AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

378.794
G43455
WP-905

Working Paper Series

WORKING PAPER NO. 905

BAYESIAN AND NON-BAYESIAN APPROACHES TO SCIENTIFIC MODELING AND INFERENCE IN ECONOMICS AND ECONOMETRICS

by

Arnold Zellner

Waite Library
Dept. of Applied Economics
University of Minnesota
1994 Buford Ave - 232 ClaOff
St. Paul, MN 55108-6040 USA

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY

BERKELEY

California Agricultural Experiment Station

University of California

378.794
C43455
WP-905

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY
DIVISION OF AGRICULTURE AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA AT BERKELEY

WORKING PAPER NO. 905

BAYESIAN AND NON-BAYESIAN APPROACHES TO SCIENTIFIC MODELING AND INFERENCE IN ECONOMICS AND ECONOMETRICS

by

Arnold Zellner

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
August, 1999

H.G.B. Alexander Research Foundation
Graduate School of Business
University of Chicago

Bayesian and Non-Bayesian Approaches to Scientific Modeling and Inference in Economics and Econometrics

by

Arnold Zellner
University of Chicago

Paper 9911 **August, 1999**

Bayesian and Non-Bayesian Approaches to Scientific Modeling and Inference in Economics and Econometrics

by

Arnold Zellner*

U. of Chicago

Abstract

After brief remarks on the history of modeling and inference techniques in economics and econometrics, attention is focused on the emergence of economic science in the 20th century. First, the broad objectives of science and the Pearson-Jeffreys' "unity of science" principle will be reviewed. Second, key Bayesian and non-Bayesian practical scientific inference and decision methods will be compared using applied examples from economics, econometrics and business. Third, issues and controversies on how to model the behavior of economic units and systems will be reviewed and the structural econometric modeling, time series analysis (SEMTSA) approach will be described and illustrated using a macro-economic modeling and forecasting problem involving analyses of data for 18 industrialized countries over the years since the 1950s. Point and turning point forecasting results and their implications for macro-economic modeling of economies will be summarized. Last, a few remarks will be made about the future of scientific inference and modeling techniques in economics and econometrics.

1. Introduction

It is an honor and a pleasure to have this opportunity to share my thoughts with you at this Ajou University Conference in honor of Professor Tong Hun Lee. He has been a very good friend and an exceptionally productive scholar over the years. We first met in the early 1960s at the U. of Wisconsin in Madison and I was greatly impressed by his intellectual ability, serious determination and genuine modesty. As stated in his book, Lee (1993),

> "I was originally drawn to the study of economics because of my concern over the misery and devastation of my native country. . . I hoped that what I learned might help to improve living conditions there. As a student, however, I encountered numerous

*Research financed in part by the U.S. National Science Foundation and by the H.G.B. Alexander Endowment Fund, Graduate School of Business, U. of Chicago. e-mail: arnold.zellner@gsb.uchicago.edu
http://gsbwww.uchicago.edu/fac/arnold.zellner/index.html

conflicts between economic theory and real world phenomena. Over time I acquired a deep conviction that economic research should be rigorous but policy relevant and that it must reflect an appreciation of empirical evidence as well as of economic theory." (p. IX)

Here we have a statement of Lee's objectives that are reflected in his many research publications on econometric methods, economic theory and applications to monetary, fiscal, regional and other problems that reflect his expertise in economic theory, econometrics and applied economic analysis. Indeed, it is the case that Lee has achieved great success in his research and career. When we contrast Lee's knowledge of economics, econometric methods and empirical economic analysis with that of an average or even an outstanding economist in the 19th century or in the early 20th century, we can appreciate the great progress in economic methodology that has been made in this century. This progress has led to the transformation of economics from an art into a science with annual Nobel Prize awards. Recall that in the 19th century and early 20th century we did not have the extensive national income and product and other data bases that are currently available for all countries of the world. In addition, scientific survey research methods have been utilized to provide us with extensive survey and panel data bases. Also, much experimental economic data are available in economics, marketing, medical and other areas.

Not only does the current data situation contrast markedly with the situation in the past but also the use of mathematics in economics was quite limited then. Further, good econometric and statistical techniques were not available. For example, as late as the 1920s, leading economists did not know how to estimate demand and supply functions satisfactorily. A leading issue was, "Do we regress price on quantity or quantity on price and do we get an estimate of the demand or supply function?" Further, in the 1950s, Tinbergen mentioned to me that he estimated parameters of his innovative macro-econometric models of various industrialized countries' economies by use of "ordinary least squares" (OLS) since this was the only method that he knew. And of course, satisfactory statistical inference techniques, that is estimation, testing, prediction, model selection and policy analysis techniques for the multivariate, simultaneous equation, time series models that Tinbergen, Klein, Schultz, and many others built in the first half of this century were unavailable.

As is well known, econometric modeling, inference and computing techniques were in a very primitive state as late as the 1940s. Some believed that fruitful quantitative, mathematical analyses of economic behavior are impossible. Then too, there were violent debates involving Tinbergen, Keynes, Friedman, Koopmans, Burns, Mitchell and many others about the methods of economics and econometrics. These included charges of "measurement without theory" and "theory without measurement." Others objected to the use of statistical sampling theory methods in analyzing non-experimental data generated by economic systems. There were heated debates about how to build and evaluate empirical econometric models of the type that Schultz, Tinbergen, Klein, Haavelmo, Koopmans, Tobin and others studied and developed. Some issues

involved simplicity versus complexity in model-building, explanatory versus forecasting criteria for model performance, applicability of probability theory in the analysis of economic data, which concept of probability to utilize, quality of available data, etc. Finally, there were arguments about which statistical approach to use in analyzing economic data, namely, likelihood, sampling theory, fiducial, Bayesian, or other inference approaches.

Indeed there were many unsettled, controversial issues regarding economic and econometric methodology in the early decades of this century. However, economics was not alone in this regard. Many other social, biological and physical areas of research faced similar methodological issues. Indeed, Sir Harold Jeffreys wrote his famous books, *Scientific Inference* (1973, 1st ed. 1931) and *Theory of Probability* (1998, 1st ed.1939) to instruct his fellow physicists, astronomers, and other researchers in the methodology of science. In I. J. Good's review of the 3rd edition of the latter book, he wrote that Jeffreys' book ". . .is of greater importance for the philosophy of science, and obviously of greater immediate practical importance, than nearly all the books on probability written by professional philosophers lumped together." See also, articles in Kass (1991) and Zellner (1980) by leading authorities that summarize Jeffreys' contributions to scientific methodology that are applicable in all fields and his analyses of many applied scientific problems. It is generally recognized that Jeffreys provided an operational framework for scientific methodology that is useful in all the sciences and illustrated its operational nature by using it in applied analyses of many empirical problems in geophysics, astronomy and other areas in his books and six volumes of collected papers, Jeffreys (1971).

Jeffreys, along with Karl Pearson (1938), emphasized the "unity of science" principle, namely that any area of study, e.g., economics, business, physics, psychology, etc., can be scientific if scientific methods are employed in analyzing data and reaching conclusions. Or as Pearson (1938, p. 16) states, "The unity of all science consists alone in its method, not in its material." With respect to scientific method, Jeffreys in his book, *Theory of Probability*, provides an axiom system for probability theory that is useful to scientists in all fields in their efforts to learn from their data, to explain past experience and make predictions regarding as yet unobserved data, fundamental objectives of science in all areas. He states that scientific induction involves (1) observation and measurement and (2) generalization from past experience and data to explain the past and predict future outcomes. Note the emphasis on measurement and observation in scientific induction. Also, generalization or theorizing is critical but deductive inference is not sufficient for scientific work since it just provides statements of proof, disproof or ignorance. Inductive inference accommodates statements less extreme than those of deductive inference. For example, with an appropriate "reasonable degree of belief" definition of probability (see Jeffreys, 1998, Ch. 7 for a penetrating analysis of various, alternative concepts of probability), inductive inference provides a quantitative measure of the degree of belief that an individual has in a proposition, say the quantity theory of money model or the Keynesian macroeconomic model. As more empirical data become available, these degrees of beliefs in propositions can be updated formally by use of probability theory, in particular Bayes' theorem, and this constitutes a formal, operational way of "learning

from experience and data," a fundamental objective of science. Since other statistical approaches do not permit probabilities to be associated with hypotheses or models, this learning process via use of Bayes' Theorem is not possible using them. See Jeffreys (1957, 1998), Cook (1990), Hill (1986), Jaynes (1983,1984), Kass (1982,1991) and Zellner (1980, 1982,1988, 1996) for further discussion of these issues. That Jeffreys was able to formulate an axiom system for probability theory and show how it can be used formally and operationally to implement learning from experience and data in all areas of science is a remarkable achievement that he illustrated in many applied studies. Currently, Jeffreys' approach is being utilized in many fields, including business, economics and econometrics and thus these fields are currently viewed as truly scientific.

Before turning to the specific operations of Bayesian inductive inference, it seems important to point out that the origins of generalizations or theories is an important issue. Some, including C. S. Pierce, cited by Hanson (1958, p.85), refer to this area as "reductive inference." According to Pierce, ". . .reduction suggests that something may be: that is, it involves studying facts and devising theories to explain them." Unfortunately, this process is not well understood. Work by Hadamard (1945) on the psychology of invention in the field of mathematics is helpful. He writes, "Indeed, it is obvious that invention or discovery, be it in mathematics or anywhere else, takes place by combining ideas."(p. 29). Thus thinking broadly and taking account of developments in various fields provides for useful input along with an esthetic sense for producing fruitful combinations of ideas. Often major breakthroughs occur, according to the results of a survey of his fellow mathematicians conducted by Hadamard, when unusual facts are encountered. In economics, e.g., the constancy of the U.S. savings rate over the first part of this century during which real income increased considerably was discovered empirically by S. Kuznets and contradicted Keynesian views that the savings rate should have increased given the large increase in income. This surprising empirical finding led various economists including Friedman, Modigliani, Tobin and others to propose new theories of consumer behavior to explain Kuznets' unusual finding. Similarly, the striking empirical fact that the logarithm of output per worker and the log of the wage rate are found to be linearly related empirically caused Arrow, Chenery, Minhas and Solow to formulate the CES production function to explain this unusual linear relation. Since unusual facts are often important in prompting researchers to produce new breakthroughs, I thought it useful to bring together various ways, some rather obvious, to help produce new and unusual facts rather than dull, humdrum facts. See the list in Zellner (1984, pp.9-10) that includes (1) study of incorrect predictions and forecasts of models, (2) study of existing models under extreme conditions, (3) strenuous simulation experiments with current models and theories, (4) observing behavior in unusual historical periods, say periods of hyperinflation or major deflation,(5) observing behavior of unusual individuals, e.g., extremely poor individuals, etc. By producing unusual facts, current models are often called into question and work is undertaken to produce better models or theories. All of this leads to the following advice for empirical economists and econometricians, namely, PRODUCE UNUSUAL FACTS.

With these brief, incomplete remarks regarding how to produce new models and theories, it is relevant to remark that when no useful model or theory is available, many

recommend that we assume all variation is random unless shown otherwise as a good starting point for analysis. Note that Christ, Friedman, Cooper, Nelson, Plosser, and others used random walk and other relatively simple time series, benchmark models to appraise the predictive performance of large-scale macro-econometric models put forward by Klein, Goldberger, the U.S. Federal Reserve System and others. Also, such benchmark models have been utilized in financial economics to evaluate proposed models that purport to predict and explain the variation of stock prices and in work by Hong (1989) and Min (1992) to evaluate the performance of complicated models for forecasting growth rates of real GDP for industrialized countries. If such work reveals that a complicated model with many parameters and equations can not perform better than a simple random walk model or a simple univariate time series model, then the model, at the very least, needs reformulation and probably should be labeled UNSAFE FOR USE. Indeed, in the last few years, we have seen the scrapping of some complicated macroeconometric models. While the issue of simplicity versus complexity is a difficult one, many in various theoretical and applied fields believe that keeping theories and models sophisticatedly simple is worthwhile. In industry, there is the principle KISS, that stands for Keep it Simple Stupid. However, since some simple models are stupid, I reinterpreted KISS to mean Keep It Sophisticatedly Simple. Indeed it is hard to find a single complicated model in science that has performed well in explanation and prediction. On the other hand, there are many sophisticatedly simple models that have performed well, e.g. demand and supply models in economics and business, Newton's laws, Einstein's laws, etc. For more on these issues of simplicity and complexity, see Jeffreys (1998) for discussion of his and Dorothy Wrinch's "simplicity postulate" and the papers and references in Kuezenkamp, McAleer and Zellner (1999).

2. Bayesian Inference and Decision Techniques

As regards statistical and econometric inference and decision techniques, in general since the 1950s and 1960s, there has been an upswing in the use and development of Bayesian inference and decision techniques in business, statistics, econometrics and many other disciplines by building on the pioneering work of Bayes, Laplace, Edgeworth, Jeffreys, de Finetti, Savage, Box, Good, Lindley, Raiffa, Schlaifer, Dreze and many others. By now almost all general econometrics textbooks include material on the Bayesian approach. In addition there are a number of Bayesian statistics, business, engineering and econometrics texts available. In 1992, the International Society for Bayesian Analysis (ISBA:www.bayesian.org) and the Bayesian Statistical Science Section of the American Statistical Association (www.amstat.org) were founded and since then have held many successful meetings and produced annual proceedings volumes, published by the American Statistical Association. Then too, for many years the NBER-NSF Seminar on Bayesian Inference in Econometrics and Statistics, the Valencia Conference, the Bayes-Maxent Workshop, the Workshop on Practical Applications of Bayesian Analysis and the Bayesian Decision Analysis Section of the Institute for Operations Research and Management Science (INFORMS) have sponsored many research meetings, produced a large number of Bayesian publications, and sponsored various awards for outstanding work in Bayesian analysis. Further, the current

statistical and econometric literature abounds with Bayesian papers. Indeed some have declared that a Bayesian Era has arrived and that the next century will be the century of Bayes.

To understand these developments, it is necessary to appreciate that Bayesian methods have been applied in analyses of all kinds of theoretical and applied problems in many fields. Bayesian solutions to estimation, prediction, testing, model selection, control and other problems have been as good as or better than those provided by other approaches, when they exist. In addition, Bayesian methods have been utilized to reproduce many non-Bayesian solutions to problems. For example, as Jeffreys pointed out many years ago, in large samples posterior densities for parameters generally assume a normal form with a posterior mean equal to the maximum likelihood estimate and posterior covariance matrix equal to the inverse of the estimated Fisher information matrix which he regarded as a Bayesian justification for the method of maximum likelihood.

As explained in Bayesian texts, e.g. Jeffreys (1998), Bernardo and Smith (1994), Berger (1985), Berry et al (1996), Box and Tiao (1993),Gelman et al (1995), Press (1989), Raiffa and Schlaifer (1961), Robert (1994), Zellner (1996), etc., Bayes' theorem, Bayes (1763) can be used to analyze estimation, testing, prediction, design, control and other problems and provides useful finite sample results as well as excellent asymptotic results. In estimation problems, we have in general via Bayes' theorem that the posterior density for the parameters is proportional to a prior density times the likelihood function. Thus information contained in a prior density for the parameters is combined with sample information contained in a likelihood function by use of Bayes' theorem to provide a posterior density that contains all the information, sample and prior. See Zellner (1988) for a demonstration that Bayes' theorem is a 100 per cent efficient information processing rule, invited discussion of this result by Jaynes, Hill, Kullback and Bernardo and further consideration of it in Zellner (1991). The works, cited above, provide many applications of Bayes' theorem to the models used in business, economics, econometrics and other areas.

Investigators can use a posterior density to compute the probability that a parameter's value lies between any two given values, e.g. the probability that the marginal propensity to consume lies between 0.60 and 0.80 or that the elasticity of a firm's sales with respect to advertising outlays lies between 0.9 and 1.1. As regards point estimation, given a convex loss function, say a quadratic loss function , it is well known that the optimal Bayesian estimate that minimizes posterior expected loss is the posterior mean while for absolute error loss and for "zero-one" loss, the optimal Bayesian estimates that minimize posterior expected loss are the median and the modal value of the posterior density, respectively. These and other results for other loss functions, e.g. asymmetric loss functions, are exact, finite sample results that are extremely useful in connection with, e.g., real estate assessment, time series and simultaneous equation models where optimal sampling theory finite sample estimators are not available; see Berry et al (1996) for many examples and references. Also, as Ramsey, Friedman, Savage and others have emphasized, this minimal expected loss or equivalently maximal expected utility action in choosing an estimate is in accord with the expected utility

theory of economics; see, e.g. Friedman and Savage (1948, 1952). Further, these Bayesian optimal estimates, viewed as estimators, have been shown to minimize Bayes' risk, when it is finite, and are admissible. For more on these properties of Bayesian estimators, see, e.g., Berger (1985), Judge et al. (1987), Greene (1998) and the other texts cited above.

As regards some Bayesian econometric estimation results, see Hong (1989) who used the Bayesian approach to analyze time series, third order autoregressive-leading indicator (ARLI) models for forecasting annual growth rates of real GDP. He not only produced finite sample Bayesian posterior distributions for the parameters of the model but also computed the probability 0.85 that the process has two complex roots and one real root. Also, he computed the posterior densities for the period and amplitude of the oscillatory component of the model. He found a posterior mean for the period of about 4 to 5 years and a high probability that the amplitude is less than one. Also, the posterior density for the amplitude of the real root was centered over values less than one. These results were computed for each of 18 industrialized countries' data in Hong's sample. From a non-Bayesian point of view, it is not possible to make such probabilistic statements regarding the properties of solutions to time series processes and, indeed, it appears that just asymptotic, approximate sampling theory procedures are available for such problems.

Another area in which Bayesian procedures have produced improved results is in the area of estimation of parameters of simultaneous equations models. For example, in estimating the parameters of the widely-used Nerlove agriculture supply model, Diebold and Lamb (1997) showed that use of easily computed Bayesian minimum expected loss (MELO) estimators led to large reductions in the mean-squared error (MSE) of estimation relative to use of the most widely used sampling theory technique. Similarly, in Park (1982), Tsurumi (1990), Gao and Lahiri (1999) and Zellner (1997, 276-287, 1998), Bayesian MELO estimators' finite sample performance was found to be generally better than that of non-Bayesian estimators including maximum likelihood, Fuller's modified maximum likelihood, two-stage least squares, ordinary least squares, etc. In addition to these fine "operating characteristics" of Bayesian procedures in repeated trials, for a given sample of data, they provide optimal point estimates, finite sample posterior densities for parameters and posterior confidence intervals, all unavailable in non-Bayesian approaches that generally rely on asymptotic justifications, e.g., consistency, asymptotic normality and efficiency, properties also enjoyed by Bayesian estimators.

Various versions of Bayes-Stein shrinkage estimation techniques, described in Stein (1956), James and Stein (1961), Berger (1985), Zellner and Vandaele (1975) and other references, have been employed with success by Garcia-Ferrar et al. (1987), Hong (1989), Min (1992), Zellner and Hong (1989), Quintana et al.(1995), Putnam and Quintana (1995) and many others. Here in say a dynamic seemingly unrelated regression equation system for countries' growth rates or for a set of stock returns, the coefficient vectors in each equation are assumed randomly distributed about a common mean vector. By adding this assumption, Bayesian analysis provides posterior means for the coefficient vectors that are "shrunk" towards an estimate of the common mean. This added information provides much improved estimation and prediction results, theoretically and

empirically. Indeed, Stein showed that many usual estimators are inadmissible relative to his shrinkage estimator using a standard quadratic loss function. See Zellner and Vandaele (1975) for various interpretations of Stein shrinkage estimators that have been extremely valuable in many empirical estimation and forecasting studies and Quintana et al (1995) for their use in solving financial forecasting and portfolio selection problems.

Further, for a wide range of dichotomous and polytomous random variable models, e.g. logit, probit, multinomial probit and logit, sample selection bias models, etc., new integration techniques, including importance function Monte Carlo numerical integration, Markov Chain, Monte Carlo (MCMC) techniques and improved MCMC techniques have permitted Bayesian finite sample analyses of these difficult models to be performed. Many applications using data from marketing, education, labor markets, etc. have been reported. See, e.g. Albert and Chib (1993), selected articles in Berry et al. (1996), Gelman et al. (1995), Geweke (1989), McCulloch and Rossi (1990,1994), Pole, West and Harrison (1994), Tobias (1999), and Zellner and Rossi (1984). It is the case that use of these new numerical techniques, described in Geweke (1989), Chib and Greenberg (1996), Gelman et al (1995) and the other references above, has permitted Bayesian analyses of problems that were considered intractable just a few years ago.

As regards prediction, the standard procedure for obtaining a predictive density function for unobserved data, either past or future, is to write down the probability density for the future, as yet unobserved data, denoted by y, given the parameters, θ, $f(y|\theta)$. By multiplying this density by a proper density for the parameters, say a posterior density, derived from past observed data via Bayes' theorem, we can integrate over the parameters to get the marginal density of the as yet unobserved data, say $h(y|I)$, where I denotes the past sample and prior information. In this case, and in many others, the integration over the parameters to obtain a marginal predictive density is a very useful way to get rid of parameters by averaging the conditional densities using the posterior density as a weight function. Given that we have the predictive density, $h(y|I)$, we can use it to make probability statements regarding possible values of y. For example, we can compute the probability that next year's rate of growth of GDP is between 3 and 5 per cent or the probability that next year's growth rate will be below this year's growth rate. Further, if we have a predictive loss function, we can derive the point prediction that minimizes expected predictive loss for a variety of loss functions. For example, for a squared error loss function, the optimal point prediction is the mean of the predictive density. See, e.g. Varian (1975), Zellner (1987) and articles in Berry et al (1996) for theoretical and applied analyses using various symmetric and asymmetric loss functions. As emphasized in this literature, symmetric loss functions, e.g. squared error or absolute error loss functions are not appropriate for many important problems. Thus it is fortunate that in estimation and prediction, Bayesian methods can be employed to obtain optimal point estimates and predictions relative to specific, relevant asymmetric loss functions such as are used in real estate assessment, bridge construction, medicine and other areas.

The predictive density has been shown to be very useful in developing optimal turning point forecasting techniques; see, e.g. Zellner and Hong (1991), Zellner, Hong and Min (1991), LeSage (1996), Zellner and Min (1999) and Zellner, Tobias and Ryu

(1998). Given that the current value of a variable, say the rate of growth of real GDP, is known, using the predictive density for next year's rate of growth, the probability, P, that it is less than this year's value can be computed and interpreted as the probability of a downturn (DT) and 1-P as the probability of no downturn (NDT). Given a two by two loss structure associated with the acts, forecast DT or forecast NDT and the possible outcomes, DT or NDT, the optimal forecast that minimizes expected loss can be easily determined. For example, if the loss structure is symmetric and $P>1/2$, the optimal forecast is DT whereas if $P<1/2$, the optimal forecast is NDT. Similar analysis can be used to obtain optimal upturn and no upturn forecasts. Using these techniques, in the papers cited above, about 70 per cent of 211 turning point outcomes for 18 industrialized countries' rates of growth of real GDP, 1974-1995 were correctly forecast. This performance was much better than that yielded by using benchmark techniques, e.g., coin-flipping, "eternal optimist," "eternal pessimist" and deterministic four-year cycle approaches. Also, LeSage (1996) used these techniques and obtained similarly satisfactory results in forecasting turning points in U.S. regional employment data.

Another area in which predictive densities play an important role is in optimal portfolio analysis in theoretical and applied finance; see, e.g., Brown (1976), Bawa, Brown and Klein (1979), Jorion (1983, 1985), Markowitz (1959, 1987), Quintana, Chopra and Putnam (1995) and Zellner and Chetty (1965). Given a predictive density for a vector of future returns, a portfolio is a linear combination of these future returns, denoted by R, with the weights on individual returns equal to the proportion of current wealth assigned to each asset. Maximizing the expected utility of R, EU(R) with respect to the weights subject to the condition that they add up to one provides an optimal portfolio. In recent work by Quintana, Chopra and Putnam (1995), a Bayesian dynamic, state space, seemingly unrelated regression model with time varying parameters has been employed to model a vector of returns through time. By use of iterative, recursive computational techniques, the model is updated period by period and its predictive density for future vectors of returns is employed to solve for period-by-period optimal portfolios. In calculations with past data, the cumulative returns, net of transaction costs, associated with these Bayesian portfolios have compared favorably with the cumulative returns associated with a hold the S&P five hundred index stocks strategy. Currently, the CDC Investment Management Corporation in New York is employing such Bayesian portfolio methods. Also, as reported at a workshop meeting at the U. of Chicago several years ago, Fisher Black and Robert Litterman reported that they use Bayesian portfolio methods at Goldman-Sachs in New York.

Last, there are many other areas in which Bayesian predictive densities are important since fundamentally induction has been defined to be generalization or theorizing to explain and predict. Further, the philosophers, according to a review paper by Feigl (1953), have defined causality to be "predictability according to a law or set of laws." Also practically, forecasting and prediction are very important in all areas and thus Bayesian predictive densities have been widely employed in almost all areas of science and application including marketing, business and economic forecasting, clinical trials, meteorology, astronomy, physics, chemistry, medicine, etc.

Bayes' theorem is also very useful in comparing and testing alternative hypotheses and models by use of posterior odds that are equal to the prior odds on alternative hypotheses or models, nested or non-nested, times the Bayes factor for the alternative hypotheses or models. The Bayes factor is the ratio of the predictive densities associated with the alternative hypotheses or models evaluated with the given sample observations. This approach to "significance testing" was pioneered by Jeffreys (1998) and applied to almost all the standard testing problems considered by Neyman, Pearson and Fisher. Indeed, Jeffreys considered the Bayesian approach to testing to be much more sensible than the Neyman-Pearson approach or the Fisher p-value approach and provided many empirical comparisons of results associated with alternative approaches. Note that in the non-Bayesian approaches, probabilities are not associated with hypotheses. Thus within these approaches, one can not determine how the information in the data change our prior odds relating to alternative hypotheses or models. See the references above and Kass and Raftery (1995) for further discussion of Bayes' factors and references to the voluminous Bayesian literature involving their use.

If, for example, we have two variants of a model, say a model to forecast GDP growth rates, as explained in Min and Zellner (1993) and Zellner, Tobias and Ryu (1998, 1999), we can employ prior odds and Bayes' factors to determine which variant of the model is better supported by the data. For example, we might start with prior odds one to one on the two variants, say a fixed parameter model versus a time-varying parameter model. Then after evaluating the Bayes' factor for the two models and multiplying by the prior odds, here equal to one, we obtain the posterior odds on the two models, say 3 to 1 in favor of the time-varying parameter model. Also, the posterior odds on alternative models can be employed to average estimates and forecasts over models, a Bayesian forecast combination procedure that has been compared theoretically and empirically to non-Bayesian forecast combination procedures in Min and Zellner (1993). Also, in Palm and Zellner (1992), the issue of whether it is always advantageous to combine forecasts is taken up. As might be expected, it is not always the case that combining forecasts leads to better results; however, many times it does.

To close this brief summary of Bayesian methods and applications, note that many formal procedures for formulating diffuse or non-informative and informative prior densities have been developed; see Kass and Wasserman (1996) and Zellner (1997) for discussion of these procedures. It should also be appreciated that the Bayesian approach has been applied in analyses of almost all parametric, nonparametric and semiparametric problems. Indeed, at this point in time, it is probably accurate to state that most, if not all, the estimation, testing and prediction problems of econometrics and statistics have been analyzed from the Bayesian point of view and the results have been quite favorable from the Bayesian viewpoint. With this said, let us turn to a comparison of some Bayesian and non-Bayesian concepts and procedures.

3. Comparison of Bayesian and Non-Bayesian Concepts and Procedures

Shown in Table 1 are 12 issues and summary statements with respect to Bayesian and Non-Bayesian positions on these issues. First we have the fundamental issue as to

whether a formal learning model is used. Bayesians use Bayes' theorem as a learning model whereas non-Bayesians do not appear to use a formal learning model. In effect,

Table 1

Some Bayes-Non-Bayes Issues and Responses

Issues	Responses	
	Bayes	Non-Bayes
1. Formal learning model?	Yes	No
2. Axiomatic support?	Yes	?
3. Probabilities associated with hypotheses and models?	Yes	No
4. Probability defined as measure of degree of confidence in a proposition?	Yes	No
5. Uses Pr(a<θ<b given data)?	Yes	No
6. Uses Pr(c<y, <d given data)?	Yes	No
7. Minimization of Bayes risk?	Yes	No
8. Uses prior distributions?	Yes	?
9. Uses subjective prior information?	Yes	Yes
10. Integrates out nuisance parameters?	Yes	No
11. Good asymptotic results?	Yes	Yes
12. Exact, good finite sample results?	Yes	Sometimes

non-Bayesians are learning informally. As mentioned above, use of the Bayesian learning model has led to many useful results. However, this does not mean that the Bayesian learning model can not be improved and indeed several researchers including Diaconis, Goldstein, Hill, Zabell, Zellner and others have been involved in research designed to extend the applicability of the Bayesian learning model.

Secondly, there are several Bayesian axiom systems that have been put forward by Jeffreys, Savage and others to provide a rationale for Bayesian inference procedures. As regards axiom systems for non-Bayesian inference and decision procedures, I do not know of any.

Third, as stated above, Bayesians use probabilities to express degrees of confidence in hypotheses and models. Non-Bayesians who use axiomatic and frequency definitions of probability do not do so formally. However, many times non-Bayesians informally do so and incorrectly associate p-values with degrees of confidence in a null hypothesis. While it is true that some posterior odds expressions are monotonic functions of p-values, the p-value is not equal to the posterior probability on the null hypothesis nor was it ever meant to be.

Fourth, non-Bayesians pride themselves in their axiomatic and frequency "objective" concepts of probability and are critical of Bayesians for their "subjective" concepts of probability. In this regard most of these non-Bayesians have not read Jeffreys'(1998, Ch. 7) devastating critiques of axiomatic and frequency definitions of probability. For example, on the long run frequency or Venn limit definition of probability, Jeffreys writes, "No probability has ever been assessed in practice, or ever will be, by counting an infinite number of trials or finding the limit of a ratio in an infinite series...A definite value is got on them only by making a hypothesis about what the result would be. On the limit definition,...there might be no limit at all....the necessary existence of the limit denies the possibility of complete randomness, which would permit the ratio in an infinite series to tend to no limit." (p.373). Many other examples and considerations are presented to show the inadequacies of the axiomatic and limiting frequency definitions of probability for scientific work. As far as I know, Jeffreys' arguments have not been rebutted, perhaps because as some have noted, they are irrefutable. He further writes, "The most serious drawback of these definitions, however, is the deliberate omission to give any meaning to the probability of a hypothesis." (p.377) See also Jeffreys (1998, pp.30-33). Kass (1982) and Zellner (1982) for discussion of Jeffreys' definition of probability as compared to the "personalistic" or "moral expectation" or "betting" definitions put forward by Ramsey, de Finetti, Savage, and others.

Under issue 5 in Table 1, we have the probability that a parameter's value lies between two given numbers, a and b given the data, a typical Bayesian posterior probability statement, first derived analytically by Bayes (1763). Non-Bayesians can not validly make such statements even though many practitioners misinterpret sampling theory confidence intervals in this manner. The same is true with respect to the Bayesian prediction interval for the future random variable under issue 6 in Table 1. For example, a Bayesian might state that the probability that this random variable lies between the given numbers c and d is 0.95. On the other hand if c and d are the realized values of the endpoints of a 95% sampling theory confidence interval, then it is incorrect to say that the future value lies between c and d with probability 0.95. Rather one should state that the interval c to d is the realized value of a random interval that has probability 0.95 of covering the random variable.

With respect to issue 7 in Table 1, non-Bayesians do not minimize Bayes risk since they don't introduce a prior density for the parameters, an essential element in the definition of Bayes' risk given in Bayesian texts. Bayesians minimize Bayesian risk in choosing estimators and predictors in order to insure that they have good operating characteristics. However in situations involving a single set of data, averaging over

unobserved outcomes may not be relevant and use of the criterion minimization of expected loss given the one sample of data is more appropriate.

As regards issue 8, Bayesians use diffuse or non-informative and informative prior densities quite broadly. Non-Bayesians generally say they do not. However in analyses of hierarchical models, state space models, random effects models, and random initial conditions for time series models, distributions are often introduced for parameters that are usually considered to be "part of the model" and not prior densities. As Good (1991) and others have recognized, this distinction is rather thin and for Good represents a possible compromise between Bayesians and non-Bayesians. Note too that in econometrics, some non-Bayesians have attempted to introduce subjective prior information using the "mixed estimation" procedure of Theil and Goldberger explained in Theil (1971), inequality restrictions on estimators and predictors, ridge regression and other approaches. In addition, many have recognized that prior subjective information is used extensively in model formulation by Bayesians and non-Bayesians.

That subjective prior information is used quite broadly is noted by several prominent non-Bayesians. For example, Tukey (1978, p. 52) writes, "It is my impression that rather generally, not just in econometrics, it is considered decent to use judgment in choosing a functional form but indecent to use judgment in choosing a coefficient . If judgment about important things is quite all right, why should it not be used for less important ones as well? Perhaps the real purpose of Bayesian techniques is to let us do the indecent thing while modestly concealed behind a formal apparatus." Also, another prominent non-Bayesian Freedman (1986, p. 127) has remarked, "When drawing inferences from data, even the most hard-bitten objectivist usually has to introduce assumptions and use prior information. The serious question is how to integrate that information into the inferential process and how to test the assumptions underlying the analysis." Last, Lehmann (1959, p. 62) writes in connection with non-Bayesian hypothesis testing, "Another consideration that frequently enters into the specification of a significance level is the attitude toward the hypothesis before the experiment is performed. If one firmly believes the hypothesis to be true, extremely convincing evidence will be required before one is willing to give up this belief, and the significance level will accordingly be set very low." From these quotations it is clearly the case that non-Bayesian, so-called objective analysts use considerable subjective information in their analyses, usually informally in a non-reproducible fashion.

On issue 10 in Table 1, Bayesians with a posterior density involving parameters of interest and nuisance parameters usually integrate out the nuisance parameters, a beautiful solution to the nuisance parameter problem. This integration has been mathematically interpreted as an averaging over conditional posterior densities of the parameters of interest given the nuisance parameters. However, non-Bayesians have no such solution to the nuisance parameter problem. For example, when a generalized least squares estimator involves nuisance parameters, say elements of a disturbance term covariance matrix, it is usual practice to insert estimates of the nuisance parameters and give the resulting "operational" estimator an asymptotic justification. Often times, the finite sample properties of such "operational" estimators are unknown and sometimes far from optimal. Bayesians, by integrating out nuisance parameters obtain a finite sample

posterior density and can use it to derive optimal, finite sample estimates of parameters of interest and to make exact finite sample probability statements about parameters' possible values.

With respect to issue 11, generally Bayesian and non-Bayesian methods produce good asymptotic results. See, e.g., Jeffreys (1998), Heyde and Johnstone (1979) and Chen (1985) for Bayesian asymptotic results for iid and stochastically dependent observations. In the former case, assumptions needed to derive asymptotic normality are the same in Bayesian and non-Bayesian cases; however, in the case of stochastically dependent observations, Heyde and Johnstone (1979) state that conditions needed for the asymptotic normality of posterior densities centered at the maximum likelihood estimate are weaker than those required for the asymptotic normality of maximum likelihood estimators. Also, in cases in which the number of parameters grows with the sample size, the incidental parameter case, both maximum likelihood and Bayesian estimators are inconsistent as emphasized by Neyman, Scott, Freedman, Diaconis and others. With just one observation per parameter, it is indeed unreasonable to have estimators' densities become degenerate as the sample size grows. By adding more information, e.g. by assuming a hyper distribution for the parameters and integrating out the incidental parameters, both Bayesian and maximum likelihood techniques yield consistent results.

Last, with respect to issue 12, generally Bayesian methods produce exact finite sample results in general whereas in many time series problems, simultaneous equations model problems, etc., non-Bayesian methods do not yield optimal finite sample estimators, exact confidence intervals and test statistics with known, finite sample distributions When this is the case, usually non-Bayesian approximate large sample inference techniques are employed as in analyses of cointegrated time series models, generalized method of moments problems, selection bias models, and many others. As stated above. Bayesian methods have been employed to obtain exact finite sample results for these and many other "difficult" models.

To close this section, a simple binomial problem will be considered to illustrate some general points, one that I have used for many years in my lectures. Suppose that in five trials. five sons are born and that the trials are considered independent with a probability θ of a male birth on each trial. How does one make inferences about the possible values of this parameter given the outcome, five sons in five trials? For example. what is a good estimate? How can one get a confidence interval? How can one test the hypothesis that $\theta = 1$? What are the odds on this hypothesis versus the hypothesis that $\theta = 1/2$? Or versus the hypothesis that the parameter is uniformly distributed? Note that the likelihood function is θ^5 and thus the maximum likelihood estimate is equal to 1! What is a good way to compute a confidence interval to accompany this estimate? Also, what test statistic is available to test the null hypothesis that the parameter's value = 1? Note that under this null hypothesis, the process is deterministic and thus there will be difficulty deriving the probability distribution of a test statistic under the null. This problem was analyzed years ago by Laplace who put a uniform prior on the parameter, and used Bayes' theorem to obtain the normalized posterior density that is proportional to the prior times the above likelihood function, that is the normalized posterior density is. $6\theta^5$. The modal value is 1, an optimal point

estimate relative to a zero-one loss function while the posterior mean is 6/7, an optimal point estimate relative to a squared error loss function and a special case of Laplace's Rule of Succession. Also, given whatever loss function that is appropriate, an optimal Bayesian point estimate can be derived that minimizes posterior expected loss. Further, posterior probability intervals giving the probability that the parameter's value lies between any two given values, say 1/2 and 1, are easily computed using the above posterior density. Also, the posterior odds on the hypotheses that the parameter's value is 1 versus that its value is 1/2 is easily evaluated. If the prior odds are 1:1 on these two hypotheses, the posterior odds in favor of 1 versus ½ is 32 to 1. Such problems are important not only regarding sex birth ratios but also in testing effectiveness of drugs, quality of products, the validity of scientific theories, etc. See Jeffreys (1998) and Zellner (1997, 1997a) for further analysis of the Laplace Rule of Succession.

4. Information Theory and Bayesian Analysis

My personal conclusion given the above considerations is that IT PAYS TO GO BAYES, to quote an old colleague of mine. However this should not be interpreted to mean that the Bayesian approach can not be improved. See, for example Soofi (1996, 1997), Jaynes (1988), Hill (1986,1988) and Zellner (1988,1991,1997) where it is recognized that inference involves information processing. In the Bayesian framework, the input information is information in a likelihood function, the data information, and information in a prior density. The output information is the information in a post data density for the parameters and a marginal density for the observations. By putting information measures on the inputs and outputs, we can seek the form of a proper output density for the parameters, say g, that minimizes the difference between the output information and the input information. Given the entropic measures of information employed, when this calculus of variations problem was solved, it was found that the solution is Bayes' theorem, namely take g proportional to the product of the prior density and likelihood function. Further, when g is taken in this form, it is the case that the output information equals the input information and none is lost in the process. Thus information-processing when Bayes' theorem is employed is 100% efficient. Jaynes (1988, p. 280-281) commented as follows on this result:

"...entropy has been recognized as part of probability theory since the work of Shannon (1948)...and the usefulness of entropy maximization...is thoroughly established...This makes it seem scandalous that the exact relation of entropy to other principles of probability theory is still rather obscure and confused. But now we see that there is, after all, a close connection between entropy and Bayes's theorem. Having seen such a start, other such connections may be found, leading to a more unified theory of inference in general. Thus in my opinion, Zellner's work is probably not the end of an old story but the beginning of a new one."

As part of the "new story," Zellner (1991) has considered the prior and sample information inputs to be of differing quality in deriving an information processing rule that minimizes the difference between output and input information subject to the output post data density for the parameters being proper. The result is a modified form of Bayes'

theorem that equates the quality adjusted input information to the quality adjusted output information. Similarly, when the information in a prior density is weighted differently from the sample information in a likelihood function, the optimizing information processing rule is different in form from Bayes' theorem, namely the post data density for the parameters is proportional to the prior raised to a power times the likelihood function raised to a power. When dynamic information processing is considered with possible costs of obtaining and adjusting to new information, from work in progress it is found that the dynamic optimization solution is different from the static solution, Bayes' theorem, just as static and dynamic maximization solutions differ in engineering, physics and the economic theory of the firm. Much work remains to be done in this area of information processing.

Another area in which information theory is useful is the problem of what to do when the form of the likelihood function is unknown. Of course for many years maxent or information theory has been employed to produce models for observations in physics and chemistry. For such work in economics, econometrics, finance and statistics, see, e.g., Davis (1941), Cover and Thomas (1991), Ryu (1990,1993), Stutzer (1996), Soofi (1996), Fomby and Hill (1997) and Zellner (1997). In addition, information criterion functionals have been employed to produce diffuse or non-informative as well as informative prior densities; see e.g. the review article on prior densities by Kass and Wasserman (1996) and results on various methods for producing prior densities in Bernardo and Smith (1994) and Zellner (1997, 127-153).

While maxent results are helpful in producing models for the observations when sampling properties of systems are known, e.g. sampling moment side conditions and other restrictions, when such sampling properties and restrictions are unknown, then a problem arises in the derivation of sampling densities for the observations using maxent. In such situations, some have resorted to empirical likelihood methods and bootstrapped likelihood functions; see, e.g., Boos and Monahan (1986) while others have introduced moment side conditions directly on functions of realized error terms of a model for the given data and from these have deduced implied post data moments of the model's parameters. For example, if $y_i = \theta + u_i$, i = 1,2,..., n, are n observed times to failure, $0 < y_i, \theta < \infty$. $\bar{y} = \theta + \bar{u}$ is the relation connecting the mean of the y's, $\bar{y}$, to the parameter θ and the mean of the realized error terms, $\bar{u}$. Then if we apply a subjective expectation operator to both sides of this last relation, we have for the given observation mean, $\bar{y} = E\theta + E\bar{u}$. If the measurements have been made properly with no outliers, no left out variables and departures from the linear form, we can then assume that $E\bar{u} = 0$. Given this moment assumption, we have $E\theta = \bar{y}$, that is the post data mean of the parameter is equal to the sample mean. Using this moment side condition, the proper probability density function with this mean that maximizes entropy is the exponential density, $f(\theta|D) = (1/\bar{y})\exp\{-\theta/\bar{y}\}$, where D denotes the given sample data and background information. This is an example of a Bayesian method of moments (BMOM) post data density for a parameter. It is called Bayesian since the density can be employed to compute the post data probability that the parameter lies between any two numbers, i.e. $\Pr\{a<\theta<b|D\}$, where D denotes the given data and prior assumptions, a solution to the

problem posed by Bayes (1763). See, e.g., Green and Strawderman (1996), Tobias and Zellner (1997), Zellner, Tobias and Ryu (1998, 1999), LaFrance (1999), van der Merwe and Viljoen (1998) and Zellner (1994, 1997, 1997a,1998) for additional applications of the BMOM to location, dichotomous random variable, univariate and multivariate regression, semi-parametric, time series and other models. In addition to moments for models' parameters, by making assumptions about future, as yet unrealized error terms and given the post data moments of parameters, it is possible to obtain moments of future, as yet unobserved values of future observations and use them as side conditions in deriving maxent probability densities for future observations as shown and applied in several of the papers cited above. Also, these predictive densities can be used to form Bayes' factors to discriminate between or among models. The use of maxent densities here is justified by their well known property of being the least informative densities that incorporate the information in the moment side conditions as explained in Jaynes (1983), Cover and Thomas (1991), Soofi (1996) and other works on information theory.

This emerging synthesis of probability theory and information processing is indeed exciting from a scientific point of view and also from an economic theory point of view in terms of the economics of information. For example, using the definition of the information provided by an experiment in Zellner (1997), with a value of a unit of information given, it becomes possible to value the information provided by an experiment. It is then possible to design experiments so as to maximize the net value of information, namely, the value of the output information minus the costs of the input information with respect to various control variables, e.g. the sample size, the number of strata to sample, etc. This represents an extension of some of the economic considerations bearing on the design of surveys and experiments described in the literature on sample survey and experimental design.

Even though the BMOM approach does not require an assumed form for the likelihood function, it does require a mathematical form for the relation satisfied by the observations and error terms. Obtaining the form and relevant input variables for such relations is a problem in reductive inference, as mentioned earlier. Unfortunately, formal procedures for obtaining satisfactory forms for such relations are not available. In the next section, an approach called the structural econometric modeling, time series analysis approach (SEMTSA) will be briefly presented and an application of it in macroeconomic modeling and forecasting will be described.

5. Formulating Models for Explanation and Prediction

The difficult problem of model formulation has been mentioned above. In this section, we describe the SEMTSA approach that has been formulated and is in the process of being applied to produce useful macroeconometric models that explain the past and are useful for prediction and policy analysis. As explained in previous work, Garcia-Ferrer et al. (1987), Palm (1976, 1977), Zellner (1979, 1984), and Zellner and Palm (1974), it is possible to derive univariate transfer function models from dynamic, multivariate time series macroeconomic and other models. Such transfer functions can be tested with data to determine whether their formulations and forecasting performance are

satisfactory. See Zellner and Palm (1975) for one example of this approach. However, if no satisfactory multivariate model is available, an alternative approach is to formulate univariate transfer functions using heuristic economic considerations and check to determine how well they perform in point and turning point forecasting. If a satisfactory transfer function equation, say for the rate of growth of real GDP for a country is obtained, it may be asked can a macroeconomic theoretical model be specified that algebraically implies a transfer function for the growth rate of real GDP that is close in form to that derived empirically from the data. Then the process is continued by producing other components, transfer functions for other variables, that perform well in terms of fitting past data and are successful in point and turning point forecasting. Thus our approach is to get components that work well in forecasting and then put them together to form a reasonable, economically motivated model for the observations. This approach contrasts markedly with the "general to specific " modeling approach employed by some in the macroeconometric literature. Note that there are many general models and if the wrong one is chosen, users of the "general to specific" modeling strategy will be disappointed.

In Garcia-Ferrer et al (1987), we began our analyses using an AR(3) model for annual real GDP growth rates since such a model could have two complex roots associated with a oscillatory component and a real root associated with a local trend. It didn't take long to find out that an AR(3) model did not work well in explaining variation in past data and in forecasting new data. A fundamental problem was that it was missing cyclical turning points by overshooting at the top of the business cycle and continuing to go down when the economy was recovering from downturns. Given that Burns and Mitchell had found in their research using pre-World War II data for the US, UK, German and French economies that money and stock prices tended to lead in business cycles that they studied, we decided to introduce the lagged rates of change of real money and of real stock prices as leading indicator variables. Earlier research has shown that changes in real money affect aggregate demand through real balance effects. Also, changes in real stock prices reflect all sorts of shocks hitting an economy, say oil shocks, war news, etc. Finally, by introducing the median rate of change of real stock prices for the countries in our sample in each country's forecasting equation, a "world return" variable, this led to contemporary error terms in countries' equations to be practically uncorrelated. Thus we could use a diagonal contemporaneous covariance matrix for our 18 countries' error terms rather than a non-diagonal matrix containing many parameters.

Thus our transfer function model or autoregressive-leading indicator (ARLI) model for the annual rate of change of real GDP for each of N countries in our sample was formulated as follows:

$$y_i = X_i \beta_i + u_i \qquad i = 1,2,3,\ldots,N \qquad (1)$$

where, for the i'th country, y_i is a Txl vector of observed annual growth rates, X_i is a Txk matrix of observations, of rank k, on input variables, namely $(1, y_{i,t-1}, y_{i,t-2}, y_{i,t-3}, SR_{i,t-1}, SR_{i,t-2}, GM_{i,t-1}, WR_{t-1})$. Here we have included three lagged growth rates to incorporate allowance for endogenous oscillatory behavior, two lagged values of the rate of growth of real stock prices, SR, one lagged value of the growth rate

of real money, GM, and the median of the N countries' one period lagged SR variables, a proxy for the world return, denoted by WR.

Initially, the model in (1) was implemented with data for nine industrialized countries and later for eighteen countries. The model and variants of it were fitted using data, 1954-73, and forecasting tests were performed using data for 1974-1981. Later the forecast period was extended to 1995. See Figure 1 for boxplots of the data. Of great value in improving our point forecasts' RMSEs was the use of Bayesian shrinkage or pooling techniques. See Table 2 for a dramatic demonstration of the effects of the use of shrinkage or pooling on countries' root mean squared errors of forecast. In addition to fixed parameter models, several time varying parameter models were also employed in point and turning point forecasting experiments.

Table 2

Root Mean Squared Errors of One Year Ahead Forecasts, 1974-1987,

Using Pooled and Unpooled ARLIWI Models*

RMSE (%)	Countries
	(a) Pooled
1.00-1.49	FRN GER NET SPN
1.50-1.99	AUR BEL CAN FIN ITY NOR SWD UK US
2.00-2.49	AUL DEN JAP SWZ
2.50-2.99	IRE
	Median = 1.74 Minimum = 1.17 Maximum = 2.53
	(b) Unpooled
1.00-1.49	UK
1.50-1.99	BEL FRN GER NET SPN SWD
2.00-2.49	AUR US
2.50-2.99	CAN DEN ITY NOR
3.00-3.49	AUL FIN IRE JAP SWZ
	Median = 2.37 Minimum = 1.39 Maximum = 3.32

*See equation (2) below and Zellner (1994) for more information regarding the autoregressive-leading indicator-world income (ARLIWI) models employed, results and references. Observations from the U.of Chicago's IMF International Financial Statistics data base were employed to fit the models, 1954-73 and to calculate one-year-ahead forecasts, 1974-87, updating estimates year by year.

Another variant of the model involved adding the current median growth rate of the 18 countries, denoted by $w' = (w_1, w_2, \ldots, w_T)$ as a variable in equation (1) as follows:

$$y_i = w\gamma_i + X_i\beta_i + u_i \qquad (2)$$

and adding an ARLI equation to explain the variation in the median growth rate, namely,

$$w_t = \alpha_o + \alpha_1 w_{t-1} + \alpha_2 w_{t-2} + \alpha_3 w_{t-3} + \alpha_4 MSR_{t-1} + \alpha_5 MGM_{t-1} + \varepsilon_t \quad (3)$$

t = 1,2,..., T, where MGM is the median annual growth rate of real money, and MSR is the median annual growth rate of real stock prices for the 18 countries in our sample; see Fig.1 for a plot of these variables. By combining the analysis of (2) and (3), it was possible to improve point and turning point forecasting performance by use of this autoregressive, leading indicator, world income, or ARLIWI model.

Given these models, one can use an economic aggregate supply and demand model to derive an equation in the form of (2); for details, see Appendix A. Further, Hong (1989) derived an equation in the form of (2) from a macroeconomic Hicksian IS-LM model while Min (1992) also did so using a generalized real business cycle model that he formulated. Thus equation (2) and variants of it are compatible with certain macroeconomic theoretical models.

The models described above in equations (1)- (3) and variants of them, including time-varying state space formulations, performed better in forecasting than various random walk, AR(3), Barro's and the Nelson-Plosser ARIMA benchmark models and about the same in terms of RMSEs of point forecasts as OECD models and probably better in terms of turning point forecasting. See Zellner, Hong and Gulati (1990), Zellner and Hong (1991), Zellner, Hong and Min (1991), Zellner, Tobias and Ryu (1998) and Zellner and Min (1999) for methods and results for forecasting turning points in 18 countries' real GDP growth rates and LeSage (1996) for use of similar methods in forecasting turning points in regional employment series. An editor of the International J. of Forecasting, R. Fildes (1994) commented as follows in his published review of one of our papers, Min and Zellner (1993):

The alternative models and methods "...were carefully compared based on their individual country performance measured by root mean squared errors for the years 1974-1987, and the distribution of these (particularly the median) The results offer mild support for using time-varying parameter schemes. Pooling [shrinkage] is important in improving accuracy. Model selection schemes are not particularly helpful except in so far as they identify pooled TVP [time varying parameter] models as the most accurate forecasting models. Combining does not improve over the TVP models and with the Granger-Ramanathan unconstrained scheme for choosing the weights, led to substantially poorer accuracy. Equal weights were not considered.

This paper is an excellent example of good empirical economics where the theory is utilized effectively in analyzing the problem in hand." (pp. 163-4)

While previous results are satisfying, there is still the problem of how to achieve improved forecasting and economic theoretical results with associated reductions in RMSEs and MAEs of forecast and higher percentages of correct turning point forecasts. One possible way to achieve improvement, emphasized for many years by Guy Orcutt, Tong Hun Lee's mentor at the U. of Wisconsin when he was a graduate student, is through thoughtful dissaggregation of the GDP variable. To show that disaggregation can produce improved forecast performance, in Zellner and Tobias (1998), equation (3) was employed, as earlier in Zellner and Hong (1989) to provide point forecasts of the median growth rates of the 18 countries. As an alternative way to forecast the countries' median growth rates, equations (2) and (3) were employed to forecast each countries' one year ahead growth rate and the median of these eighteen forecasts was employed as a point forecast. It was found that this latter disaggregated approach produced better one year ahead point forecasts with RMSEs of forecast approximately 20 per cent lower for the period 1974-84, that is RMSEs of 1.22 versus 1.54 percentage points and MAEs of 1.08 versus 1.44. Thus there is some evidence that disaggregation may help in certain circumstances.

To get a meaningful, economic disaggregation, we are currently formulating Marshallian demand, supply and entry equations, see e.g. Veloce and Zellner (1985) for major industrial sectors of the U.S. economy. For each sectoral model it is possible to solve for the output transfer function equation that is a function of the growth rates of factor prices, real aggregate income, household formation, real money balances, etc. Thus we shall have a set of transfer function equations for sectoral outputs that can be combined with transfer equations for real income and factor prices derived from demand and supply models for labor, capital, money, and bond markets. Then forecasts can be obtained for sectoral output growth rates and combined to provide a forecast of aggregate output. It will be of great interest to determine whether such forecasts are more accurate than aggregate forecasts, derived from aggregate data and whether this sectoral Marshallian model will add to our understanding of how the economy operates. Note that this model includes interactions among sectors such as construction, manufacturing, retail, agriculture, mining, etc..which have different cyclical properties. Instead of thinking of the economy as a single oscillator, it may be better to consider it to be a set of oscillating sectors coupled through common factor markets and product markets and affected by macro variables such as real income, real money balances, technical change, real interest rate, exchange rates and expectations of macroeconomic variables. When such a macroeconomic model of interacting sectors is formulated, it will be subjected to simulation and forecasting experiments to determine its properties and compare them to properties of the many theoretical Keynesian, Monetarist, Neo-Keynesian, Neo-Monetarist, Real Business Cycle, Generalized Real Business Cycle and currently operating macroeconometric models.

In this work, data are being collected and will be analyzed using Bayesian estimation, testing and forecasting techniques, described above. Such techniques have been found very useful in forecasting in our work with aggregate data for 18 countries and will also be extremely valuable in our future work with models using data for various interrelated sectors of economies.

Given Tong Hun Lee's keen interest in macroeconomic theory, forecasting and policy, it will be a pleasure to keep him and his colleagues informed of progress in our continuing research. Also, it is very satisfying to recognize the great progress in economic research and analysis that has been made in Korea and many other countries of the world that has resulted in economics and business as coming to be recognized as progressive, fruitful and useful sciences.

Appendix

Aggregate Demand and Supply Macroeconomic Model

Here we consider a short-run aggregate demand and supply model for total real output, Y, assumed given, that is equal to aggregate demand, the sum of real consumption, C, investment, I, government expenditures, G, and export, X minus import, IM, demands, i.e. $Y = C + I + G + X - IM$ is the equilibrium condition. Now, if we write $y = \log Y$ and assume a semi-logarithmic representation of total demand, incorporating a monetarist real balance effect, we have:

$$y_t = \gamma + \alpha_0 y_t + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \alpha_4 r_t^a + \alpha_5 m_{t-1} + \alpha_6 W_t + \alpha\ ex_t + \alpha_8 t \quad \text{(A1)}$$

where r^a = anticipated real rate of interest, m = log of real money balances, W = log of real world income, and ex = real exchange rate. The left side of

(A1) is given supply and the right side is a representation of total demand that includes lagged real income, real balance, interest rate, world income, exchange rate and trend effects. If the change in the anticipated real rate of interest is assumed given by the empirical relation,

$$\Delta r_t^a = \beta_0 + \beta_1 SR_{t-1} + \beta_2 SR_{t-2} + \beta_3 WR_{t-1} \quad \text{(A2)}$$

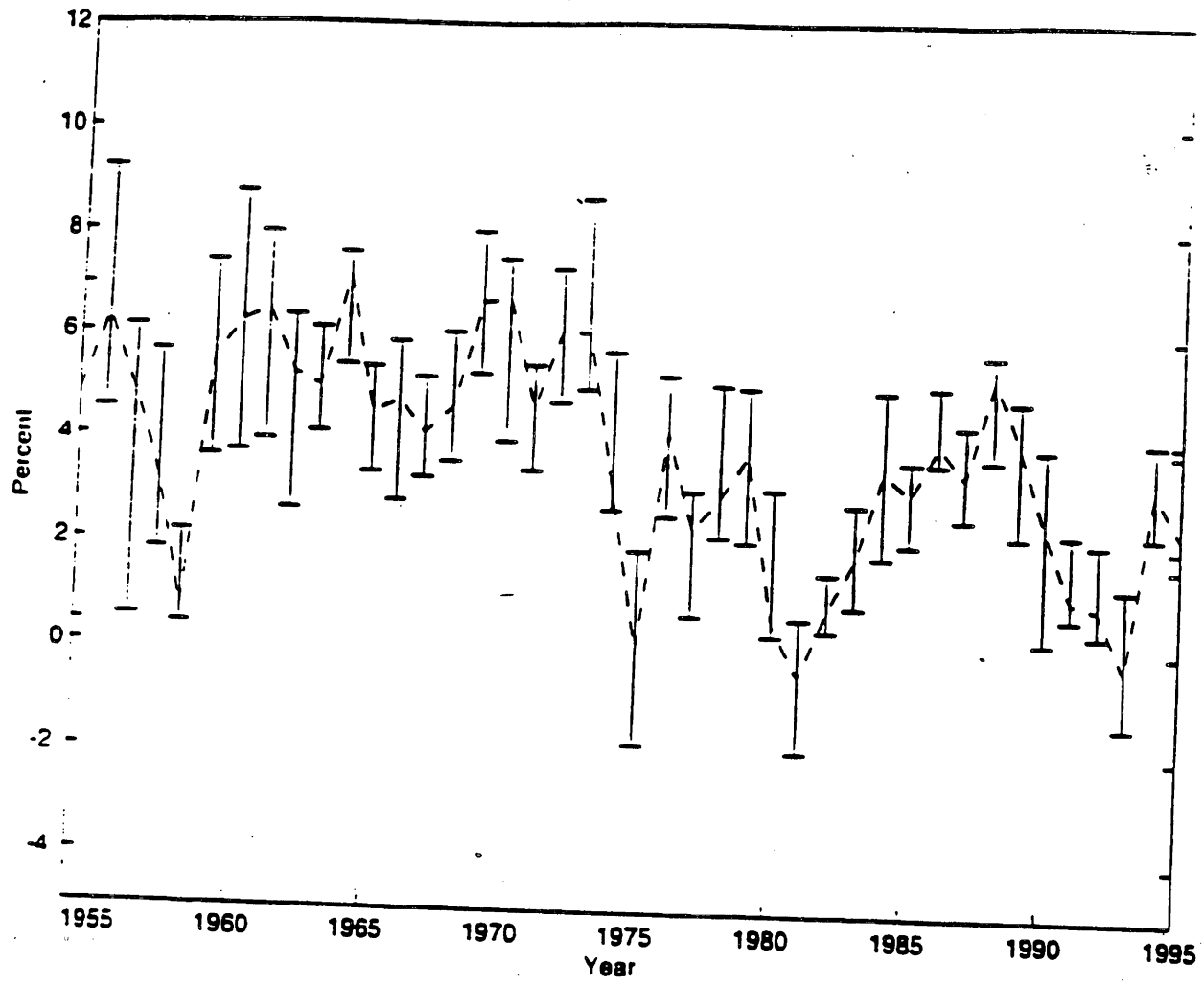
where SR = rate of change of real stock prices and WR = rate of change of real world stock prices, then we can first difference (A1) and substitute from (A2) to obtain:

$$\Delta y_t = \delta_0 + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \delta_3 \Delta y_{t-3} + \delta_4 SR_{t-1} + \delta_5 SR_{t-2} + \delta_6 WR_{t-1} + \delta_7 \Delta m_{t-1} + \delta_8 \Delta W_t + \delta_9 \Delta ex_t \quad \text{(A3)}$$

where $\Delta y_t = \log Y_t - \log Y_{t-1}$, the growth rate of Y, measured as real GDP in our empirical work. Note that (A3) is in the form of our ARLIWI model except for the inclusion of an exchange rate variable that has as yet not been included in our forecasting equations. Since the change in the exchange rate is close to being white noise, it is included in our error terms. Taking it out of the error term and including it in our forecasting equation may help to improve our forecasting results, as Granger noted many years ago in a general context. See also Hong (1989) and Min (1992) for more detailed derivations of our ARLIWI model from Hicksian IS-LM and generalized real business cycle models.

Figure 1: Medians and Interquartile Ranges for Growth Rates of Real Output (A). Real Money (B) and Real Stock Prices (C): for 18 Industrialized Countries: 1954-1995.[1]

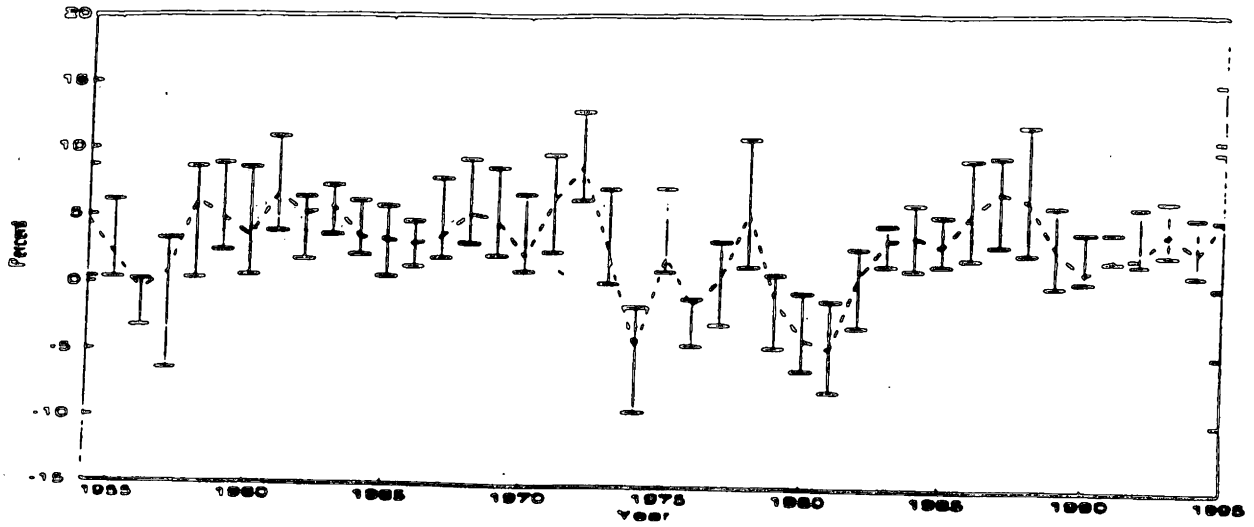
(A) Growth Rates of Real Output



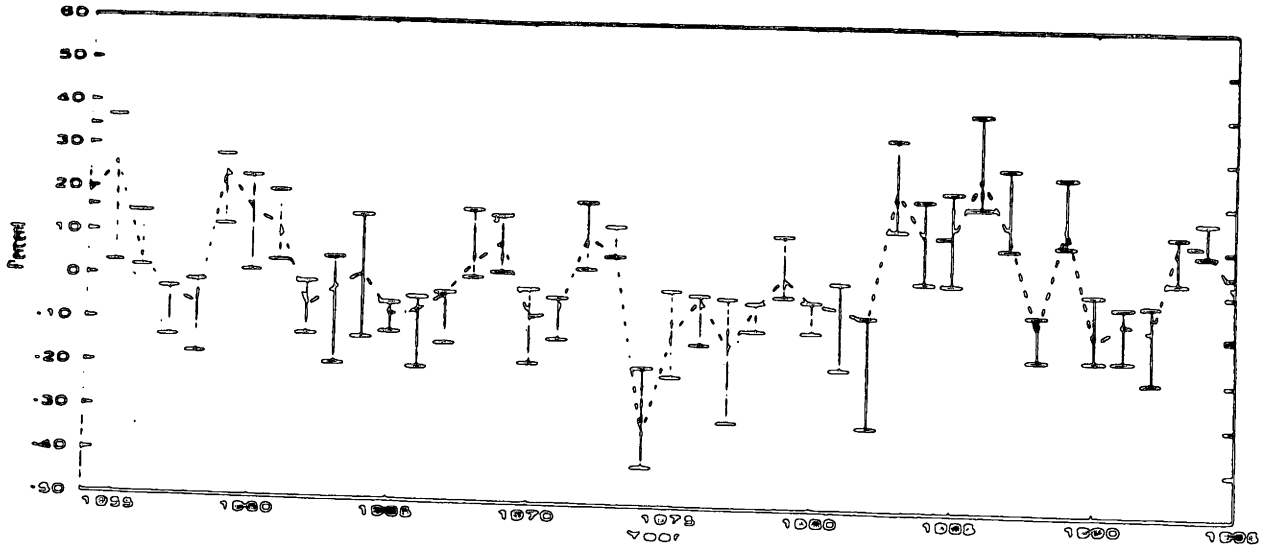
[1]The dashed line connects the annual median growth rates (the w_t's) and the vertical lines give the interquartile ranges.

Figure 1 Continued

(B) Growth Rates of Real Money



(C) Growth Rates of Real Stock Prices



References

Albert, J. and S. Chib (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *J. of the American Statistical Association*, 88, 669-679.

Bawa, V., Brown, S. and Klein, R. (1979), *Estimation Risk and Optimal Portfolio Choice*, Amsterdam: North-Holland.

Bayes, T. (1763), "An Essay Towards Solving a Problem in the Doctrine of Chances," *Phil. Trans. Royal Soc. Of London*, 53, 370-418, reprinted in *Press* (1989).

Berger, J. (1985), *Bayesian Decision Theory*, (2nd ed.), New York: Springer-Verlag.

Bernardo, J. and Smith, A. (1994), *Bayesian Theory*, New York: Wiley.

Berry, D., Chaloner, K. and Geweke, J. (1996), *Bayesian Analysis in Statistics and Econometrics*, New York: Wiley.

Boos, D. and Monahan, J. (1986), "Bootstrap Methods Using Prior Information," *Biometrika*, 73, 77-83.

Box, G. and Tiao, G. (1993), *Bayesian Inference in Statistical Analysis*, in *Wiley Classics Library*, reprint of 1973 ed., New York: Wiley.

Brown, S. (1976), "Optimal Portfolio Choice Under Uncertainty," Ph.D. Thesis, Graduate School of Business, U. of Chicago.

Chen, C. (1985), "On Asymptotic Normality of Limiting Density Functions with Bayesian Implications," *J. Royal Statistical Society*, Ser. B, 47, 540-546.

Chib, S. and Greenberg, E. (1996), "Markov Chain Monte Carlo Simulation Methods In Econometrics," *Econometric Theory*, 12, 409-431.

Cover, T. and Thomas, J. (1991), *Elements of Information Theory*, New York: Wiley.

Davis, H. (1941), *The Theory of Econometrics*, Bloomington, IN: Principia.

Diebold, F. and Lamb, R. (1997), "Why Are Estimates of Agricultural Supply So Variable?" *J. of Econometrics*, 357-373.

Feigl, H. (1953), "Notes on Causality," in *Readings in the Philosophy of Science*, Feigl, H. and Brodbeck, M. (eds.), New York: Appleton-Century-Crofts, 408-418.

Fildes, R. (1994), "Research on Forecasting," *International J. of Forecasting*, 10, 161-164.

Fomby, T. and Hill, R. (1997), *Advances in Econometrics: Applying Maximum Entropy To Econometric Problems*, Vol 12, Greenwich: Jai Press.

Freedman, D. (1986), "Reply," *J. of Business and Economic Statistics*, 4, 126-127.

Friedman, M. and Savage, J. (1948), "The Utility Analysis of Choices Involving Risk," *J. of Political Economy*, 56, 279-304.

Friedman, M. and Savage, J. (1952), "The Expected Utility Hypothesis and the Measurability of Utility," *J. of Political Economy*, 60, 463-474.

Gao, C. and Lahiri, K. (1999), "A Comparison of Some Recent Bayesian and Non-Bayesian Procedures for Limited Information Simultaneous Equation Models," 39 pp., paper presented to American Statistical Association Meeting, Baltimore, Maryland, August, 1999.

Garcia-Ferrer, A., Highfield, R., Palm, F. and Zellner, A. (1987), "Macroeconomic Forecasting Using Pooled International Data," *J. of Business and Economic Statistics*, 5, 53-67.

Gelman, A., Carlin, J., Stern, H. and Rubin, D. (1995), *Bayesian Data Analysis*, London: Chapman & Hall.

Geweke, J. (1989), "Bayesian Inference in Econometric Models Using Monte Carlo Integration," *Econometrica*, 57, 1317-1339.

Good, I.J. (1991), "The Bayes/Non-Bayes Compromise: A Review," invited paper presented at the General Methodology Session, American Statistical Association Meeting, August 1991, Atlanta Georgia.

Green, E. and Strawderman, W. (1996), "A Bayesian Growth and Yield Model for Slash Pine Plantations," *J. of Applied Statistics*, 23, 285-299.

Greene, W. (1998), *Econometric Analysis*, New York: Macmillan.

Hadamard, J. (1945), *The Psychology of Invention in the Mathematical Field*, New York: Dover.

Hanson, N. (1958), *Patterns of Discovery*, New York: Cambridge U. Press.

Heyde, C. and Johnstone, I. (1979), "On Asymptotic Posterior Normality for Stochastic Processes," *J. of the Royal Statistical Society*, Ser. B, 41, 184-189.

Hill, B. (1986), "Some Subjective Bayesian Considerations in the Selection of Models," *Econometric Reviews* 4, No. 2, 191-245 (with discussion).

_____ (1988), "Discussion," *The American Statistician*, 42, No. 4, 281-2.

Hong, C. (1989), "Forecasting Real Output Growth Rates and Cyclical Properties of Models: A Bayesian Approach," Ph.D. Doctoral Dissertation, Dept. of Economics, U. of Chicago.

Jacquier, E., Polson, N. and Rossi, P. (1994), "Bayesian Analysis of Stochastic Volatility Models," *J. of Business and Economic Statistics*, 12, 371-417 (with discussion).

James, W. and Stein, C. (1961), "Estimation with Quadratic Loss," in Neyman, J. (ed.), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, Berkeley: U. of California Press, 361-79.

Jorion, P. (1983), "Portfolio Analysis of International Equity Investments," Ph.D. Thesis, Graduate School of Business, U. of Chicago.

_____ (1985), "International Portfolio Diversification with Estimation Risk," *J. of Business*, 58, July.

Jaynes, E. (1983), *Papers on Probability, Statistics and Statistical Physics*, R. Rosenkrantz (ed.), Dordrecht, Netherlands: D. Reidl.

_____ (1984), "The Intuitive Inadequacy of Classical Statistics," *Epistemologia VII* (special issue on probability, statistics and inductive logic), 43-74.

_____ (1988), "Discussion," *The American Statistician*, 42, No. 4, 280-1.

Jeffreys, H. (1973), *Scientific Inference*, 3rd ed., Cambridge: Cambridge U. Press.

_____ (1971), *Collected Papers of Sir Harold Jeffreys on Geophysical and Other Sciences*, Vols. 1-6, London: Gordon & Breach Publishers.

_____ (1998), *Theory of Probability*, reprinted 3rd ed. In *Oxford Classic Texts*, Oxford: Oxford U. Press.

Judge, G., et al (1987), *Theory and Practice of Econometrics* (2nd ed.), New York: Wiley.

Kass, R. (1982), "Comment on "Is Jeffreys a 'Necessarist'?", *American Statistician*, 36, No. 4, 390-391.

_____ (1991), ed., *Sir Harold Jeffreys*, Special issue of Chance, 4, No. 2. Kass, R. and Raftery, A. (1995), "Bayes Factors," *J. American Statistical Assoc.*, 90, 773-795.

_____ And Wasserman, L. (1996), "The Selection of Priors by Formal Rules," *J. of the American Statistical Association*, 91, 1343-1370, corrections, 1998.

Kuezenkamp, H., McAleer, M. and Zellner, A. (eds.), to appear (1999), *Simplicity, Inference and Econometric Modeling*, Cambridge: Cambridge U. Press.

Lee, T. (1993), *The Collected Papers of Tong Hun Lee*, Seoul: Yonsei U. Press.

LaFrance, J. (1999), "Inferring the Nutrient Content of Food with Prior Information," *American J. of Agricultural Economics*, 81, August, 728-734.

Lehmann, E. (1959), *Testing Statistical Hypotheses*, New York: Wiley.

LeSage, J. (1996), "A Comparison of Techniques for Forecasting Turning Points in Regional Employment Activity," in Berry, D., Chaloner, K. and Geweke, J. (1996), 41-51.

Markowitz, H. (1959), *Portfolio Selection: Efficient Diversification of Investments*, New York: Wiley.

_____ (1987), *Mean-Variance Analysis in Portfolio Choice and Capital Markets*, Oxford: Basil Blackwell.

McCulloch, R. and Rossi, P. (1990), "Posterior, Predictive and Utility-based Approaches To Testing the Arbitrage Pricing Theory," *J. of Financial Economics*, 7-38.

McCulloch, R. and Rossi, P. (1994), "An Exact Likelihood Analysis of the Multinomial Probit Model," *J. of Econometrics*, 64, 207-240.

Min, C. (1992), "Economic Analysis and Forecasting of International Growth Rates Using Bayesian Techniques," Ph.D. dissertation, Dept. of Economics, U. of Chicago.

Min, C. and Zellner, A. (1993), "Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates," *J. of Econometrics*, 56, 89-118.

Palm, F. (1976), "Testing the Dynamic Specification of an Econometric Model with an Application to Belgian Data," *European Economic Review*, 8, 269-289.

_____ (1977), "On Univariate Time Series Methods and Simultaneous Equation Econometric Models," *J. of Econometrics*, 5, 379-388.

_____ (1983), "Structural Econometric Modeling and Time Series Analysis: An Integrated Approach," in Zellner, A. (ed.), *Applied Time Series Analysis of Economic Data*, Washington, DC: Bureau of the Census, U.S. Dept. of Commerce, 199-233.

Palm, F. and Zellner, A. (1992), "To Combine or Not to Combine? Issues of Combining Forecasts," *J. of Forecasting*, 11, 687-701.

Park, S. (1982), "Some Sampling Properties of Minimum Expected Loss (MELO) Estimators of Structural Coefficients," *J. of Econometrics*, 18, 295-311.

Pearson, K. (1938), *The Grammar of Science*, London: Everyman.

Pole, A., West, M. and Harrison, P. (1994), *Applied Bayesian Forecasting and Time Series Analysis*, London: Chapman & Hall.

Press, S. (1989), *Bayesian Statistics*, New York: Wiley.

Putnam, B. and Quintana, J. (1995), "The Evolution of Bayesian Forecasting Models," in Putnam, B. (ed.), *Applying Quantitative Discipline to Asset Allocation*, London: Euromoney Publications.

Quintana, J., Chopra,V. and Putnam, B. (1995), "Global Asset Allocation: Stretching Returns by Shrinking Forecasts," in *Proceedings of the Section on Bayesian Statistical Science, American Statistical Association.*

Robert, C. (1994), *The Bayesian Choice*, New York: Springer-Verlag.

Ryu, H. (1990), "Orthogonal Basis and Maximum Entropy Estimation of Probability Density and Regression Functions," Ph.D. Dissertation, Dept. of Economics, U. of Chicago.

_____ (1993), "Maximum Entropy Estimation of Probability Density and Regression Functions," *J. of Econometrics*, 56, 397-440.

Soofi, E. (1996), "Information Theory and Bayesian Statistics," in Berry, Chaloner and Geweke (1996), 179-189.

Stein, C. (1956), "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," in Neyman, J. (ed.), *Proc. Of Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, Berkeley: U. of California Press, 197-206.

Stutzer, M. (1996), "An Information-theoretic Index of Risk in Financial Markets," in Berry, Chaloner and Geweke (1996), 191-200.

Theil, H. (1971), *Principles of Econometrics*, New York: Wiley.

Tobias, J. (1999), Ph.D. thesis, Dept. of Economics, U. of Chicago.

Tobias, J. and Zellner, A. (1997), "Further Results on Bayesian Method of Moments Analysis of the Multiple Regression Model," to appear in *International Economic Review*.

Tsurumi, H. (1990), "Comparing Bayesian and Non-Bayesian Limited Information Estimators," in Geisser, S., Hodges, S., Press, J. and Zellner, A. (eds.), *Bayesian and Likelihood Methods in Statistics and Econometrics*, Amsterdam: North-Holland, 179-207.

Tukey, J. (1978), "Discussion of Granger on Seasonality," in Zellner, A., ed., *Seasonal Analysis of Economic Time Series*, Washington, DC: U.S. Govt. Printing Office, 50-53.

van der Merwe, A. and Viljoen, C. (1998), "Bayesian Analysis of the Seemingly Unrelated Regression Model," Dept. of Math. Statistics, U. of the Orange Free State U., presented to the South African Statistical Association Meeting, Nov. 1998.

Varian, H. (1975), "A Bayesian Approach to Real Estate Assessment," in Fienberg, S. And Zellner, A. (eds.), *Studies in Bayesian Econometrics and Statistics*, Amsterdam: North-Holland, 195-208.

Veloce, W. and Zellner, A. (1985), "Entry and Empirical Demand and Supply Analysis For Competitive Industries," *J. of Econometrics*, 30, 459-471.

Zellner, A. (1980), editor, *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys*, Amsterdam: North-Holland.

_____(1982), "Is Jeffreys a 'Necessarist'?" *American Statistician*, 36, No. 1, 28-30, and "Reply," 36, No. 4, 302-303.

_____ (1984). *Basic Issues in Econometrics*. Chicago: U. of Chicago Press, reprinted in 1987.

_____ (1986). "Bayesian Estimation and Prediction Using Asymmetric Loss Functions," *J. of the American Statistical Association*, 81, 446-451.

_____ (1988). "Optimal Information Processing and Bayes' Theorem," *The American Statistician*, 42, No. 4, 278-284 (with invited discussion).

_____ (1991). "Bayesian Methods and Entropy in Economics and Econometrics," in Grandy, W. and Schick L. (eds.). *Maximum Entropy and Bayesian Methods*, Netherlands: Kluwer cad. Publ., 17-31.

_____ (1994), "Time Series Analysis, Forecasting and Econometric Modeling: The Structural Econometric Modeling, Time Series Analysis Approach," *J. of Forecasting*, 13, 215-233, invited paper with discussion.

_____ (1996), *An Introduction to Bayesian Inference in Econometrics*, *Wiley Classics Library*, reprint of 1971 edition, New York: Wiley.

_____ (1997), *Bayesian Analysis in Econometrics and Statistics: The Zellner View and Papers*, Cheltenham, UK: Edward Elgar.

_____ (1997a), "The Bayesian Method of Moments (BMOM): Theory and Applications," *Advances in Econometrics*, Vol. 12, *Applying Maximum Entropy to Econometric Problems*, Fomby, T.B. and Hill, R.C., eds., 85-105.

Zellner, A. and Chetty, K. (1965), "Prediction and Decision Problems in Regression Models from the Bayesian Point of View," *J. of the American Statistical Association*, 60, 608-616.

_____ and Palm, F. (1974), "Time Series Analysis and Simultaneous Equation Models," *J. of Econometrics*, 2, 17-54.

_____and Palm, F. (1975), "Time Series Analysis of Structural Monetary Models of the U.S. Economy," *Sankhya*, Series C, 37, 12-56.

_____ and Tobias, J. (1998), "A Note on Aggregation, Disaggregation and Forecasting Performance," H.G.B. Alexander Research Foundation, Grad. School of Business, U. of Chicago, to appear in *J. of Forecasting*.

_____ and Vandaele, W. (1975), "Bayes-Stein Estimators for k-Means, Regression and Simultaneous Equation Models," in Fienberg, S. and Zellner, A. (eds.), *Studies in Bayesian Econometrics and Statistics*, Amsterdam: North-Holland, 627-653.

_____ and Rossi, P. (1984), "Bayesian Analysis of Dichotomous Quantal Response Models," *J. of Econometrics*, 25, 365-393.

_____ and Hong, C. (1989), "Forecasting International Growth Rates Using Bayesian Shrinkage and Other Procedures," in Schmidt, P. (ed.), "Issues in Econometric Forecasting, Annals Issue," *J. of Econometrics*, 40, 183-202.

_____ and Hong, C. (1991), "Bayesian Methods for Forecasting Turning Points in Economic Time Series: Sensitivity of Forecasts to Asymmetry of Loss Structures," in Lahiri, K. and Moore, G. (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records*, Cambridge: Cambridge U. Press, 129-140.

_____, Hong, C. and Min, C. (1991), "Forecasting Turning Points in International Output Growth Rates Using Bayesian Exponentially Weighted, Autoregression, Time Varying Parameter and Pooling Techniques," *J. of Econometrics*, 49, 275-304.

_____, Tobias, J. and Ryu, H. (1999a), "Bayesian Method of Moments Analysis of Semiparametric Regression Models," to appear in the *J. of the South African Statistical Association.*

_____, Tobias, J. and Ryu, H. (1999b), "Bayesian Method of Moments Analysis of Time Series Models with an Application to Forecasting Turning Points in Output Growth Rates," to appear in *Estadistica* with discussion.

_____ and Min, C. (1999), "Forecasting Turning Points in Countries' Output Growth Rates: A Response to Milton Friedman," *J. of Econometrics*, 88, 203-206.