ON INFORMATION, PRIORS, ECONOMETRICS, AND ECONOMIC MODELING

Francisco Venegas-Martínez
Centro de Investigación y Docencia Económicas

Enrique de Alba
Instituto Tecnológico Autónomo de México

Manuel Ordorica-Mellado*
El Colegio de México

Resumen: Se busca reconciliar aquellos métodos inferenciales que a través de la maximización de una funcional producen distribuciones *a priori* no-informativas e informativas. En particular, las distribuciones *a priori* de Evidencia Minimax (Good, 1968), las de Máxima Información de los Datos (Zellner, 1971) y las de Referencia (Bernardo, 1979) son vistas como casos especiales de la maximización de un criterio más general. Bajo un enfoque unificador se presentan las distribuciones *a priori* de Good-Bernardo-Zellner, que aplicamos en varios métodos de inferencia Bayesiana útiles en investigación económica. Asimismo, utilizamos las distribuciones de Good-Bernardo-Zellner en varios modelos económicos.

Abstract: This paper attempts to reconcile all inferential methods which by maximizing a criterion functional produce *non-informative* and *informative* priors. In particular, Good's (1968) Minimax Evidence Priors, MEP, Zellner's (1971) Maximal Data Information Priors, MDIP, and Bernardo's (1979) Reference Priors, RP, are seen as special cases of maximizing a more general criterion functional. In a unifying approach Good-Bernardo-Zellner priors are introduced and applied to a number of Bayesian inference procedures which are useful in economic research, such as the Kalman Filter and the Normal Linear Model. We also use the Good-Bernardo-Zellner distributions in several economic models.

* We are indebted to Arnold Zellner, José M. Bernardo, Jim Berger, George C. Tiao, Manuel Mendoza, and David Mayer for valuable comments and suggestions on earlier drafts of this paper. The authors bear sole responsibility for opinions and errors.

1. Introduction

The distinctive task in Bayesian analysis of deriving priors so that the inferential content of the data is minimally affected in the posterior distribution, has been of great interest for more than 200 years since the early work of Bayes (1763). More current approaches to this problem, based on the maximization of a specific criterion functional, have been suggested by Good (1969), Zellner (1971) and Bernardo (1979).

When modeling economic systems or conducting empirical research, prior information from previous research or from our knowledge of economic theory is always available. In either case, the estimates of the parameters of a regression model or the estimates of the time-varying parameters of a state-space model can usually be improved by incorporating any information about the parameters beyond that contained in the sample. In this work, we provide a broad class of priors that are likely to be useful in a variety of situations in economic modeling.

The principle of maximum invariantized negative cross-entropy is introduced in Good's (1969) minimax evidence method of deriving priors. There, the initial density is taken as the square root of Fisher's information. Zellner (1971) presents, for the first time, a method to obtain priors through the maximization of the *total* information about the parameters provided by independent replications of an experiment (prior average information in the data minus the information in the prior). Bernardo (1979) proposed a procedure to produce *reference* priors by maximizing the expected information about the parameters provided by independent replications of an experiment (average information in the posterior minus the information in the prior).

All of the above methods have certain advantages:

i) While Zellner's method is based on an exact finite sample criterion functional, Good's approach uses a limiting criterion functional, and Bernardo's procedure is based on asymptotic results. In Bernardo's proposal a reference prior (posterior) is defined as the limit of a sequence of priors (posteriors) that maximize finite-sample criteria. Many reference prior algorithms have been developed in a pragmatic approach in which results are most important. See, for instance, Berger, Bernardo and Mendoza (1989), and Berger and Bernardo (1989), (1992a), (1992b), Bernardo and Smith (1994), and Bernardo and Ramón (1997).

ii) The criterion functional used by Bernardo is cross-entropy, which satisfies a number of remarkable properties; in particular, it is invariant with respect to one-to-one transformations of the parameters (Lindley, 1956). In contrast, the total information functional employed by Zellner is invariant only for the location-scale family and under linear transformations of the parameters. Additional side conditions are needed to generate invariance under more general transformations.

iii) The way in which these methods have been tested is by seeing how well they perform in particular examples.

The evaluation is often based on contrasting the derived priors with Jeffreys' (1961), usually improper, priors which are somewhat arbitrary and inconsistent. In fact, there are cases in which one can strongly recommend avoiding Jeffreys' priors. See, for instance: Box and Tiao (1973), p. 314; Akaike (1978), p. 58; and Berger and Bernardo (1992a), p. 37.

In this paper, we attempt to reconcile all inferential methods that produce *non-informative* and *informative* priors. In our unifying approach, Minimax Evidence Priors (Good, 1968 and 1969), Maximal Data Information Priors (Zellner, 1971, 1977, 1991, 1993, and 1995) and Reference Priors (Bernardo, 1979 and 1996) are seen as special cases of maximizing an indexed criterion functional. Hence, properties of the derived priors will depend on the choice of indexes from a wide range of possibilities, instead of on a few personal points of view with *ad hoc* modifications. In the spirit of Akaike (1978) and Smith (1979), we can say that this will look more like Mathematics than Psychology —without denigrating the importance of the latter in the Bayesian framework. This unifying approach will enable us to explore a vast range of possibilities for constructing priors. Needless to say, a good choice will depend on the specific characteristics of the problem we are concerned with. It is worthwhile mentioning that our general method extends Soofi's (1994) pyramid in a natural way by adding more vertices and including their convex hull.

This paper is organized as follows. In section 2, we will introduce an indexed family of information functionals. In section 3 we will state a relationship between Bernardo's (1979) criterion functional and some members of the indexed family, on the basis of asymptotic normality. In section 4, we will study a Bayesian inference problem associated with convex combinations of relevant members of the proposed indexed family. Here, we will introduce the Good-Bernardo-Zellner priors and their *controlled* versions as solutions to the problem of maximizing discounted entropy. We will pay special attention to the existence and uniqueness of the solution to the corresponding optimization problems. In section 5, we will study the Good-Bernardo-Zellner priors as Kaiman Filtering priors. In section 6, we will apply Good-Bernardo-Zellner priors to the normal linear model, In section 7, we apply Good-Bernardo-Zellner priors to a variety of situations in economic modeling. Finally, in section 8, we present conclusions, acknowledge limitations, and make suggestions for future research.

2. An Indexed Family of Information Functionals

In this section, we define an indexed family of information functionals and study some distinguished members. For the sake of simplicity, we will remain in the single parameter case.

Suppose that we wish to make inferences about an unknown parameter $\theta \in \Theta \subseteq \mathbb{R}$ of a distribution P_θ, from which an observation, say, X, is available. Assume that P_θ has density $f(x \mid \theta)$ (Radon-Nikodym derivative) with respect to some fixed dominating σ–finite measure λ on $\mathbb{R}$ for all $\theta \in \Theta \subseteq \mathbb{R}$. That is, $dP_\theta/d\lambda = f(x \mid \theta)$ for all $\theta \in \Theta \subseteq \mathbb{R}$ and thus $P_\theta(A) = \int_A f(x \mid \theta)d\lambda(x)$ for all Borel sets $A \in \mathbb{R}$.

The Bayesian approach starts with a prior density, $\pi(\theta)$, to describe initial knowledge about the values of the parameter, θ. We will assume that $\pi(\theta)$ is a density with respect to some σ-finite measure μ on $\mathbb{R}$. Once a prior distribution has been prescribed, then the information about the parameter provided by the data, x is used to modify the initial knowledge, via Bayes' theorem, to obtain a posterior distribution of θ, namely, $f(\theta \mid x) \propto f(x \mid \theta)\pi(\theta)$ for every in $x \in \mathbb{R}$ (We use f generically to represent densities). The normalized posterior distribution is then used to make inferences about θ.

Let us define an infinite system of nesting functionals (cf. Venegas-Martínez, 1997):

$$V_{\gamma,\alpha,\delta}(\pi) = \frac{1}{1-\gamma}\int \pi(\theta)G(\mathcal{I}(\theta), \mathcal{F}(\theta), \gamma, \alpha, \delta)d\mu(\theta), \qquad (2.1)$$

where

$$G(\mathcal{I}(\theta), \mathcal{F}(\theta), \gamma, \alpha, \delta) =$$

$$\log\left\{\frac{\exp\left\{[\mathcal{F}(\theta)/\mathcal{I}(\theta)]^{1-\delta}[\mathcal{I}(\theta)]^{\frac{1-\gamma}{1+\alpha}} - \delta[\mathcal{I}(\theta)]^{1-\alpha}\right\}}{\pi(\theta)^{1-\gamma}}\right\},$$

$0 \leq \gamma < 1, \alpha \in \{0, 1\}, \delta \in \{0, 1\}$, and

$$\mathcal{I}(\theta) = \int \left(\frac{\partial}{\partial \theta} \log f(x \mid \theta) \right)^2 f(x \mid \theta) d\lambda(x) \tag{2.2}$$

is Fisher's information about θ provided by an observation X with density $f(x \mid \theta)$, and

$$\mathcal{F}(\theta) = \int f(x \mid \theta) \log f(x \mid \theta) d\lambda(x) \tag{2.3}$$

is the negative Shannon's information of $f(x \mid \theta)$, provided $\mathcal{I}(\theta)$ and $\mathcal{F}(\theta)$ exist. In the case that n independent observations of X are drawn from P_θ, say, $(X_1, X_2, ..., X_n)$, then $\mathcal{I}(\theta)$ and $\mathcal{F}(\theta)$ will still stand for the average Fisher's information and the average negative Shannon's information of $f(x \mid \theta)$, respectively. It is not unusual to deal with indexed functionals in inference problems about a distribution; see Good (1968).

In particular, note that for the location parameter family

$$f(x \mid \theta) = f(x - \theta), \ \theta \in \mathbb{R},$$

with the properties $\int [f'(x)]^2 / f(x) d\lambda(x) < \infty$ and $\int f(x) \log f(x) d\lambda(x) < \infty$, where $\lambda = \mu$ stands for the Lebesgue measure, both $\mathcal{I}(\theta)$ and $\mathcal{F}(\theta)$ are constant. Observe also that the scale parameter family $f(x \mid \theta) = (1/\theta) f(x/\theta)$, $\theta > 0$, with the above properties, satisfies the following relationship between $\mathcal{F}(\theta)$ and $\mathcal{I}(\theta)$:

$$\mathcal{F}(\theta) = \frac{1}{2} \log \mathcal{I}(\theta) + \text{constant}. \tag{2.4}$$

Throughout this paper, we will be concerned with the following indexed family:

$\mathcal{A} = conv[\overline{\{ V_{\gamma, \alpha, \delta}(\pi) \}}]$ = convex hull of the closure of the family $\{ V_{\gamma, \alpha, \delta}(\pi) \}$.

We readily identify a number of distinguished members of $\mathcal{A}$:

(*i*) Criterion for Maximum Entropy Priors, MAXENTP:

$$V_{0, 0, 1}(\pi) = - \int \pi(\theta) \text{iog} \pi(\theta) d\mu(\theta),$$

which is just Shannon's information measure of a density $\pi(\theta)$, or Jaynes' (1957) criterion functional to derive maximum entropy priors. Notice also that (2.3) can be rewritten in a simpler way as $\mathcal{F}(\theta) = - V_{0,0,1}(f(x \mid \theta))$.

(ii) Criterion for Minimax Evidence Priors, MEP:

$$V_{1,1,1}(\pi) \stackrel{\text{def}}{=} \lim_{\gamma \to 1} V_{\gamma,1,1}(\pi) = - \int \pi(\theta) \log \frac{\pi(\theta)}{p(\theta)} d\mu(\theta) - \log C, \qquad (2.5)$$

which is Good's invariantized negative cross-entropy, taking as initial density $p(\theta) = C[\mathcal{I}(\theta)]^{1/2}$ with $C = \{\int [\mathcal{I}(\theta)]^{1/2} d\mu(\theta)\}^{-1}$, provided that $\int [\mathcal{I}(\theta)]^{1/2} d\mu(\theta) < \infty$. We can also write (2.5) as:

$$V_{1,1,1}(\pi) - V_{0,0,1}(\pi) = \int \pi(\theta) \log [\mathcal{I}(\theta)]^{1/2} d\mu(\theta). \qquad (2.6)$$

(iii) Criterion for Maximal Data Information Priors, MDIP:

$$V_{0,0,0}(\pi) = \iint f(x) f(\theta|x) \log \frac{l(\theta|x)}{\pi(\theta)} d\mu(\theta)\, d\lambda(x), \qquad (2.7)$$

which is Zellner's criterion functional. Here, as usual,

$$f(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{f(x)}, \quad f(x) = \int f(x|\theta)\pi(\theta) d\mu(\theta),$$

and $l(\theta|x) = f(x|\theta)$ is the likelihood function. An alternative formulation of (2.7), which is useful, is given by

$$V_{0,0,0}(\pi) - V_{0,0,1}(\pi) = \int \pi(\theta) \mathcal{F}(\theta) d\mu(\theta). \qquad (2.8)$$

Some other members of $\mathcal{A}$ define new criterion functionals in which the information provided by the sampling model, $\mathcal{I}(\theta)$, plays an important role:

(iv) Criterion for Maximal Modified Data Information Priors, MMDIP:

$$V_{0,1,0}(\pi) = \iint f(x) f(\theta|x) \log \frac{[l(\theta|x)]^{[\mathcal{I}(\theta)]^{1/2}}}{\pi(\theta)} d\mu(\theta) d\lambda(x), \qquad (2.9)$$

which is the prior average information in the data *modified* by Fisher's information minus the information in the prior. Note that when $\mathcal{I}(\theta)$ is constant, (2.9) reduces to Zellner's criterion functional (up to a constant factor).

(*v*) Criterion for Maximal Fisher Information Priors, MFIP:

$$V_{0,1,1}(\pi) = -\int \pi(\theta) \log \frac{\pi(\theta)}{\exp\{[\mathcal{I}(\theta)]^{1/2}\}} d\mu(\theta) - 1, \tag{2.10}$$

which is the prior average Fisher's information minus the information in the prior.

3. Revisiting Bernardo's Reference Priors

The maximization of Bernardo's (1979) criterion is usually difficult. In order to obtain a simpler alternative procedure under specific conditions, we will derive a useful asymptotic approximation between Bernardo's criterion functional (or Lindley's information measure, 1956) and some members of the class $\mathcal{A}$. As stated in Bernardo (1979), the concept of reference prior is very general. However, in order to keep the analysis tractable, we will restrict ourselves to the continuous one-dimensional parameter case.

Suppose that there are n independent observations, $X_1, X_2, ..., X_n$, from a distribution P_θ, $\theta \in \Theta \subseteq \mathbb{R}$. Accordingly, the random vector $(X_1, X_2, ..., X_n)$, has density $d\mathcal{P}_\theta/d\nu = f(\xi|\theta) = \prod_{k=1}^n f(x_k|\theta)$ for all $\xi = (x_1, x_2, ..., x_n)$ and all $\theta \in \Theta \subseteq \mathbb{R}$, where

$$\mathrm{P}_\theta = \underbrace{P_\theta \otimes P_\theta \otimes \ldots \otimes P_\theta}_{n} \text{ and } \nu = \underbrace{\lambda \otimes \lambda \otimes \ldots \otimes \lambda}_{n}$$

Following Lindley (1956), a measure of the expected information about θ in a sampling model $f(x|\theta)$ provided by a random sample of size n, when the prior distribution of θ is $\pi(\theta)$, is defined to be

$$\mathcal{L}^{(n)}(\pi) = \int f(\xi) \int f(\theta|\xi) \log \frac{f(\theta|\xi)}{\pi(\theta)} d\mu(\theta) d\nu(\xi). \tag{3.1}$$

In order to obtain an asymptotic approximation of (3.1) in terms of $V_{1,1,1}$ and $V_{0,0,1}$, we state a limit theorem which justifies the passage of the limit under the integral signs in (3.1). The theorem rules out the possibility that the *essentials* of the statistical model, $f(\xi|\theta)$, change when samples grow in size. Let us rewrite (3.1) as:

$$\int_{|\omega| > n^{\rho}} \left(T_n(\omega) W_n(\omega) - T_n(\omega) \right) d\mu(\omega) \xrightarrow{P} 0; \tag{3.2}$$

(IX) The sequence of random variables $\{\log\ U_n\}_{n=1}^{\infty}$ where $U_n = \int T_n(\omega) W_n(\omega) d\mu(\omega)$ satisfies

$$\lim_{\varepsilon \to \infty} \sup_{n \geq 1} \int_{|\log U_n| \geq \varepsilon} |\log U_n| dP = 0, \tag{3.3}$$

where

$$P\{\xi \in A, \theta \in B\} = \int_B \pi(\theta) \int_A f(\xi|\theta) d\nu(\mathbf{x}) d\mu(\theta) \tag{3.4}$$

for all $A \in \mathbb{R}^n$ and $B \in \Theta$.

Then, as $n \to \infty$,

$$\mathcal{L}^{(n)}(\pi) - V_{1,1,1}(\pi) = - V_{0,0,1}(\varphi) + \log C\sqrt{n} + o(1), \tag{3.5}$$

where $\varphi(z)$ is the density of $Z \sim N(0, 1)$, and C is taken as in (2.4).

Some comments are in order: (I)-(IV) are standard regularity conditions, (V) states desirable properties for $\mathcal{I}(\theta)$, (VI) is a bounded variance condition, (VII) is a smoothness condition, (VIII) is a convergence condition, and (IX) says that the sequence $\{\log\ U_n\}_{n=1}^{\infty}$ is uniformly integrable with respect to P.

It can be shown (details can be found in Venegas-Martínez, 1990a) that (I)-(VI) lead to

$$T_n(\omega) \xrightarrow{\mathcal{L}} \exp\{\omega\sqrt{\mathcal{I}(\theta)}[Z - \frac{1}{2}\,\omega\sqrt{\mathcal{I}(\theta)}]\}, \tag{3.6}$$

where $Z \sim N(0, 1)$, and (3.6) along with (VII)-(IX) imply

$$\log\ U_n = \log \int T_n(\omega) W_n(\omega) d\mu(\omega) \xrightarrow{\mathcal{L}} \log \sqrt{2\pi / \mathcal{I}(0)} + \frac{1}{2} Z^2.$$

The conclusion of the theorem follows. Note that the right-hand side of (3.5) is independent of π. Thus, if conditions (I)-(IX) are fulfilled, instead of maximizing $\mathcal{L}^{(\infty)}(\pi)$, which is usually difficult, we can maximize $V_{1,1,1}(\pi)$, which is independent of n. Note that for maximization purposes the right-hand side of (3.5) becomes a constant.

Finally, it is worthwhile to note that the location parameter family $f(x|\theta) = f(x-\theta)$, with $\sqrt{f(x)}$ absolutely continuous in $\mathbb{R}$, and

$$\int [f'(x)]^2/f(x)\,d\lambda(x) < \infty,$$

fully satisfies the conditions of Theorem 3.1.

4. Good-Bernardo-Zellner Priors

In this section, we introduce Good-Bernardo-Zellner priors as solutions to the problem of maximizing convex combinations of elements of $\mathcal{A}$. We emphasize the existence and uniqueness of the solutions to the corresponding variational problems.

Very often, priors exist for which entropy becomes infinite, especially when dealing with the non-informative case. To overcome this difficulty, we propose the concept of discounted entropy and introduce Good-Bernardo-Zellner *controlled* priors as solutions to the problem of maximizing discounted entropy.

Throughout this section, we will discuss a number of Bayesian inferential problems associated with convex combinations of distinctive elements of $\mathcal{A}$. We begin considering

$$\mathcal{M}_\phi(\pi) \stackrel{\text{def}}{=} \phi \mathcal{V}_{1,1,1}(\pi) + (1-\phi)\mathcal{V}_{0,0,0}(\pi), \quad 0 \le \phi \le 1.$$

Clearly, $\mathcal{M}_\phi(\pi) \in \mathcal{A}$. To see that $\mathcal{M}_\phi(\pi)$ is concave with respect to π, it is enough to observe, as in Zellner (1991), that

$$\mathcal{V}_{0,0,0}(\pi(\theta)) = \mathcal{L}^{(1)}(\pi(\theta)) + \mathcal{V}_{0,0,1}(\pi(\theta)) - \mathcal{V}_{0,0,1}(f(x)),$$

is the sum of concave functions with respect to π (up to the constant $\mathcal{V}_{0,0,1}(f(x))$). Since $\mathcal{V}_{1,1,1}(\pi)$ is concave with respect to π, $\mathcal{M}_\phi(\pi)$ is also concave with respect to π.

Usually, in the absence of data, *supplementary* information, in terms of expectations about the parameter, comes from additional knowledge of the experiment, or from the experience of the researcher, namely,

$$\int a_k(\theta)\pi(\theta)d\mu(\theta) = \bar{a}_k, \; k = 1, 2, \ldots, s, \tag{4.1}$$

where both the functions a_k and the constants $\bar{a}_k$, $k = 1, 2, ..., s$, are known. Hereafter, we will assume that (4.1) does not lead to any contradiction with respect to $\pi(\theta)$.

In the rest of the paper, we will leave out the details in deriving the necessary conditions for the maximization problems. These conditions follow from very standard results in calculus of variations and optimal control (see, for instance, Kamien and Schwartz, 1991).

PROPOSITION 4.1. *Consider the Good-Bernardo-Zellner problem:*

$$\textit{Maximize} \quad M_\phi(\pi) \quad (\textit{with respect to } \pi)$$

subject to $\mathcal{C}$: $\int a_k(\theta)\pi(\theta)d\mu(\theta) = \bar{a}_k$, $k = 0, 1, 2, ..., s$, $a_0 \equiv 1 = \bar{a}_0$.

Then a necessary condition for a maximum is

$$\pi_\phi^*(\theta) \propto [\mathcal{I}(\theta)]^{\phi/2}\exp\{(1-\phi)\mathcal{F}(\theta) + \sum_{k=0}^{s} \lambda_k a_k(\theta)\}, \tag{4.2}$$

where λ_k, $k = \theta, 1, ..., s$, *are the Lagrange multipliers associated with the constraints* $\mathcal{C}$ *(cf. Zellner, 1995).*

Note that when no supplementary information is available, $\pi_\phi^*(0)$ is appropriate for an unprejudiced experimenter. Otherwise it will be suitable for an informed experimenter who is in favor of $\mathcal{C}$. Observe also that $\pi_1^*(\theta)$ is a Good-Bernardo prior, and $\pi_0^*(\theta)$ is a Zellner prior. In particular, consider the Bernoulli distribution, $f(x|\theta) = \theta^x(1-\theta)^{1-x}$, $\theta \le \theta \le 1$. In such a case, $\pi_1^*(\theta) = \theta^{-1/2}(1-\theta)^{-1/2}$ and $\pi_0^*(\theta) = \theta^\theta(1-\theta)^{1-\theta}$ for $\theta \in [\theta, 1]$, which are quite different.

COROLLARY 4.1. *Consider the location and scale parameter families,* $f(x|\theta) = f(x-\theta)$, $\theta \in \mathbb{R}$, *and* $f(x|\theta) = (1/\theta)f(x/\theta)$, $\theta > 0$, *respectively, both satisfying* $\int[f'(x)]^2/f(x)d\lambda(x) < \infty$ *and* $\int f(x)\log f(x)d\lambda(x) < \infty$. *Then, Good-Bernardo and Zellner priors agree regardless of the value of* $\phi \in (\theta, 1)$.

It is important to point out that when there is no supplementary information, we require $\mu(\Theta) < \infty$. Of course, the parameter space Θ can have limits as wide as needed to include the range where the likelihood for θ is relevant.

Note that Proposition 4.1 can be used recursively when there is additional supplementary information, say

$$\int a_k(\theta)\pi(\theta)d\mu(\theta) = \bar{a}_k, \quad k = s+1, s+2, \ldots, t. \tag{4.3}$$

In this case, using a cross-entropy formulation (Kullback 1959), we take (4.2) as the initial density, and (4.3) as the additional information. Hence,

$$\begin{aligned}\pi^*_\phi(\theta) &\propto [I(\theta)^{\phi/2}\exp\{(1-\phi)F(\theta) + \sum_{k=0}^{s} \lambda_k a_k(\theta)\} \exp\{\sum_{k=s+1}^{t} \lambda_k a_k(\theta)\} \\ &= [I(\theta)]^{\phi/2}\exp\{(1-\phi)F(\theta) + \sum_{k=0}^{t} \lambda_k a_k(\theta)\}.\end{aligned}$$

To deal with the (local) uniqueness of the solution to the problem stated in Proposition 4.1, we rewrite the constraints, C, as a function of the multipliers in the form $A(\Lambda) = [\int a_k(\theta)\pi^*_\phi(\theta)d\mu(\theta)]^s_{k=0} = \bar{A}$, where

$$\bar{A}^T = (\bar{a}_0, \bar{a}_1, \ldots, \bar{a}_s), \quad \Lambda^T = (\lambda_0, \lambda_1, \ldots, \lambda_s)$$

and T denotes the usual vector or matrix transposing operation.

PROPOSITION 4.2. *Let $\pi^*_\phi(\theta)$ be as in (4.2), and suppose that a_k, $k = \theta, 1, \ldots, s$, are linearly independent continuous functions in $L^2[\Theta, \pi^*_\phi\, d\mu]$, the space of all $\pi^*_\phi\, d\mu$-measurable functions $a(\theta)$ defined on Θ such that $|a(\theta)|^2$ is $\pi^*_\phi\, d\mu$-integrable. Suppose that $A(\Lambda)$ is defined on an open set $\Delta \subset \mathbb{R}^{s+1}$, and let Λ_0, be a solution to $A(\Lambda) =$ for a fixed value of $\bar{A} = \bar{A}_0$. Then there exists a neighborhood of Λ_0, $N(\Lambda_0)$, in which Λ_0 is the unique solution to $A(\Lambda) = \bar{A}_0$ in $N(\Lambda_0)$.*

The proof follows from the fact that $A(\Lambda)$ is continuously differentiable in Δ, with nonsingular derivative $A'(\Lambda) = [\int a_\iota(\theta)a_l(\theta)\pi^*_\phi(\theta)d\mu(\theta)]_{0\le \iota, l\le s}$, and from a straightforward application of the inverse function theorem (cf. Venegas-Martínez, 199θa).

From (4.1) we may derive the following necessary condition, which is useful in practical situations.

PROPOSITION 4.3. *The multipliers* $\Lambda^T = (\lambda_0, \lambda_1, ..., \lambda_s)$ *appearing in* (4.2) *satisfy the following non-linear system of* $s + 1$ *equations:*

$$1 = \lambda_0 + \log\left\{\int[\mathcal{I}(\theta)]^{\phi/2} e^{(1-\phi)\mathcal{F}(\theta)} \prod_{k=1}^{s} e^{\lambda_k a_k(\theta)} d\mu(\theta)\right\},$$

$$1 = \lambda_0 - \log \bar{a}_k + \log\left\{\int a_k(\theta)[\mathcal{I}(\theta)]^{\phi/2} e^{(1-\phi)\mathcal{F}(\theta)} \prod_{u=1}^{s} e^{\lambda_u a_u(\theta)} d\mu(\theta)\right\},$$

$$k = 1, 2, ..., s.$$

Moreover,

(*i*) *if the integral in the first equality has a closed-form solution, then the rest of the multipliers can be found from the relations:*

$$\frac{\partial \lambda_0}{\partial \lambda_k} = \bar{a}_k, \quad k = 1, 2, ..., s,$$

(*ii*) *the formula*

$$\phi V_{1,1,1}(\pi_\phi^*) + (1-\phi)[V_{0,0,0}(\pi_\phi^*) - 2V_{0,0,1}(\pi_\phi^*)] = 1 - \sum_{k=0}^{s} \lambda_k \bar{a}_k,$$

holds for all $0 \le \phi \le 1$.

Very often, researchers are concerned with assigning weights $\bar{a}_k$, $k = 1, 2, ..., s$ to regions A_k, $k = 1, 2, ..., s$, to express how likely it is that θ belongs to each region, based on past experience. The following result, based on Proposition 4.3, characterizes Good-Bernardo-Zellner priors when such supplementary information comes in the form of quantiles, and both $\mathcal{I}(\theta)$ and $\mathcal{F}(\theta)$ are constant. Under these assumptions, the non-linear system of $s + 1$ equations given in Proposition 4.3 is transformed into an homogeneous linear system of the same dimension as shown below:

PROPOSITION 4.4. *Suppose that the sets* $A_k = (b_k, b_{k+1}]$, $k = 1, 2, ..., s-1$ *and* $A_s = (b_s, b_{s+1})$ *with* $b_1 < b_2 < \ldots b_{u+1}$, $u \ge 2$, *constitute a partition of* Θ, $0 < \mu(\Theta) < \infty$. *Suppose also that both* $\mathcal{I}(\theta)$ *and* $\mathcal{F}(\theta)$ *are constant. Let* $\bar{a}_1, \bar{a}_2, ..., \bar{a}_s > 0$ *be such that* $\sum_{k=1}^{s} \bar{a}_k = 1$, *and* $\int l_{A_k}(\theta)\pi(\theta)d\mu(\theta) = \bar{a}_k$, $k = 1, 2, ..., s$. *If we define new multipliers:*

$$\omega_0 = e^{1-\lambda_0}/D \text{ where } D = [I(0)]^{\phi/2} e^{(1-\phi)\bar{H}(\theta)},$$

and $\omega_k = e^{\lambda_k}$, $k = 1, 2, \ldots, s$. *Then,* $\Omega = (\omega_0, \omega_1, \ldots, \omega_s)$ *can be found from the following homogeneous linear system:*

$$\begin{pmatrix} -1 & u_1 & u_2 & \cdots & u_s \\ -1 & v_1 & 0 & \cdots & 0 \\ -1 & 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & v_s \end{pmatrix} \begin{pmatrix} \omega_0 \\ \omega_1 \\ \omega_2 \\ \vdots \\ \omega_s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{4.4}$$

where $u_k = \mu(A_k)$, *and* $v_k = \bar{a}_k^{-1} u_k$, $k = 1, 2, \ldots, s$.

Observe that the determinant, Δ, of the matrix in (4.4) is given by

$$\Delta = \left(\frac{\sum_{k=1}^{s} \bar{a}_k - 1}{\prod_{k=1}^{s} \bar{a}_k} \right) \prod_{k=1}^{s} u_k,$$

which guarantees that there exists a unique nontrivial solution since $\sum_{k=1}^{s} \bar{a}_k = 1$. In this case, the solution is $\Omega^{*T} = (1, v_1^{-1}, v_2^{-1}, \ldots, v_s^{-1})$, and $\pi_\phi^* = \sum_{k=1}^{s} v_k^{-1} I_{A_k}$ (cf Venegas-Martínez, 1990b, 1990c, and 1992).

The following proposition extends the Good-Bernardo-Zellner priors to a richer family by using the MMDIP and MFIP criteria:

PROPOSITION 4.5. *Let*

$$N_{\phi,\psi}(\pi) \stackrel{\text{def}}{=} \phi V_{1,1,1}(\pi) + (1-\phi)(1-\psi) V_{0,0,0}(\pi)$$

$$+ (\psi(1-\phi)/2)[V_{0,1,1} + V_{0,1,0}]$$

$0 \le \phi, \psi \le 1$. *Then*

(i) $N_{\phi,\psi}(\pi) \in \mathcal{A}$ *and is concave with respect to* π.
(ii) A necessary condition for π *to be a maximum of the problem*

$$\textit{Maximize} \quad N_{\phi,\psi}(\pi)$$

subject to $\mathcal{C}$: $\int a_k(\theta)\pi(\theta)d\mu(\theta) = \bar{a}_k$, $k = 0, 1, 2, \ldots, s$, $a_0 \equiv 1 = \bar{a}_0$, *is given by*

$$\pi^*_{\phi,\psi}(\theta) \propto [I(\theta)]^{\phi/2}\exp\Bigg\{(1-\phi)\,(1-\psi)\;\mathcal{F}(\theta) + \frac{\psi(1-\phi)}{2}\left[[\mathcal{I}(\theta)]^{1/2} + \frac{\mathcal{F}(\theta)}{[\mathcal{I}(\theta)]^{1/2}}\right] + \sum_{k=0}^{s}\lambda_k a_k(\theta)\Bigg\}, \tag{4.5}$$

where λ_k, $k = 0, 1, \ldots, s$ are the Lagrange multipliers associated with the constraints $\mathcal{C}$.

The second term inside the exponential of (4.5) is the average between Fisher's information and the negative relative Shannon-Fisher information. Note that $\pi^*_{\phi,0}(\theta)$ is just the Good-Bernardo-Zellner prior.

In the following proposition, Good-Bernardo-Zellner type priors are derived as MAXENTP solutions by treating (2.5) and (2.8) as constraints (for the rationale of MAXENTP methods see Jaynes 1982).

PROPOSITION 4.6. *Consider the Jaynes-Good-Bernardo-Zellner problem:*

$$\textit{Maximize} \quad V_{0,0,1}(\pi)$$

$$\textit{subject to:}\begin{cases} V_{1,1,1}(\pi) - V_{0,0,1}(\pi) = \bar{b}_1, \\ V_{0,0,0}(\pi) - V_{0,0,1}(\pi) = \bar{b}_2, \\ \int a_k(\theta)\pi(\theta)d\mu(\theta) = \bar{a}_k, \ k = 0, 1, 2, \ldots, s, \ a_0 \equiv 1 = \bar{a}_0. \end{cases}$$

Then a necessary condition for a maximum is

$$\pi^*(\theta) \propto [I(\theta)]^{p_1/2}\exp\{p_2\mathcal{F}(\theta) + \sum_{k=0}^{s}\lambda_k a_k(\theta)\}, \tag{4.6}$$

where p_j, $j = 1, 2$, *and* λ_k, $k = 0, 1, \ldots, s$, *are the Lagrange multipliers associated with the constraints.*

Unlike the coefficients ϕ and $1 - \phi$ appearing in (4.6), the multipliers p_j, $j = 1, 2$, do not necessarily add up to 1.

Typically, priors exist for which the Shannon-Jaynes entropy becomes infinite. One way to remedy this situation consists in discounting entropy at a constant rate $\nu > 0$. The following proposition introduces Good-Bernardo-Zellner controlled priors as solutions to the problem of maximizing discounted entropy.

PROPOSITION 4.7. *Consider the discounted version of the problem stated in the preceding proposition:*

$$\textit{Maximize} \quad -\int e^{-\nu\theta}\pi(\theta)\log \pi(\theta) d\mu(\theta),$$

subject to:

$$\begin{cases} \dfrac{1}{\pi(\theta)}\dfrac{dh_1(\theta)}{d\mu(\theta)} = \log[\mathcal{I}(\theta)]^{1/2},\ h_1(-\infty) = 0, \\ h_1(\infty) = V_{1,1,1}(\pi) - V_{0,0,1}(\pi) < \infty, \\ \dfrac{1}{\pi(\theta)}\dfrac{dh_2(\theta)}{d\mu(\theta)} = \mathcal{F}(\theta),\ h_2(-\infty) = 0,\ \ h_2(\infty) = V_{0,0,0}(\pi) - V_{0,0,1}(\pi) < \infty, \\ \dfrac{1}{\pi(\theta)}\dfrac{dg_k(\theta)}{d\mu(\theta)} = a_k(\theta),\ g_k(-\infty) = 0,\ g_k(\infty) < \infty,\ k = 0, 1, 2, \ldots, s \end{cases}$$

where $a_0 \equiv 1 = \bar{a}_0$. *Then, a necessary condition for* $\pi^*(\theta)$ *to be an optimal control is given by*

$$\pi^*(\theta) \propto [\mathcal{I}(\theta)]^{\frac{p_1(\theta)}{2}} \exp\{p_2(\theta)\mathcal{F}(\theta) + \sum_{k=0}^{s} \lambda_k(\theta) a_k(\theta)\}, \tag{4.7}$$

where $p_j(\theta) = p_{j0}e^{\nu\theta}$, $j = 1, 2$, *and* $\lambda_k(\theta) = \lambda_{k0}e^{\nu\theta}$, $k = 0, 1, \ldots, s$ *are the co-state variables associated with the state variables* $h_j(\theta)$, $j = 1, 2$, *and*

$g_k(\theta)$, $k = 0, 1, \ldots, s$, respectively. Furthermore, the constants ρ_{j0}, $j = 1, 2$, and λ_{k0}, $k = 0, 1, \ldots, s$ can be computed from the following non-linear system of $s + 3$ equations:

$$1 + \log h_1(\infty) =$$

$$\log\left\{\int \log[\mathcal{I}(\theta)]^{1/2} m(\rho_{10}, \rho_{20}, \lambda_{00}, \lambda_{10}, \ldots, \lambda_{s0}; \theta)d\mu(\theta)\right\},$$

$$1 + \log h_2(\infty) = \log\left\{\int \mathcal{F}(\theta) m(\rho_{10}, \rho_{20}, \lambda_{00}, \lambda_{10}, \ldots, \lambda_{s0}; \theta)d\mu(\theta)\right\},$$

$$1 + \log g_k(\infty) = \log\left\{\int a_k(\theta) m(\rho_{10}, \rho_{20}, \lambda_{00}, \lambda_{10}, \ldots, \lambda_{s0}; \theta)d\mu(\theta)\right\},$$

$$k = 0, 1, 2, \ldots, s,$$

where

$$m(\rho_{10}, \rho_{20}, \lambda_{00}, \lambda_{10}, \ldots, \lambda_{s0}; 0) = \left([\mathcal{I}(\theta)]^{\frac{\rho_{10}}{2}} e^{\rho_{20}\mathcal{F}(\theta)} e^{\lambda_{00}} \prod_{u=1}^{s} e^{\lambda_{u0} a_u(\theta)} \right)^{e^{\nu\theta}}.$$

5. Kalman Filtering Priors

In this section, we will study Good-Bernardo-Zellner priors as Kalman Filtering priors (Kalman 1960, and Kalman and Bucy 1961). We will continue to work with the single parameter case, and focus our attention on the location parameter family.

Let $Y_1, Y_2, \ldots, Y_t$ be a set of indirect measurements, from a polling system or a sample survey, of an unobserved state variable β_t. The objective is to make inferences about β_t. The relationship between Y_t and β_t is specified by the measurement or observation equation:

$$Y_t = A_t \beta_t + \varepsilon_t, \tag{5.1}$$

where $A_t \neq 0$ is known, and ε_t is the observation error distributed as $\mathcal{N}(0, \sigma^2_{\varepsilon_t})$ with $\sigma^2_{\varepsilon_t}$ known. Note that the main difference between the measurement equation and the linear model is that, in the former, the coefficient β_t changes with time. Furthermore, we suppose that β_t is driven by a first order autoregressive process, that is,

$$\beta_t = Z_t\, \beta_{t-1} + \eta_{t-1}, \tag{5.2}$$

where $Z_t \neq \theta$ is known, and $\eta_t \sim \mathcal{N}(\theta, \sigma^2_{\eta_t})$ with $\sigma^2_{\eta_t}$ known. In what follows, we will assume that β_0, ε_t, and η_t are independent random variables. We could state nonlinear versions of (5.1) and (5.2), but this would not make any essential difference in the subsequent analysis.

Suppose now that at time $t = \theta$, supplementary information is given by $\hat{\beta}_0$ and $\hat{\sigma}_0^2$, the mean and variance of β_0 respectively. That is,

$$\mathcal{C}: \begin{cases} \int_{-\infty}^{\infty} \pi(\beta_0)d\beta_0 = 1, \\ \int_{-\infty}^{\infty} \beta_0 \pi(\beta_0)d\beta_0 = \hat{\beta}_0, \\ \int_{-\infty}^{\infty} (\beta_0 - \hat{\beta}_0)^2 \pi(\beta_0)d\beta_0 = \hat{\sigma}_0^2. \end{cases} \tag{5.3}$$

In this case, the Good-Bernardo-Zellner prior is given by

$$\pi^*_\phi(\beta_0) \propto [\mathcal{I}(\beta_0)]^{\phi/2} \exp\{(1-\phi)\mathcal{F}(\beta_0) + \lambda_0 + \lambda_1\beta_0 + \lambda_2(\beta_0 - \hat{\beta}_0)^2\}, \tag{5.4}$$

where λ_j, $j = \theta, 1, 2$, are Lagrange multipliers (cf. Venegas-Martínez *et al.*, 1995, and Ordorica-Mellado, 1995).

Suppose that, at time t, we wish to make inferences about the conditional state variable $\theta_t = \beta_t \mid I_t$, where $I_t = \{Y_1, Y_2, ..., Y_{t-1}\}$. To obtain a posterior distribution of θ_t, the information provided by the measurement Y_t, with density $f(Y_t \mid \theta_t)$, is used to modify the initial knowledge in $\pi^*_\phi(\theta_t)$ according to Bayes' theorem:

$$f(\theta_t \mid Y_t) \propto f(Y_t \mid \theta_t)\pi^*_\phi(\theta_t). \tag{5.5}$$

We are now in a position to state the Bayesian recursive updating procedure of the Kalman Filter (KF) for the location parameter family $f(Y_t \mid \theta) = f(Y_t - \theta)$, $\theta \in \mathbb{R}$. To start off the KF procedure, we substitute (5.4) in (5.3), obtaining the result that the Good-Bernardo-Zellner prior at time $t = \theta$, is given by $\mathcal{N}(\hat{\beta}_0, \hat{\sigma}_0^2)$, which describes the initial knowledge of the system. Proceeding inductively, at time t, $\hat{\beta}_{t-1}$ and $\hat{\sigma}^2_{t-1}$ become *supplementary* information, and therefore the Good-Bernardo-Zellner prior at time t is given by

$$\theta_t = \beta_t \mid \mathcal{I}_t \sim N(Z_t \hat{\beta}_{t-1}, M_t), \tag{5.6}$$

where

$$M_t = Z_t^2 \hat{\sigma}^2_{t-1} + \sigma^2_{\eta t-1}. \tag{5.7}$$

The sampling model (or likelihood function) is determined by

$$Y_t \mid \theta_t \sim N(A_t \beta_t, \sigma^2_{\varepsilon_t}). \tag{5.8}$$

The posterior distribution, at time t, is then obtained by substituting both (5.6) and (5.7) in (5.5):

$$f(\theta_t \mid Y_t) \propto \exp\{-\frac{1}{2}[(A_t \beta_t - Y_t)^2 \sigma_{\varepsilon_t}^{-2} + (\beta_t - Z_t \hat{\beta}_{t-1})^2 M_t^{-1}]\}.$$

Noting that $\pi^*_\phi(0_t)$ is a natural conjugate prior, it follows that

$$\theta_t \mid Y_t \sim N[Z_t \hat{\beta}_{t-1} + K_t(Y_t - A_t Z_t \hat{\beta}_{t-1}), M_t - K_t A_t M_t],$$

where

$$K_t = M_t A_t(\sigma^2_{\varepsilon_t} + A_t^2 M_t)^{-1}. \tag{5.9}$$

This, of course, means that

$$\begin{cases} \hat{\beta}_t = Z_t \hat{\beta}_{t-1} + K_t(Y_t - A_t Z_t \hat{\beta}_{t-1}), \\ \hat{\sigma}^2_t = M_t - K_t A_t M_t. \end{cases} \tag{5.10}$$

We then proceed with the next iteration. Equations (5.7), (5.9), and (5.10) are known in the literature as the KF. The previous analysis can be summarized in the following proposition:

PROPOSITION 5.1. *Consider the state-space representation:*

$$\begin{cases} Y_t = A_t \beta_t + \varepsilon_t, \\ \beta_t = Z_t \beta_{t-1} + \eta_{t-1}, \end{cases}$$

defined as in (5.1) and (5.2). Suppose that supplementary information about the mean and variance of β_0 *is available. Let* $\theta_t = \beta_t \mid I_t$, *where* $I_t = \{Y_1, Y_2, ..., Y_{t-1}\}$, *and consider the location parameter family,* $f(Y_t \mid \theta) = f(Y_t - \theta)$, $\theta \in \mathbb{R}$, *along with the properties stated in Corollary 4.1. Then, under the Good-Bernardo-Zellner prior,* $\pi^*_\phi(\theta_t)$, *the posterior estimate of* β_t, $\hat{\beta}_t$, *is given by*

$$\hat{\beta}_t = \omega_t Z_t \hat{\beta}_{t-1} + (1 - \omega_t)(Y_t / A_t),$$

where $\omega_t = \sigma^2_{\varepsilon_t}(\sigma^2_{\varepsilon_t} + A_t^2 M_t)^{-1}$.

6. Normal Linear Mode

The results on Good-Bernardo-Zellner priors given so far can be easily extended to the multi-dimensional parameter case, namely,

$$\theta = (\theta_1, \theta_2, ..., \theta_m) \in \Theta \subseteq \mathbf{R}^m, \ m > 1.$$

Consider a vector of independent and identically distributed normal random variables $(X_1, X_2, ..., X_n)$ with common and known variance σ^2 satisfying

$$\mathrm{E}(X_k) = a_{k1}\theta_1 + a_{k2}\theta_2 + \dots + a_{km}\theta_m, \quad k = 1, 2, ..., n \tag{6.1}$$

where $A = (a_{ij})$ is a matrix of known coefficients for which $(A^T A)^{-1}$ exists.

Let X and θ stand for the column vectors of variables X_k and parameters θ_j, respectively. Then (6.1) can be written in matrix notation as, $\mathrm{E}(X) = A\theta$. In this case, we have

$$f(\xi|\theta) = (\frac{1}{2\pi\sigma^2})^{n/2} \exp\{-\frac{1}{2\sigma^2} \|\xi - A\theta\|^2\}, \tag{6.2}$$

where $\xi = (x_1, x_2, ..., x_n)$. Since σ^2 has been assumed known, only the location parameter is unknown. The analogue of (2.2) is now given by the matrix:

$$\mathcal{I}_n(\theta) \equiv \left(\int \left(\frac{\partial}{\partial \theta_\iota} \log f(x|\theta) \right) \left(\frac{\partial}{\partial \theta_l} \log f(x|\theta) \right) f(x|\theta) d\lambda(x) \right)_{1 \leq \iota, l \leq m}$$

$$= \frac{1}{\sigma^2} A^T A \tag{6.3}$$

and so $\det[\mathcal{I}_n(\theta)]$ is constant,which implies that the Good-Bernardo-Zellner prior distribution $\pi^*_\phi(\theta)$, describing a situation of vague information on θ, must be a locally uniform prior distribution.

Let $\hat{\theta}$ be the least squares estimate for θ. Then it is known that $A^T A\hat{\theta} = A^T X$, $E(\hat{\theta}) = \theta$, and $\mathrm{Var}(\hat{\theta}) = \sigma^2(A^T A)^{-1}$. Noting from equation (6.2) that

$$f(\xi|\theta) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} \exp\{ - \frac{1}{2\sigma^2} (\|\xi - A\hat{\theta}\|^2 + \langle A^T A(\theta - \hat{\theta}), \theta - \hat{\theta} \rangle)\},$$

and applying Bayes' theorem, we get as the posterior distribution of θ

$$f(\theta|\xi) = (2\pi)^{-\frac{m}{2}} (\det[\frac{1}{\sigma^2} A^T A])^{\frac{1}{2}} \exp\{ - \frac{1}{2} \langle \frac{1}{\sigma^2} A^T A(\theta - \hat{\theta}), \theta - \hat{\theta} \rangle\}.$$

If supplementary information about the mean, c, and the variance-covariance matrix, D, is now incorporated, then the (informative) Good-Bernardo-Zellner prior is given by

$$\pi^*_\phi(\theta) = (2\pi)^{-\frac{m}{2}} (\det[D])^{-\frac{1}{2}} \exp\{ - \frac{1}{2} \langle D^{-1}(\theta - c), \theta - c \rangle\}.$$

The posterior distribution is now

$$f(\theta|\xi) = (2\pi)^{-\frac{m}{2}} (\det[B])^{\frac{1}{2}}$$
$$\times \exp\{ - \frac{1}{2} \langle B[\theta - ((DB)^{-1} c + \frac{1}{\sigma^2} B^{-1} A^T A\hat{\theta})] \theta$$
$$- ((DB)^{-1} c + \frac{1}{\sigma^2} B^{-1} A^T A\hat{\theta}) \rangle\},$$

where $B = D^{-1} + \frac{1}{\sigma^2} A^T A$.

7. Good-Bernardo-Zellner Priors in Economic Modeling

In this section we apply Good-Bernardo-Zellner priors to a variety of situations in economic modeling.

Example 7.1

Let us examine the behavior of an individual who learns about the parameters of her/his utility function under inflation. If we think of the parameters as random variables, then the information gained from experience (consumption) is incorporated into a prior distribution. Once a prior is available, the agent makes consumption decisions. To illustrate this process, we shall borrow some ideas from Calvo (1986). Let us consider a small open economy with a single infinitely-lived consumer in a world with a single perishable consumption good. Suppose that the good is freely traded, and its domestic price level, P_t, is determined by the purchasing power parity condition, namely $P_t = P_t^* E_t$, where P_t^* is the foreign-currency price of the good, and E_t is the nominal exchange rate. Throughout the paper, we will assume, for the sake of simplicity, that P_t^* is equal to 1. We also assume that the exchange-rate initial value, E_0, is known and equal to 1.

The expected utility function of a representative individual at the present, $t = 0$, has the following separable form:

$$V = \int_0^\infty \left\{ \int_0^\infty u(c_t; \theta) e^{-rt} dt \right\} \pi(\theta) d\theta \tag{7.1}$$

where $u(c_t; \theta)$ is the utility of consumption; c_t is consumption; $\theta > 0$ is a parameter related to the utility index; r is the subjective rate of discount; $\pi(\theta)$ is a prior distribution describing initial knowledge of θ coming from the experience of the consumer before $t = 0$ (the present).

Let us assume that: 1) the representative individual has perfect foresight of the inflation rate so $\dot{P}_t / P_t = q = q^e$, that is, she/he accurately perceives the rate at which inflation is proceeding, the value $P(0)$ is assumed to be known, 2) there are no barriers to free trade, 3) the international interest rate is equal to r, 4) capital mobility is perfect. If i is the nominal interest rate then $r = i + q^e$. Denoting income and government lump-sum transfers by y_t and g_t respectively, we can write the consumer's budget constraint, at time $t = 0$, as

$$a_0 + \int_0^{\infty} (y + g_t)e^{-rt}dt = \int_0^{\infty} (c_t + im_t)e^{-rt}dt, \tag{7.2}$$

where for the sake of simplicity we have chosen $y_t \equiv y = \text{constant}$. The consumer holds two assets: cash balances, $m_t = M_t/P_t$, where M_t is the nominal stock of money; and an international bond, k_t. The bond pays a constant interest rate r (i.e., pays r units of the consumption good per unit of time). Thus, the consumer's wealth, a_t, is defined by

$$a_t = m_t + k_t, \tag{7.3}$$

where a_0 is exogenously determined. Furthermore, we suppose that the rest of the world does not hold domestic currency.

Consider a cash-in-advance constraint of the Clower-Lucas-Feenstra form, $m_t \geq \alpha c_t$, where c_t is consumption, and $\alpha > 0$ is the time that money must be held to finance consumption. Given that $i > 0$, the cash-in-advance constraint will hold with equality,

$$m_t = \alpha c_t \tag{7.4}$$

For the sake of concreteness, let us suppose that $u(c_t; \theta) = -e^{-\theta c}$. Plainly, $u_c > 0$ and $u_{cc} < 0$. Moreover, let us assume that there is supplementary information about $\theta > 0$ in terms of the mean value $E[\theta] = 1/\lambda$. We also assume that $\mathcal{I}(\theta)$ and $\mathcal{F}(\theta)$ are constant, i.e., before supplementary information becomes available, initial knowledge is vague. In such a case, following Proposition 4.1, the Good-Bernardo-Zellner prior is given by $\pi^*(\theta) = e^{-\lambda\theta}$, $\theta > 0$, and (7.1) can be written as

$$V = \int_0^{\infty} \left\{ \int_0^{\infty} -e^{-\theta(c+\lambda)} d\theta \right\} e^{-rt}dt = \int_0^{\infty} -\left(\frac{1}{c_t + \lambda} \right) e^{-rt}dt. \tag{7.5}$$

In maximizing (7.5) subject to (7.2) the first-order condition for an interior solution is:

$$\frac{1}{(c+\lambda)^2} = \lambda(1 + \alpha i), \tag{7.6}$$

where λ is the Lagrange multiplier associated with (7.2).We assume a government budget constraint of the form

$$\int_0^\infty g_t e^{-rt}dt = b_0 + \int_0^\infty (\dot{m}_t + qm_t)e^{-rt}dt, \tag{7.7}$$

where b denotes the government's holding of international bonds. Let us denote by f the total bond holding of the economy, i.e., $f = k_t + b_t$. Then by (7.2) and (7.7) we get

$$f_0 + \int_0^\infty ye^{-rt}dt = \int_0^\infty c_t e^{-rt}dt. \tag{7.8}$$

Suppose that expected inflation (depreciation) takes the values q_1^e in $[0, T]$ and q_2^e in (T, ∞), where $T > 0$ and $q_1^e < q_2^e$. Since λ is time-invariant, we have

$$c_2 = Ac_1 + \lambda(A - 1), \quad 0 < A = \sqrt{\frac{1 + \alpha(r + q_1^e)}{1 + \alpha(r + q_2^e)}} < 1 \tag{7.9}$$

where c_1 is consumption in $[0, T]$ and c_2 is consumption in (T, ∞).

On the other hand, from (7.8), we obtain

$$f_0 + \int_0^\infty ye^{-rt}dt = \int_0^T c_1 e^{-rt}dt + \int_T^\infty c_2 e^{-rt}dt \tag{7.10}$$

which leads to

$$c_2 = (y + rf_0)e^{rT} + c_1(1 - e^{rT}) \tag{7.11}$$

The perfect foresight equilibrium consistent with the consumer's optimal decisions and government behavior is the intersection point, (c_1, c_2), between (7.9) and (7.11). Observe that, in (7.9), a once-and-for-all increase in λ, which results in a decrease in the mean value, $E[\theta] = 1/\lambda$, will decrease the value of the intercept, $\lambda(A - 1)$, which in turn increases c_1. In other words, λ reinforces the effect of the rate of time preference. Thus, an increase in λ causes a rise in present consumption and a fall in future consumption.

Other possibilities of supplementary information, using the notation in (4.1), are listed below. In some cases, however, it might only be possible to analyze the equilibrium via numerical methods.

(i) If $a_1(\theta) = |\theta|$ and $\bar{a}_1 = \alpha$, $\alpha > 0$, then the Good-Bernardo-Zellner prior is

$$\pi^*(\theta) = \frac{1}{2\alpha} e^{-\frac{1}{\alpha}|\theta|},$$

which is a Laplace distribution.

(ii) If

$$\begin{Bmatrix} a_1(\theta) = \theta \\ a_2(\theta) = e^{\theta} \end{Bmatrix} \quad \text{and} \quad \begin{Bmatrix} \bar{a}_1 = -\dfrac{\kappa}{\alpha} + \beta,\ \beta \in \mathbb{R} \\ \bar{a}_2 = e^{\beta}\Gamma(1 + \dfrac{1}{\alpha}) \end{Bmatrix},$$

where κ is Euler's constant, then $\pi^*(\theta) = \alpha e^{\alpha(\theta - \beta)}\exp\{-e^{\alpha(\theta - \beta)}\}$, which is a Gumbel (or extreme value) distribution.

(iii) If

$$\begin{Bmatrix} a_1(\theta) = I_{(0,\infty)}(\theta) \\ a_2(0) = \theta \\ a_3(\theta) = \log\theta \end{Bmatrix} \quad \text{and} \quad \begin{Bmatrix} \bar{a}_1 = 1 \\ \bar{a}_2 = \dfrac{\alpha}{\beta},\ \alpha > 0,\ \beta > 0 \\ \bar{a}_3 = \psi(\alpha) - \log\beta \end{Bmatrix},$$

where $I_{(0,\infty)}$ is the usual indicator function and, as before, $\psi(\alpha)$ is the *psi* function, then $\pi^*(\theta) = \dfrac{1}{\Gamma(\alpha)}(\beta\theta)^{\alpha - 1}\beta e^{-\beta\theta}$, which is a Gamma distribution (or Erlang distribution, if α is a positive integer).

(iv) If

$$\begin{Bmatrix} a_1(\theta) = I_{(0,\infty)}(\theta) \\ a_2(\theta) = \theta^{\beta},\ \beta > 0 \\ a_3(\theta) = \log\theta \end{Bmatrix} \quad \text{and} \quad \begin{Bmatrix} \bar{a}_1 = 1 \\ \bar{a}_2 = \dfrac{1}{\alpha},\ \alpha > 0 \\ \bar{a}_3 = -\dfrac{\kappa}{\beta} - \dfrac{\log\alpha}{\beta} \end{Bmatrix},$$

where κ is Euler's constant, then $\pi^*(\theta) = \alpha\beta\theta^{\beta - 1}e^{-\alpha\theta^{\beta}}$, which is a Weibull distribution.

Example 7.2

We will develop Good-Bernardo-Zellner interval estimates to test convergence of rational expectations. Consider a simple macroeconomic model

We may also write β_t as $\beta_t = \lambda^t x_t$, where x_t is any martingale, that is, x is any stochastic process that satisfies

$$\mathrm{E}\{x_{t+1} \mid \hat{I}_t\} = x_t.$$

Therefore, there are infinitely many divergent forward rational expectations solutions. Convergence will require $\beta_t = 0$ for all t.

Note now that from successive substitution of (7.13) into (7.14), we can show that $E_t\, m_{t+j-1} = \rho^j m_{t-1}, j = 0, 1, \ldots$, and therefore (7.14) becomes

$$p_t = \frac{\rho m_{t-1}}{1-\gamma(1-\rho)} - \delta\bar{y} + \beta_t + \frac{\nu_t}{1-\gamma} \tag{7.16}$$

There are many stochastic processes (bubbles) consistent with (7.15), for instance,

$$\beta_{t+1} = \begin{cases} \dfrac{\lambda\beta_t}{q} & \text{with probability } q,\ 0 < q \leq 1, \\ 0 & \text{with probability } 1-q, \end{cases}$$

or

$$\beta_{t+1} = \lambda\beta_t + \eta_t, \tag{7.17}$$

where the η_t's are independent Gaussian variables with mean zero and variance σ^2_η.

We suppose that the β_t's are unobserved location parameters satisfying (7.17). We also assume that there is supplementary information in terms of the two first moments on the initial β_0, namely $E\{\beta_0\} = \hat{\beta}_0$ and $E\{\beta_0^2\} = \hat{\sigma}_0^2 + \hat{\beta}_0^2$. Then, according to Proposition 4.1, the Good-Bernardo-Zellner prior compatible with such a information is $\mathcal{N}(\hat{\beta}_0, \hat{\sigma}_0^2)$. We suppose that the random variables β_0, ε_t and η_t are independent. Hence, under normally distributed errors, the rational expectations system is given by

$$\begin{cases} \beta_t = \lambda\beta_{t-1} + \eta_{t-1}, \\ m_t = \rho m_{t-1} + \nu_t, \\ p_t = \dfrac{\rho m_{t-1}}{1-\gamma(1-\rho)} - \delta\bar{y} + \beta_t + \dfrac{\nu_t}{1-\gamma}, \end{cases}$$

or equivalently, in terms of (5.1) and (5.2),

$$\begin{cases} p_t + \delta\bar{y} - \dfrac{m_t}{1-\gamma(1-\rho)} = \beta_t + \varepsilon_t \\ \beta_t = \lambda\beta_{t-1} + \eta_{t-1} \end{cases}$$

where

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2), \text{ and } \sigma_\varepsilon^2 = \left[\frac{\gamma\rho\sigma_v}{[1-\gamma(1-\rho)](1-\gamma)}\right]^2.$$

To test the common assumption of convergence with available data on p_t, m_t, γ, δ, ρ, and $\bar{y}$, and under normally distributed errors we use equations (5.7), (5.9) and (5.10) with univariate error terms. In such a case, the posterior distribution of $\beta_t \mid I_{t-1}$ is $N(\hat{\beta}_t, \hat{\sigma}_t^2)$, where

$$\begin{cases} \hat{\beta}_t = \theta_t\lambda\hat{\beta}_{t-1} + (1-\theta_t)\left[p_t + \delta\bar{y} - \dfrac{\rho m_t}{1-\gamma(1-\rho)}\right], \\ \hat{\sigma}_t^2 = (1-\theta_t)\sigma_\varepsilon^2, \\ \theta_t = \sigma_\varepsilon^2(\sigma_\varepsilon^2 + \lambda^2\hat{\sigma}_{t-1}^2 + \sigma_\eta^2)^{-1}, \end{cases}$$

The null hypothesis to be tested is $H_0 : \beta_t = 0$ for all $t \geq 1$. Proceeding recursively and starting off at $t = 1$, we reject H_0 if a t appears for which $\beta_t = 0$ does not lie within a highest posterior density interval with a given uniform significance level α, namely $(\hat{\beta}_t - z_{\alpha/2}\hat{\sigma}_t, \hat{\beta}_t + z_{\alpha/2}\hat{\sigma}_t)$ where, as usual, $P\{Z > z_{\alpha/2}\} = \alpha/2$ and $Z \sim N(0, 1)$.

Example 7.3

Finally, we will apply Good-Bernardo-Zellner priors to consumption decisions under uncertain inflation. We assume that there is a large number of identical consumers, each of whom makes consumption decisions in $T - 1$ periods ($t = 0, 1, ..., T - 1$), and has the following budget constraint:

$$w_{t-1}M_t = w_{t-1}M_{t-1} + g_{t-1} + y_{t-1} - c_{t-1}, \qquad (7.18)$$

$$t = 1, ..., T, \ M_0 > 0 \text{ given}, M_T > 0,$$

where M_t is the stock of currency owned at the beginning of period t, w_t is the value of the currency measured in goods at t (the reciprocal of the price level), g_t stands for government lump-sum transfers at t, y_t is real income at t, and c_t is consumption at t. Equation (7.18) can be rewritten, in terms of the inflation rate

$$\pi_t = \frac{w_{t-1}}{w_t} - 1$$

as

$$(1 + \pi_t)m_t = (1 + \pi_{t-1})m_{t-1} + g_{t-1} + y_{t-1} - c_{t-1} - \pi_{t-1}m_{t-1}, \quad (7.19)$$

$$t = 1, ..., T,$$

where $m_t = w_t M_t$ represents money balances and the last term on the right-hand side stands for depreciation of money balances from inflation. Note, however that the above budget constraint requires additional information on w_{-1} and w_T.

Private agents have no knowledge of $w_{-1}, w_0, ..., w_T$, and therefore, they do not know the inflation rate, π_t. However, we assume they have partial information on the distribution of w_{-1}, in terms of the first two moments, say, $E\{w_{-1}\} = \hat{w}_{-1}$ and $E\{w_{-1}^2\} = \hat{\sigma}_{-1}^2 + \hat{w}_{-1}^2$.

By using Proposition 4.1, with $\mathcal{I}(\theta)$ and $\mathcal{F}(\theta)$ constant (i.e., before supplementary information becomes available, initial knowledge is vague), we find that the Good-Bernardo-Zellner prior compatible with the available information for w_{-1} is $\mathcal{N}(\hat{w}_{-1}, \hat{\sigma}_{-1}^2)$. Therefore,

$$w_{-1}M_0 = (1 + \pi_0)m_0 \sim \mathcal{N}(\hat{w}_{-1}M_0, \hat{\sigma}_{-1}^2 M_0^2).$$

Of course, we assume that $\hat{w}_{-1} > 0$.

Suppose also that private agents are capable of making indirect measurements , $\bar{\pi}_t$, of π_t, according to the rule

$$(1 + \bar{\pi}_t)\bar{m} = (1 + \pi_t)m_t + \varepsilon_t, \quad t = 1, ..., T, \quad (7.20)$$

where $\overline{m}$ is a constant target chosen by the monetary authority at $t = 1$. We assume that the observation errors, ε_t, are independent normal random variables with mean zero, variance σ_ε^2 and $\mathrm{E}\{w_{-1}\varepsilon_t\} = 0$.

The representative individual's objective is to maximize, at the present ($t = 0$), his total expected utility of consumption over $T - 1$ periods, namely,

$$\mathrm{E}\left\{\sum_{t=1}^{T} u(c_{t-1}) + v(w_{T-1}M_T)\right\}. \tag{7.21}$$

Note that, for simplicity, no discount factor has been included in the overall utility, and money services provide no utility. The utility function is expressed as the quadratic function

$$u(c_t) = a_1 c_t - \frac{a_2}{2} c_t^2, \quad t = 0, \ldots, T-1. \tag{7.22}$$

Here, $a_1, a_2 > 0$, and the ratio a_1/a_2 determines the level of satiation. Note that $u(0) = u(2a_1/a_2) = 0$, $u(c_t) > 0$ for $0 < c_t < 2a_1/a_2$, $u(c_t) < 0$ for $c_t > 2a_1/a_2$, $u'(c_t) \geq 0$ for $0 \leq c_t \leq a_1/a_2$, and $u'(c_t) < 0$ for $c_t > a_1/a_2$. The salvage value is chosen as $v(w_{T-1}M_T) = -(a_2/2)[w_{T-1}M_T]^2$.

We assume that the income of the individual fluctuates randomly around his income satiation level following

$$y_t = \frac{a_1}{a_2} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2), \quad t = 0, \ldots, T-1, \tag{7.23}$$

where the η_t's are independent endowment shocks satisfying $\mathrm{E}\{\varepsilon_s\eta_t\} = 0$ for all t, s, and $\mathrm{E}\{w_{-1}\eta_t\} = 0$.

In order to keep monetary experiments as separate as possible from the effect of other government activities, we suppose that at each time $t = 0, 1, \ldots, T-1$, the government consumes nothing, has no debt and is committed to provide a lump-sum subsidy to compensate for depreciation of money balances whatever the rate of inflation is. Thus, the government budget constraint is given by

$$g_t = \pi_t m_t, \quad t = 0, \ldots, T-1. \tag{7.24}$$

After incorporating government behavior, (7.24), and income fluctuations, (7.23), into the representative individual's budget constraint, (7.19), we get the consolidated constraint for the economy

$$(1+\pi_t)m_t = (1+\pi_{t-1})m_{t-1} - (c_{t-1} - \frac{a_1}{a_2}) + \eta_{t-1}, \quad t = 1, \ldots, T \qquad (7.25)$$

Let us denote $\beta_t = (1+\pi_t)m_t$, $\hat{\beta}_0 = \hat{w}_{-1}M_0$ and $\hat{\sigma}_0^2 = \hat{\sigma}_{-1}^2 M_0^2$. Note that π_t is unobserved, and therefore β_t is unobserved. The social planner problem is thus stated as

$$\text{Minimize E}\left\{\sum_{t=1}^{T}\left(c_{t-1} - \frac{a_1}{a_2}\right)^2 + \beta_T^2\right\},$$

$$\text{subject to:}\begin{cases} \beta_t = \beta_{t-1} - (c_{t-1} - \frac{a_1}{a_2}) + \eta_{t-1}, \quad t = 1, \ldots, T \\ (1+\bar{\pi}_t)\bar{m} = \beta_t + \varepsilon_t, \quad t = 1, \ldots, T \\ \beta_0 \sim N(\hat{\beta}_0, \hat{\sigma}_0^2), \\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2),\ \eta_t \sim N(0, \sigma_\eta^2), \text{ with } \beta_0, \varepsilon_t \text{ and } \eta_t \text{ independent.} \end{cases}$$

The above constraints determine the state-space representation of the dynamics of β_t with control c_{t-1}. It is worthwhile to note that such constraints collapse into $\tilde{y}_{t-1} = c_{t-1}$, where $\tilde{y}_{t-1} = y_{t-1} + \xi_{t-1} - \xi_t$, $\xi_0 = \beta_0 - \bar{m}$, $\xi_t \sim N(\bar{\pi}_t \bar{m}, \sigma_\varepsilon^2)$ for $t = 1, \ldots, T-1$, and $\xi_T = \beta_T - \bar{m}$. The optimal planned consumption path, $\{\hat{c}_t\}_{t=0}^{T-1}$, satisfies

$$\hat{c}_t = \frac{a_1}{a_2} + \frac{1}{T-t+1}\hat{\beta}_t, \quad t = 0, \ldots, T-1, \qquad (7.26)$$

where the estimates $\hat{\beta}_t$ are computed through the equations (5.7), (5.9) and (5.10) with univariate error terms, as

$$\hat{\beta}_t = \theta_t \hat{\beta}_{t-1} + (1-\theta_t)(1+\bar{\pi}_t)\bar{m}, \quad t = 1, \ldots, T-1, \qquad (7.27)$$

$$\theta_t = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \hat{\sigma}_{t-1}^2 + \sigma_\eta^2}, \quad t = 1, \ldots, T-1 \qquad (7.28)$$

$$\hat{\sigma}_t^2 = (1 - \theta_t)\sigma_e^2, \quad t = 1, \ldots, T-1 \tag{7.29}$$

Moreover, the optimal salvage value is reached at

$$\hat{\beta}_T = \theta_T \hat{\beta}_{T-1} + (1 - \theta_T)(1 + \bar{\pi}_T)\overline{m} > \theta.$$

8. Summary and Conclusions

We have presented, in a unifying framework, a number of well-known methods that maximize a criterion functional to obtain non-informative and informative priors. Our general procedure is, by itself, capable of dealing with a range of interesting issues in Bayesian analysis. However, in this paper, we have limited our attention to Good-Bernardo-Zellner priors as well as their application to Bayesian inference.

The choice of a prior distribution depends on experience and knowledge. Thus, it is impossible to choose a prior that will always be applicable to all circumstances. In our approach the Good-Bernardo-Zellner priors provide a broad class of prior distributions that are appropiate for use in a variety of situations in economic theory and applied econometrics.

Throughout the paper, we have emphasized the existence and uniqueness of the solutions to the corresponding variational and optimal control problems. There are, of course, many other members of the class $\mathcal{A}$ that deserve much more attention than what we have attempted here. Needless to say, more work will be required in this direction. Results will be reported in future work.

References

Akaike, H. (1978). "A New Look at the Bayes Procedure", *Biometrika*, no. 65, pp. 53-59.

Bayes, T. (1763). "An Essay Towards Solving a Problem in the Doctrine of Chances", reprinted in *Biometrika* (1958). no. 45, pp. 243-315.

Berger, J. O. and J. M. Bernardo (1992a). "On the Development of Reference Priors", in J. M. Bernardo *et al.*(eds.), *Bayesian Statistics*, no. 4, pp. 35-60.

——— (1992b). "Ordered Group Reference Priors with Application to a Multinomial Problem", *Biometrika*, no. 79, pp. 25-37.

——— (1989). "Estimating a Product of Means: Bayesian Analysis with Reference Priors", *Journal of the American Statistical Association*, no. 84, pp. 200-207.

——— and M. Mendoza (1989). "On Priors that Maximize Expected Information", in J. Klein and J. Lee (eds.), *Recent Developments in Statistics and their Applications*, pp. 1-20.

Bernardo, J. M. (1996). "Noninformative Priors Do Not Exist: A Discussion", *Journal of Statistical Planning and Inference.*

——— (1979). "Reference Posterior Distributions for Bayesian Inference", *Journal of the Royal Statistical Society*, B41, pp. 113-147.

——— and J. M. Ramón (1997). *An Introduction to Bayesian Reference Analysis: Inference on the Ratio of Multinomial Parameters*, Tech. Rep. 3-97, Universitat de València, Spain.

——— and A. F. M. Smith (1994). *Bayesian Theory*, John Wiley & Sons, Wiley Series in Probability and Mathematical Statistics.

Box, G. E. P. and G. C. Tiao (1973). *Bayesian Inference and Statistical Analysis*, Addison-Wesley Series in Behavioral Science, Quantitative Methods.

Calvo, G. A. (1986). "Temporary Stabilization: Predetermined Exchange Rates", *Journal of Political Economy*, no. 94, pp. 1319-1329.

Good, I. J. (1969). "What is the Use of a Distribution?", Krishnaiah (ed.), *Multivariate Analysis*, Vol. II, New York, Academic Press, pp. 183-203.

——— (1968). "Utility of a Distribution", *Nature*, no. 219.

Jaynes, E. T. (1982). *On the Rationale of Maximum-Entropy Methods*, Procedures of the IEEE, no. 70, pp. 939-952.

——— (1957). "Information Theory and Statistical Mechanics", *Physical Review*, no. 106, pp. 620-630.

Jeffreys, H. (1961). *Theory of Probability*, 3rd. edition, Oxford University Press.

Kaiman, R. E. and R. Buey (1961). "New Results in Linear Filtering and Prediction Theory", *Transactions ASME, Series D. J. of Basic Engineering*, no. 83, pp. 95-108.

Kaiman, R. E. (1960). "A New Approach to Linear Filtering and Prediction Problems", *Transactions ASME, Series D. J. of Basic Engineering*, no. 82, pp. 35-45.

Kamien, M. I. and N. L. Schwartz (1991). *Dynamic Optimization, the Calculus of Variations and Optimal Control in Economics and Management*, 2nd. edition, Amsterdam, North Holland.

Kullback, S. (1959). *Information Theory and Statistics*, New York, John Wiley & Sons.

Lindley, D. V. (1956). "On a Measure of Information Provided by an Experiment", *Annals of Mathematical Statistics*, no. 27, pp. 986-1005.

Ordorica-Mellado, M. (1995). *El Filtro de Kaiman en la planeación demográfica*, Ph. D. Dissertation, Department of Operations Research, Engineering Graduate School, UNAM.

Smith, C. A. B. (1979). "Discussion of Professor Bernardo's paper", *Journal of the Royal Statistical Society*, B41, pp. 134-135.

Soofi, E. S. (1994). "Capturing the Intangible Concept of Information", *Journal of the American Statistical Association*, no. 89, pp. 1243-1254.

Venegas-Martínez, F. (1997). *On Information Functionals and Priors*, Memoria del XII Foro Nacional de Estadística, Asociación Mexicana de Estadística, INEGI, pp. 183-188.

——— (1992). "Entropy Maximization and Cross-Entropy Minimization: A Matrix Approach", *Agrociencia*, Serie Matemáticas Aplicadas, Estadística y Computación, pp. 71-77.

——— (1990a). "On Regularity and Optimality Conditions for Maximum Entropy Priors", *The Brazilian Journal of Probability and Statistics*, no. 4, pp. 105-136.

——— (1990b). "Información suplementaria a priori, aspectos computacionales y clasificación", *Estadística*, Interamerican Statistical Institute, no. 42, pp. 64-80.

——— (1990c). "Información suplementaria a priori", *Contributions to Probability and Mathematical Statistics*, Acts of the IV Latin-American Meeting in Probability and Mathematical Statistics, no. 4, pp. 228-237.

———, E. de Alba, and M. Ordorica-Mellado (1995). "An Economist's Guide to the Kaiman Filter", *Estudios Económicos*, El Colegio de México, vol. 10, no. 2, pp. 123-145.

Zellner, A. (1995). "Past and Recent Results on Maximal Data Information Priors", Graduate School of Busmess, University of Chicago (manuscript).

——— (1993). "Models, Prior Information and Bayesian Analysis", presented at the Conference on Informational Aspects of Bayesian Statistics in Honor of H. Akaike, to appear in a special issue of *The Journal of Econometrics.*

——— (1991). "Bayesian Methods and Entropy in Economics and Econometrics", in W. T. Grandy and L. H. Schick (eds.), *Maximum Entropy and Bayesian Methods*, Netherlands, Kluwer, pp. 17-31.

——— (1977). "Maximal Data Information Prior Distributions", in A. Aykac and C. Brumai (eds.), *New Developments in the Application of Bayesian Methods*, Amsterdam, North-Holland, pp. 201-232.

——— (1971). *An Introduction to Bayesian Inference in Econometrics*, New York, John Wiley & Sons.