Priors for Unit Root Models [1]

By

Joseph B. Kadane, Ngai Hang Chan,
and Lara J. Wolfson

Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213
U.S.A.

[1] Research supported in part by the Office of Naval Research Grant No. N00014-89-J-1851.

ABSTRACT

In the recent debate about unit roots, part of the discussion now centers on what prior is correct to use. From the perspective of subjective Bayesians, this question seems ill posed. To the extent that the likelihood is sharply peaked, the prior will not matter much. To the extent that the likelihood is flat, the posterior depends principally on the prior. A special consideration in unit roots models is that the spaces $\rho < 1$, $\rho = 1$, and $\rho > 1$ have very different economic implications. A better way to asses the role of the prior is to study its connections with these implications through prior elicitations.

In each of the ρ intervals, a conjugate prior (i.e. normal-inverse gamma) is fit to the elicited opinion of an expert. Since the likelihood in each interval is normal, the posterior is of the same form as the prior (i.e. piece-wise normal-inverse gamma), so this family of priors is closed under sampling.

A modification of existing methods for the normal linear model permits elicitation of this expanded family of prior distributions. An example elicitation of a macro-economist is discussed.

1 Introduction

Critical phenomena occur frequently in stochastic modeling. A typical example arises in time series. When the parameters of a stationary time series approach the boundary of the stationary regime, the series becomes non-stationary and its underlying behavior exhibits a sudden change. Similar situations also exist in branching processes. Since critical phenomena lead to abrupt changes in the underlying mechanisms, new inferential tools are usually required to analyze such transitions. The monograph of Basawa and Scott (1983) discusses some of these unconventional asymptotic tools for analyzing non-ergodic processes.

Owing to its diversified applications in economics, one specific critical phenomenon for non-stationary time series, the so-called unit root problem, has been receiving considerable attention in the literature recently. In its simplest form, this problem concerns the behavior of the sample estimate of the autoregressive coefficient ρ for an AR(1) model

$$y_t = \rho y_{t-1} + \epsilon_t, \quad t = 1, \ldots, n, \tag{1}$$

where $\{\epsilon_t\}$ is an i.i.d. or a weakly dependent innovation sequence. When $\rho = 1$, (1) is known as the random walk model. Much empirical research in macro-economic time series aims at distinguishing the case where ρ equals to one from the case where ρ lies in the vicinity of one (hence, the term unit root problem), see Nelson and Plosser (1982). This literature, together with the subsequent developments on co-integration, has led to theoretical as well as empirical research in this area. The recent article by Stock and Watson (1988) illustrates some of these empirical problems in regression theory of non-stationary series.

On the other hand, studies on the unit root problem in statistics started as early as in White (1958). It was subsequently analyzed in Anderson (1959), Dickey and Fuller (1979) and more recently in Chan and Wei (1987). These papers concern the asymptotic behavior of a least squares statistic τ_n, based on the least squares estimate $\hat{\rho} = \sum_{t=1}^{n} y_t y_{t-1} / \sum_{t=1}^{n} y_{t-1}^2$.

Specifically, if $\{y_t\}$ follows (1) with $\{\epsilon_t\}$ i.i.d. $N(0, \sigma^2)$, then

$$\tau_n := \left(\sum_{t=1}^{n} y_{t-1}^2 \right)^{\frac{1}{2}} (\hat{\rho} - \rho) \Rightarrow \begin{cases} Z & |\rho| < 1, \\ \mathcal{L}(0) & \rho = 1, \\ \Gamma & |\rho| > 1, \end{cases} \tag{2}$$

where Z is a standard normal random variable, Γ is a standard Cauchy random variable, $\mathcal{L}(0) = \int_0^1 W(t)dW(t) / (\int_0^1 W^2(t)dt)^{\frac{1}{2}}$, $W(t)$ is a standard Brownian motion, and "$\Rightarrow$" designates convergence in distribution as the sample size $n \to \infty$. Since both Z and Γ are symmetric while $\mathcal{L}(0)$ is asymmetric at zero, (2) indicates the critical role played by ρ and its effect on the abrupt transition of the distribution of τ_n from $|\rho| < 1$ to $\rho = 1$. By considering the order of magnitude of the observed Fisher's information number, $\sum_{t=1}^{n} y_{t-1}^2$, for ρ near 1, Chan and Wei (1987) propose a reparameterization of ρ so that a smooth transition in (2) can be restored asymptotically. These results indicate a unique feature of critical phenomena, viz., instead of standard asymptotic normality, limiting distributions involving non-linear functionals of diffusion processes are usually obtained. Numerical percentiles for these functionals can be calculated, of which special cases can be found in Dickey and Fuller (1979) and Chan (1988). Several empirical studies based on these asymptotics have been conducted in macroeconomics.

¿From a classical perspective, (2) provides a reasonable statistical means to determine the value of ρ. However, the classical approach is not the only way to provide inferential procedures to analyze time series. It has long been known among statisticians and economists that a Bayesian approach to linear model also provides useful information, see for example, Zellner (1971). It is, therefore, important to explore a Bayesian analysis of the random walk model (1).

This paper is organized as follows. Recent developments on the Bayesian analysis of the unit root problem are reviewed in section 2. In section 3, the subjective Bayesian approach to elicitation is introduced. A family of conjugate priors for elicitation is given in section 4, a multivariate case is considered in section 5, an example of multivariate elicitation is presented in section 6, while concluding remarks are discussed in section 7.

2 Bayesian settings

Although classical asymptotics such as (2) provide one basis for inferential procedures empirically, it has not been universally accepted as the only means to analyze macro-economic series. Sims (1988) put forward a Bayesian approach to the random walk model (1) and argues that a Bayesian approach provides a more reasonable inferential scheme than (2). By taking a uniform prior on the autoregressive coefficient ρ, Sims and Uhlig (1991) show that the posterior distribution of ρ given the data $\{y_t\}$ has a t distribution which is symmetric at $\hat{\rho}$. This is identical to the traditional normal linear model result, because the likelihood is identical. On the empirical side, DeJong and Whiteman (1991) re-analyze the series originally studied in Nelson and Plosser (1982) using a Bayesian approach and obtain substantially different conclusions. Those results indicate not only the operational difference between the Bayesian and the classical approaches, but also that the discrepancy resulted from the presence of critical phenomenon (1). In a traditional normal linear model setting, a Bayesian and a classical approach yield inferential procedures that can be regarded as similar. This is not the case for the random walk model.

Phillips (1991a,b) introduces a Jeffreys-type prior for ρ in applying the Bayesian method. Contrary to Sims' result, the posterior distribution of ρ given the data $\{y_t\}$ is not a t-distribution under this prior. It is a distribution which is asymmetric at $\hat{\rho}$ and is expressed as a non-linear functional of a standard Brownian motion such as (2). Further remarks about the pros and cons of different priors for ρ are given in Phillips (1991b).

All these observations indicate a general message. To the extent that the likelihood is sharply peaked, the prior will not matter much; this is often the case for a traditional normal linear model setting. To the extent that the likelihood is flat, the posterior depends principally on the prior; this is often the case for a random walk model. It seems prudent to reconsider the issue of a prior from a different perspective for the latter case.

3 A subjective Bayesian approach

One of the difficulties with a debate about the "right" prior for the unit root problem is the lack of criteria for success. What would distinguish the "right" prior from other "wrong" ones?

This issue has been discussed extensively in the Bayesian literature. Following deFinetti and Savage, a view has developed that the question of the "right" prior for the autoregressive model is badly posed in two senses: the answer should depend on the specific application addressed and on the analyst involved. In this view, adopted in this paper, a prior is "right" because an analyst says it represents his or her opinion about some specific application. This view thus admits the possibility that a given analyst might think differently about different applications of the autoregressive model and that the same application might be viewed differently by different analysts. Both possibilities have a certain verisimilitude.

To say that a prior is subjective and depends on the applied problem does not give explicit guidance as to how to write it down. The study of tools to help a person do this task is known as "elicitation". There are fundamentally two types of elicitation, structural and predictive (see Kadane (1980) and Winkler (1980)). In structural elicitation, the analyst is presumed to understand the parameterization of the model and be able to state a prior directly. While econometricians may find this an attractive thought, it should be pointed out that most real applications involve a highly parameterized model, for which structural elicitation would be quite difficult.

The predictive method assumes a form for the likelihood function (here normal, conditioned on the covariates, including y_{t-1}), and integrates over the prior family. The application of predictive elicitation to the problem of autoregressive models is taken up in the next section.

4 Predictive elicitation for autoregressive models

The general idea of predictive elicitation is that a parameterized class of prior distributions is chosen that represents a reasonable range of possibilities. The parameters of the prior are conventionally called hyperparameters to distinguish them from the parameters of the likelihood. Integrating out the latter parameters yields the predictive distribution, which is the distribution of the observations as a function of the hyperparameters.

The obvious class to begin with is the conjugate family. In the case of the normal linear model, if σ^2 is known, the conjugate family is the normal distribution, and if unknown, the conjugate family is the normal-inverse gamma distribution. In the latter case, the predictive distribution is a multivariate-t distribution; in the former case, it is a multivariate normal distribution. The use of these distributions to carry out elicitation is demonstrated in the paper of Kadane et al. (1980).

A possible line of criticism of this work is that it relies on a special and restrictive class of priors. One response is to cite theorems (Diaconis and Ylvisaker (1985), Dallal and Hall (1983)) to the effect that convex combinations of conjugate prior distributions are dense in the space of all proper prior distributions. Hence whatever one's "true" prior, there is a convex combination of conjugate priors that is arbitrarily close. These theorems offer some comfort, but, as is usual with limit theorems, they do not specify how many conjugates are necessary for a given degree of approximation.

One important matter is whether special cases of interest are contained in the conjugate family. Certainly two special cases of concern are the flat prior of Sims and the Jeffreys-type prior of Phillips, having the form

$$\pi(\rho) = \begin{cases} \frac{n}{1-\rho^2} - \frac{1-\rho^{2n}}{(1-\rho^2)^2} & \rho \neq 1, \\ \frac{n(n-1)}{2} & \rho = 1, \end{cases} \tag{3}$$

for $y_0 = 0$.

Taken over the whole real line, both are improper in that they integrate to infinity. Each could be interpreted as proper, by truncating its domain, say from $-m$ to m, for some suitable bound m. Presumably in this problem it makes sense to take $m \geq 1$. One hopes in doing such a truncation that the posterior distribution that results from it is insensitive to the choice of m, at least for large m. For the Sims flat prior this is obvious, as the likelihood times flat prior will behave as the likelihood, which is normal in ρ. Hence a truncation far from the mean $\hat{\rho}$ has a negligible effect. The situation is not so clear for the Phillips-Jeffreys prior, since it is bowl-shaped with minimum at $\rho = 0$. As ρ gets large, $\pi(\rho)$ is dominated by the term ρ^{2n}. Whatever be n, the prior is still dominated by the normal likelihood as the latter decreases at the rate $e^{-k\rho^2}$. Hence, again, the posterior is not sensitive to the exact value of m as m gets large. Thus for this likelihood either prior may be taken as improper or as proper-by-truncation, with assurance that the resulting posterior is insensitive to the truncation value m.

The Sims flat prior can be approximated by a normal conjugate prior in which the variance goes to infinity, or, equivalently, the precision goes to zero. Hence in this extended sense, it is reasonable to consider the Sims prior to be a member of the conjugate family. The same cannot be said for the Phillips-Jeffreys prior. Its shape, bowl-shaped with a minimum at 0, is opposite to the shape of the normal conjugate priors that decrease from a maximum. Hence one cannot expect a single conjugate prior to approximate a Phillips-Jeffreys prior well.

In the unit root problem in particular, a prior that is a single normal over the entire range of ρ may well not be flexible enough to represent an appropriate range of views. One way to extend the normal conjugate family but still have a family closed under sampling is to allow a piece-wise normal family (where each piece is conditional on a known variance, σ_i^2. For example one could allow one piece to be a normal distribution over the range $(-\infty, 1)$, a second normal distribution might pertain to $(1, \infty)$, and conceivably there might be a lump of prior at $\rho = 1$. Thus the density of ρ given σ_0^2 and σ_1^2 is

$$f(\rho \mid \sigma_0^2, \sigma_1^2) = \frac{\omega_0}{\sigma_0} \frac{\varphi\left(\frac{\rho-\rho_0}{\sigma_0}\right) I(\rho < 1)}{\Phi\left(\frac{1-\rho_0}{\sigma_0}\right)} + \frac{\omega_1}{\sigma_1} \frac{\varphi\left(\frac{\rho-\rho_1}{\sigma_1}\right) I(\rho > 1)}{1 - \Phi\left(\frac{1-\rho_1}{\sigma_1}\right)} + \omega_2 I(\rho = 1), \tag{4}$$

where $\varphi(.)$ and $\Phi(.)$ denote the probability density function (pdf) and the cumulative distribution function (cdf) of a standard normal random variable. The hyperparameters here are ω_0, ω_1, and ω_2 (constrained to be non-negative and to sum to 1), ρ_0, ρ_1, σ_0 and σ_1.

This prior is quite flexible; if one wishes to restrict inference to the interval $[k^-, 1]$, where k^- is some number less than 1, this can be achieved by setting $\omega_1 = 0$ (and thus ω_0 and ω_2 must sum to 1), and replacing the denominator in the first term of (4) with

$$f(\rho \mid \sigma_0^2, \sigma_1^2) = \frac{\omega_0}{\sigma_0} \frac{\varphi\left(\frac{\rho-\rho_0}{\sigma_0}\right) I(\rho < 1)}{\Phi\left(\frac{1-\rho_0}{\sigma_0}\right) - \Phi\left(\frac{k^- -\rho_0}{\sigma_0}\right)} + \omega_2 I(\rho = 1). \tag{5}$$

An analogous modification can be made to the second term in (4) if the desired modification is to restrict inference to the region $[1, k^+]$

The posterior marginal distribution with the likelihood (4), given $\sigma_i^2, i =$

0, 1 and prior is again piece-wise normal, in the same form as (4). The updated values of the hyperparameters are:

$$\left.\begin{array}{lcl} \rho_i' & = & (\rho_i/\sigma_i^2 + \sum_{t=1}^T y_{t-1}y_t)(\sigma_i'^2), \\ \frac{1}{\sigma_i'^2} & = & \frac{1}{\sigma_i^2} + \frac{1}{t^2}, \end{array}\right\} \; i = 0, 1$$

$$\text{where } t^2 = 1/\sum_{t=1}^{T} y_{t-1}^2 \ , \tag{6}$$

$$\omega_i' \propto \omega_i \ \frac{\sigma_i'}{\sigma_i} \ \frac{c_i'}{c_i} \ \exp\left\{-\frac{1}{2}\left[\frac{(\hat{\rho} - \rho_i)^2}{\frac{1}{t^2} + \sigma_i^2}\right]\right\}, \quad i = 0, 1$$

$$\begin{array}{lcllcl} c_o & = & \Phi\left((1-\rho_0)/\sigma_0\right), & c_0' & = & \Phi\left((1-\rho_0')/\sigma_0'\right), \\ c_1 & = & 1 - \Phi\left((1-\rho_1)/\sigma_1\right), & c_1' & = & 1 - \Phi\left((1-\rho_1')/\sigma_1'\right), \end{array}$$

and $\omega_2' \propto \omega_2 \exp\{-\frac{1}{2}\left[(\hat{\rho}_i - 1)^2\right] t^2\}$.

Elicitation for this truncated mixture family can be accomplished conditionally. Thus, we would ask the analyst to imagine that a guarantee had been given that $\rho < 1$. Then answering the Kadane et al. (1980) elicitation questions would give σ_0 and ρ_0. Similarly imagining a guarantee of $\rho > 1$ would permit elicitation of σ_1 and ρ_1. The weights would have to be requested directly. Thus elicitation for such a truncated mixture family seems feasible.

If the variance σ^2 is taken to be stochastic with an inverse gamma distribution, nothing essential changes. The prior is piece-wise normal-inverse gamma, as is the posterior. Elicitation continues to be done piece-wise, with student-t predictive distributions in each piece.

5 Multivariate Case

A more realistic and general likelihood has the usual multivariate normal linear model structure:

$$y_t = \rho y_{t-1} + \boldsymbol{\beta}' \mathbf{x}_t + \epsilon_t \quad (y_0 = 0), \tag{7}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)'$, $\mathbf{x}_t = (x_{1t}, \ldots, x_{kt})'$ and ϵ_t iid $N(0, \sigma^2)$. Analogous to the previous section, a conjugate prior is specified for (7), in each of the pieces of the ρ-space.

In each piece of ρ-space, the elicited joint prior distribution of $\boldsymbol{\beta}, \rho$ given σ^2, is a multivariate normal distribution with elicited mean vector $\boldsymbol{b}$, and elicited covariance matrix $\Sigma = \frac{\sigma^2}{\delta + r} R^{-1}$, where δ is the degrees of freedom parameter for the prior predictive t-distribution, r is the dimension of the partitioned vector $\boldsymbol{\beta}^{*\prime} = [\rho \mid \boldsymbol{\beta}']$, and R^{-1} is one of the elicited hyperparameters.

To find the multivariate analogue of equation(4), $\boldsymbol{\beta}$ is integrated out, conditional on ρ, in each piece. Thus, conditional on σ^2, the distribution of ρ in piece i ($i = 0, 1, 2$) is a normal distribution with mean b_{1_i}, the first element of $\boldsymbol{b}_i$, and variance Σ_{11}, the $(1,1) - th$ element of Σ_i (as ρ is taken to be the first element of $\boldsymbol{b}$). This results, once again, in a piecewise normal distribution for $\rho \mid \sigma^2$ in exactly the form (4). Unconditionally (i.e., in this case the variance is stochastic), letting

$$\sigma_i^2 = \frac{\mathrm{w}_i}{\delta_i + r} R_i^{-1}, \tag{8}$$

where w_i is the elicited scale factor of the prior distribution for σ^2, the distribution of ρ will be

$$f(\rho) = \frac{\omega_0}{\sigma_0} \frac{g_{\delta_0}\left(\frac{\rho - b_{1_0}}{\sigma_0}\right) I(\rho < 1)}{G_{\delta_0}\left(\frac{1-b_{1_0}}{\sigma_0}\right)} + \frac{\omega_1}{\sigma_1} \frac{g_{\delta_1}\left(\frac{\rho - b 1_1}{\sigma_1}\right) I(\rho > 1)}{1 - G_{\delta_1}\left(\frac{1-b_{1_1}}{\sigma_1}\right)} + \omega_2 I(\rho = 1), \tag{9}$$

where g_{δ_i} and G_{δ_i} are, respectively, the pdf and cdf of a standard t random variable with δ_i degrees of freedom.

The constants in this multivariate version of equation (4) are found by integrating out $\boldsymbol{\beta}$, conditional on ρ, in each piece. So for $i = 0, 1$, ρ has a non-central t-distribution with δ_i degrees of freedom, with center given by b_{1_i} and spread given by σ_i.

An interesting special case of the general multivariate case is given by Schotman and Van Dijk (1991), as

$$y_t = \rho y_{t-1} + \mu(1-\rho) + \epsilon_t. \tag{10}$$

This model differs from (7) in that the constant term, $c = \mu(1-\rho)$ is zero if $\rho = 1$. Such a model is easily incorporated into the elicitation scheme proposed here. In the space $\rho = 1$, the analyst either believes or does not believe that $c = 0$, as evidenced by an elicited prior satisfying (or close enough to satisfying) $E(c) = 0$ and Var $(c) = 0$. If $c = 0$ with probability one when $\rho = 1$, then model (10) can be said to hold (assuming of course that (7) is already conceded). In this case, joint distributions for (c, ρ) are elicited for the spaces $(\rho < 1)$ and $(\rho > 1)$, from which a transformation to (μ, ρ) can be calculated if desired. The joint distribution of (c, ρ) has the conjugate structure described above. Given σ^2, the prior distribution will be a multivariate normal distribution, and the unconditional prior distribution will be a t-distribution. Since $\mu = c/(1-\rho)$, the conditional joint prior distribution of (μ, ρ) will not have a multivariate normal distribution. It is possible to derive the prior numerically given the elicited hyperparameters for (c, ρ). In this case what is at issue, in some sense, is the likelihood itself. From a subjective Bayesian perspective, the likelihood is as subjective as is the prior, and hence is equally appropriate as an object of elicitation.

Both models (7) and (10) can include a time trend as a special case. This is important, since it allows for consideration of difference stationarity versus trend stationarity, and shows that the elicitation methods presented here are feasible for whichever parameterization (7) or (10) is easiest for the expert to answer questions about.

6 An Example

The opinion of a macro-economist on the projected real domestic GNP for 5 quarters (i.e, 2nd quarter of 1993 to the 2nd quarter of 1994) is given for the model

$$\log(GNP)_t = \mu + \rho \log(GNP)_{t-1} + \epsilon_t, \tag{11}$$

where $\epsilon_t \sim N(0, \sigma^2)$, and σ^2 is taken to be stochastic.

The hyperparameters elicited from the expert are given in Table 1.

These elicitations were performed using a variant of the methods described in Kadane *et. al.* (1980). The Fortran program *elicit-normlin.f* available on StatLib can be used for elicitation with a normal linear model. The program *elicit-norm-ts.f* is a modification of these methods to perform the elicitation described in sections 4 and 5.

The basic procedure for the elicitation was as follows: The economist

Given:	ω_i	ρ_i	σ_{ρ_i}
$\rho < 1$	0.85	0.98	0.0045
$\rho > 1$	0.05	0.9976	0.0082
$\rho = 1$	0.10	1.0	NA

Table 1: Elicited Hyperparameters

was asked to assess 50th, 75th, and 90th percentiles of the prior predictive distribution for $\log(GNP)$, given various values of $\log(GNP)_{t-1}$. Thinking of the model as stating that $(\rho, \mu) \mid \sigma^2 \sim N(\boldsymbol{b}, \sigma^2 R^{-1})$ and $1/\sigma^2 \sim \mathrm{w}\delta\chi^2_\delta/\mathrm{w}\delta$, then the medians are used to estimate $\boldsymbol{b}$, and the relationship between the median and the 75th and 90th percentiles is used to derive the degrees of freedom hyperparameter, δ. Once these hyperparameters have been estimated, hypothetical data is generated. Based on this, the economist is asked to update the opinion conditional on having observed this hypothetical data, and re-assesses some of the previously assessed percentiles. This information is used to estimate the scale factor, w, and the covariance matrix, R^{-1}.

Figure 1A shows the marginal prior distribution for ρ which is a weighted mixture of two t-distributions, and a point mass at 1. Although most of the prior concentrates in the area where ρ is very close to one, much of the tail weight lies below $\rho = 1$.

Figure 1B shows what the marginal prior distribution would look like if the distribution for $\rho > 1$ was taken to be the same as the distribution of ρ when ρ is known to be stationary (i.e., $\rho_o = \rho_1$, and $\sigma_0 = \sigma_1$). In practice, some experts may be unable to provide different priors for the two ranges of

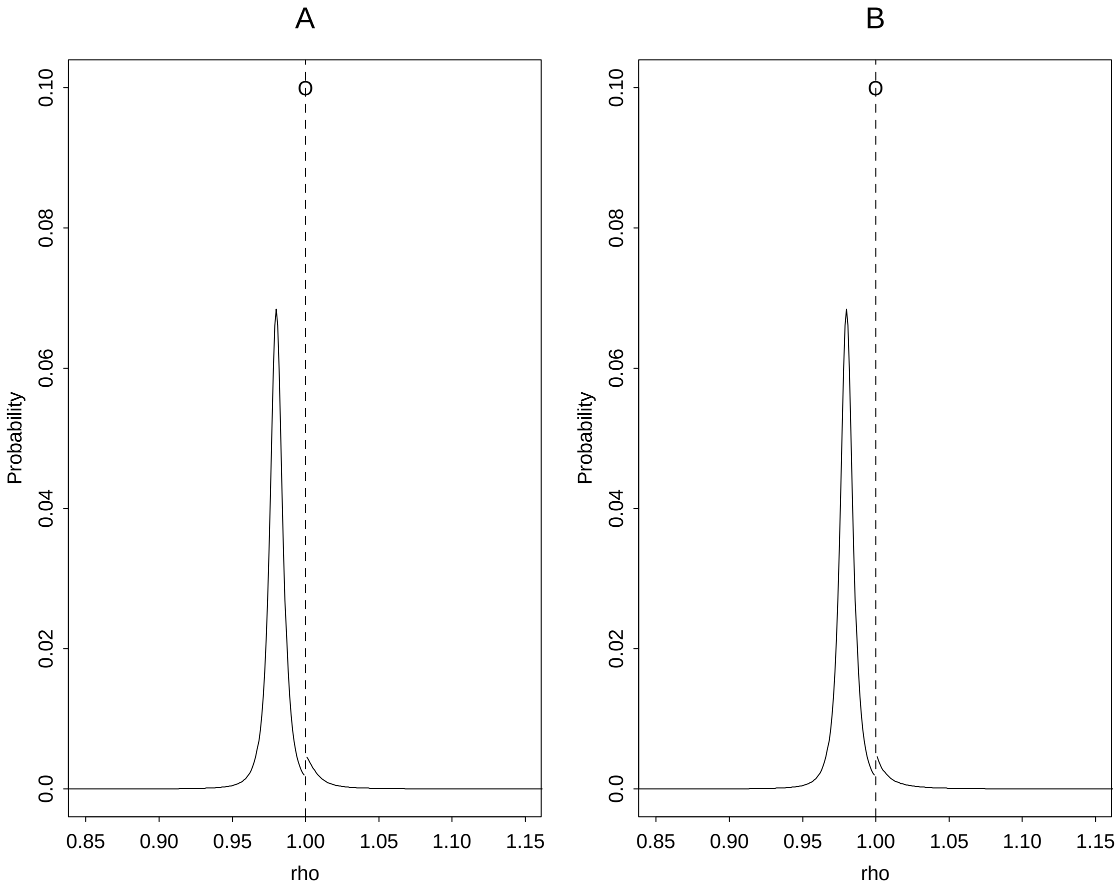

Figure 1: Prior Distributions for ρ.

ρ, which is the reason to consider this limiting case. The similarity between Figures 1A and 1B shows that in this case, although the two distributions specified were different in A and not in B, very little difference can be seen in the resulting prior distributions.

In the unit root problem in particular, there is often interest in calculating the odds that $\rho = 1$. Clearly, in both cases described above, the prior odds are given by $\omega_2/(\omega_0 + \omega_1) = 0.111$.

Data on real domestic GNP from the third quarter of 1993 to the 3rd quarter of 1994 was obtained from the Statistics in Education database server. The posterior odds of a unit root may be calculated as

$$Odds = \frac{P(\rho = 1 \mid data)}{P(\rho < 1 \mid data) + P(\rho > 1 \mid data)}. \tag{12}$$

The posterior odds of a unit root when the distributions for $\rho < 1$ and $\rho > 1$ are different are 2.098; when the distributions are the same, the posterior odds of a unit root are 0.489. The posterior distributions of ρ for these two cases are shown in Figures 2A and 2B.

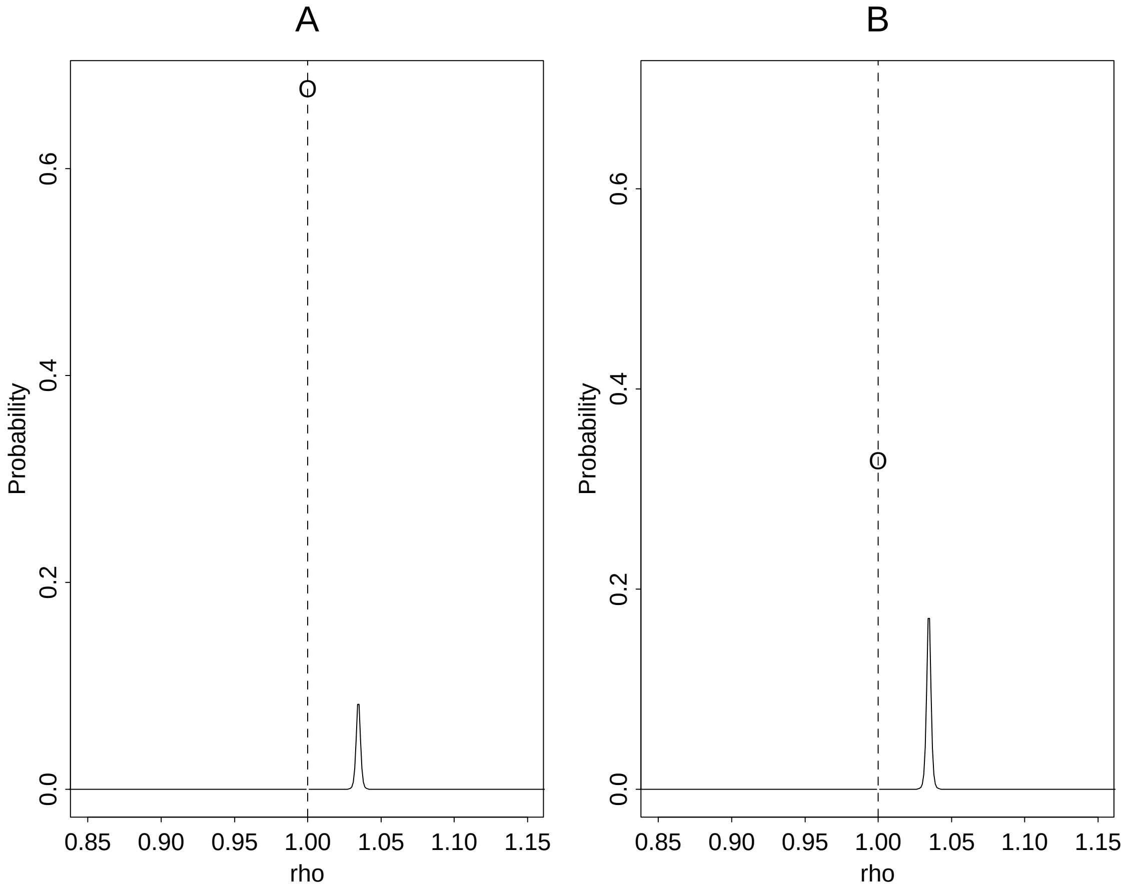

Figure 2: Posterior Distributions for ρ.

The likelihood has been fairly informative; the majority of the weight has been shifted away from the stationary root into the unit and explosive roots.

Although the model considered in this example has been examined in the context of model (7), it would have also been possible to consider it in the context of model (10). The marginal prior and posterior distributions of ρ would have remained unchanged, although the marginal prior distribution for μ would have no longer been a t-distribution, and thus conjugacy would not have held, and the posterior distribution for μ would have to be calculated numerically.

7 Conclusions

Posterior opinions, as an amalgam of prior opinions and likelihoods, can be dominated by either. In the example presented here, the likelihood shifts probability away from stationarity and toward the unit root and explosive cases. The extent to which it does so is partly a matter of the relationship between the likelihood and where in the stationarity ($\rho < 1$) and explosive ($\rho > 1$) regions the prior has most of its weight.

Elicitation of this kind is feasible for many econometric and seems especially important in unit root models. It requires economists to be very explicit about their assumptions, an exercise not always comfortable. Nonetheless, we believe that the process of stating and justifying prior distributions in econometrics can help communication among economists.

REFERENCES

Anderson, T.W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations. *Ann. Math. Statist.*, **30**, 676-687.

Basawa, I.V. and Scott, D.J. (1983). *Asymptotic Optimal Inference for Non-Ergodic Models.* Lecture Notes in Statistics, Springer-Verlag, New York.

Chan, N.H. (1988). The parameter inference for nearly non-stationary time series. *J. Amer. Statist. Assoc.*, **83**, 857-862.

Chan, N.H. and Wei, C.Z. (1987). Asymptotic inference for nearly non-stationary AR(1) processes. *Ann. Statist.*, **15**, 1050-1063.

Dallal, S.R. and Hall, W.J. (1983) Approximating priors by mixtures of natural conjugate priors. *J.R. Statist. Soc.*, B, **45**, 278-286.

DeJong, D.N. and Whiteman, C.H. (1991). Reconsidering trends and random walks in macro-economic time series. *J. Monetary Economics*, **28**, 221-254.

Diaconis, P. and Ylvisaker, D. (1985) Quantifying prior opinion. In *Bayesian Statistics 2* (eds J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith). Amsterdam: North-Holland.

Dickey, D.A. and Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.*, **74**, 427-431.

Kadane, J.B. (1980). Predictive and structural methods for eliciting prior distributions. *Bayesian Analysis in Econometrics and Statistics: Essays in honor of Harold Jeffreys*, A. Zellner, ed., North Holland Publishing Company, 89-93.

Kadane, J.B., Dickey, J.M., Winkler, R.L., Smith, W.S. and Peters, S.C. (1980). Interactive elicitation of opinion for a normal linear model. *J. Amer. Statist.*

Assoc., **15**, 845-854.

Nelson, C.R. and Plosser, C.I. (1982). Trends and random walks in macroeconomic time series. *J. Monetary Economics*, **10**, 139-162.

Phillips, P.C.B. (1991a). To criticize the critics: an objective Bayesian analysis of stochastic trends. *J. Applied Econometrics*, **6**, 333-364.

Phillips, P.C.B. (1991b). Bayesian routes and unit roots: de rebus prioribus semper est disputandum. *J. Applied Econometrics*, **6**, 435-473.

Schotman, P. and Van Dijk, H.K. (1991). A Bayesian analysis of the unit root in real exchange rates. *Journal of Econometrics*, **49**, 195-238.

Sims, C.A. (1988). Bayesian scepticism on unit root econometrics. *J. Economic Dynamics and Control*, **12**, 436-474.

Sims, C.A. and Uhlig, H. (1992). Understanding unit rooters: a helicopter tour. *Econometrica*, **60**, 121-124.

Stock, J.H. and Watson, M.W. (1988). Variable trends in economic time series. *J. Economic Perspective*, **2**, 147-174.

Uhlig, H. (1992). What macroeconomists should know about unit roots as well: a Bayesian perspective. Technical Report, Princeton University, New Jersey.

White, J.S. (1958). The limiting distribution of the serial coefficient in the explosive case. *Ann. Math. Statist.*, **29**, 1188-1197.

Winkler, R.L. (1980). Prior information, predictive distributions, and Bayesian model-building. *Bayesian Analysis in Econometrics and Statistics: Essays in honor of Harold Jeffreys*, A. Zellner, ed., North Holland Publishing Company, 95-109.

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