ERASMUS

9082/A

ECONOMETRIC INSTITUTE

POSTERIOR ANALYSIS OF POSSIBLY INTEGRATED TIME SERIES WITH AN APPLICATION TO REAL GNP

P. SCHOTMAN AND H.K. VAN DIJK

REPORT 9082/A

Erasmus

ERASMUS UNIVERSITY ROTTERDAM - P.O.BOX 1738 - 3000 DR ROTTERDAM-THE NETHERLANDS

POSTERIOR ANALYSIS OF POSSIBLY INTEGRATED TIME SERIES

WITH AN APPLICATION TO REAL GNP

by

Peter Schotman and Herman K. van Dijk

Econometric Institute
Erasmus University
P.O. Box 1738
3000 DR Rotterdam
Netherlands

January 1990

Revised: October 1990

Abstract

We consider Bayesian statistical inference for univariate time series models where one of the autoregressive roots is close to or equals unity. Classical sampling theory for this type of models is hampered by the vast differences between asymptotic approximations in the stationary case and under the unit root hypothesis. Because of this dichotomy one has to decide early on in an empirical study whether a given time series is stationary or not. The present paper shows that a Bayesian approach allows for a smooth continuous transition between stationary and integrated time series models. Empirical results are presented for time series of annual real per capita GNP for 16 OECD countries.

Contents

1. INTRODUCTION.

Economic time series like real Gross National Product (GNP) have a tendency to grow over time. One of the purposes of the analysis of such time series is the decomposition into a trend and a cyclical component. If a deterministic linear trend is extracted, the resulting cyclical component typically has a first order autocorrelation that is close to unity. The deviations from the trend tend to be long-lasting. An example of such a stylized fact is shown in figure 1a for U.S. postwar data on real gross national product. This motivated Nelson and Plosser [1982] to test formally the hypothesis that macro economic time series show no tendency at all to return to a linear trendline. They implemented this test as a test for a unit root in an autoregressive model representation (see Fuller [1976]). Using long time series for 14 major economic variables they were unable to reject the unit root hypothesis.

The presence of a unit root however invalidates the extraction of a deterministic trend. The first differencing induced by the unit root implies that information on the level of the series is lost. Statistically, the intercept in the autoregressive representation is no longer identified. In the case of a unit root Beveridge and Nelson [1981] proposed a different decomposition in trend and cycle, in which the trendline changes stochastically over time. Figure 1b shows the two different trends. TS denotes the linear trend derived from the Trend Stationary I(0) model; DS is the stochastic trend from the Difference Stationary model. Empirically it appears that a random walk with drift can account for almost all fluctuations in a major macroeconomic time series like U.S. real GNP.

As a result figure 1c shows that the "DS" cyclical component has a much smaller amplitude and shorter duration than the "TS" cycle. Blanchard and Fischer [1989, Ch. 1] discuss the economic implications of alternative

FIGURE 1a: U.S. real GNP and trend

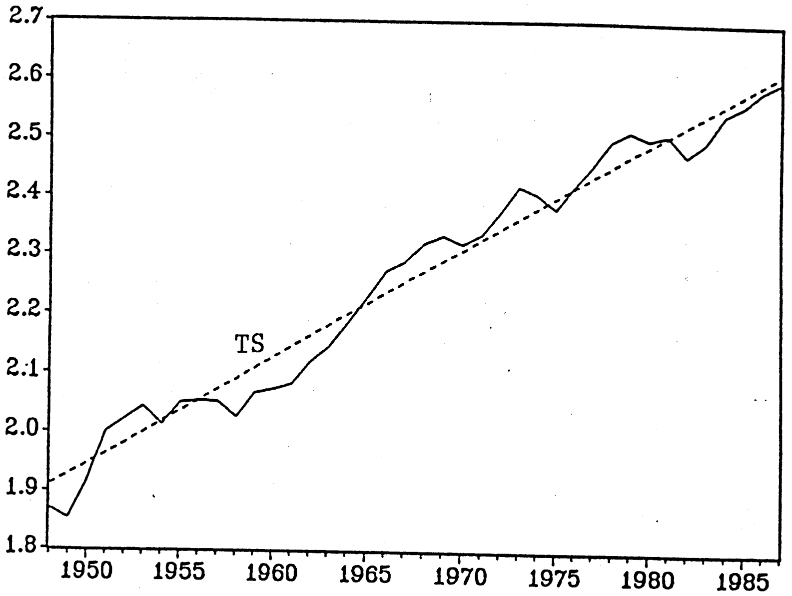


FIGURE 1b: Trend components of U.S. real GNP

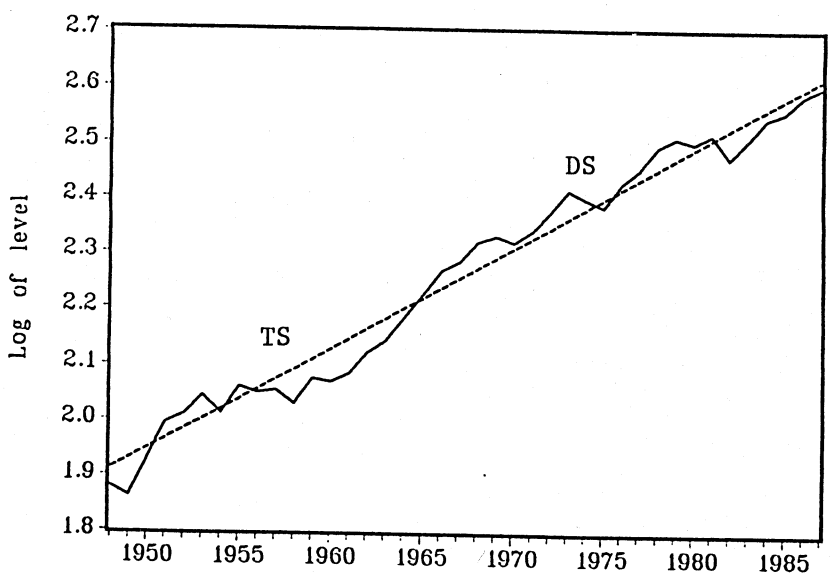
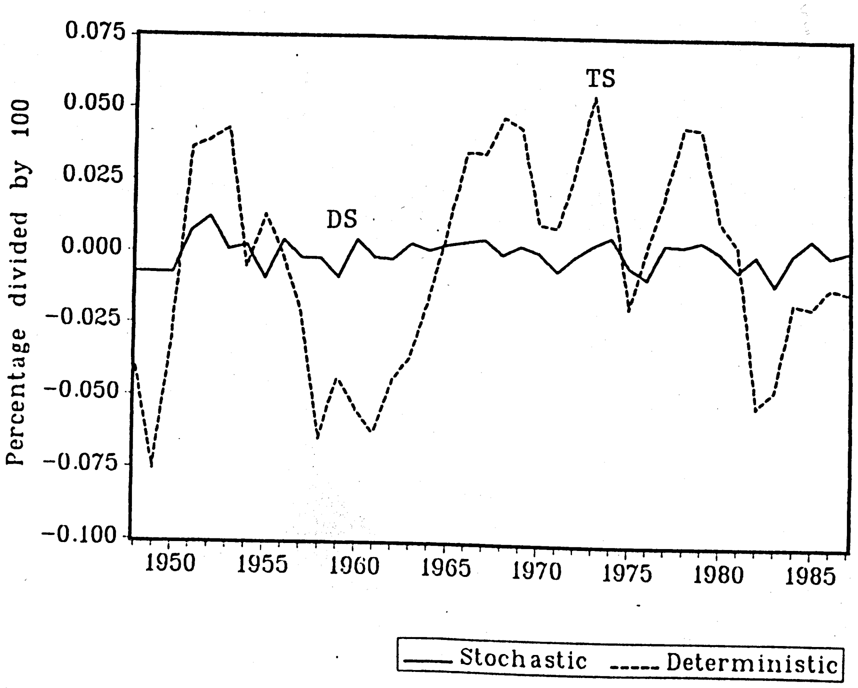


FIGURE 1c: Cyclical components of U.S. real GNP



trend/cycle decompositions. With a smooth deterministic trend (and the associated long and large cycles) traditional (Keynesian) business cycle theories are most relevant. Under the alternative trend and cycle are highly correlated; all shocks have both transitory and permanent effects. Growth models stressing supply factors like technology provide the more natural type of economic theory.

The widely divergent implications of the "DS" and "TS" view of the world have stimulated the econometric investigation of the unit root hypothesis. Classical sampling theory has mostly taken the unit root as the statistical null hypothesis.[1] In the trend/cycle decomposition there is, however, no clear null hypothesis. A priori both views of the world are equally likely. Therefore a Bayesian approach seems more suitable for discriminating between the competing hypotheses.

Some authors (Sims [1988], Christiano and Eichenbaum [1990]) have questioned whether unit roots really matter. Many economic theories can be formulated irrespective of the presence of unit roots. An example is the discussion on the permanent income hypothesis. Also, some parameters of interest are well defined irrespective of unit roots.

The long-run growth rate of real GNP is an example. For the data in figure 1a the trends were estimated using an AR(3) model in levels (TS) and an AR(2) in first diferences (DS):

TS $$D_t = \underset{(0.026)}{1.911} + \underset{(0.001)}{0.018t} \qquad (1)$$

DS $$D_t = D_{t-1} + \underset{(0.004)}{0.018} + \underset{(0.221)}{1.030}\,\varepsilon_t, \qquad (2)$$

---

[1] See Fuller [1976], and the survey by Diebold and Nerlove [1988]. See Cochrane [1988] and Campbell and Mankiw [1987].

The augmented Dickey-Fuller test on the unit root hypothesis gives $\hat{\tau}_{\tau} = -2.7$ which is not significant at the 10% level (see Fuller [1976, page 373]. Hence the "DS" model could very well have generated the data, in which case the inference on the trend will be invalid. Although the point estimates of the growth rate do not differ between (1) and (2), the asymptotic standard error quadruples going from "TS" to "DS". If the unit root hypothesis is true, one has to impose the unit root condition to obtain correct standard errors. The "TS" model leads to spuriously low standard errors in that case. Clearly the problem stems from the different asymptotic theories. Hence it is critical to make the correct decision on the presence of a unit root before proceeding with further inference on parameters of interest. No such discontinuities would arise if finite sample distribution theory would be available.

Although there might in some cases be no direct economic interest in the unit root hypothesis, still asymptotic econometric inference on parameters of interest is widely different for stationary and integrated time series.[2] A general example is the linear regression model when some of the explanatory variables are possibly integrated. Because of the dichotomy in asymptotic approximations one has to decide early on in an empirical study whether a given time series is stationary or not. A wrong decision can lead to serious pretest bias. A Bayesian approach can avoid this dichotomy because it takes account of the uncertainty on the critical autoregressive root. Through a careful specification of the prior in accordance with the time series representation Bayesian inference allows for a smooth continuous transition between a stationary and an integrated time series model.

Specifically, in this paper we will use a normal prior on the unconditional mean of the time series; the variance of this normal distribution continuously increases as an autoregressive root approaches

[2] See, *e.g.*, Durlauf and Phillips [1988].

unity. The posterior distribution of the autoregressive parameters also allows for a Bayesian posterior odds test of the unit root hypothesis.

The organization of the paper is as follows. In section 2 we review some of the issues in the representation of a time series model. We write an autoregressive model in a form that is explicit in our parameters of interest. In section 3 we specify our prior and derive the posterior distribution for an AR(p) model with trend and constant term. We also discuss a posterior odds test for the unit root hypothesis within this model. Section 4 contains empirical results for time series of annual real per capita GNP for 16 OECD countries. Some final remarks are given in section 5.

2. MODEL REPRESENTATION.

In the statistical analysis we will consider autoregressive models around a linear trend,

$$\alpha(L)(y_t - \delta t - \mu) = \varepsilon_t \tag{3}$$

In (1) $\alpha(L) = \sum_{i=0}^{p} \alpha_i L^i$ is a lag polynomial with $\alpha_0 = 1$. All roots λ_j (j=1,...,p) of $\alpha(z)$ lie outside the unit circle with the possible exception of a single unit root, $\lambda_1 = 1$. The deterministic component of the series $\{y_t\}$ is defined as $D_t = \mu + \delta t$, which we will call the "trend". The errors ε_t are white noise with variance σ^2.

To isolate the parameter that determines the presence of a unit root we decompose the lag polynomial as

$$\alpha(L) = \alpha(1)L + \alpha^*(L)(1-L), \tag{4}$$

where $\alpha^*(L) = \sum_{i=0}^{p-1} \alpha_i^* L^i$ with $\alpha_i^* = -\sum_{j=i+1}^{p} \alpha_j$ for $i>0$, and $\alpha_0^* = \alpha_0 = 1$. Using (4) the model (3) can be written explicitly as

$$\Delta y_t - \delta = -\alpha(y_{t-1} - \delta(t-1) - \mu) + \sum_{i=1}^{p-1} \alpha_i^*(\Delta y_{t-i} - \delta) + \varepsilon_t, \tag{5}$$

where we have used the shorthand α for $\alpha(1)$. The model can also be written in two forms that are linear in the parameters:

$$\Delta y_t = \beta_0 + \beta_1 t + \beta_2 y_{t-1} + \sum_{i=1}^{p-1} \beta_{i+2} \Delta y_{t-i} + \varepsilon_t, \tag{6a}$$

$$y_t = \beta_0 + \beta_1 t + \sum_{i=1}^{p} \alpha_i y_{t-i} + \varepsilon_t \tag{6b}$$

where the β's are functions of the parameters μ, δ, α and α_i^* $(i=1,\ldots,p-1)$. Equations (4) and (5) show that in the unit root case, $\alpha = 0$, the trend term cancels, meaning that in both (6a) and (6b) there is the additional constriant $\beta_1 = 0$. From (5) it is seen that μ drops out if $\alpha=0$. The parameter δ represents the "equilibrium" or "natural" growth rate of the time series $\{y_t\}$. The parameter μ can be interpreted as the intercept of the trendline $D_t = \mu + \delta t$.[3] Since we consider α, μ and δ as parameters of interest, we prefer to work with representation (5) instead of representations (6a,b) where β_0 and β_1 are unrestricted.

The choice of parameterization is critical in the Bayesian analysis with uninformative priors. This is one primary difference between the analysis presented below and related work by DeJong and Whiteman [1989a,b], who use the unrestricted linear representation (6b). Although the likelihood functions for the two models ((5) and (6b)) are identical, the prior on the

[3] If there is no trend growth, $\delta = 0$, the parameter μ denotes the unconditional mean of a series.

parameters of interest differs in an essential way. A flat prior on the β's in (6b) implies an informative prior on the parameters in (5) and vice versa. The second important difference between our approach and DeJong and Whiteman concerns the treatment of the constant term; see section 3 below.

The decomposition of a given time series in a trend component and a cyclical component is often the main purpose of time series analysis. But as μ does not appear in (5) when $\alpha=1$, the intercept of the trendline can not be estimated in that case. In fact, if the series contains a unit root the intercept is not defined, and the deterministic trend D_t does not exist.

For series with a unit root (called I(1), integrated of order one), the trend component is defined differently. Under the unit root hypothesis representation (5) simplifies to

$$\alpha^*(L)(\Delta y_t - \delta) = \varepsilon_t \tag{7}$$

Since the unit root has been extracted, the resulting polynomial $\alpha^*(L)$ is stationary and can be inverted, implying

$$\Delta y_t = \delta + \phi(L)\varepsilon_t = \delta + \phi(1)\varepsilon_t + \phi^*(L)(1-L)\varepsilon_t, \tag{8}$$

where $\phi(L) = \alpha^*(L)^{-1} = \sum_{i=0}^{\infty}\phi_i L^i$, and $\phi^*(L) = \sum_{i=0}^{\infty}\phi_i^* L^i$ with $\phi_i^* = -\sum_{j=i+1}^{\infty}\phi_j$. A stochastic trend is defined as

$$D_t = \delta + D_{t-1} + \phi(1)\varepsilon_t \tag{9}$$

If $\phi(1) = 0$, the trend is deterministic and coincides with the earlier definition.[4,5] The parameter δ is well defined both under the I(0) as well as the I(1) hypothesis.

3. PRIOR SPECIFICATION AND THE DERIVATION OF THE POSTERIOR DENSITY.

Given a sample of T observations on $\{y_t\}$ and p pre-sample observations we write the AR(p) model in matrix notation. Define the following functions of α, δ, and the data:

$\mathbf{y} = \mathbf{y}(\alpha,\delta)$: a T-vector with t^{th} element $(\Delta y_t - \delta + \alpha y_{t-1} - \alpha\delta(t-1))$,

$\mathrm{Z} = \mathrm{Z}(\delta)$: a $(T\times(p-1))$ matrix with t^{th} row $(\Delta y_{t-1}-\delta, \ldots, \Delta y_{t-p+1}-\delta)$,

$\mathrm{X} = \mathrm{X}(\alpha,\delta) = (\iota\alpha \mid \mathrm{Z})$, with ι a T vector of ones.

In matrix notation (5) than becomes

$$\mathbf{y} = \mathrm{X}\beta + \mathbf{e}, \tag{10}$$

where $\beta' = (\mu \mid \alpha^{*\prime})$, and $\mathbf{e}$ is a T vector of independently and identically distributed errors. Conditional on α and δ the model is linear in μ and α^*. Given the p pre-sample observations all rows of $\mathbf{y}$ and X are well defined functions of α and δ. Note that $\mathrm{X}(\alpha,\delta)$ is of reduced column rank if $\alpha=0$. The likelihood function for this model reads

---

[4] Alternative decompositions in trend and cycle exist. See for example Watson [1986].

[5] The implied range of $\phi(1)$ shows the limitations of a low order autoregressive representation. In an AR(1) the persistence measure $\phi(1)$ can take only two values: zero and unity. For low order AR models $\phi(1)$ is either zero, or it can take on values in an interval that is bounded away from zero. The latter property follows from the stationarity conditions on $\alpha^*(L)$. For an ARMA representation the range of $\phi(1)$ is not restricted.

$$L(\mathbf{y}|\alpha,\delta,\beta,\sigma^2,\underset{\sim}{y}) = (2\pi)^{-T/2}\sigma^{-T}\exp[-\frac{1}{2\sigma^2}\mathbf{e}'\mathbf{e}], \qquad (11)$$

where $\underset{\sim}{y} = (y_{-p+1},\ldots,y_0)'$ is the vector of pre-sample observations, and e is defined in (10).

For most parameters we can use a non-informative uniform prior. Only α and μ require a careful prior specification. The complete prior is given as

$$\begin{aligned}
&f(\sigma) \propto \sigma^{-1}, && \text{flat prior on } \ln\sigma \\
&f(\alpha) = (1-A)^{-1}, && \text{uniform prior on } S = (0,A],\ A\le 1, \\
&f(\delta,\alpha^*) \propto 1 && \text{uninformative flat prior,} \\
&f(\mu|\alpha,\sigma) = (2\pi)^{-1/2}\sigma^{-1}(1-(1-\alpha)^2)^{1/2}\exp[-\frac{1-(1-\alpha)^2}{2\sigma^2}(\mu-y_0)^2], && \\
& && \text{normal prior of } \mu \text{ conditional on } \alpha \text{ and } \sigma.
\end{aligned} \qquad (12)$$

The prior on α is specified to exclude explosive time series behaviour. The upper bound A will be close to zero for time series with a near unit root. The prior on (α,α^*) does not guarantee that the model is stationary. We will, however, always consider data that are informative enough to ensure that the posterior probability mass in the non-stationary region is negligible. The uniform prior on α^* greatly simplifies the derivations.

One would also like to be non-informative on μ. But a uniform prior defined on $\mu\in\mathbb{R}$ leads to an improper marginal posterior density for α. The intuitive reason is that the data will contain almost no information on μ when α is close to zero, so that the improper prior on μ will not be revised sufficiently to obtain a proper bivariate marginal density on (α,μ) that can be integrated; see Schotman and Van Dijk [1990] for technical details.

The prior on μ in (12) is more concentrated than the uniform density. It is a proper density for all $\alpha\in S$. The prior is centered around y_0. The variance is a function of α and σ, and increases as α tends to zero.

Approaching the limit $\alpha=0$ the prior will become improper, consistent with the specification of the time series model. Under the unit root one can not learn about the unconditional mean of a time series (in deviation from a trend).[6]

The prior can be expressed as an additional observation generated by the model

$$y_0 = \mu + \left(1 - (1-\alpha)^2\right)^{-1/2} e_0, \tag{13}$$

where e_0 is assumed normally distributed $n(0,\sigma^2)$. Therefore we write the augmented linear model

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\beta + \tilde{\mathbf{e}}, \tag{14}$$

where

$$\tilde{\mathbf{y}} = \begin{pmatrix} \sqrt{1 - (1-\alpha)^2}\; y_0 \\ \mathbf{y} \end{pmatrix}, \qquad \text{and} \qquad \tilde{\mathbf{X}} = \left(\begin{array}{c:c} \sqrt{1 - (1-\alpha)^2} & 0' \\ \iota\alpha & \mathbf{Z} \end{array}\right).$$

The posterior density $p(\alpha,\delta,\beta,\sigma|\text{data})$ can be written in a straightforward way as

$$p(\alpha,\delta,\beta,\sigma^2|\text{data}) \propto \sqrt{\alpha(2-\alpha)}\; \sigma^{-(T+2)} \exp[-\frac{1}{2\sigma^2}\, \tilde{\mathbf{e}}'\tilde{\mathbf{e}}] \tag{15}$$

To compute the posterior of the parameters of interest we integrate the posterior density over the nuisance parameter σ using the integration formula

---

[6] The prior can be further sharpened by exploiting all pre-sample observations. In that case the prior will also become dependent on δ and α^*. The current prior is chosen for analytical convenience.

$$\int_0^\infty \sigma^{-(T+1)} \exp(-\frac{1}{2\sigma^2}\tilde{e}'\tilde{e})d\sigma = (\tilde{e}'\tilde{e})^{-T/2} \tag{16}$$

Using (16) the marginal posterior of (α,δ,β) takes the form

$$p(\alpha,\delta,\beta|\text{data}) \propto \sqrt{\alpha(2-\alpha)}\ (\tilde{e}'\tilde{e})^{-(T+1)/2} \tag{17}$$

Conditional on α and δ the posterior of β is of the Student-t type. The marginal posterior of α and δ can thus be obtained by straightforward integration over β. To perform the integration step we write $\tilde{e}'\tilde{e}$ as an explicit quadratic form in β.

$$\tilde{e}'\tilde{e} = Q + (\beta - \hat{\beta})'\tilde{X}'\tilde{X}(\beta - \hat{\beta}), \tag{18}$$

where $Q = \tilde{y}'(I - \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}')\tilde{y}$, and $\hat{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y}$. Using the form of the multivariate Student-t density the marginal posterior of (α,δ) becomes

$$p(\alpha,\delta|\text{data}) \propto \sqrt{\alpha(2-\alpha)}\ |\tilde{X}'\tilde{X}|^{-1/2}Q^{-(T-p+1)/2} \tag{19}$$

This density is not of any standard type, since Q and $|\tilde{X}'\tilde{X}|$ depend in a highly non-linear way on α and δ through the data functions $\tilde{y}$ and $\tilde{X}$.

Our interest is in the limiting behaviour of the posterior when α tends to zero. We will therefore write $\tilde{X}'\tilde{X}$ and Q more explicit as functions of α. First we have that

$$\tilde{X}'\tilde{X} = \begin{pmatrix} \alpha(2-\alpha) + \alpha^2\iota'\iota & \alpha\iota'Z \\ \alpha Z'\iota & Z'Z \end{pmatrix}, \tag{20}$$

from which it follows that

$$|\tilde{X}'\tilde{X}| = q|Z'Z|, \tag{21}$$

where $q = \alpha(2 - \alpha + \alpha\iota'M\iota)$ with $M = I - Z(Z'Z)^{-1}Z'$. The product of the two Jacobians in (19) thus can be written as

$$\sqrt{\alpha(2-\alpha)}\ |\tilde{X}'\tilde{X}|^{-1/2} = |Z'Z|^{-1/2}\left(\frac{\alpha(2-\alpha)}{\alpha(2-\alpha) + \alpha^2\iota'M\iota}\right)^{1/2}$$

$$= |Z'Z|^{-1/2}\left(1 + \left(\frac{\alpha}{2-\alpha}\right)\iota'M\iota\right)^{-1/2}, \tag{22}$$

Since Z and therefore M only depend on δ (22) has the well defined limit $|Z'Z|^{-1/2}$ as $\alpha\to 0$. For the limiting behaviour of Q we need the partitioned inverse

$$(\tilde{X}'\tilde{X})^{-1} = \left(\begin{array}{c|c} 1/q & -(\alpha/q)\iota'Z(Z'Z)^{-1} \\ -(\alpha/q)(Z'Z)^{-1}Z'\iota & (Z'Z)^{-1} + (\alpha^2/q)(Z'Z)^{-1}Z'\iota\iota'Z(Z'Z)^{-1} \end{array}\right),$$

Expanding the quadratic form Q using this inverse one can obtain

$$Q = \tilde{y}'(I - \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}')\tilde{y} \tag{23}$$

$$= y'My + \alpha(2-\alpha)y_0^2 - \frac{\alpha^2}{q}\left((2-\alpha)y_0 + \iota'My\right)^2$$

Using the definition of q this can be simplified to the weighted average of two sums of squares

$$Q = \frac{1}{2-\alpha + \alpha\iota'M\iota}\left(\alpha((\iota'M\iota)(y'My) - (\iota'My)^2) + \frac{2-\alpha}{\iota'M\iota}(y - \alpha y_0\iota)'M(y - \alpha y_0\iota)\right)$$

$$= wQ_1 + (1-w)Q_2, \tag{24}$$

where

$$w = \frac{\alpha}{\left(\frac{2-\alpha}{\iota'\mathbf{M}\iota}\right) + \alpha}$$

$$Q_1 = \left(\mathbf{y}'\mathbf{M}\mathbf{y} - \frac{(\iota'\mathbf{M}\mathbf{y})^2}{\iota'\mathbf{M}\iota}\right)$$

$$Q_2 = (\mathbf{y} - \alpha y_0\iota)'\mathbf{M}(\mathbf{y} - \alpha y_0\iota)$$

In (24) Q_1 can be interpreted as the residual sum of squares, as a function of α and δ, from the regression of $\mathbf{y}$ on a constant after partialling out Z. It is a measure of the precision of determining the intercept of the trend. Likewise Q_2 is the sum of squared residuals of the regresion of $\mathbf{y}-\alpha y_0\iota$ on Z, which for given α and δ measures the precision by which the prior determines the intercept of the trend. The weight w depends critically on α. If $\alpha\to 0$, then $w\to 0$ and the sum of squares function Q is completely determined by Q_2, which itself tends to the sum of squared residuals of the regression of $(\Delta y_t-\delta)$ on $(p-1)$ lags. The latter is also the sum of squared residuals under the unit root hypothesis.

Notice that $\iota'\mathbf{M}\iota = Ts^2$, where s^2 is the residual variance of a regression of a constant on $(p-1)$ lags of $\{\Delta y_t\}$. Hence for all $\alpha>0$ the weight function $w(\alpha,\delta) = ((2-\alpha)s^2T + 1)^{-1}$ tends to unity as the sample size T increases. The posterior will then be dominated by Q_1.

Posterior odds

The univariate marginal posterior of α is obtained after numerical integration over δ. This marginal posterior $p(\alpha|\text{data})$ is of particular interest, as it can be used for a Bayesian test of the unit root hypothesis. For the purpose of the test we will extend the domain of the random variable α to include the endpoint $\alpha=0$. Since the unit root hypothesis is a sharp null, we must assign a discrete prior probability to the event $\alpha=0$. Then one can perform this test using posterior odds K_1 defined as (see Leamer [1978, sec 4.3] or Zellner [1971, p. 297-298]):

$$K_1 = \frac{\Pr(\alpha=0 \mid \text{data})}{\Pr(0<\alpha<a \mid \text{data})} \tag{25}$$

In (24) the stationarity hypothesis is represented as the probability that α is in some interval $(0,a]$ in the stationary region, that may be smaller than the full interval $(0,A]$ over which the posterior is defined. The denominator is proportional to $\frac{1}{a}\int_0^a p(\alpha \mid \text{data})d\alpha$. To derive the posterior probability of a unit root we set up an AR($p-1$) model with a constant growth rate δ for the differenced series Δy. Using the same prior on (α^*,δ,σ) as in (12) the results in Schotman and Van Dijk [1990] imply that the numerator in (25) is proportional to $\lim_{\alpha\to 0} p(\alpha \mid \text{data}) \equiv p(\alpha=0 \mid \text{data})$.[7] As shown above this limit is well defined. Further the constant of proportionality is the same as for the denominator. We treat the unit root hypothesis and the hypothesis of stationarity in a symmetrical way and give both a prior probability of one half. Then the posterior odds are equal to

$$K_1 = \frac{p(\alpha=0 \mid \text{data})}{\frac{1}{a}\int_0^a p(\alpha \mid \text{data})d\alpha} \tag{26}$$

The length of the interval $(0,a]$ is chosen in such a way that it contains 99% of the posterior probability mass in the stationary region. This choice of a avoids Jeffreys' paradox by restricting attention to the domain of α where

---

[7] Because μ is unidentified under the hypothesis $\alpha=0$, we can take any proper prior density on μ.

the posterior density is non-negligible.[8] It implies that the unit root is always compared to a local alternative with interval length of order $T^{-1/2}$.

4. EMPIRICAL RESULTS.

Bivariate marginal posterior densities of α and δ have been computed for annual data on the log of real per capita GNP (or GDP if GNP was unavailable) in 16 OECD countries for the sample period 1948-1987. Results are summarised in figures 2A to 17A. Tables 1 provides some summary statistics of the data. Tables 2 and 3 give classical parameter estimates and posterior moments of parameters of interest for comparison. All results have been obtained with an AR(3) model. The AR(3) specification is the simplest autoregressive model to be characterised by a near nonstationary dominant real root, while also allowing for cyclical behaviour due to complex roots.

The sixteen figures of bivariate posteriors are very similar. The mode of (α,δ) is always attained for positive α and δ. The mode of δ seems independent of the value of α, since the ridge of conditional modes of δ is almost orthogonal to the δ-axis. There is no full independence between α and δ, though. The conditional posterior of δ given $\alpha=0$ has much wider tails than the conditional posterior of δ for any positive value of α; the variance of the conditional distribution decreases as α gets larger. This is the effect alluded to in the example in the introduction. In several cases the location estimate of δ is not affected by the unit root hypothesis, but the estimated scale measures (standard deviations) differ.

---

[8] The value of a directly enters the posterior odds ratio. If a is set at a large value there will be a large interval where the posterior density contains almost no mass. Hence the averaged posterior in the denominator of the odds becomes small. On the other hand one should also avoid taking a too close to zero. See Schotman and Van Dijk [1990] for more discussion.

The effect of the prior can be seen by comparing the posterior densities with the classical results in table 2. Consider, for example, the case of Sweden. The point estimate of δ (under the hypothesis of stationarity) is −5.3% with an extremely large estimate of its standard error. These estimates are obtained as functions of the estimated parameters from the linear AR(3) regression (6b): $\hat{\delta} = \hat{\beta}_1/(1 - \sum_{i=2}^{p+1}\hat{\alpha}_i)$. This estimate has very poor small sample properties close to the unit root. The posterior density in figure 12A, however, has all its mass concentrated on a small positively valued region for δ. The asymptotic standard error in table 2 shows further that the regression model cannot cope with the non-linearity of the model. The Bayesian posterior moments are more sensible and accurate.

Numerically integrating over δ one obtains the marginal posterior $p(\alpha|\text{data})$ shown in figures 2B to 17B.[9] The wide conditional densities of δ close to the unit root can cause irregular behaviour of the univariate marginal posterior of α close to α=0. A relatively clear case is the U.K.. Although the bivariate posterior of (α,δ) has a single mode in the interior of the parameter region, the marginal density $p(\alpha|\text{data})$ attains its mode at the boundary, and is approximately flat in the neighbourhood of α=0. The shift of the mode of α towards the boundary also occurs for other countries. This is another way to express the nonlinear dependence between α and δ.

Entertaining the unit root hypothesis as the null the classical augmented Dickey-Fuller test does not reject the null hypothesis on the 10% level for any series, see table 3. By treating the null and the alternative symmetrically the posterior odds illustrate the uncertainty with regard the conclusion that the series really contain a unit root. The posterior probabilities of the unit root are greater than or equal to 0.7 for eight

[9] For the figures a grid of 61×41 points was chosen for (α,δ). The numerical integrations take a finer grid of 120×60 points.

countries. For these countries the posterior means of δ agree with the sample means and the classical point estimates under the null. The classical point estimates of δ under the stationary alternative are much less plausible. For two countries (USA and NOR) the posterior probability of a unit root is less than 0.5. The point estimates for δ are in these cases very similar but the standard deviations differ. In the remaining six cases the posterior odds and the parameter estimates show that both models are equally likely. Summarizing, the evidence does not favor a particular hypothesis except for a few cases. However, mechanical application of the classical testing procedure would lead to non-rejection of the unit root hypothesis in all cases. This latter conclusion was also reached (with one exception) by Nelson and Plosser for the time series of 14 major macro-economic variables of the USA. The analysis presented in this paper casts serious doubts on such a strong conclusion.

Table 3 also reports estimates of the persistence $\phi(1)$ which is computed from the AR(3) model after imposing the unit root condition. All estimates are close to unity. This partly reflects the limitations of the estimated ARI(2,1) model: $\phi(1)$ must be larger than 0.25 due to the stationarity conditions of an AR(2). The conditional posterior of $\phi(1)$ under the unit root hypothesis is shown for the case of the USA in Figure 18. It is fairly concentrated around one and skew to the right. Since $\phi(1)$ is a function of all parameters one has to make use of a particular numerical procedure to compute its marginal posterior density. In our case we proceeded as follows. The marginal posterior density of δ is not of a known form which allows one to generate random drawings from it in an easy way. Therefore, draw a uniform random number in the interval [0, 1] and through a numerical inversion using the distribution function of δ, generate a random drawing of δ. Conditional upon a given value of δ, generate a value of the remaining parameters of the model from a conditional multivariate t distribution and compute finally a value of $\phi(1)$ as a function of all generated parameters. These steps are repeated 8000 times and the set of 8000 drawings is used to plot the marginal posterior of $\phi(1)$. For illustrative purposes we confine ourselves to the marginal posterior of $\phi(1)$ for the USA. A more detailed analysis of the posterior of $\phi(1)$ is a matter of further research.

5. FINAL REMARKS.

The paper has provided a Bayesian analysis of autoregressive time series around a linear trend, where one of the roots is close to or equal to unity. The application to time series of real GNP has provided some first empirical results. The analysis needs to be extended in several directions.

First, the sensitivity of the posterior odds test of the unit root hypothesis with respect to the specification of the model, in particular the presence of MA components, must be investigated. The inclusion of MA parameters overcomes the restrictions on the persistence measure $\phi(1)$ implied by a low order AR model. If $\phi(1)$ can take on values arbitrarily close to zero, it will be possible to obtain a meaningful marginal posterior density. The posterior density of $\phi(1)$, conditional on the unit root hypothesis, forms a natural complement to the posterior of μ if the time series is stationary. We anticipate a connection between these two posteriors. If the position of the trend is very ill-determined by the data, a unit root is likely, and a sharp estimate of $\phi(1)$ will obtain. Conversely, if $\phi(1)$ has a mode close to zero, the trend is not shifting around much, and stationarity will be a plausible hypothesis with a fixed intercept μ. Another point of research is the sensitivity with respect to a change in the fixed trendline.

Second, the empirical results have been obtained for relatively short time series. With few observations the finite sample Bayesian results differ much from the asymptotic classical results. It is therefore of interest to augment the dataset to longer time series, which for some countries are readily available.

Third, one has to investigate the sensitivity with respect to the prior specification. Our prior on (α,α^*) implies a prior on the roots λ_j $(j=1,\dots,p)$. What is the effect on posterior odds and the marginal distributions of μ and δ if we specify a prior directly on the roots? A related issue is the specification of a prior on $\phi(1)$, treating persistence as a parameter of interest.

TABLE 1: GROWTH RATES OF REAL GNP/GDP PER CAPITA OF 16 OECD COUNTRIES

Country	Period	Mean (percentage)	Standard deviation
United States USA (*)	1948-87	1.88	2.68
Great Britain GBR	1948-87	2.16	1.86
Austria AUT	1950-87	3.64	2.61
Belgium BEL (*)	1950-87	2.88	2.15
Denmark DEN	1950-87	2.65	2.43
Federal Republic of Germany: FRG	1950-87	3.93	3.32
France FRA	1950-87	3.35	1.99
Canada CAN (*)	1948-86	2.58	2.69
Italy ITA	1950-87	3.81	3.17
Netherlands NED (*)	1950-87	2.80	2.65
Norway NOR	1950-87	3.40	1.92
Sweden SWE	1950-87	3.52	1.73
Switserland SWI	1948-86	2.14	3.03
Japan JAP (*)	1952-86	5.94	3.92
Ireland IRL	1948-86	2.66	2.18
Australia AUS	1949-87	2.14	2.29

[1] The data for * countries refer to Gross National Product (GNP). For the other countries the data refer to Gross Domestic Product (GDP).

TABLE 2: ESTIMATED GROWTH RATES OF REAL GNP/GDP

Country	$\hat{\delta}$	$\hat{\delta}_0$	$E(\delta)$	$\hat{\mu}$
USA	0.018 *(0.001)*	0.018 *(0.004)*	0.019 *(0.002)*	1.95 *(0.03)*
GBR	0.020 *(0.002)*	0.021 *(0.002)*	0.021 *(0.002)*	0.82 *(0.06)*
AUT	0.021 *(0.014)*	0.039 *(0.006)*	0.039 *(0.006)*	4.30 *(0.59)*
BEL	0.029 *(0.005)*	0.028 *(0.007)*	0.030 *(0.005)*	4.99 *(0.18)*
DEN	0.019 *(0.021)*	0.028 *(0.004)*	0.028 *(0.003)*	3.88 *(1.21)*
FRG	0.015 *(0.011)*	0.036 *(0.008)*	0.037 *(0.007)*	2.27 *(0.45)*
FRA	0.034 *(0.005)*	0.037 *(0.007)*	0.038 *(0.006)*	2.99 *(0.23)*
CAN	0.026 *(0.002)*	0.025 *(0.004)*	0.026 *(0.003)*	1.68 *(0.08)*
ITA	0.014 *(0.029)*	0.037 *(0.006)*	0.038 *(0.006)*	8.63 *(1.57)*
NED	0.005 *(0.037)*	0.029 *(0.005)*	0.029 *(0.005)*	3.23 *(2.17)*
NOR	0.035 *(0.001)*	0.033 *(0.003)*	0.035 *(0.002)*	3.23 *(0.03)*
SWE	-0.053 *(0.483)*	0.025 *(0.004)*	0.025 *(0.003)*	13.27 *(111.7)*
SWI	0.010 *(0.013)*	0.022 *(0.004)*	0.022 *(0.004)*	3.03 *(0.68)*
JAP	-0.002 *(0.150)*	0.060 *(0.014)*	0.061 *(0.012)*	8.72 *(10.2)*
IRL	0.029 *(0.004)*	0.026 *(0.004)*	0.027 *(0.003)*	7.00 *(0.13)*
AUS	0.023 *(0.003)*	0.021 *(0.003)*	0.021 *(0.003)*	1.45 *(0.08)*

Notes to TABLE 2:

1) $\hat{\mu}$ and $\hat{\delta}$ are estimated from the linear AR(3) regression reparameterized as

$$\Delta y_t - \delta = -\alpha(y_{t-1} - \delta(t-1) - \mu) + \sum_{i=0}^{p-1} \alpha_i^*(\Delta y_{t-i} - \delta) + \varepsilon_t.$$

2) $\hat{\delta}_0$ is estimated from the ARI(2,1) regression model reparameterized as

$$\Delta^2 y_t = -\frac{1}{\phi(1)}(\Delta y_{t-1} - \delta_0) - \alpha_1^* \Delta^2 y_{t-1} + \varepsilon_t.$$

3) Standard errors of parameters estimated from the autoregressive regression model are in parentheses.

4) $E(\delta)$ denotes the posterior mean of δ under the alternative hypothesis. Posterior standard deviation are given in parenthesis.

TABLE 3: UNIT ROOTS IN REAL GNP.

Country	$\hat{\alpha}$	$E(\alpha)$	$\hat{\phi}(1)$	$Pr(\alpha=0)$	a	$\hat{\tau}_\tau$
USA	0.36	0.26	1.03	0.44	0.60	-2.7
	(0.13)	*(0.14)*	*(0.22)*			
GBR	0.21	0.14	0.80	0.63	0.42	-1.7
	(0.12)	*(0.10)*	*(0.15)*			
AUT	0.08	0.03	1.55	0.74	0.13	-1.5
	(0.06)	*(0.03)*	*(0.46)*			
BEL	0.09	0.08	1.85	0.68	0.23	-1.3
	(0.07)	*(0.06)*	*(0.72)*			
DEN	0.05	0.06	0.90	0.70	0.22	-0.6
	(0.08)	*(0.05)*	*(0.19)*			
FRG	0.11	0.04	1.41	0.74	0.14	-2.1
	(0.05)	*(0.04)*	*(0.39)*			
FRA	0.09	0.07	1.72	0.64	0.21	-1.6
	(0.06)	*(0.05)*	*(0.65)*			
CAN	0.19	0.14	0.88	0.63	0.40	-1.6
	(0.12)	*(0.10)*	*(0.18)*			
ITA	0.07	0.05	1.11	0.72	0.19	-0.9
	(0.07)	*(0.04)*	*(0.29)*			
NED	0.05	0.05	1.17	0.73	0.18	-0.7
	(0.07)	*(0.04)*	*(0.29)*			
NOR	0.48	0.32	0.99	0.42	0.71	-2.8
	(0.17)	*(0.18)*	*(0.24)*			
SWE	0.01	0.03	1.25	0.72	0.13	-0.2
	(0.05)	*(0.03)*	*(0.30)*			
SWI	0.07	0.05	1.05	0.71	0.18	-1.0
	(0.07)	*(0.04)*	*(0.21)*			
JAP	0.02	0.03	2.24	0.73	0.12	-0.5
	(0.05)	*(0.03)*	*(0.93)*			
IRL	0.11	0.08	1.23	0.69	0.27	-1.1
	(0.10)	*(0.06)*	*(0.31)*			
AUS	0.15	0.101	0.86	0.66	0.32	-1.4
	(0.10)	*(0.08)*	*(0.18)*			

Notes to table 3:

1) $\hat{\alpha}$ is estimated from the linear AR(3) regression reparameterized as

$$\Delta y_t - \delta = -\alpha(y_{t-1} - \delta(t-1) - \mu) + \sum_{i=0}^{p-1} \alpha_i^*(\Delta y_{t-i} - \delta) + \varepsilon_t.$$

2) $\hat{\phi}(1)$ is estimated from the ARI(2,1) regression model reparameterized as

$$\Delta^2 y_t = -\frac{1}{\phi(1)}(\Delta y_{t-1} - \delta) - \alpha^* \Delta^2 y_{t-1} + \varepsilon_t.$$

3) Standard errors of parameters estimated from the autoregressive regression model are in parentheses.

4) $\hat{\tau}_\tau$ is the augmented Dickey-Fuller test against a unit root, defined as minus the regression t-statistic of α.

5) $\Pr(\alpha=0)$ is the posterior probability of a unit root.

6) a is the length of the 99% confidence interval $(0, a]$ for α. $E(\alpha)$ denotes the posterior mean under the alternative hypothesis.

Figure 2A. Bivariate posterior of (α, δ): United States

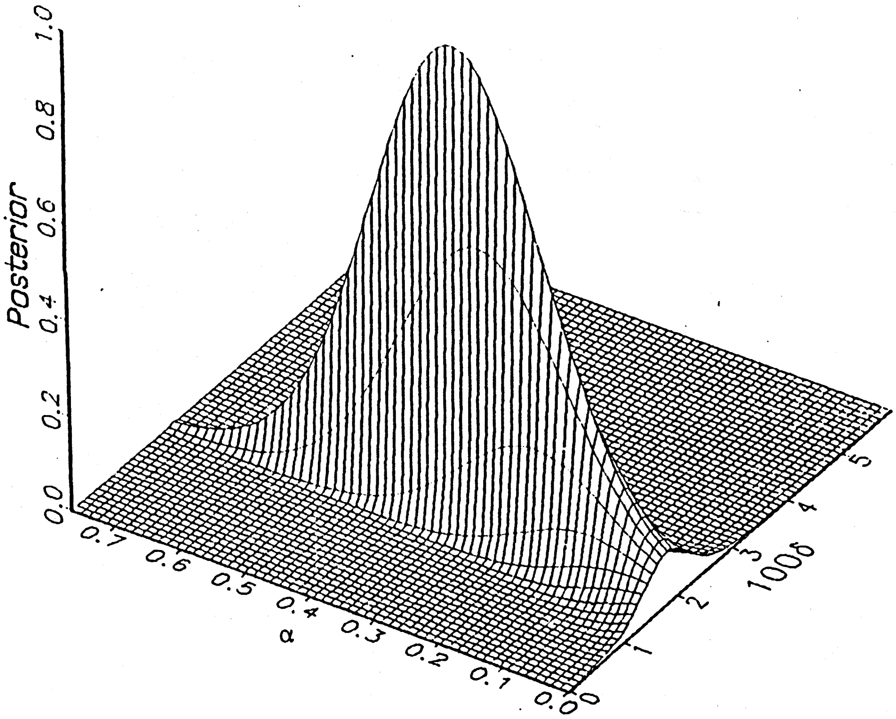


Figure 2B. Marginal posterior of α: United States

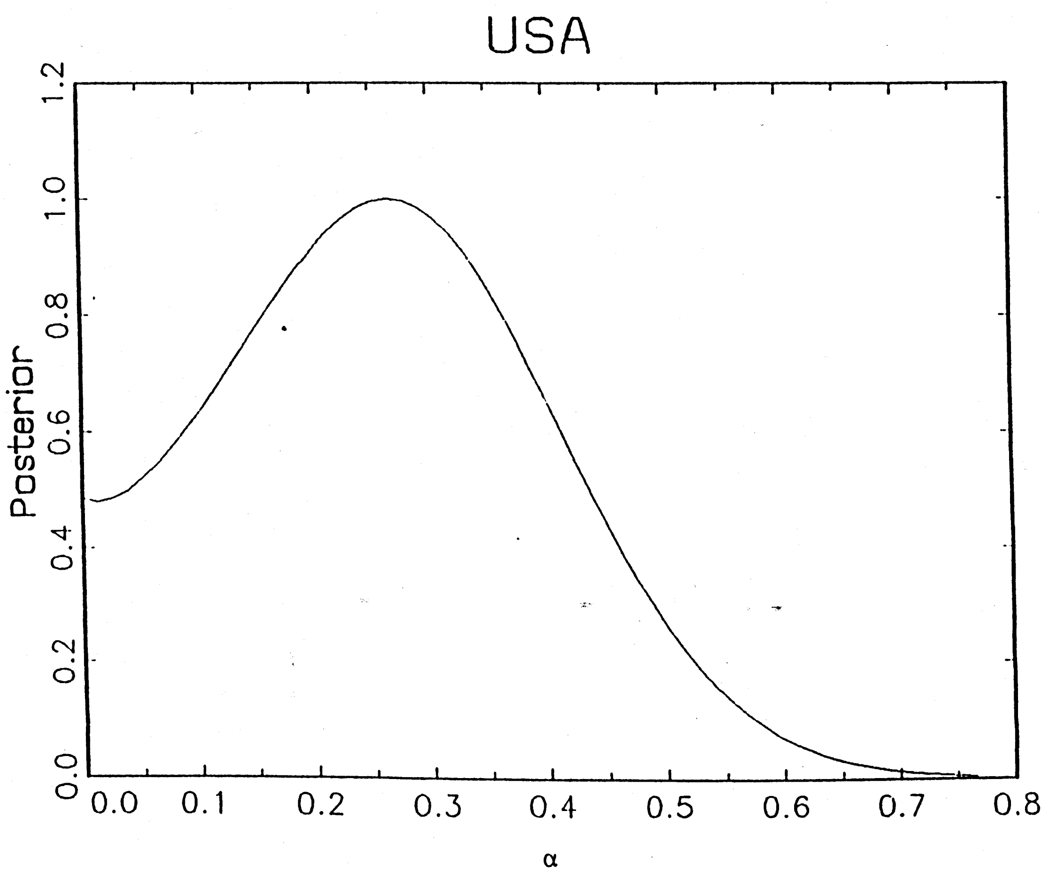


Figure 3A. Bivariate posterior of (α, δ): United Kingdom

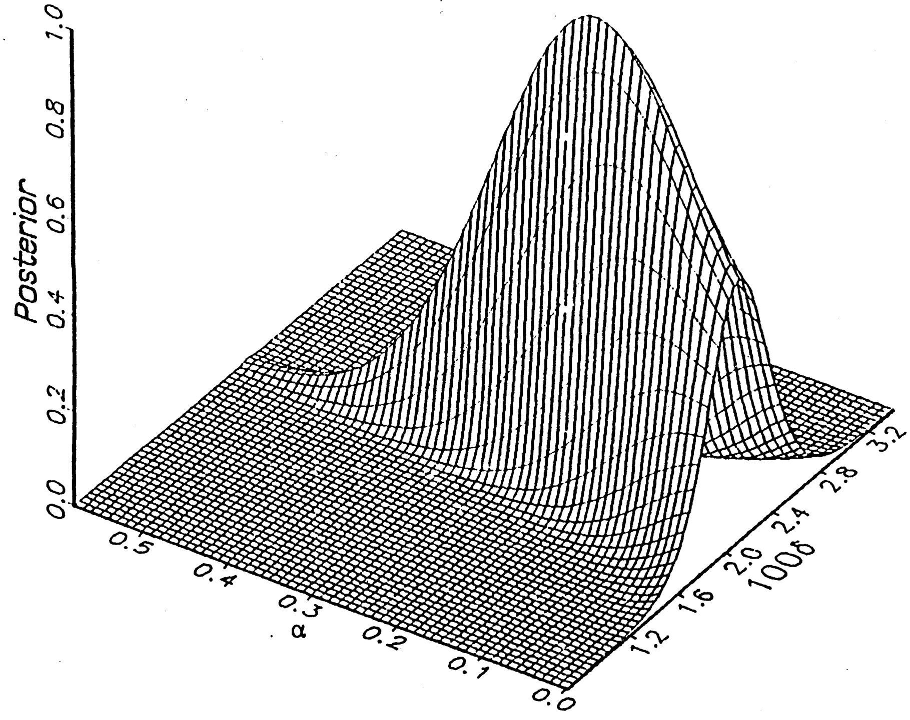


Figure 3B. Marginal posterior of α: United Kingdom

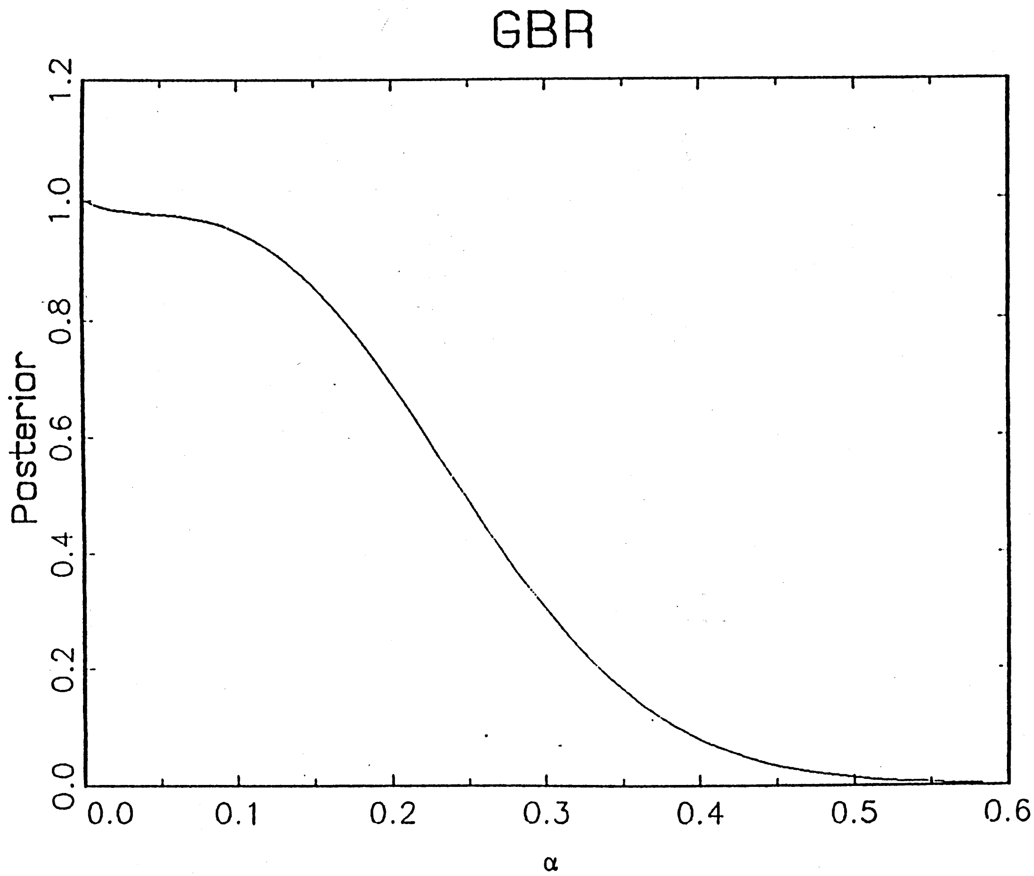


Figure 4A. Bivariate posterior of (α, δ): Austria

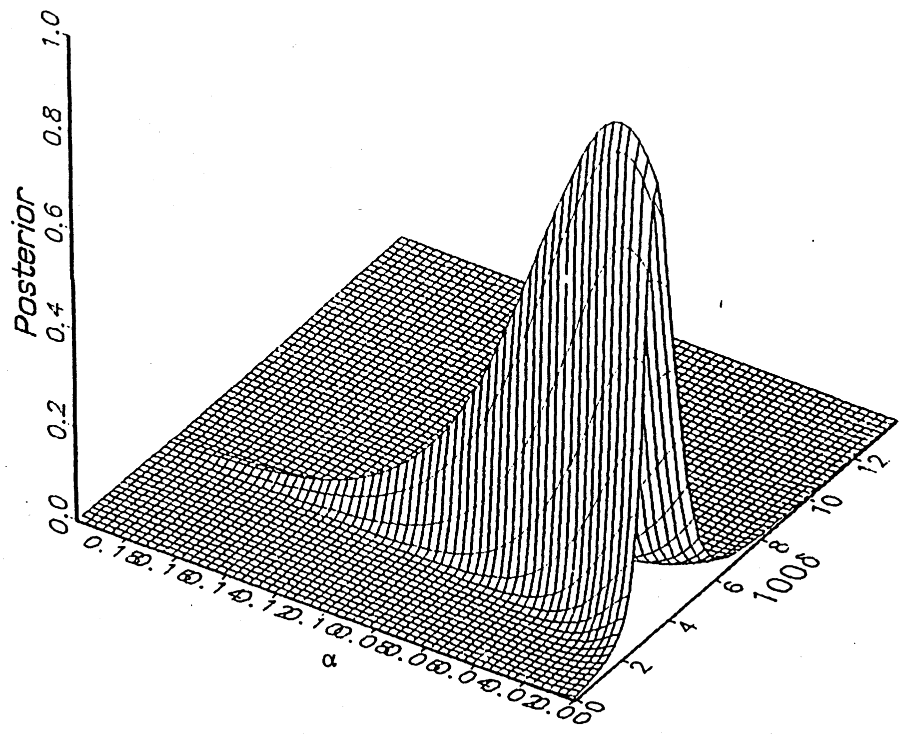


Figure 4B. Marginal posterior of α: Austria

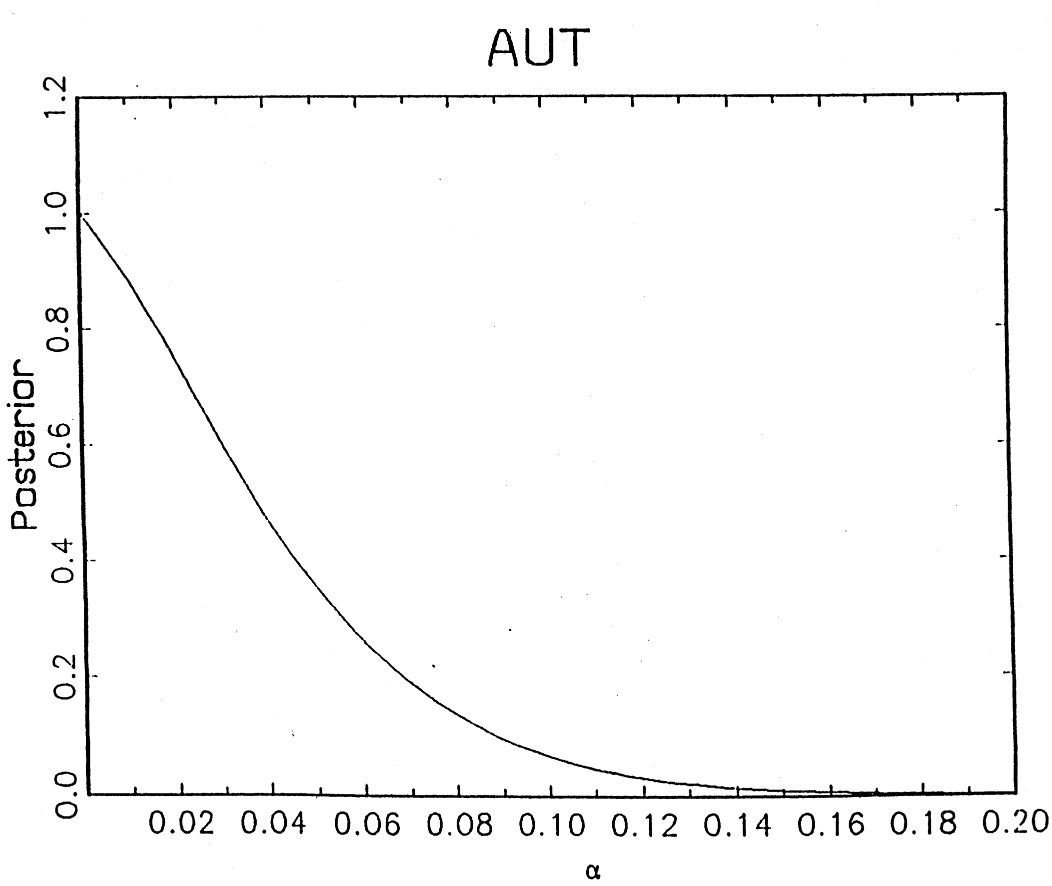


Figure 5A. Bivariate posterior of (α, δ): Belgium

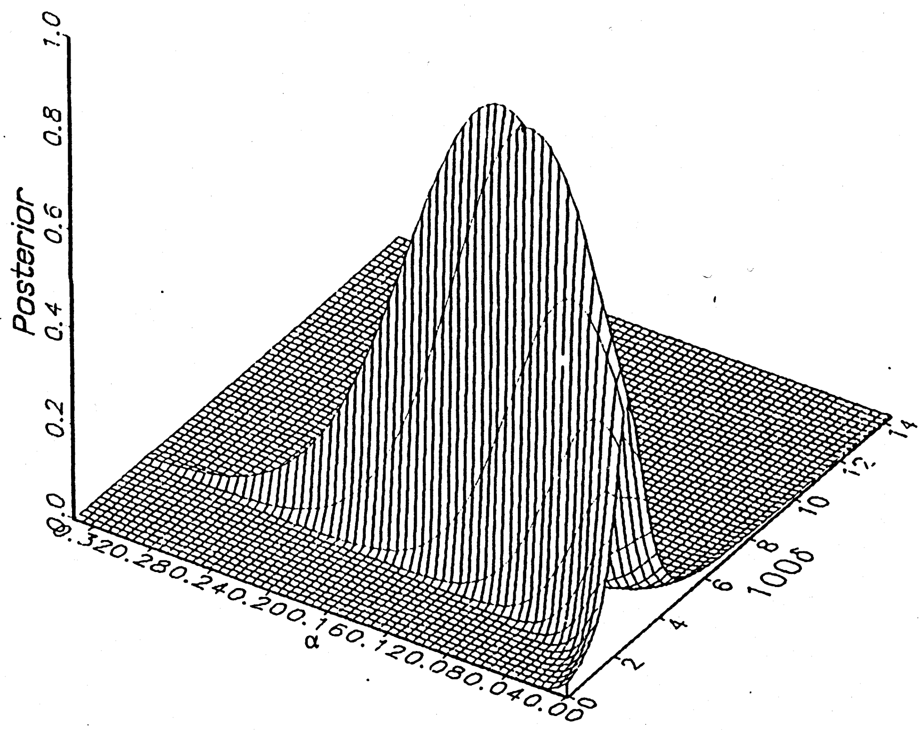


Figure 5B. Marginal posterior of α: Belgium

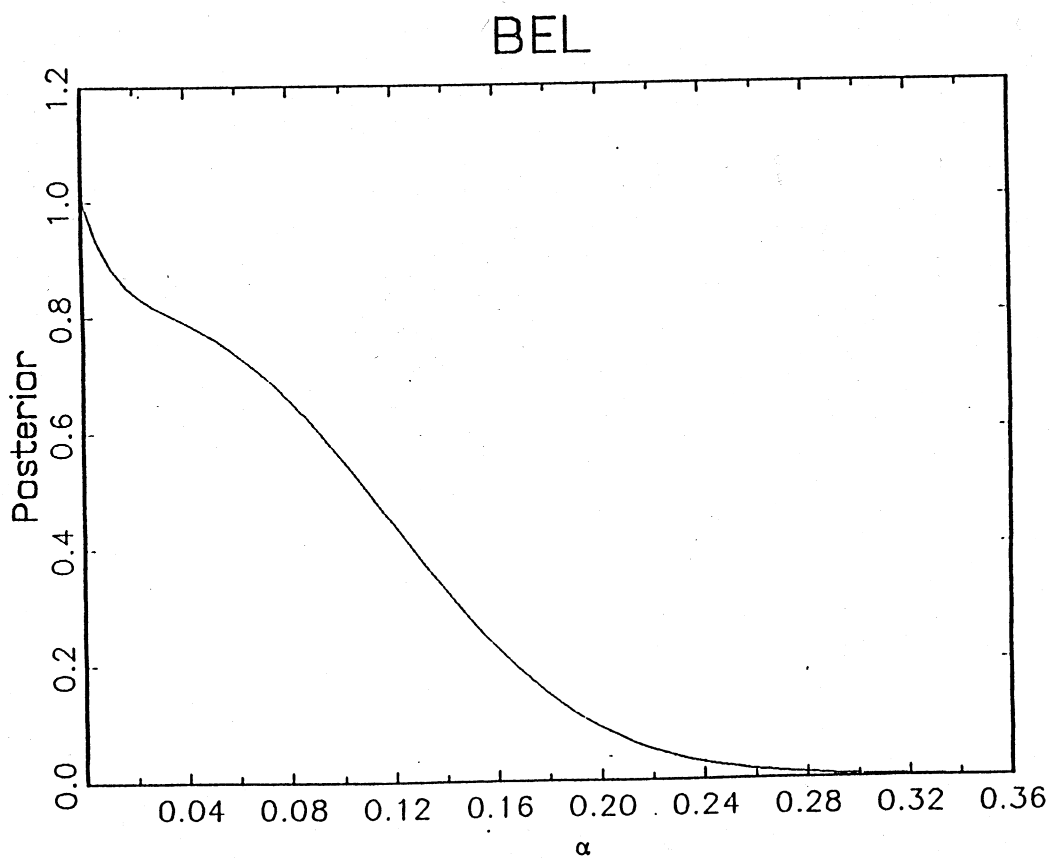


Figure 6A. Bivariate posterior of (α, δ): Denmark

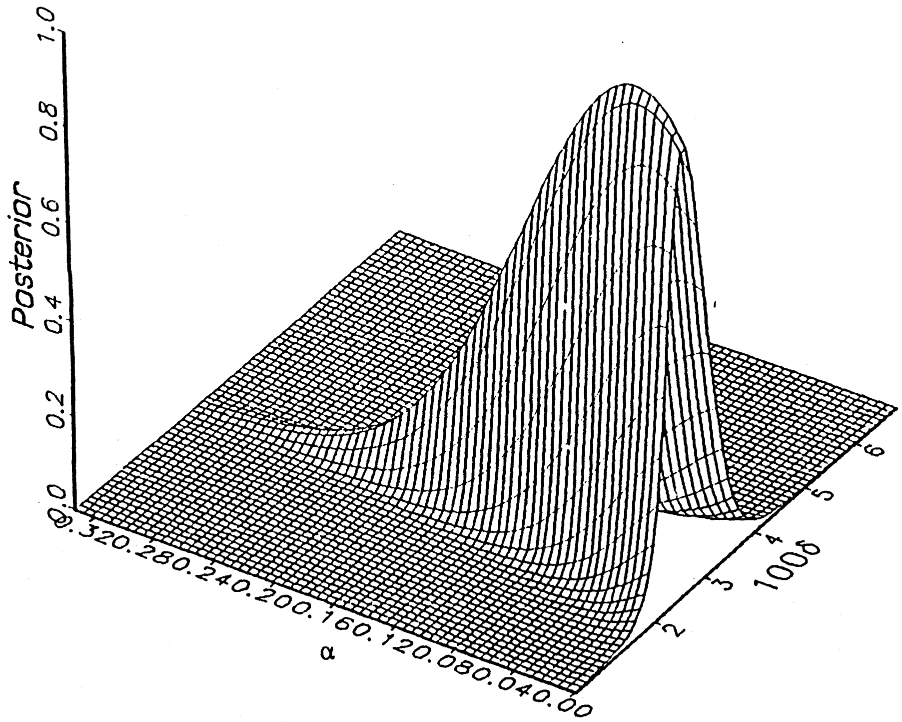


Figure 6B. Marginal posterior of α: Denmark

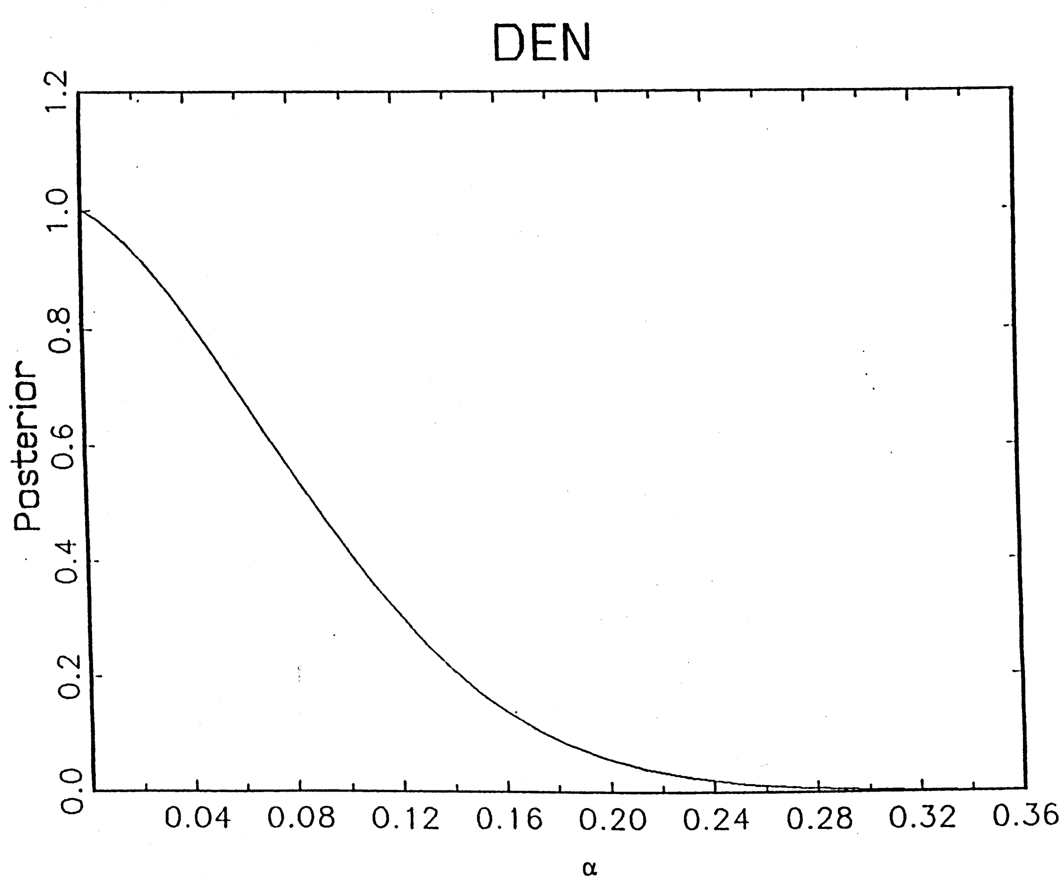


Figure 7A, Bivariate posterior of (α, δ): France

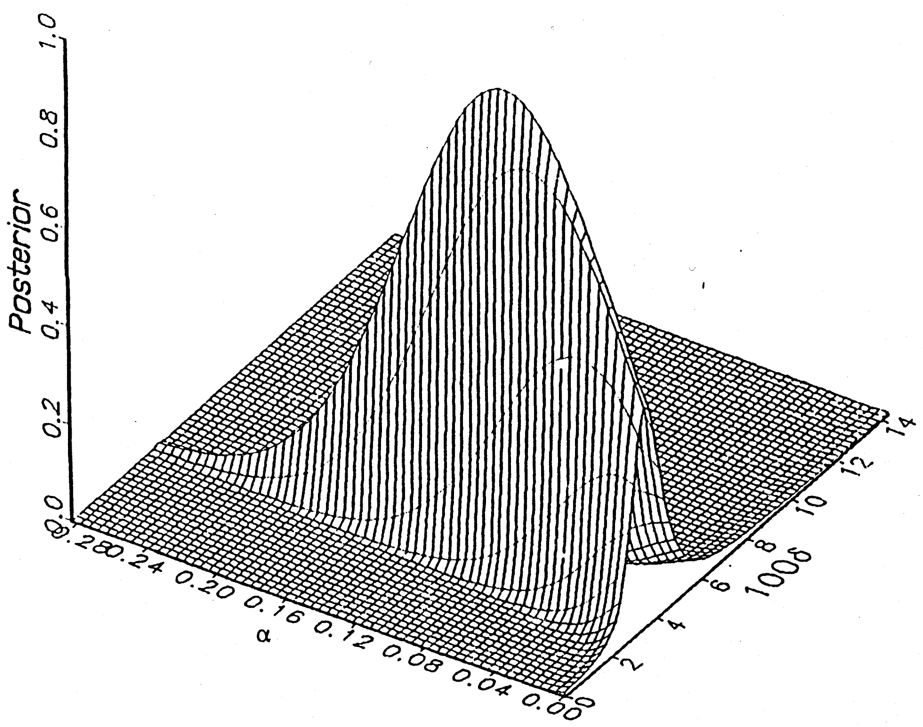


Figure 7B. Marginal posterior of α: France

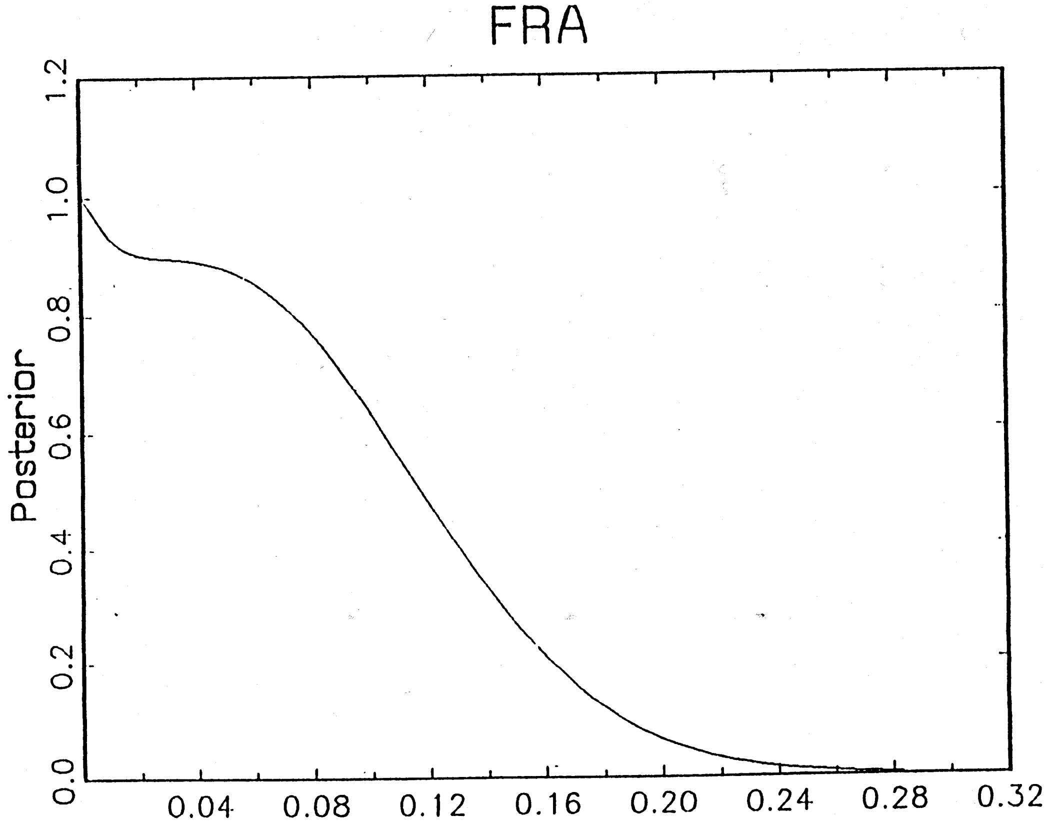


Figure 8A. Bivariate posterior of (α, δ): Germany

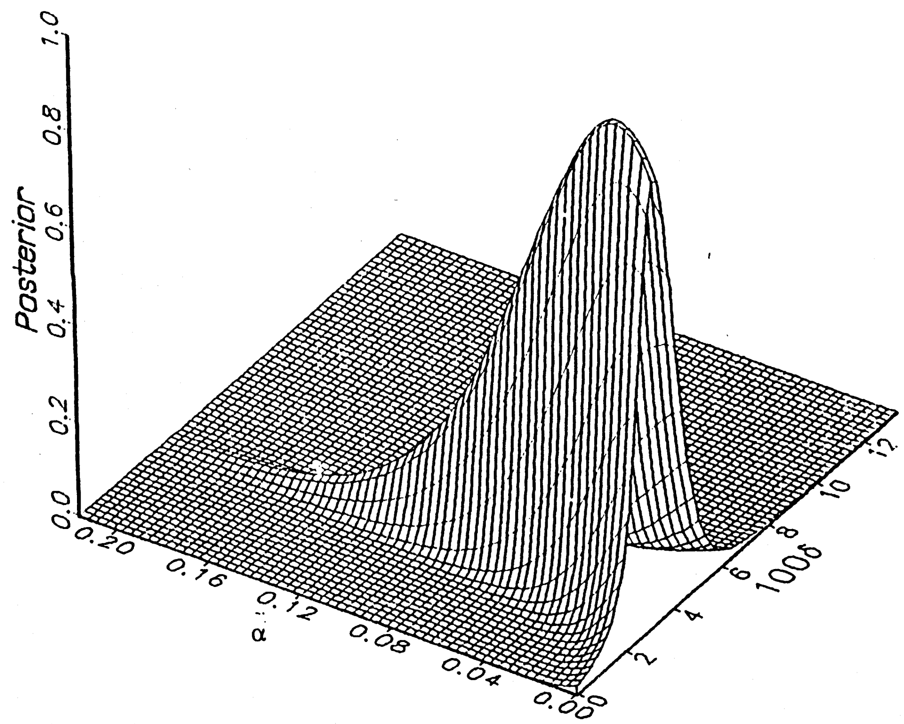


Figure 8B. Marginal posterior of α: Germany

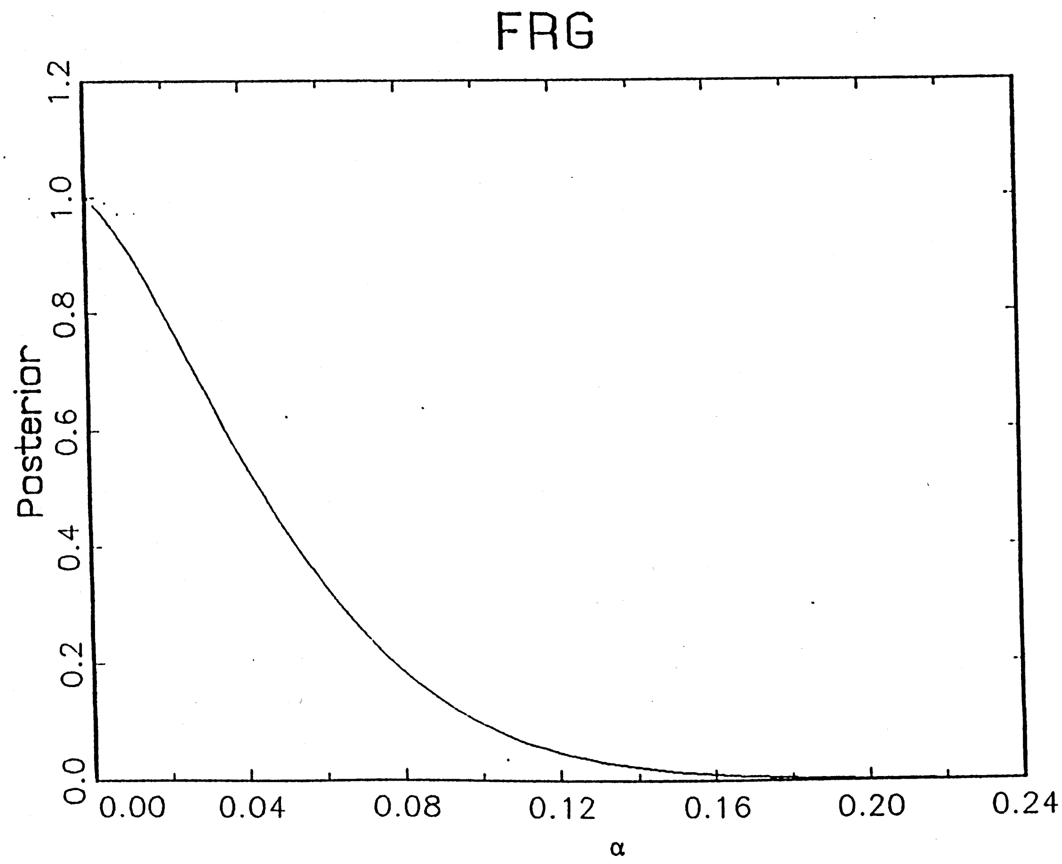


Figure 9A. Bivariate posterior of (α, δ): Italy

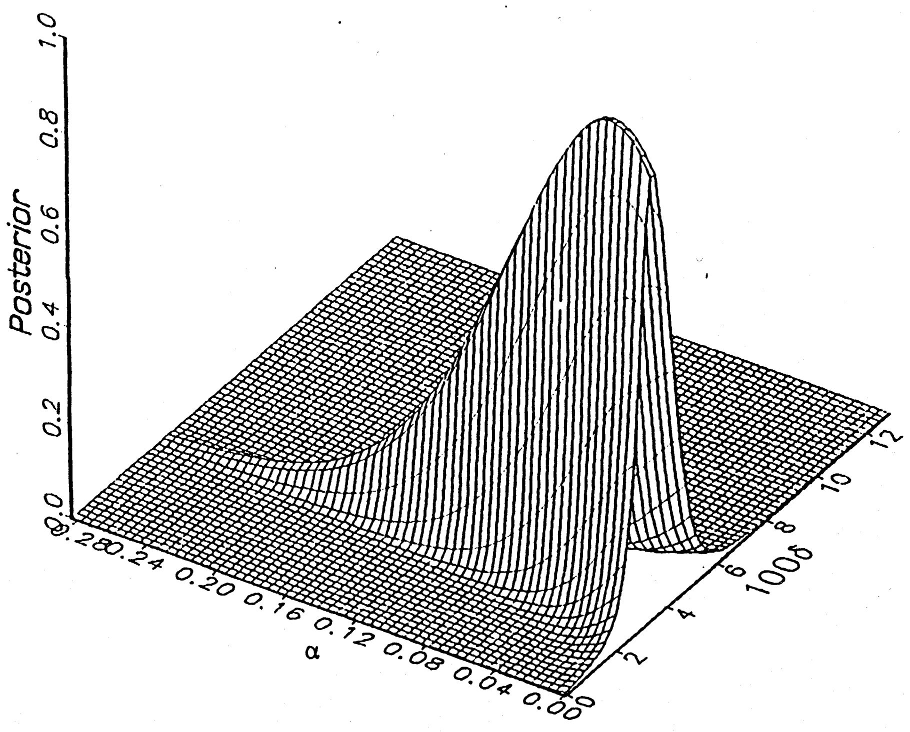


Figure 9B. Marginal posterior of α: Italy

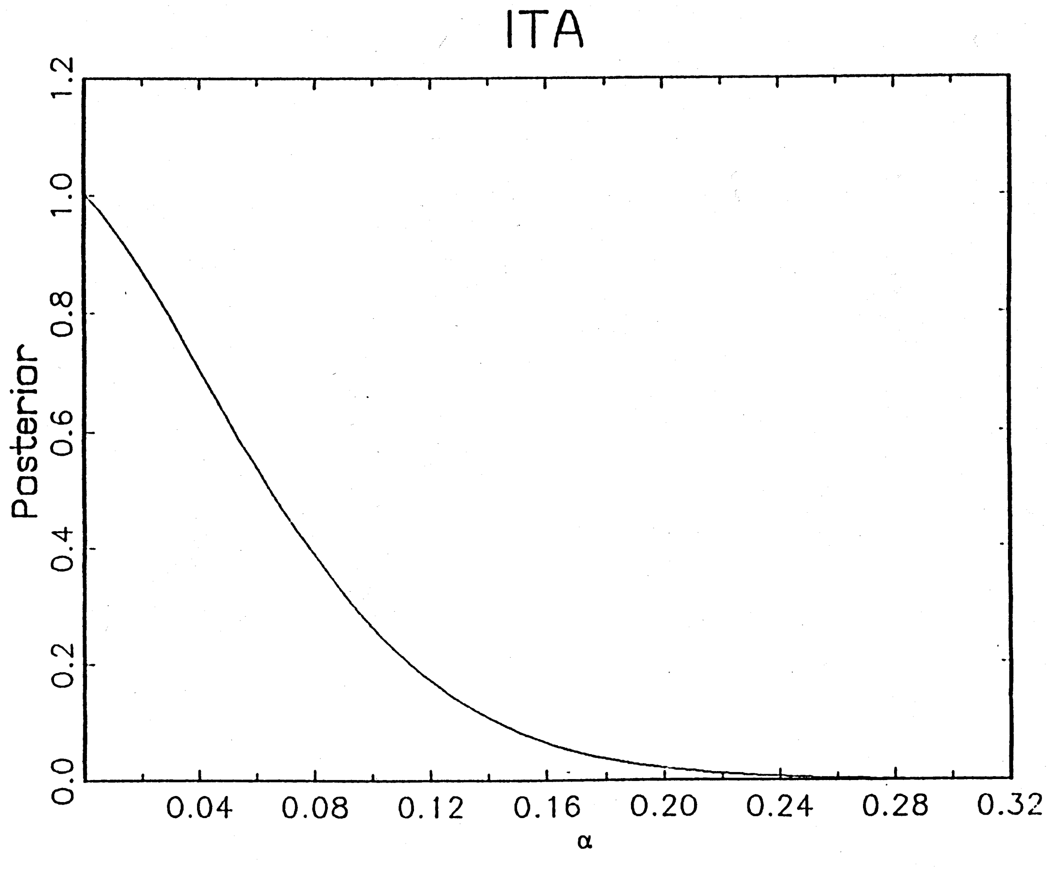


Figure 10A. Bivariate posterior of (α, δ): Netherlands

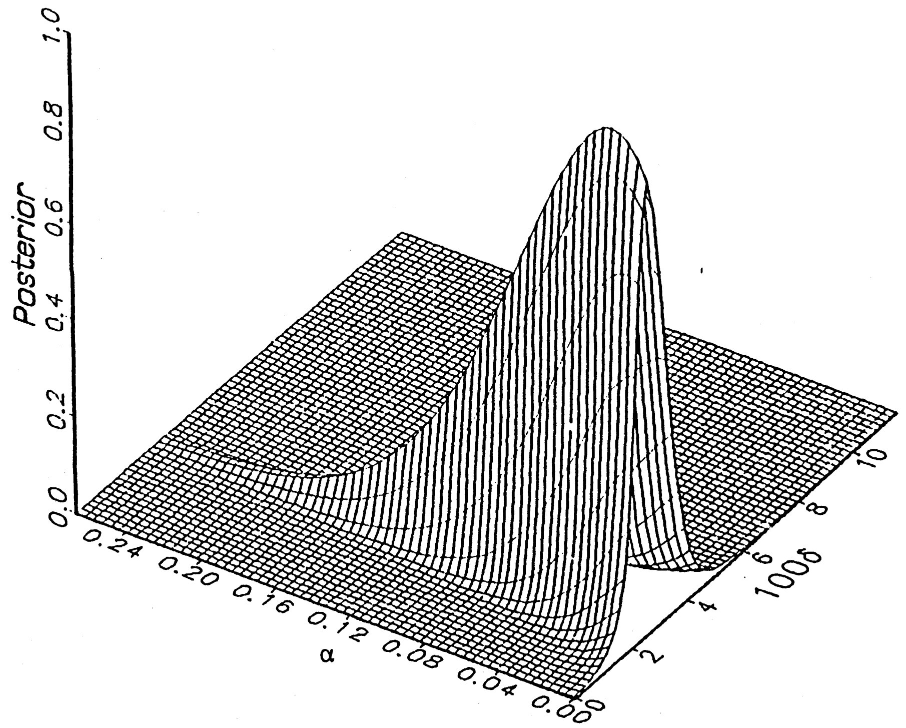


Figure 10B. Marginal posterior of α: Netherlands

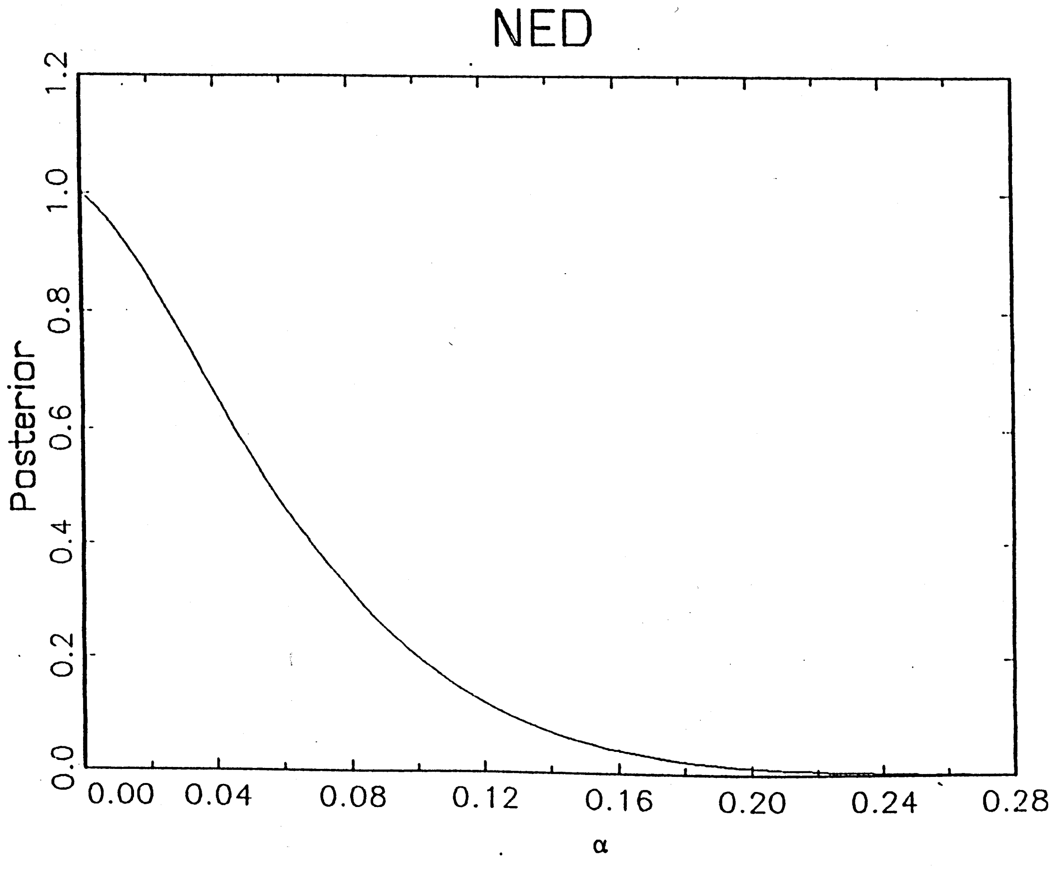


Figure 11A. Bivariate posterior of (α, δ): Norway

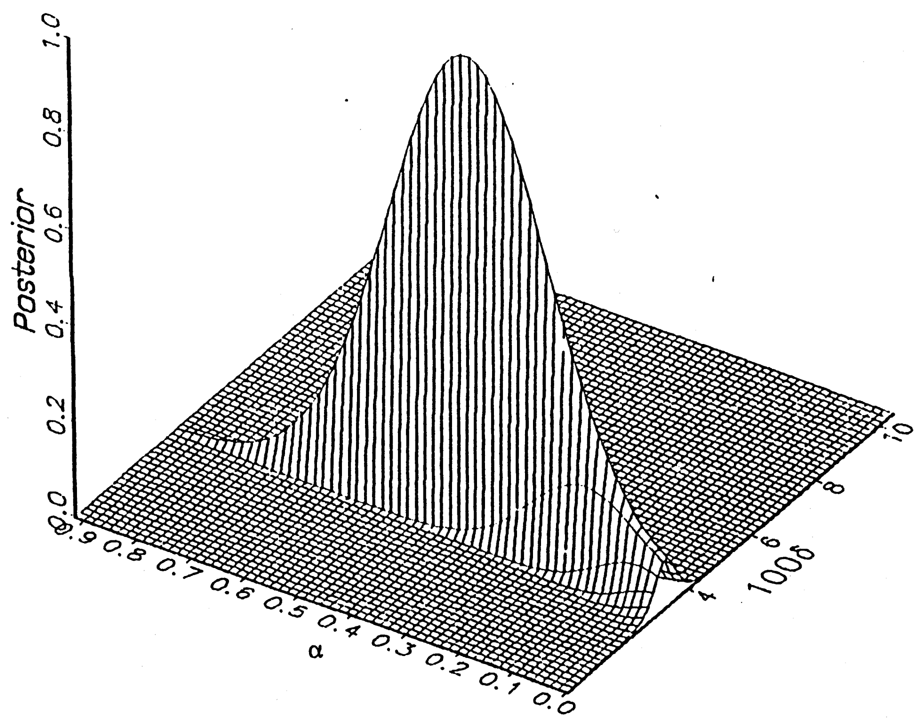


Figure 11B. Marginal posterior of α: Norway

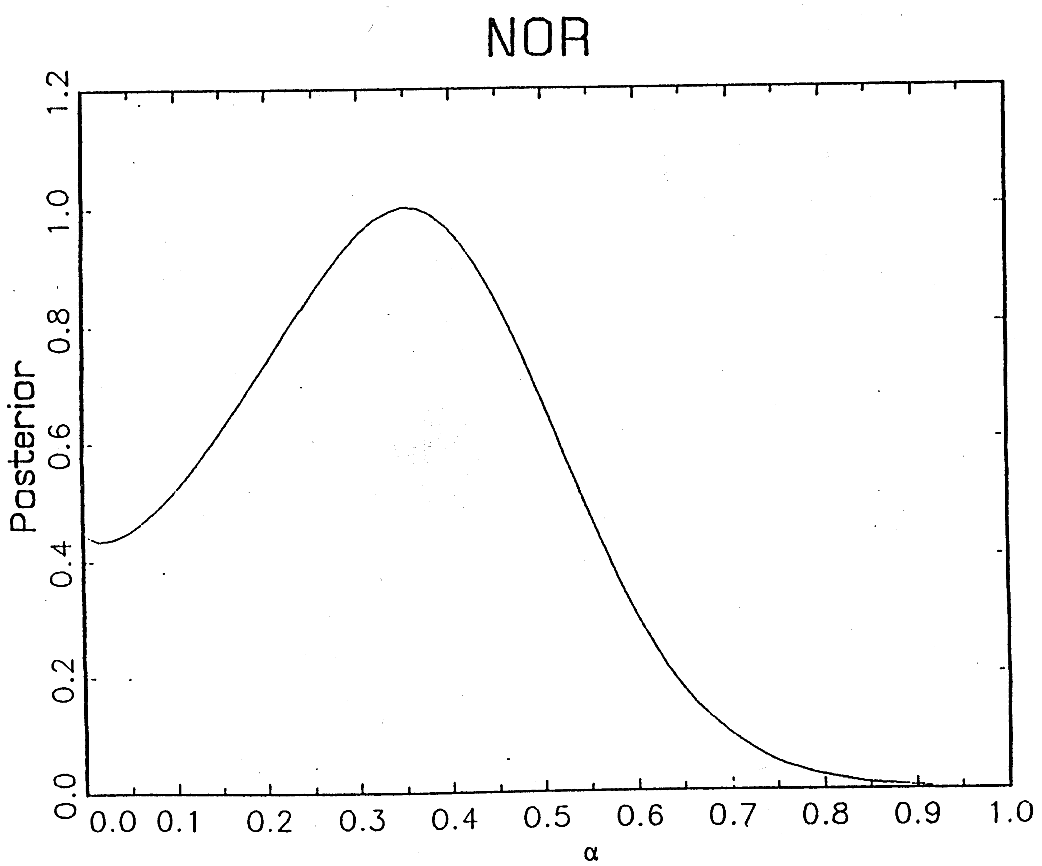


Figure 12A. Bivariate posterior of (α, δ): Sweden

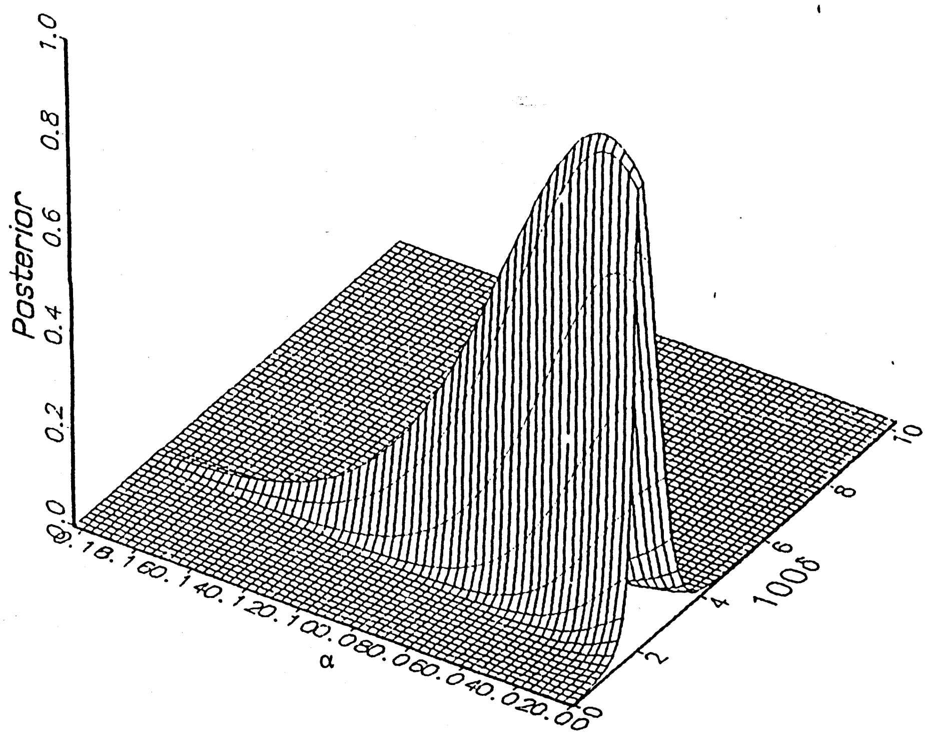


Figure 12B. Marginal posterior of α: Sweden

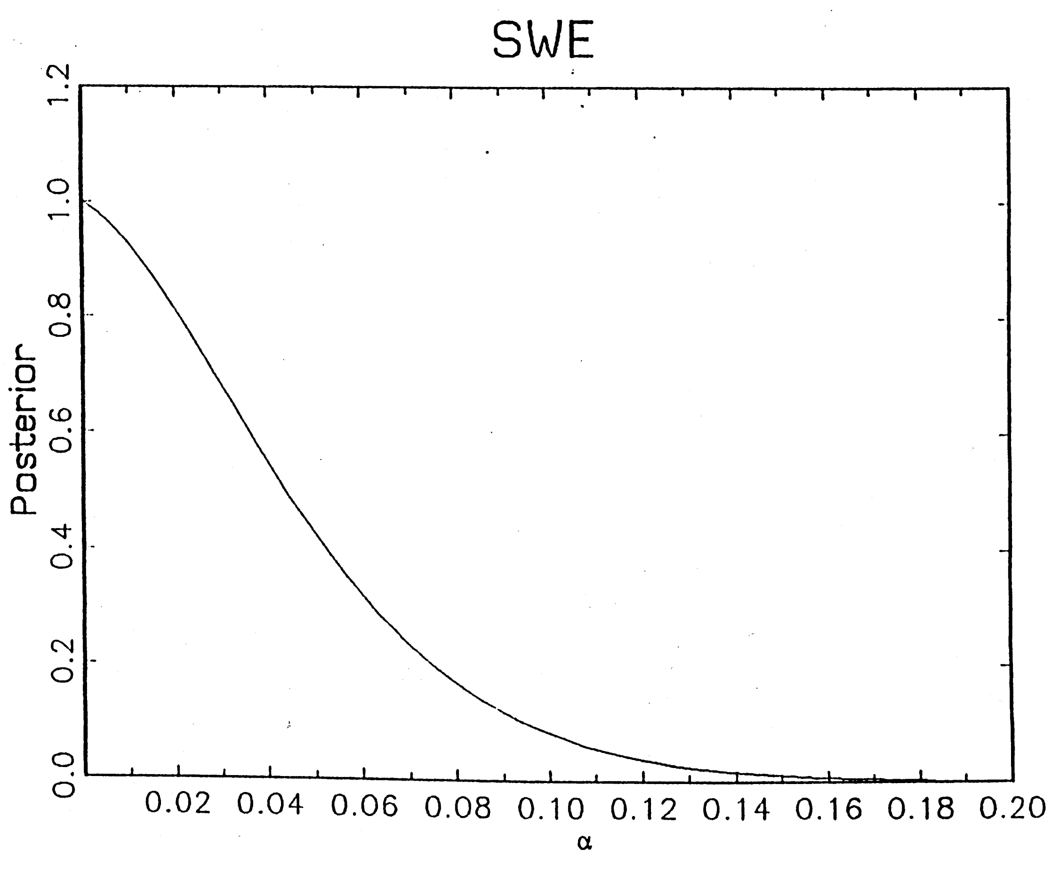


Figure 13A. Bivariate posterior of (α, δ): Switserland

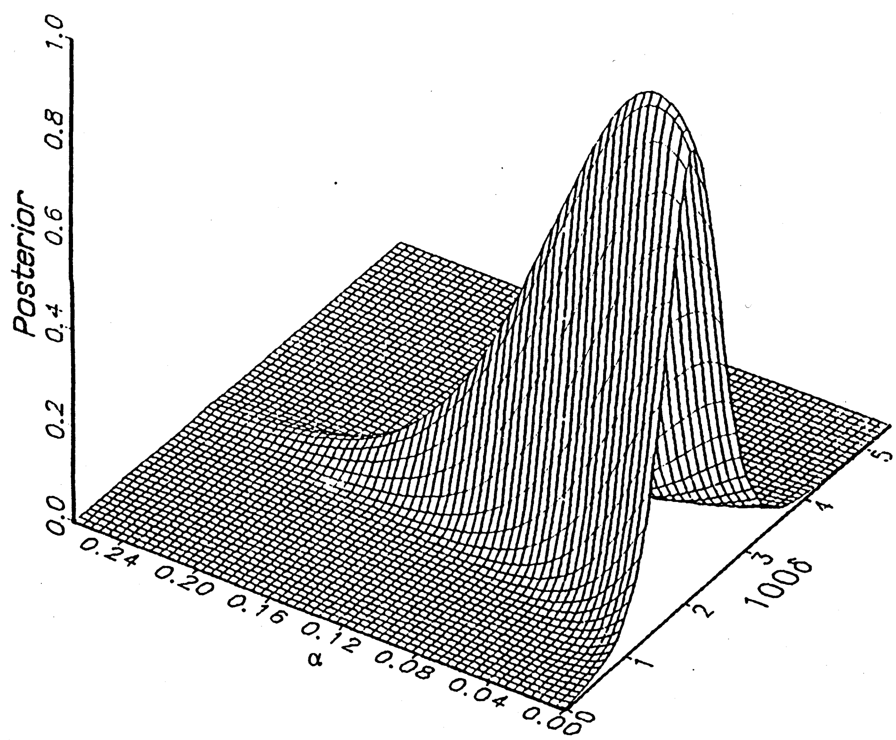


Figure 13B. Marginal posterior of α: Switserland

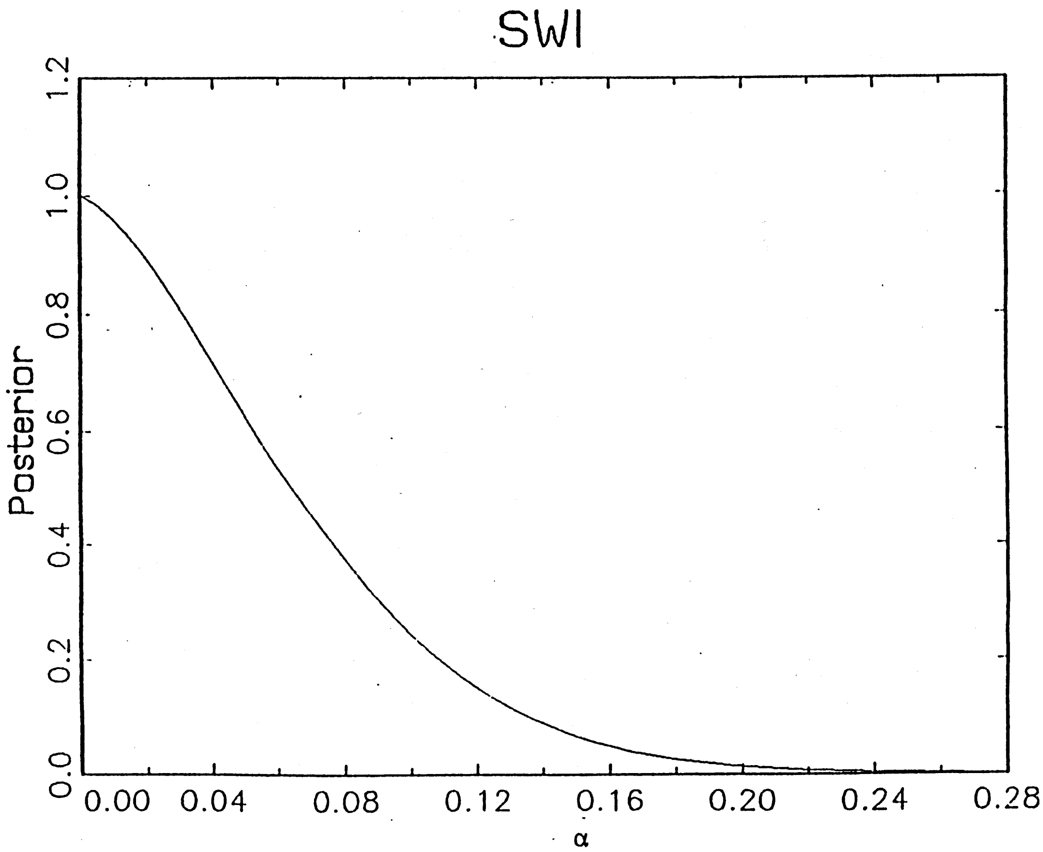


Figure 14A. Bivariate posterior of (α, δ): Canada

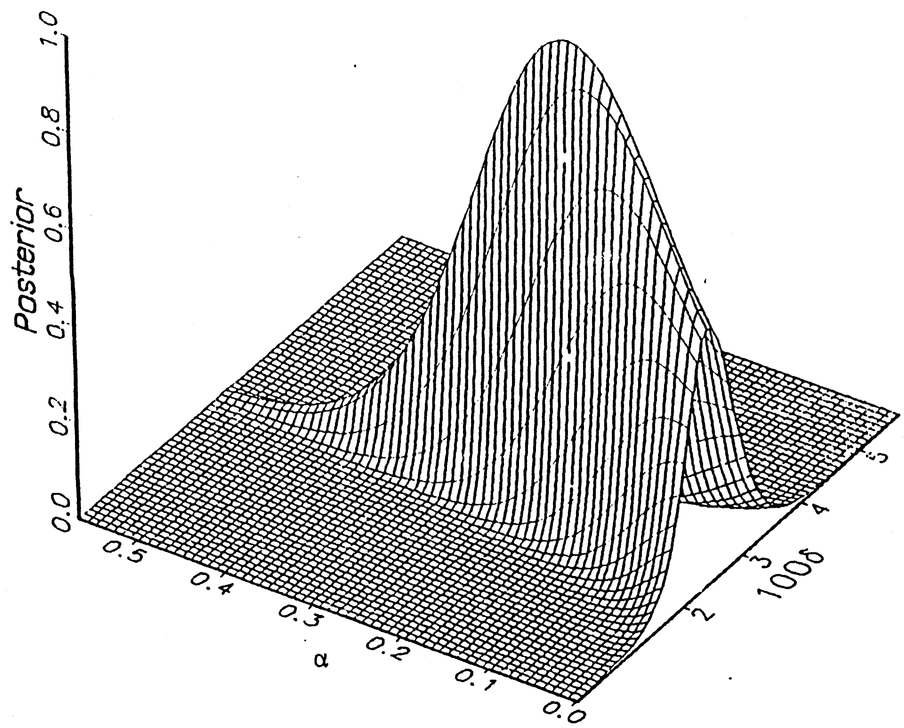


Figure 14B. Marginal posterior of α: Canada

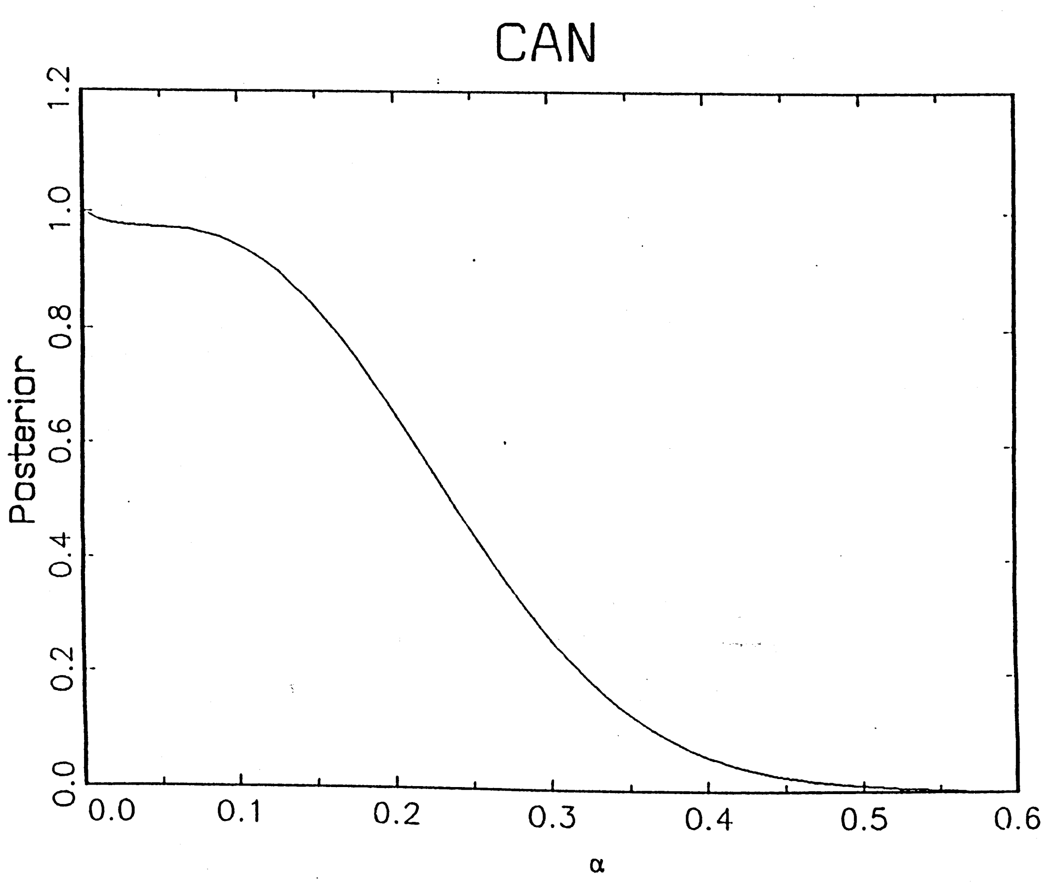


Figure 15A. Bivariate posterior of (α, δ): Japan

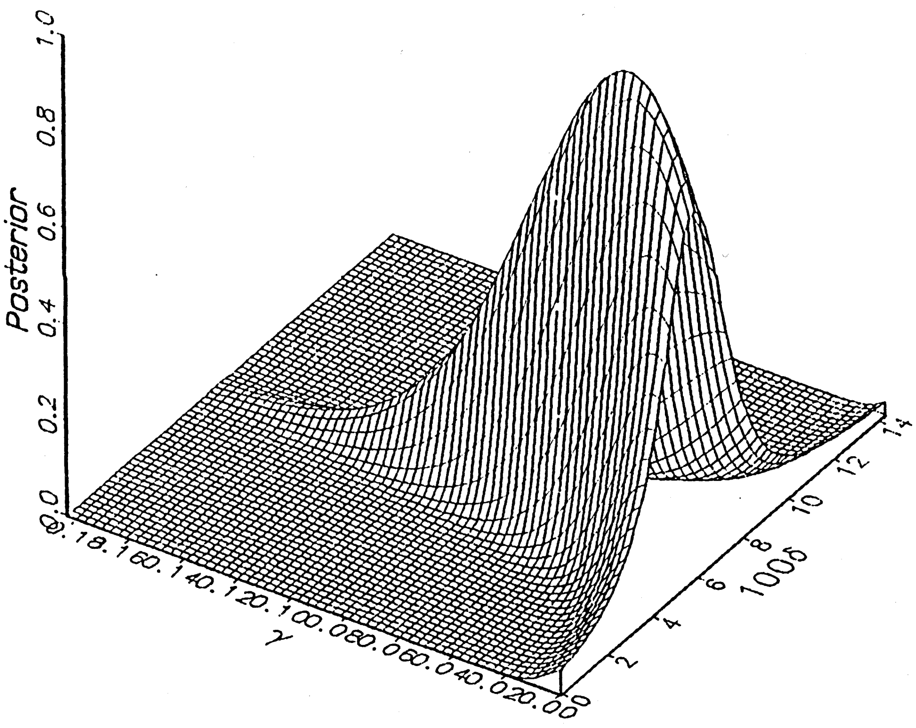


Figure 15B. Marginal posterior of α: Japan

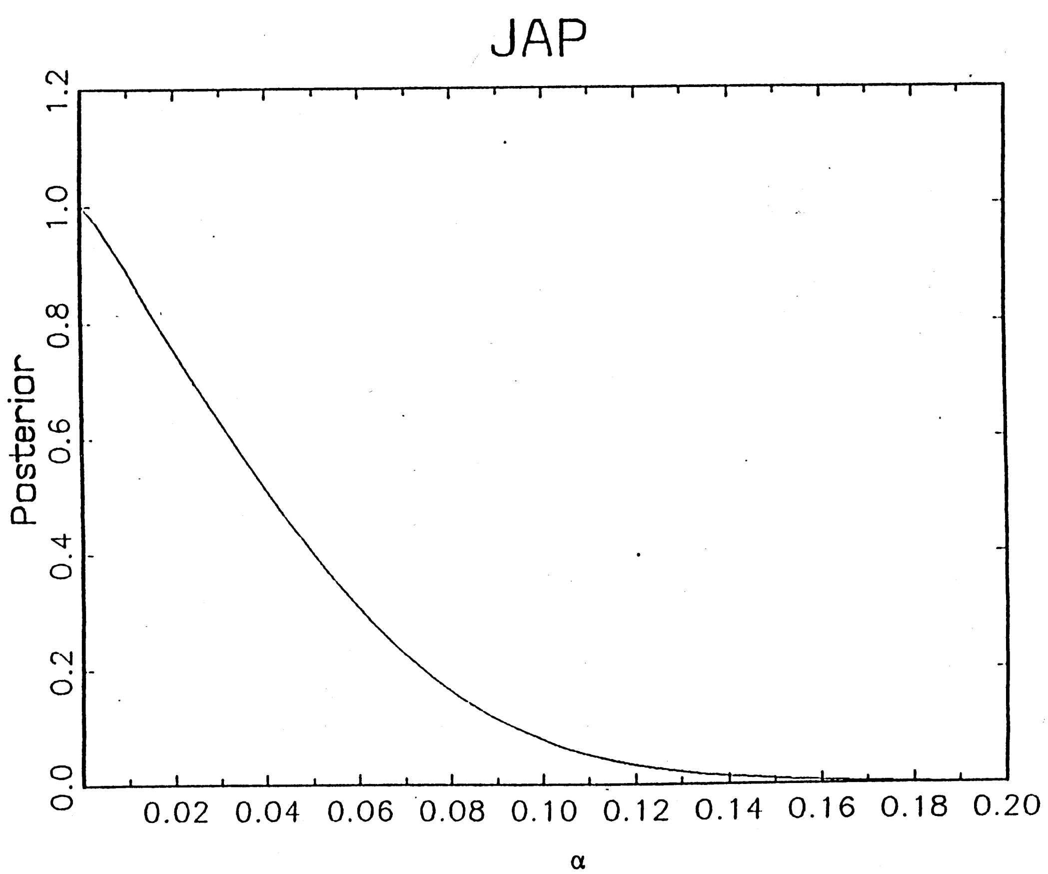


Figure 16A. Bivariate posterior of (α, δ): Ireland

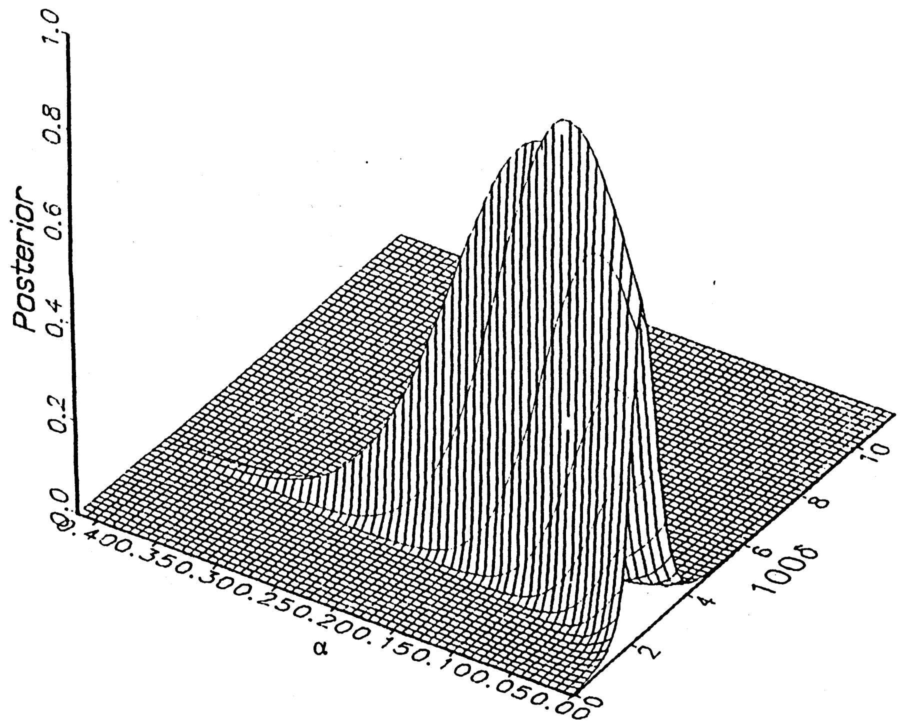


Figure 16B. Marginal posterior of α: Ireland

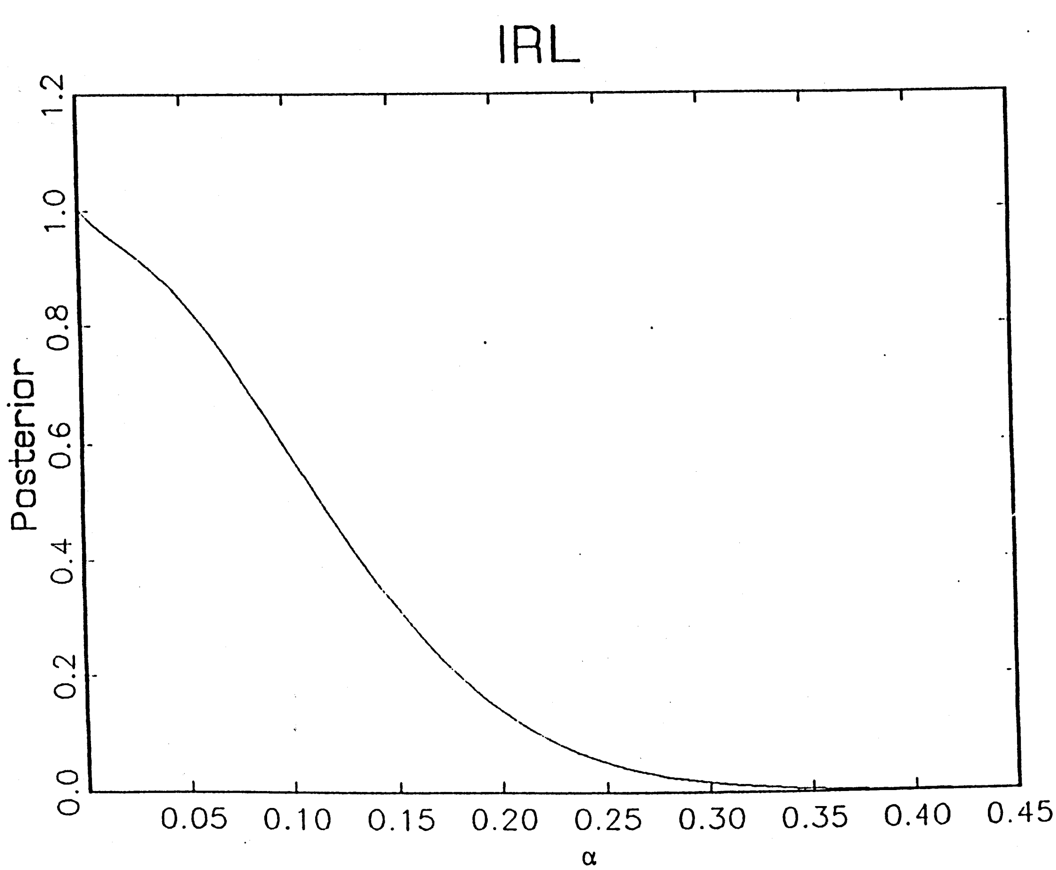


Figure 17A. Bivariate posterior of (α, δ): Australia

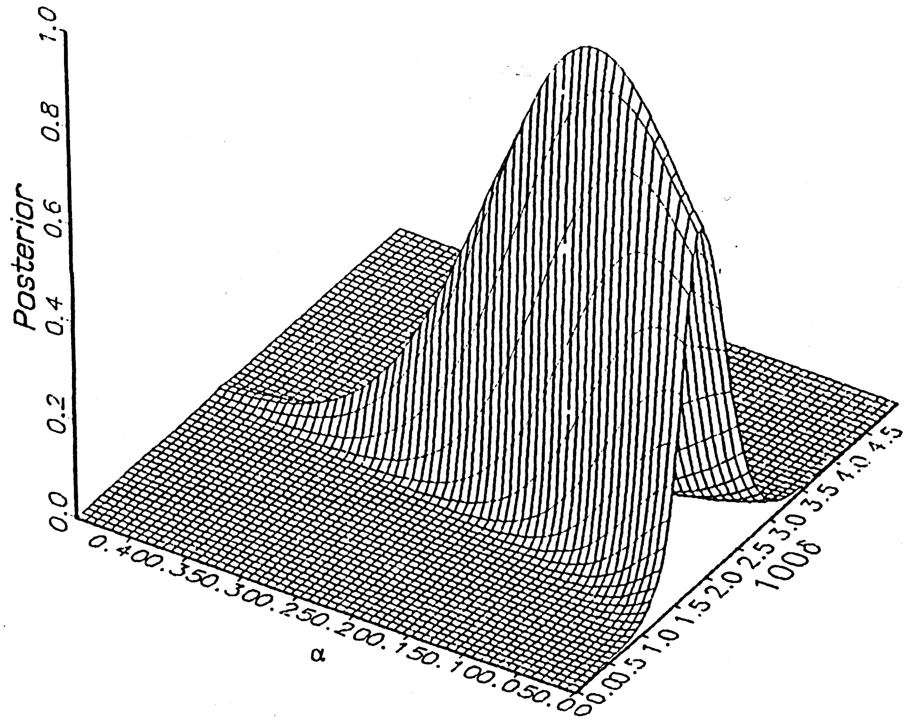
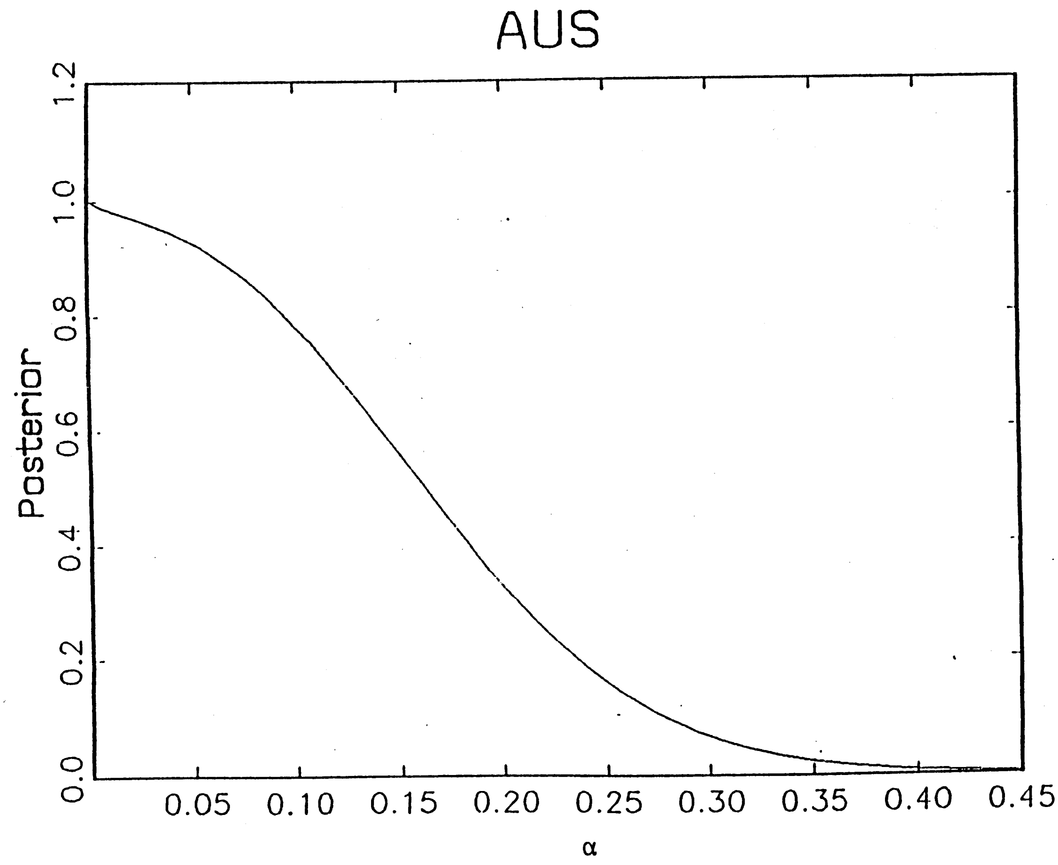


Figure 17B. Marginal posterior of α: Australia



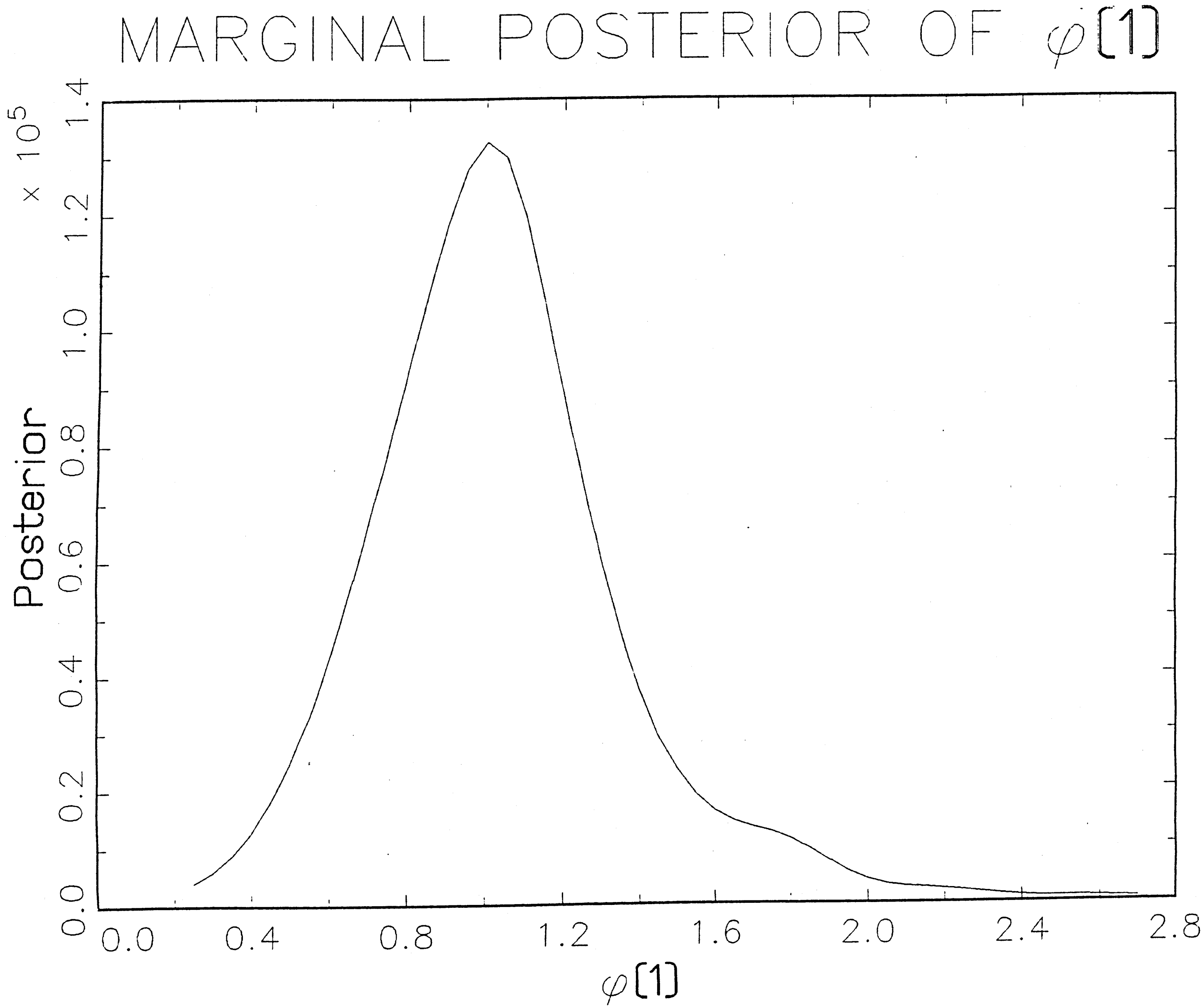

MARGINAL POSTERIOR OF φ[1]
Posterior
× 10^5
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
φ[1]

FIGURE 18

model from a conditional multivariate t distribution and compute finally a value of $\phi(1)$ as a function of all generated parameters. These steps are repeated 8000 times and the set of 8000 drawings is used to plot the marginal posterior of $\phi(1)$. For illustrative purposes we confine ourselves to the marginal posterior of $\phi(1)$ for the USA. A more detailed analysis of the posterior of $\phi(1)$ is a matter of further research.

5. FINAL REMARKS.

The paper has provided a Bayesian analysis of autoregressive time series around a linear trend, where one of the roots is close to or equal to unity. The application to time series of real GNP has provided some first empirical results. The analysis needs to be extended in several directions.

First, the sensitivity of the posterior odds test of the unit root hypothesis with respect to the specification of the model, in particular the presence of MA components, must be investigated. The inclusion of MA parameters overcomes the restrictions on the persistence measure $\phi(1)$ implied by a low order AR model. If $\phi(1)$ can take on values arbitrarily close to zero, it will be possible to obtain a meaningful marginal posterior density. The posterior density of $\phi(1)$, conditional on the unit root hypothesis, forms a natural complement to the posterior of μ if the time series is stationary. We anticipate a connection between these two posteriors. If the position of the trend is very ill-determined by the data, a unit root is likely, and a sharp estimate of $\phi(1)$ will obtain. Conversely, if $\phi(1)$ has a mode close to zero, the trend is not shifting around much, and stationarity will be a plausible hypothesis with a fixed intercept μ. Another point of research is the sensitivity with respect to a change in the fixed trendline.

Second, the empirical results have been obtained for relatively short time series. With few observations the finite sample Bayesian results differ much from the asymptotic classical results. It is therefore of interest to augment the dataset to longer time series, which for some countries are readily available.

Third, one has to investigate the sensitivity with respect to the prior specification. Our prior on (α, α^*) implies a prior on the roots λ_j $(j=1,\ldots,p)$. What is the effect on posterior odds and the marginal distributions of μ and δ if we specify a prior directly on the roots? A related issue is the specification of a prior on $\phi(1)$, treating persistence as a parameter of interest.

REFERENCES

Beveridge, S., and C.R. Nelson [1981], A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle', *Journal of Monetary Economics*, 7, 151-174.

Blanchard, O.J., and S. Fischer [1989], *Lectures on Macroeconomics*, MIT Press (Cambridge).

Campbell, J.Y., and N.G. Mankiw, [1987], Permanent and transitory components in macroeconomic fluctuations, *American Economic Review*, 77, 111-117.

Cochrane, J.H., [1988], How big is the random walk in GNP?, *Journal of Political Economy*, 96, 893-920.

Cristiano, L.J., and M. Eichenbaum, [1990], Unit roots in real GNP: do we know and do we care?, *Carnegie-Rochester Conference Series*, 32, 7-62.

DeJong, D.N., and C.H. Whiteman [1989a], The temporal stability of dividends and stock prices: evidence from the likelihood function, University of Iowa working paper series 89-03.

DeJong, D.N., and C.H. Whiteman [1989b], Trends and random walks in mmacroeconomic time series: a reconsideration based on the likelihood principle, University of Iowa working paper series 89-04.

Diebold, F.X., and M. Nerlove [1988], Unit roots in economic time seris: a selective survey, Finance and Discussion Paper Series 49, Federal Reserve Board.

Durlauf, S.N., and P.C.B. Phillips [1988], Trends versus random walks in time series analysis, *Econometrica*, 56, 1333-1354.

Fuller, W.A., [1976], *Introduction to statistical time series*, Wiley (New York).

Leamer, E.E., [1978], *Specification Searches*, Wiley (New York).

Nelson, C.R., and C.I. Plosser [1982], Trends and random walks in macroeconomic time series: some evidence and implications, *Journal of Monetary Economics*, 10, 139-162.

Schotman, P.C., and H.K. van Dijk [1990], A Bayesian analysis of the unit root in real exchange rates, *Journal of Econometrics* (forthcoming).

Sims, C.A., [1988], Bayesian skepticism on unit root econometrics, *Journal of Economic Dynamics and Control*, 12, 463-474.

Watson, M., [1986], Univariate detrending methods with stochastic trends, *Journal of Monetary Economics*, 18, 17-33.

Zellner, A., [1971], *An introduction to Bayesian inference in econometrics*, Wiley (New York).