bayes Theory*. New York: Springer-Verlag, 1983.
[24] HAVENNER, A., AND P. A. V. B. SWAMY: "A Random Coefficient Approach to Seasonal Adjustment of Economic Time Series," *Journal of Econometrics*, 15(1981), 177-209.
[25] HEYDE, C. C., AND I. M. JOHNSTONE: "On Asymptotic Posterior Normality for Stochastic Processes," *Journal of the Royal Statistical Society, Series B*, 41(1979), 184-189.
[26] JEFFREYS, H.: *Theory of Probability*. London: Oxford University Press, 1967.
[27] JORION, P.: "Portfolio Analysis of International Equity Investments: Bayes-Stein Estimators for the Mean and Implications for Risk Reduction," Ph.D. Thesis, Graduate School of Business, University of Chicago, 1983.
[28] KANJII, G. K. (ED.): "Proceedings of the Annual Conference on Practical Bayesian Statistics" Cambridge University, 1982, published in *Journal of the Institute of Statisticians*, 32(1983), 1-278.
[29] LAUB, P. M.: "The Dividend-Earning Relationship: A Study of Corporate Quarterly Panel Data, 1945-65," Ph.D. Thesis, Graduate School of Business, University of Chicago, 1971.
[30] LAUB, P. M.: "Some Aspects of the Aggregation Problem in the Dividend-Earning Relationship," *Journal of the American Statistical Association*, 67 (1972), 552-559.
[31] LEAMER, E. E.: *Specification Searches*. New York: John Wiley, 1978.
[32] LEE, T. C., G. G. JUDGE, AND A. ZELLNER: *Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data*. Amsterdam: North-Holland, 1977.
[33] LEVEDAHL, J. W.: "The Effects of Households' Permanent and Transitory Income on Both Their Timing of Automobile Purchases and Their Net Outlay for Automobile Services," Ph.D. Thesis, Department of Economics, University of Chicago, 1976.
[34] LINDLEY, D. V.: *Introduction to Probability and Statistics from a Bayesian Viewpoint* (2 vols.). Cambridge: Cambridge University Press, 1965.
[35] LITTERMAN, R.: "A Bayesian Procedure for Forecasting with Vector Autoregressions," unpublished manuscript, Department of Economics, Massachusetts Institute of Technology, 1980.
[36] MILLER, R.: "Actuarial Applications of Bayesian Statistics," in *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys*, ed. by A. Zellner. Amsterdam: North-Holland, 1980, pp. 197-212.
[37] MONAHAN, J. F.: "Fully Bayesian Analysis of ARIMA Time Series Models," *Journal of Econometrics*, 21(1983), 307-332.
[38] MONTMARQUETTE, C.: "A Model of Inventory Holdings with Empirical Application to Canadian Manufacturing Industries" Ph.D. Thesis, Department of Economics, University of Chicago, 1973.
[39] ———: "Random Walk Behavior of Finished Goods Inventory Investment," *Révue Economique*, 28(1977), 352-375.
[40] MORRIS, C. N.: "Parametric Empirical Bayes Inference: Theory and Applications," *Journal of the American Statistical Association*, 78(1983), 47-54.
[41] MUTH, J.: "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29(1961), 315-335.
[42] NERLOVE, M.: "Lags in Economic Behavior," *Econometrica*, 40(1972), 221-251.
[43] PARK, S. B.: "Some Sampling Properties of Minimum Expected Loss (MELO) Estimators of Structural Coefficients;" *Journal of Econometrics*, 18 (1982), 295-312.

[44] PECK, S. C.: "A Test of Alternative Theories of Investment Using Data from the Electric Utilities Industry," Ph.D. Thesis, Graduate School of Business, University of Chicago, 1973.
[45] PECK, S. C.: "Alternative Investment Models for Firms in the Electric Utilities Industry," *Bell Journal of Economics and Management Science*, 5(1974), 420-458.
[46] PRESCOTT, E. C.: "Adaptive Decision Rules of Macroeconomic Planning," in *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, ed. by S. E. Fienberg and A. Zellner. Amsterdam: North-Holland, 1975, pp. 427-440.
[47] PRESS, S. J.: "Bayesian Computer Programs," in *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys*, ed. by A. Zellner. Amsterdam: North-Holland, 1980, pp. 429-442.
[48] QUENOUILLE, M. H.: *The Analysis of Multiple Time-Series*. New York: Hafner Publishing Co., 1957.
[49] SAVAGE, L. J.: *The Foundations of Statistics*. New York: John Wiley, 1954.
[50] SUREKHA, K., AND W. E. GRIFFITHS: "A Monte Carlo Comparison of Some Bayesian and Sampling Theory Estimators in Two Heteroscedastic Error Models," manuscript, Indian Statistical Institute, New Delhi, and University of New England, Armidale, Australia, 1983.
[51] SWAMY, P. A. V. B.: "A Comparison of Estimators for Undersized Samples," *Journal of Econometrics*, 14(1980), 161-181.
[52] SWAMY, P. A. V. B., AND J. S. MEHTA: "Further Results on Zellner's Minimum Expected Loss and Full Information Maximum Likelihood Estimators for Undersized Samples," *Journal of Business and Economic Statistics*, 1(1983), 154-162.
[53] THORNBER, H.: "Finite Sample Monte Carlo Studies: An Autoregressive Illustration," *Journal of the American Statistical Association*, 62(1967), 801-808.
[54] VARIAN, H. R.: "A Bayesian Approach to Real Estate Assessment," in *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, ed. by S.E. Fienberg and A. Zellner. Amsterdam: North-Holland, 1975, pp. 195-208.
[55] VELOCE, W., AND A. ZELLNER: "Modelling a Competitive Industry with Entry: Implications for Demand and Supply Analysis," *Economics Letters*, 16 (1984), 71-75.
[56] WRIGHT, R. L.: "Measuring the Precision of Statistical Cost Allocations," *Journal of Business and Economic Statistics*, 1(1983), 93-100.
[57] ZELLNER, A.: "Tests of Some Basic Propositions in the Theory of Consumption," *American Economic Review*, 50(1960), 565-573.
[58] ———: *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley, 1971.
[59] ———: *Seasonal Analysis of Economic Time Series*, U.S. Bureau of the Census, Washington, DC: U.S. Government Printing Office, 1978.
[60] ———: "Statistical Analysis of Econometric Models," *Journal of the American Statistical Association*, 74(1979), 628-641.
[61] ———: "The Current State of Bayesian Econometrics," *Basic Issues in Econometrics*, ed. by A. Bellner, Chicago: University of Chicago Press, 1984, 306-321.
[62] ———: *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys*. Amsterdam: North-Holland, 1980.
[63] ———: "Philosophy and Objectives of Econometrics," in *Macroeconomic Analysis: Essays in Macroeconomics and Econometrics*, ed. by D. Currie, R. Nobay and D. Peel. England: Croom Helm, 1981.
[64] ZELLNER, A., AND F. PALM: "Time Series Analysis and Simultaneous Equation Econometric Models," *Journal of Econometrics*, 2(1974), 17-54.
[65] ZELLNER, A., AND W. VANDAELE: "Bayes-Stein Estimators for k-Means, Regression and Simultaneous Equation Models," in *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, ed. by S.E. Fienberg and A. Zellner. Amsterdam: North-Holland, 1975, pp. 627-653.
[66] ZELLNER, A., AND P. E. ROSSI: "Bayesian Analysis of Dichotomous Quantal Response Models," *Journal of Econometrics*, 25 (1984), 365-393.