Cowles Foundation Paper 825

11.

Reprinted from Colin Hargreaves (ed.),
Macroeconomic Modelling of the Long Run, 1992

THE LONG-RUN AUSTRALIAN CONSUMPTION FUNCTION REEXAMINED: AN EMPIRICAL EXERCISE IN BAYESIAN INFERENCE

Peter C. B. Phillips [1]

1. INTRODUCTION

Most theoretical and empirical econometric work in the field of nonstationary time series has used classical statistical methods. However, interest has lately begun to shift to Bayesian alternatives. Many econometricians (e.g. Geweke, 1988, Leamer, 1978 and 1988, Poirier, 1988, Zellner, 1971) find the Bayesian paradigm well suited to problems of inference in econometrics. Some econometricians (Sims, 1988, Sims and Uhlig, 1988/1991, and DeJong and Whiteman, 1989, 1991) have even suggested that the Bayesian approach is actually superior to classical econometric methods in time-series applications. Most recently, the author (1991c, 1991d) has taken issue with this latter view and put forward a different perspective which argues for the use of impartial or objective Bayesian methods in time-series analysis, especially with respect to the empirical issue of trend determination. This perspective puts value on both classical and Bayesian approaches. Moreover, in joint work with W. Ploberger, the author has gone further in reconciling Bayesian and classical methods of time-series analysis. Phillips and Ploberger (1991) studies the impact of data conditioning, via the operation of the likelihood principle, on Bayesian analysis and provides a new class of Bayesian tests that includes a Bayes model posterior-odds test for the presence of a stochastic trend. This latter test has interesting asymptotic properties (unlike classical tests it is completely consistent in the sense that both type I and type II errors go to zero as the sample size tends to infinity); it has excellent performance in finite samples; and it is especially easy to apply in empirical research.

The present paper will provide an overview and empirical illustration of these new Bayesian methods of trend determination. Attention will focus on the use of model-based reference priors, as advocated in Phillips (1991d) and on the objective posterior-odds test developed in Phillips and Ploberger (1991). These methods will be used both for trend analyses of individual time series and for studying relations between series. The methods will be illustrated by an empirical application to the Australian consumption function which is the main subject of the paper. This application takes a new look at the consumption-income data studied by Hall and Trevor (1991) and seeks to resolve some of the puzzling empirical outcomes in their results concerning the presence of a cointegrating relationship for the nominal data but not for the real data.

2. OBJECTIVE BAYESIAN METHODS OF TIME-SERIES ANALYSIS

This section will provide a brief overview of recent work by the author (1991d and e) and joint work with Werner Ploberger (1991) and Eric Zivot (1991) concerned with the development of objective Bayesian methods of analysing time series. These papers set out to provide an objective correlative for more traditional subjective Bayesian methods. Two distinguishing features of Bayesian analysis are addressed in this work: priors and data conditioning. These will be considered in turn.

2.1. Objective Priors and Bayes Confidence Sets

Clearly, the input of information via the prior distribution, whether it is deliberate or unwitting, is a major reason for potential divergence between classical and Bayesian statistical analyses. A conventional mechanism for achieving an impartial or objective analysis is to use a flat (or diffuse) prior on the regression coefficients. In the linear regression model with fixed regressors this approach leads to Bayes confidence sets from the posterior distributions that are identical with classical confidence regions under Gaussian assumptions. The same is not true in time-series models. There are further complications in time-series models arising from the discontinuity in classical asymptotic theory between stationary and nonstationary models. No such discontinuity occurs in a Bayesian analysis because the approach is exact or approximate (depending on the numerical and analytic methods that are employed) for the given data and sample size (T). No consideration is given to ($T \rightarrow \infty$) limit theory in the derivation of Bayes confidence sets. In linear time-series models with Gaussian errors these confidence sets are actually identical to those that would apply for the conventional linear regression model (see Zellner, 1971, and Malinvaud, 1980, for expositions of traditional Bayesian analysis in the linear regression model).

To illustrate these background ideas consider the AR(1) model

$$y_t = \rho y_{t-1} + u_t, \qquad t = 1, \ldots, T \tag{1}$$

with $u_t \equiv \text{iid } N(0, \sigma^2)$. Conditioning on the initial value y_0, the Gaussian likelihood follows from the density

$$f(y|\rho, \sigma, y_0) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^T exp\left\{-\frac{1}{2\sigma^2}\sum_1^T (y_t - \rho y_{t-1})^2\right\} \tag{2}$$

Assuming a flat prior for (ρ, log σ) leads to the usual diffuse prior for (ρ, σ), viz.

$$\pi(\rho, \sigma) \propto \frac{1}{\sigma} \tag{3}$$

Bayesian analysis of (1) under this prior is identical to that of the linear regression model. The joint posterior distribution is

$$p(\rho, \sigma | y, y_0) \propto \left(\frac{1}{\sigma}\right)^{T+1} exp\left\{-\left(\frac{1}{2\sigma^2}\right)\left[m(\hat{u}) + (\rho - \hat{\rho})^2 m(y)\right]\right\} \tag{4}$$

where

$$\hat{\rho} = \sum y_t y_{t-1} / \sum y_{t-1}^2,$$

$$m(y) = \sum y_{t-1}^2$$

$$m(\hat{u}) = \sum \hat{u}_t^2$$

and

$$\hat{u}_t = y_t - \hat{\rho} y_{t-1}.$$

The marginal posterior for ρ is

$$p_F(\rho|y,y_0) \propto \left[m(\hat{u})+(\rho-\hat{\rho})^2 m(y)\right]^{-T/2} \tag{5}$$

This posterior for ρ (we use the affix *'F'* to signify that a flat prior is used) is a univariate t_{T-1} distribution. Observe that ρ is symmetrically distributed about the OLS estimate $\hat{\rho}$.

The prior (3) has in the past been taken to be 'uninformative' about ρ in the sense that ρ values in any two intervals of the real line of equal length are equally likely a priori. However, this flat prior actually ignores information that we have from the AR(1) model about the way ρ values affect sample behaviour. Sample behaviour is known to be very different for ρ values in different intervals especially between stationary ($|\rho| < 1$) and nonstationary ($|\rho| \geq 1$) regions, and this represents prior knowledge based on the postulated AR(1) model. In my (1991d) paper I argued that an objective prior for ρ in a model such as (1) that allows for stochastic nonstationarity should incorporate such model-based information. Thus, when $|\rho| \geq 1$ in (1) we anticipate confidence regions for the true value ρ to be tighter than when $|\rho| < 1$. This expectation should be represented in an objective prior on the autoregressive coefficient ρ. Thus, even though the true coefficient ρ is unknown, an objective prior on ρ will reflect the knowledge we have about the AR(1) model that, were $|\rho|$ to be large, the data would be much more informative about ρ. This generic model characteristic that confidence sets will be tighter when $|\rho|$ is large is neglected in a flat prior. In treating all values of ρ as equally likely, the flat prior unwittingly carries information that downweights large values of ρ. In so doing, Bayesian inference under a flat prior on ρ will be distorted by information that will bias the posterior towards stationary alternatives. In time-series models with deterministic trends it is therefore hardly surprising that Bayesian inference under flat priors strongly favours trend-stationary alternatives. The simulations in my (1991c) paper make these distortionary effects of flat priors quite evident.

The alternative approach that was suggested in my (1991d) paper is to use an information matrix prior. This achieves certain desirable invariance properties (such as posteriors being invariant to 1:1 transformations of the parameters) and encapsulates model characteristics such as anticipated tighter confidence sets for $|\rho|$ large. The idea of such priors goes back to Jeffreys (1946). Since then they have attracted considerable interest and come within a more general class of 'ignorance' priors that possess useful invariance properties and seek to represent the notion of 'knowing little' a priori. If we set

$$I_{\theta\theta} = -E\left\{\left(\partial^2 / \partial\theta\partial\theta'\right) log\left(f(x|\theta)\right)\right\}$$

for a family of probability densities $f(x|\theta)$, indexed by the parameter $\theta \in \Theta \subset R^k$ then Jeffreys' suggestion was the prior

$$\pi(\theta) \propto \left|I_{\theta\theta}\right|^{1/2} \tag{6}$$

For the model (1) and allowing ρ to take both stationary values, this approach leads to the prior

$$\pi(\rho,\theta) \propto \frac{1}{\sigma} I_{\rho\rho}^{1/2} \tag{7}$$

where

$$I_{\rho\rho} = \begin{cases} \dfrac{T}{1-\rho^2} - \dfrac{1}{1-\rho^2}\dfrac{1-\rho^{2T}}{1-\rho^2} + \left(\dfrac{y_0}{\sigma}\right)^2 \dfrac{1-\rho^{2T}}{1-\rho^2}, & \rho \neq 1 \\ \dfrac{T(T-1)}{2} + T\left(\dfrac{y_0}{\sigma}\right)^2, & \rho = 1 \end{cases}$$

which is continuous in ρ for $-\infty < \rho < \infty$. This prior depends on y_0, which is the given initialization of the model, and the sample size T. The latter dependence is especially important. The Jeffreys prior (7) recognizes the information content of the sample variance of the regression in this model, and it also recognizes that this information will grow as T increases and at a geometric rate when $\rho > 1$.

The marginal posterior for ρ is obtained by integrating the product of (2) and (7) with respect to σ. The (1991d) paper used a Laplace approximation to reduce this integral directly to

$$p_J(\rho|y) \propto \alpha_0(e)^{1/2}\left[m(\hat{u}) + (\rho - \hat{\rho})^2 m(y)\right]^{-T/2} \tag{8}$$

where the affix 'J' signifies the use of the Jeffreys prior (7) and where

$$\alpha_0(e)^{1/2} = \begin{cases} \dfrac{T}{1-\rho^2} - \dfrac{1}{1-\rho^2}\dfrac{1-\rho^{2T}}{1-\rho^2}, & \rho \neq 1 \\ \\ \dfrac{T(T-1)}{2}, & \rho = 1 \end{cases}$$

is the effective contribution of $I_{\rho\rho}^{1/2}$ to the posterior. Unlike (5), the posterior density (8) is not symmetric about $\hat{\rho}$. It has one mode close to $\hat{\rho}$ and, depending on the values of $\hat{\rho}$, $m(\hat{u})$ and $m(y)$, often has a second mode around some $\rho > 1$. As is shown in the (1991d) paper, posterior inferences that are based on (8) can be very different from those based on (5). In general, (8) is less subject to the downward bias that the F-posterior (5) inherits from the estimator on which it is centred.

Empirical models usually allow for a more complex dynamic structure than (1) and also for the possibility of deterministic trends. An extension of (1) that is popular in applied work is the model

$$y(t) = \mu + \beta t + \rho y_{t-1} + \sum_1^{k-1} \varphi_i \Delta y_{t-i} + u_t = \mu + \beta t + \rho y_{t-1} + \varphi' x_t + u_t \quad (9)$$

The parameter ρ in (9) is the long-run autoregressive coefficient and the key parameter in determining the long-run stochastic behaviour of y_t (note that the spectrum at the origin of detrended y_t depends only on ρ). The hypothesis

$$H_0: \rho = 1 \quad (10)$$

corresponds to the presence of a stochastic trend in y_t. To determine the support in the data for (10) by traditional Bayesian methods requires the use of the posterior-odds ratio in favour of (10) and this in turn requires the assignment of some prior mass to the sharp (or point null) hypothesis (10). Neither F-priors like (3) nor J-priors like (7) accommodate this possibility. Instead, some spike and slab prior is usually employed and it is difficult to find a suitable objective correlative in this approach. An alternative approach to 'objective' posterior-odds testing will be described in the following section.

Another possibility for assessing the empirical evidence in support of stochastic nonstationarity is to evaluate the posterior probability (or Bayes confidence) of the nonstationary set $\{\rho \geq 1\}$ using the marginal posterior of ρ. This approach is used in the (1991d) paper and was first employed by DeJong-Whiteman (1989, 1991) for a different parameterization of (9) in which ρ is replaced by the modulus of the largest root of the characteristic equation.

The computation of Bayes confidence sets involving ρ requires the assignment of a prior distribution for the parameters of (9) (i.e. k+2 coefficients and the error variance σ^2). The (1991d) paper uses an approximation to the Jeffreys information matrix prior that has the form

$$\pi(\rho,\mu,\beta,\varphi) \propto \sigma^{-k-2}\left\{\alpha_0(\rho) + \alpha_1(\rho,\mu,\beta)/\sigma^2\right\}^{1/2} \tag{11}$$

where

$$\alpha_0(\rho) = T\left(1-\rho^2\right) - \left(1-\rho^2\right)^{-2}\left(1-\rho^{2T}\right),$$

and

$$\alpha_1(\rho,\mu,\beta) = \sum_0^{t-1}\left[\mu(1-\rho)^{-1}\left(1-\rho^i\right) + \beta\left\{(1-\rho)^{-1}i - \rho(1-\rho)^{-2}\left(1-\rho^i\right)\right\}\right]^2$$

This leads to a posterior for ρ of the (approximate) form

$$p_J(\rho|y) \propto \alpha_0(e)^{1/2}\left[m(\hat{u}) + (\rho-\hat{\rho})^2 m_V(y)\right]^{-T/2} \tag{12}$$

where V is the matrix of observations of $\left(1,t,\Delta y_{t-1},\ldots,\Delta y_{t-k+1}\right)$, and in conventional regression notation

$$m_v(y) = y'_{-1}Q_V y_{-1}$$

$$Q_V = I - V(V'V)^{-1}V'$$

$$m(\hat{u}) = \sum_1^T \hat{u}_t^{\,2}$$

and

$$\hat{u}_t = y_t - \hat{\mu} - \hat{\beta}t - \hat{\rho}y_{t-1} - \sum_1^{k-t}\hat{\varphi}_i \Delta y_{t-i}$$

are the OLS residuals.

The posterior (12) has properties analogous to those of (8) in the simple AR(1) model. In empirical applications (12) often leads to results that are quite different from those obtained with a flat prior, which for the model (9) leads to the posterior

$$p_F(\rho|y) \propto \left[m(\hat{u}) + (\rho-\hat{\rho})^2 m_V(y)\right]^{-(T-k-1)/2} \tag{13}$$

analogous to (5) in the simple model.

One disadvantage of (11) is that the prior is flat for φ and thereby fails to account for interactions between the long-run parameter ρ and the transient dynamic parameters φ in determining the information matrix. Zivot and Phillips (1991) address this issue and derive a generalization of (11) that takes this dependence into account. Their prior has the general form

$$\pi(\rho,\mu,\beta,\varphi) \propto \sigma^{-3}\left|A_0(\rho,\varphi)+\sigma^{-2}A_1(\rho,\mu,\beta,\varphi)\right|^{1/2} \tag{14}$$

for certain matrices that depend on (ρ,μ,β,φ) and that can be calculated readily by recursion (the formulae are given in Section 2.3 of Zivot and Phillips, 1991). Zivot and Phillips show that direct use of (14) leads to an improper posterior for ρ and they therefore suggest a class of modified Jeffreys priors that are indexed by a single parameter ε. These ε-priors have the general form

$$\pi_\varepsilon(\rho,\varphi) \propto \left\{\prod_1^k a_{oii}(\rho,\varphi)\right\}^{1/2} exp\left\{-\rho^{2c_k(\varepsilon)}\right\} \tag{15}$$

where

$$c_k = -(4\varepsilon)^{-1} + (4\varepsilon)^{-1}\left\{1+4\varepsilon(Tk-d_k)\right\}^{1/2}$$

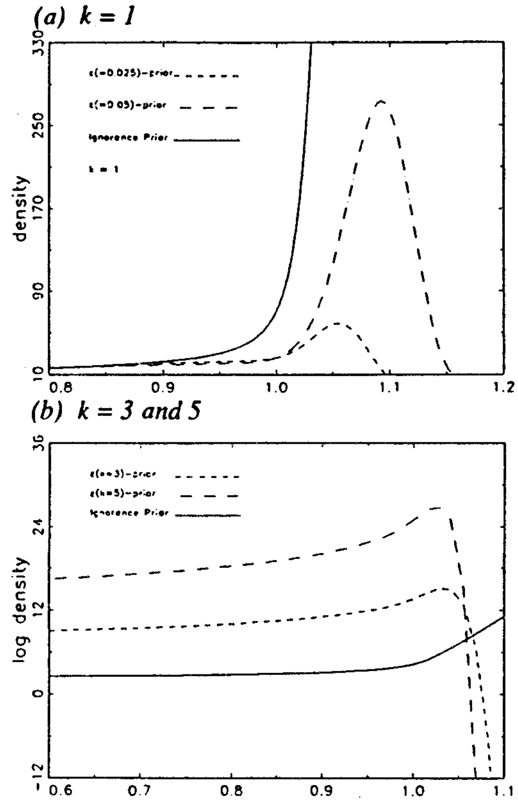
$$d_k = 1+(k+1)(k+2)/2$$

and a_{oii} is the i'th diagonal element of the matrix $A_0(\rho,\varphi)$ in (14).

The value of ε in (15) is set by the investigator. The function $c_k(\varepsilon)$ is determined as the solution of an optimization problem so that the modal value of the prior (15) occurs around $\rho = 1+\varepsilon$, after which the prior falls away rapidly. Thus, for priors in the family (15) low prior probability is attached to values of ρ much greater than $1+\varepsilon$. Suitable choices of ε in various practical applications (we mention an important one below) might be $\varepsilon = 0.001$, 0.01, 0.025, 0.05 and posteriors for a range of plausible values might be computed to indicate potential fragilities in inference. Note that (15) is also dependent on the lag parameter k and this parameter does influence the shape of the prior.

Figures 11.1(a) and (b) display these priors as a function of ρ for selected values of ε and k and for $\varphi = 0$ against the standard Jeffreys ignorance prior (7) with $\sigma = 1$ and $y_0 = 0$. As is apparent from Figure 11.1(a), larger values of ε delay the mode and, as ε increases, the prior approaches the ignorance prior from below when $k = 1$. Figure 11.1(b) reveals that for $k > 1$ (with $\varepsilon = 0.05$) the prior (15) attaches more weight

Figure 11.1: Prior densities for ρ.



to stationary alternatives than the ignorance prior but rises in the same way as ρ approaches unity from below. Again the ε-priors fall away rapidly from their peak as ρ increases beyond the modal value. The import of Figure 11.1(b) is that when $k > 1$ the ε-prior gives more weight to both stationary ρ and ρ in the vicinity of unity than the ignorance prior, but much less weight than the ignorance prior to values of ρ greatly in excess of unity.

One area of application where the ε-priors would seem to be especially useful is in the analysis of residuals. Models (1) and (9) may be used both for raw data and for residuals from regressions. For (9) in the latter case we typically set $\mu = \beta = 0$ and in this format the equation corresponds to the augmented Dickey-Fuller (ADF) regression suggested by Engle-Granger (1987) for testing a null of no cointegration (i.e. $\rho = 1$) by using the residuals from an OLS regression fitted according to the hypothesized relationship. By virtue of the construction of the data as residuals in this case, it seems sensible to attach less prior weight to extremely large (i.e. explosive) values of ρ. However, in order to properly allow for the possibility of no cointegration in the original data, for which $\rho = 1$ in the model for the residuals, we need to anticipate that the residual data may be more informative in this case so that an objective prior should increase as ρ approaches unity. The ε-priors capture these notions well and therefore seem like an appropriate family of priors when a general model like (9) is used for fitting residuals. Section 3 provides an empirical application of these priors in precisely this context.

Posterior distributions for the parameters of (9) are obtained in Zivot and Phillips using Laplace approximations for the general family of ε-priors (15). The form of the posterior for ρ, which will be our main interest in the present paper, is analogous to (12) above. From Theorem 1 of Zivot and Phillips we have the explicit expression

$$p_\varepsilon(\rho|y) \propto \left\{\prod_1^k a_{oii}\left(\rho, \tilde{\varphi}(\rho)\right)\right\}^{1/2} exp\left\{-\rho^{2c_k(\varepsilon)}\right\}\left[m(\hat{u}) + \left(\rho - \hat{\rho}\right)^2 m_V(y)\right]^{-(T-k-1)/2} \tag{16}$$

where $\tilde{\varphi}(\rho)$ is obtained from the appropriate element of $\tilde{\delta}' = \left(\tilde{\mu}, \tilde{\beta}, \tilde{\varphi}'\right)$ where

$$\tilde{\delta} = \hat{\delta} + (V'V)^{-1}V'y_{-1}\left(\hat{\rho} - \rho\right)$$

and $\left(\hat{\rho}, \hat{\delta}'\right)$ are the OLS estimates of the coefficients in (9).

Bayes confidence sets can be computed directly from the posterior density formulae (12), (13) and (16) using one dimensional numerical integration. Thus, posterior probabilities of the nonstationary set $\{\rho \geq 1\}$ are given by

$$\left.\begin{aligned} P_J(\rho \geq 1) &= \int_1^\infty p_J(\rho|y)d\rho \\ P_F(\rho \geq 1) &= \int_1^\infty p_F(\rho|y)d\rho \\ P_\varepsilon(\rho \geq 1) &= \int_1^\infty p_\varepsilon(\rho|y)d\rho \end{aligned}\right\} \tag{17}$$

These calculations typically take only a few seconds on a desktop 386 – 20. Thus, the approach is eminently feasible for empirical research.

2.2 Data Conditioning, the Bayesian Frame of Reference, and Objective Posterior Odds

Phillips and Ploberger (1991) show that classical statistical models like (1) and (9) take on a new statistical meaning when they are employed in a traditional Bayesian framework. In fact, the use of the likelihood principle that underlies all Bayesian inference involves data conditioning in the context of the specified likelihood function. This in turn implies the use of a 'Bayes model' associated with the historical time-series trajectory on which the Bayesian inference is based. The methods of the Phillips and Ploberger paper involve some stochastic process theory to elicit this 'change of geometry' or 'frame of reference' as it is referred to in that paper. The main ideas can be easily explained in the context of the AR(1) model (1).

With reference to (1), let $h = \rho - 1$, $\sigma^2 = 1$, P_n^ρ be the probability measure of $Y_n = (y_t)_1^n$ and $P_n = P_n^1$ (corresponding to the random walk case). Then the likelihood function given Y_n can be written in terms of h as

$$
\begin{aligned}
L_n &= dP_n^\rho / dP_n \\
&= exp\left\{(1/2)\hat{h}_n^2 A_n\right\} exp\left\{-(1/2)\left(\rho - \hat{\rho}_n\right)^2 A_n\right\} \\
&= exp\left\{(1/2)\hat{h}_n^2 A_n\right\} exp\left\{-(1/2)\left(h - \hat{h}_n\right)^2 A_n\right\} \qquad (18)
\end{aligned}
$$

where $\hat{h}_n = \hat{\rho}_n - 1$, $\hat{\rho}_n$ is the OLS/MLE of ρ based on Y_n and $A_n = \Sigma_1^n y_{t-1}^2$. Equation (18) is just the likelihood function (2) given earlier, standardized by its value at $\rho = 1$ (which corresponds to the reference measure). From (18) it is apparent that only the second factor for L_n is important for likelihood-based inference about ρ. This factor produces the symmetric Gaussian shape of the likelihood about $\hat{\rho}_n$, and hence that of the posterior based on a flat prior for ρ (cf. $p_F(\rho|y)$ as given in (5)). Note that L_n may be written in a more revealing manner as follows:

$$
L_n = \left[A_n^{-1/2} exp\left\{(1/2)\hat{h}_n^2 A_n\right\}\right] N\left(\hat{\rho}_n, A_n^{-1}\right) \qquad (19)
$$

$$
\propto N\left(\hat{\rho}_n, A_n^{-1}\right) \qquad (20)
$$

In (20) we get the Gaussian posterior for ρ about $\hat{\rho}_n$ that applies under a flat prior. Note from (19) and (20) that deviations of ρ from $\hat{\rho}_n$ are measured in units that are determined by $A_n = \Sigma_1^n y_{t-1}^2$. This changes the geometry of inference. Conditional on A_n the likelihood for ρ and the posterior is Gaussian, but the conditioning under which this is true is not innocuous. The passage to the posterior via the proportionality sign that appears in (20) eliminates the first factor of (19). This factor is data dependent only and does not figure in traditional Bayesian inference (it is absorbed into the proportionality sign). But, as Phillips and Ploberger (1991) shows, this factor changes the reference measure from P_n (the measure of the unit root model) to a conditional Bayes model measure in which $\hat{\rho}_n$ figures prominently, viz. the measure for the model

$$y_{n+1} = \hat{\rho}_n y_n + u_{n+1} \tag{21}$$

Writing the proportionality factor in (19) as

$$M_n = A_n^{-1/2} exp\left\{(1/2)\hat{h}_n^2 A_n\right\} \tag{22}$$

we can show that M_n is a local L_2 martingale (in fact, M_n satisfies the martingale conditional expectation property $E_{n-1}M_n = M_{n-1}$). It may be regarded as a likelihood ratio (density) process, which we write as dQ_n^B / dP_n, and as such it defines the measure, Q_n^B, of the 'Bayes model' (21) in which the parameter $\hat{\rho}_n$ evolves according to the MLE from the latest available data. The measure Q_n^B is σ-finite but induces a probability measure for (21) as soon as we condition on minimal information for the construction of $\hat{\rho}_n$ (clearly at $n = 0$ there is no data to construct $\hat{\rho}_n$ and thus the model and measure Q_n^B are undefined without suitable initialization). Phillips and Ploberger show how to use (22) for 'Bayes model' likelihood ratio (BLR) tests and posterior-odds tests and they develop an asymptotic theory for likelihood-based inference along these lines under general regularity conditions.

For the purpose of our subsequent empirical application it is the Phillips and Ploberger (BLR) posterior-odds test that is most useful. We illustrate the form this test takes in the case of the AR(1) model (1) and the AR(k) model (9). Let us start with the AR(1). Suppose we wish to test the point null $H_0{:}\rho=1$, i.e. (10) above. In the Phillips and Ploberger geometry the alternative 'Bayes model' is the evolving parameter model (21), which we call model B. The Bayes model measure of B is denoted Q_n^B in Phillips and Ploberger notation and the Bayes model posterior-odds criterion is based on the likelihood ratio dQ_n^B / dP_n which is the Radon-Nikodym (RN) derivative of Q_n^B with respect to the reference measure P_n for the random walk that applies under H_1. When $\sigma^2 = 1$ is known, this RN derivative is simply (22) as given above. When σ^2 is to be estimated, we use

$$\frac{dQ_n^B}{dP_n} = \left(\frac{A_n}{\hat{\sigma}_n^2}\right)^{-1/2} exp\left\{\frac{1}{2}\hat{h}_n^2 \frac{A_n}{\hat{\sigma}_n^2}\right\} \tag{23}$$

where $\hat{\sigma}_n^2$ is the usual OLS estimator of σ^2. The decision for testing H_0 against B is then simply

$$if \ \frac{dQ_n^B}{dP_n} > \frac{\pi_1}{\pi_\rho} \ decide \ in \ favour \ of \ model \ B \ (i.e. \ (21))$$

$$if \ \frac{dQ_n^B}{dP_n} < \frac{\pi_1}{\pi_\rho} \ decide \ in \ favour \ of \ H_1 \ (i.e. \ model \ (1) \ with \ \rho = 1)$$

where π_1 / π_ρ is the prior-odds ratio. Setting $\pi_1 / \pi_\rho = 1$ then leads to a very simple test criterion for the presence of a unit root in the model (1).

This idea is generalized to more complex models like (9). In this case it is helpful to rewrite (9) as

$$\Delta y_t = hy_{t-1} + \sum\nolimits_1^{k-1} \varphi_i \Delta y_{t-i} + \mu + \beta t + u_t \tag{24}$$

with $u_t \equiv$ iid $N(0,\sigma^2)$. Suppose we wish to employ a Phillips and Ploberger (BLR) posterior-odds test of $h = 0$, i.e. a unit root in (24). This is achieved by using the RN derivative of the Bayes model measure Q_n^B (i.e. h unrestricted) with respect to $Q_n^{B_0}$ (i.e. h restricted to $h = 0$). Calculations in Phillips and Ploberger (1991b) show that

$$\frac{dQ_n^{B_h}}{dQ_n^{B_0}} = \frac{exp\left\{(1/2)\hat{h}_n^2 y'_{-1} Q_V y_{-1} / \hat{\sigma}_n^2\right\}}{\left\{y'_{-1} Q_V y_{-1} / \hat{\sigma}_n^2\right\}^{1/2}} \tag{25}$$

where $\hat{h}_n = \hat{\rho}_n - 1$ is the MLE/OLS estimator of h in (24) and $\hat{\sigma}_n^2$ is the corresponding estimator of σ^2. Our decision rule with equal prior odds is then simply

$$if\ \frac{dQ_n^{B_h}}{dQ_n^{B_0}} < 1\ \textit{decide in favour of the hypothesis}\ H_0\text{: } h = 0\ \textit{in}\ (24)$$

$$(\textit{i.e. model}\ (9)\ \textit{with}\ \rho = 1) \qquad (26)$$

Tests based on the criterion (25) will be applied extensively in our empirical work that follows. We shall use the test for evaluating the evidence in favour of stochastic trends in the raw data and for testing whether there is cointegration in the time series by using regression residuals.

3. THE AGGREGATE AUSTRALIAN CONSUMPTION FUNCTION

This section reports an empirical application of the Bayesian methodology just described to aggregate Australian macroeconomic data. The Bayesian methods are used in conjunction with classical statistical procedures. This facilitates comparisons of the empirical results and highlights the ways in which the different methods elicit information from the data. Specifically, this application is to Australian data on household disposable income, private consumption expenditure, inflation and liquid assets. Our main focus of attention is the long-run form of the Australian consumption function.

3.1 Earlier Work and Empirical Puzzles

There have, of course, been many previous studies of the Australian consumption function such as those, in chronological order, by Arndt and Cameron (1957), Smyth and McMahon (1972), Freebairn (1976), William (1979), Anstie, Gray and Pagan (1983), Johnson (1983), McKibbon and Richards (1988) and Hall and Trevor (1991). Of these previous studies the most relevant for our present purposes is the work of Anstie, Gray and Pagan (1983) and Hall and Trevor (1991). Like the Hall and Trevor study, our work will focus on the long-run cointegrating relations between private consumption and other macroeconomic variables. An interesting outcome of the Hall and Trevor study was the finding that the null hypothesis of no cointegration could not be rejected for real variables (aggregate real consumption and household disposable income) but could be rejected for nominal variables, leading them to postulate a nominal aggregate consumption relation. Moreover, since their estimates of the long-run nominal relationship led to a unitary income elasticity, the empirical outcome of cointegration for nominal variables and no cointegration for real variables would seem prima facie to be inconsistent. Hall and Trevor advance several possible explanations of this inconsistency but do not reach any conclusion. One of the objectives of the present paper is to look further into this somewhat puzzling result.

One important explanation of the puzzle that is of economic as well as statistical interest is that inflation plays an important role in the consumption relation. In seeking to explain the rise in the Australian savings ratio during the 1970s, Anstie, Gray and Pagan (1983) suggested that household disposable income series should be adjusted to take into account the effects of inflation on wealth. According to their view real 'economic' income corresponds to income adjusted to leave real wealth intact. After making such an adjustment to income, Anstie, Gray and Pagan found that the empirical savings ratio stabilized. The idea that measured income should be adjusted downwards to account for inflation tax on nominal assets was also put forward by Hendry and von Ungern-Sternberg (1981) (hereafter, HVS). According to HVS,

> ...as inflation increases, nominal interest rates tend to rise, thereby increasing the interest component of Y. It seems appropriate to measure 'real income' as increasing in such a situation, since large nominal interest receipts are offset by capital losses on all monetary assets which are *not* being deducted from the income variable used (HVS, p. 245).

HVS go on to suggest the use in consumption relations of a real 'perceived income' measure Y^* that is designed to adjust real income Y for such capital losses as $\dot{p}L$ (where $\dot{p}$ is the rate of inflation and L is real liquid assets). Y^* is defined as

$$Y^* = Y - \beta\dot{p}L = Y(1 - \beta\dot{p}L/Y) \tag{27}$$

and the parameter β is introduced to allow for scale effects resulting from an inappropriate choice of $\dot{p}$ or L. The composite variable $\dot{p}L/Y$ may be regarded as a measure of relative capital loss (RCL).

3.2 The Data and Empirical Regularities

We shall attempt to explore the relevance of these ideas to the Australian consumption function. The data we use cover the periods 1959(3) –1988(4) and 1965(1) – 1988(4) and are described with the notation we

Table 11.1: Macroeconomic variable notation and descriptions

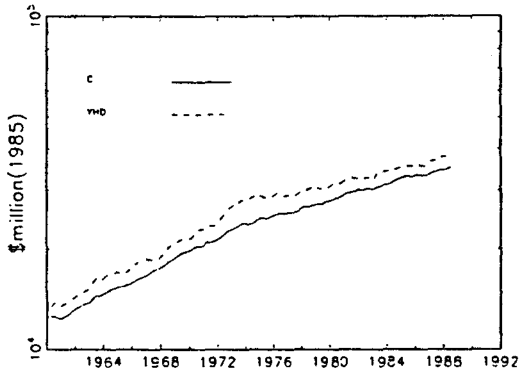
Variable	Description	Sample Period
C	aggregate real private final consumption expenditure	1959(3) – 1988(4)
YHD	real household disposable income (= $YHD/P)	1959(3) – 1988(4)
$C	aggregate nominal private final consumption expenditure	1959(3) – 1988(4)
$YHD	nominal household disposable income	1959(3) – 1988(4)
P	implicit deflator for aggregate private final consumption expenditure	1959(3) – 1988(4)
Inf_{-1}	inflation rate (= $\Delta \ln(P)$)	1959(4) – 1988(4)
Inf_{-4}	standardized annual inflation rate (= $\Delta_4 \ln(P)/4$)	1960(3) – 1988(4)
M3	real money stock (= $M3/P)	1965(1) – 1988(4)
RCL	relative capital loss (= Inf_{-4}*M3/YHD)	1965(1) – 1988(4)

employ in Table 11.1. All variables are seasonally adjusted and the constant price series are at average 1984/1985 prices.

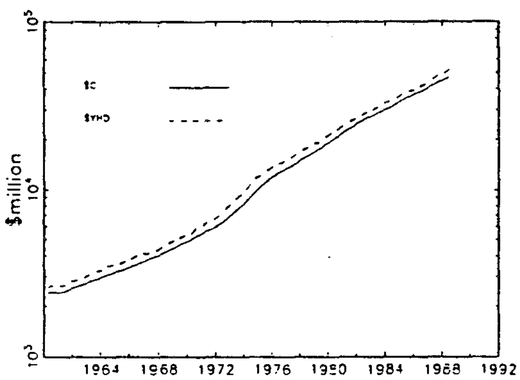
Figures 11.2(a) – (d) graph these series in various combinations to illustrate how the observed historical trajectories relate. Clearly C and YHD (in lag levels) move in consort over time, although the nominal series appear to move together more closely than the real series (however, note the difference in scale between Figures 11.2(a) and (b)). Figure 11.2(c) graphs these series (in levels form) against M3 and a scaled version of the relative capital loss variable, RCL. Apparently M3 shows some divergent behaviour with respect to YHD, especially towards the end of the sample period. The capital loss measure, RCL, has high sample variability. Note that RCL peaks over the period 1974 – 1978, which is precisely the interval when the empirical savings rate rises (as is

*Figure 11.2: Graphs of Australian real and nominal household disposable income and consumption, M3, INF_{-4}*M3/YHD and inflation rates.*

(a) YHD and C



(b) \$YHD and \$C



*(c) M3, YHD and Inf_{-4}*M3/YHD*

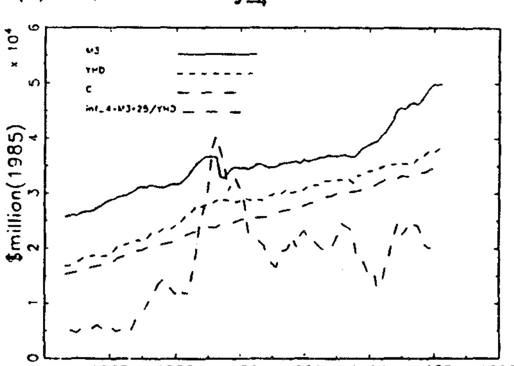
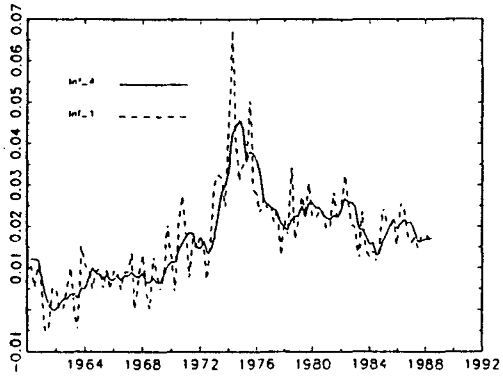


Figure 11.2: continued

(d) Inflation rates



apparent from the graphs of YHD and C). This suggests that RCL may have a useful role to play in a consumption function relating C and YHD.

Figure 11.2(d) graphs the two inflation variables Inf_{-1} and Inf_{-4}. Inf_{-1} is the quarterly inflation rate and Inf_{-4} is the average inflation rate over the past four quarters (i.e. the latest annual rate standardized to quarterly units). The rate Inf_{-4} is used because in the computation of 'perceived' measures like Y^* in (27) and the capital loss measure RCL the average prevailing inflation rate over the past year seems more appropriate than an instantaneous rate. HVS employ a two-year moving average of the quarterly rate in their empirical work (i.e. Inf_{-8} in our notation). We choose Inf_{-4} because it captures the essential idea of using longer-term inflation measures to assess capital loss and it involves less sample size reduction than Inf_{-8}.

3.3 Testing the Data for Nonstationarity

Bayes and classical tests for the presence of stochastic trends in the data were conducted. Table 11.2 reports results for the Phillips-Ploberger posterior-odds test, as given in equation (25). For all the series except Inf_{-1} and \$C the results are unambiguously supportive of the presence of a stochastic trend. For Inf_{-1} and \$C series there is uncertainty due to the variation in the outcomes for different lag lengths. The results show rejection of a unit root for Inf_{-1} when the lag length $k = 1, 2, 3$ and acceptance when $k = 4, 5, 6$. Looking at the series in Figure 11.2 it is apparent that Inf_{-1} has a choppy appearance that reduces the correlation at lag 1, yet also shows evidence of stochastic drift over the full period. This helps to explain the uncertainty in the outcome. The Inf_{-4} series retains the stochastic drift evident in Inf_{-1} but smoothing serves to eliminate the choppy character of the Inf_{-1} series. There is less uncertainty about $C. Only for $k = 1$ do the results reject a unit root in the series. For all other lag lengths there is strong support for the presence of a stochastic trend.

Bayesian posteriors support the results of Table 11.2. Figures 11.3(a) – (c) graph the Bayesian posteriors for the long-run autoregressive parameter ρ in equation (9) for different lag lengths (k = 1,3) for the two inflation series Inf_{-1}, Inf_{-4} and the aggregate series C for comparison purposes. In Figure 11.3(a) the uncertainty over the value of ρ for the Inf_{-1} series is manifested by the elongated second mode of the posterior based on the ignorance prior (hereafter *I*-posterior). By contrast for the Inf_{-4} series both the *I*-posterior (for k=1) and ε-posterior (k=3) give strong support to the presence of a stochastic trend. Only the flat prior posterior (hereafter *F*-posterior) suggests stationarity and these posteriors are known to suffer from substantial bias, as explained earlier in Section 2. Figure 11.3(c) shows that for the data series ln(C) the posterior evidence is unambiguously supportive of a stochastic trend, thereby concurring with the posterior-odds outcome given in Table 11.2. Similar results were obtained for the other aggregate series. Table 11.3 shows the Bayes posterior confidence associated with these figures for each of the three priors. Again the evidence in favour of stochastic nonstationarity for the series Inf_{-4} and ln(C) is very strong.

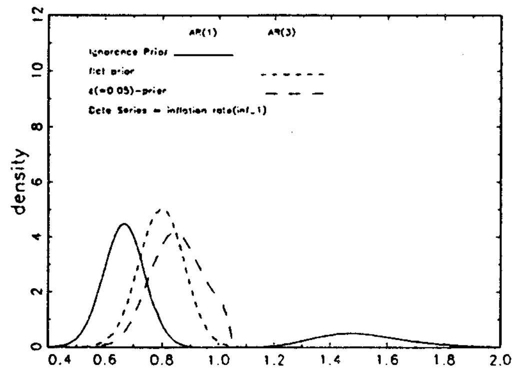
Table 11.2: Results of Phillips and Ploberger posterior-odds unit root tests for Australian macroeconomic data; values of the BLR criterion for different lag lengths

Variable	Lag Length					
	1	2	3	4	5	6
Trend degree = 1						
C	0.18	0.022	0.023	0.039	0.044	0.049
YHD	0.060	0.040	0.053	0.077	0.063	0.068
$C	39.010	0.133	0.048	0.029	0.033	0.049
$YHD	0.500	0.757	0.172	0.105	0.094	0.072
Inf_{-1}	5359.400	87.527	1.723	0.367	0.306	0.713
Inf_{-4}	0.058	0.165	0.317	0.696	0.169	1.147
Trend degree = 0						
Inf_{-1}	282.979	14.864	0.787	0.298	0.349	0.841
Inf_{-4}	0.045	0.089	0.133	0.312	0.153	0.771

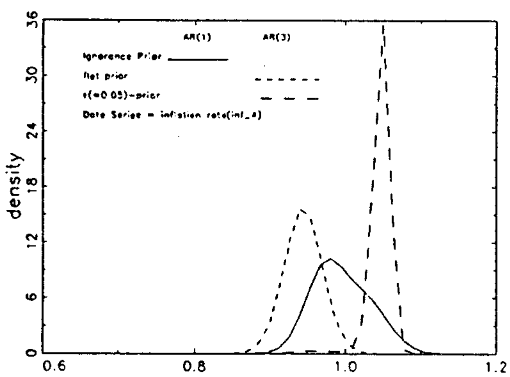
Legend:reject unit root if 'BLR > 0.1'

Figure 11.3: Bayesian posteriors for the long-run autoregressive parameter, ρ, for the inflation rates and real consumption

(a) Inf_{-1}



(b) Inf_{-4}



(c) Real Consumption

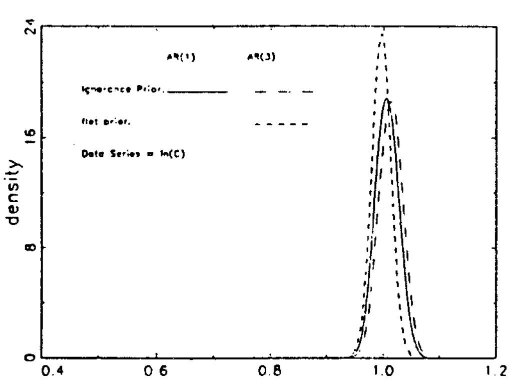


Table 11.3: Posterior probabilities of stochastic nonstationarity

Series	AR(1) + trend model		AR(3) + trend model	
	$P_I(\rho \geq 1)$	$P_F(\rho \geq 1)$	$P_\varepsilon(\rho \geq 1)$	$P_F(\rho \geq 1)$
Inf_{-1}	0.193	0.000	0.064	0.005
Inf_{-4}	0.433	0.113	0.980	0.015
ln(C)	0.646	0.356	0.756	0.453

Legend: P_I = ignorance posterior; P_ε = ε-posterior (ε=0.05); P_F = flat posterior

Table 11.4: Classical unit root tests

Test	lag	Inf_{-1}	Inf_{-4}	ln(C)	5%CV
ρ	3	0.842	0.935	0.983	
ADF	3	−1.902	−2.521	−0.671	−3.500
	5	−2.074	−2.822	−1.194	−3.500
	7	−2.546	−1.784	−1.371	−3.500
$Z(a)$	3	−34.577	−5.216	−1.048	−21.099
	5	−32.056	−6.727	−1.082	−21.099
	7	−33.709	−7.853	−1.255	−21.099
$Z(t)$	3	−4.577	−1.539	−0.495	−3.500
	5	−4.445	−1.767	−0.507	−3.500
	7	−4.532	−1.919	−0.567	−3.500

These outcomes can be compared with classical unit root tests. In Table 11.4, we provide results for the ADF, $Z(a)$ and $Z(t)$ tests applied to the same group of variables. Each of the $Z(a)$ and $Z(t)$ tests rejects the presence of a unit root for Inf_{-1} at both the 1 per cent and 5 per cent levels, while the ADF test does not reject the hypothesis at the 5 per cent. This conflict in outcomes mirrors the Bayesian results. By contrast all of the tests confirm (i.e. 'do not reject') the presence of a unit root for Inf_{-4} and ln(C), again corroborating the Bayesian evidence. Similar results to those of ln(C) were obtained for the ln(YHD) series and these are not reported.

3.4 Testing for Cointegration

Our first step is to analyse the aggregate consumption and disposable income data for cointegration. We examine relationships in both real and nominal variables of the form

$$\ln(C) = a + b \ln(YHD) + \text{error}, \tag{28}$$

and

$$\ln(\$C) = a + b \ln(\$YHD) + \text{error}. \tag{29}$$

Again, both Bayesian and classical methods are employed.

Table 11.5 gives the results of the posterior-odds test (26) applied to the residuals of the cointegrating regressions (28) and (29). Estimation of these equations was by OLS but the results given in the table are robust to alternative methods of estimation that would be asymptotically efficient if the variables were cointegrated (Section 3.5 below reports estimates from a range of different efficient methods for augmented regressions based on (28) and (29)). For each equation and for every lag selection except $k = 1$ the posterior-odds test favours the presence of a unit root in the residuals of (28) and (29). According to this Bayes test, therefore, the evidence does not support the hypothesis that aggregate consumption and household disposable income are cointegrated either in real or in nominal terms. Clearly, this conflicts with the earlier empirical finding of cointegration for the nominal variables \$C and \$YHD by Hall and Trevor (1991).

Table 11.5: Results of Phillips and Ploberger posterior odds tests for cointegration; values of the BLR criterion

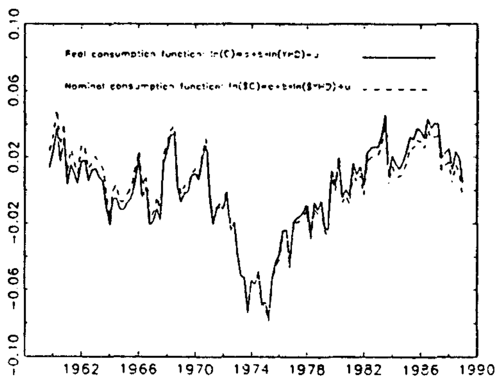
Consumption Function	Lag Length in Residual Regression					
	1	2	3	4	5	6
(28) real	2.077	0.442	0.624	0.436	0.228	0.157
(29) nominal	3.060	0.631	0.925	0.616	0.352	0.204

Legend: reject unit root in residuals if 'BLR>1.0'
(i.e. accept the presence of cointegration if BLR>1.0)

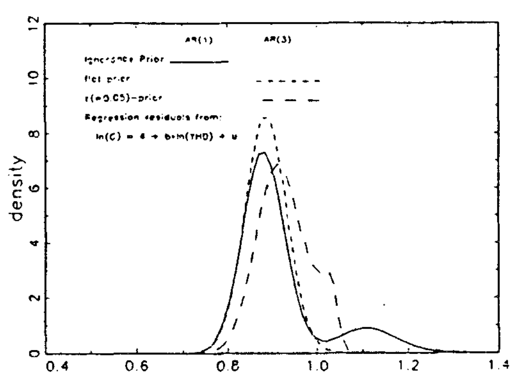
Figure 11.4(a) gives plots of the residuals from the two regressions (28) and (29). These graphs show that the behaviour of the residuals from these two regressions is very similar over the sample period. The test outcomes from Table 11.5 therefore seem consistent with the visual

Figure 11.4: Graphs of the residuals for regressions (28) and (29) and their posterior priors for ρ.

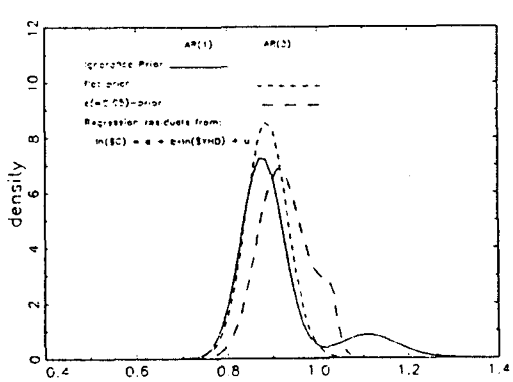
(a) Cointegrating residuals for (28) and (29)



(b) Posterior densities for ρ on equation (28)'s residuals



(c) Posterior densities for ρ on equation (29)'s residuals



evidence. Thus, failure of cointegration in one case seems compatible with the other. In both cases the residual graphs display drift away from the origin over the period 1974 – 78 indicating a persistent overprediction of consumption expenditure by the two regressions (28) and (29) during this period. This is in fact the period where the Australian savings ratio appears to rise, which we have already had occasion to discuss. The failure of cointegration therefore seems to be closely connected with this particular phenomena. The next section attempts to address this issue by suitable augmentation of the regression equation.

Additional evidence is provided by an analysis of the Bayesian posteriors for the parameter ρ in a model such as (9) fitted with the regression residuals from (28) and (29). These posteriors are shown in Figures 11.4(b) and (c). The results are very similar for the real and nominal data regressions. In both figures, the I-posteriors and ε-posteriors indicate a non-negligible probability of nonstationarity in the residuals. Table 11.6 gives the Bayes confidence probabilities assigned by these posteriors to the nonstationary set $\{\rho \geq 1\}$. Only in the case of the flat prior is this posterior probability negligible. Thus, a Bayes posterior analysis leads to results that support the posterior-odds tests.

Table 11.6: Posterior probabilities of stochastic nonstationarity in regression residuals of equations (28), (29), (31) and (32)

Regression Equation	AR(1) + trend model $P_I(\rho \geq 1)$	AR(1) + trend model $P_F(\rho \geq 1)$	AR(3) + trend model $P_\varepsilon(\rho \geq 1)$	AR(3) + trend model $P_F(\rho \geq 1)$
(28)	0.143	0.002	0.142	0.007
(29)	0.138	0.002	0.137	0.008
(31)	0.004	0.000	0.008	0.000
(32)	0.001	0.000	0.003	0.000

Legend: P_I = ignorance posterior; P_ε = ε-posterior (ε=0.05); P_F = flat posterior

Again, it useful to compare these Bayesian results with classical tests. Table 11.7 gives results for the ADF, $Z(a)$ and $Z(t)$ tests applied to the same regression residuals. The empirical results are unambiguous. At the 5 per cent level all of these tests confirm nonstationarity in the residuals. Thus, classical residual-based tests accord well with the Bayesian results. The evidence against cointegration for real and nominal C and YHD variables is therefore rather persuasive.

Table 11.7: Classical residual based tests for cointegration

Test	lag	(28)	(29)	5%CV	(31)	(32)	5%CV
ρ	2	0.897	0.924		0.743	0.673	
ADF	2	–2.292	–2.449	–3.466	–3.732	–4.135	–3.903
	6	–1.668	–1.836	–3.466	–2.972	–3.587	–3.903
$Z(a)$	2	–12.869	–13.775	–19.614	–30.056	–37.030	–25.820
	6	–11.123	–11.733	–19.614	–31.848	–39.758	–25.820
$Z(t)$	2	–2.639	–2.764	–3.466	–4.453	–4.954	–3.903
	6	–2.469	–2.576	–3.466	–4.543	–5.077	–3.903

3.5 Alternative Forms of the Long-Run Consumption Function Involving Inflation Rates

Since (28) and (29) fail classical and Bayesian tests as cointegrating equations we now consider alternative forms for these aggregate consumption relationships. Arguments discussed earlier in Section 3.1 suggest the use of a 'perceived income' measure in place of YHD in (28) and (29). If we employ a definition based on equation (27) that takes account of capital losses and use the M3 variable for liquid assets then we have in our notation

$$\begin{aligned} \text{YHD*} &= \text{YHD} - \beta\, \text{Inf}_{-4}\text{*M3} \\ &= \text{YHD}\{1 - \beta(\text{Inf}_{-4}\text{*M3})/\text{YHD}\} \\ &= \text{YHD}\{1 - \beta\,\text{RCL}\}\ . \end{aligned}$$

Taking logarithms we have the approximation (since Inf_{-4} is small)

$$\ln(\text{YHD*}) = \ln(\text{YHD}) - \beta\text{RCL}. \qquad (30)$$

If βM3/YHD is treated as effectively constant in the long run (it would be in traditional steady-state theory if the income elasticity of M3 demand were unity) then (30) may in turn be replaced by the approximation

$$\ln(\text{YHD*}) = \ln(\text{YHD}) - \gamma\text{Inf}_{-4}.$$

Since

$$\$YHD^* = \$YHD(1 - \beta RCL),$$

we also have the same approximations in nominal terms, viz.

$$\ln(\$YHD^*) \sim \ln(\$YHD) - \gamma \text{Inf}_{-4},$$

and

$$\ln(\$YHD^*) \sim \ln(\$YHD) - \beta RCL.$$

Using these approximations in versions of (28) and (29) in which YHD is replaced by the perceived measure YHD* we obtain the following alternative forms of the aggregate consumption relation

$$\ln(C) = a + b \ln(YHD) + c\text{Inf}_{-4} + \text{error}, \tag{31}$$

$$\ln(\$C) = a + b \ln(\$YHD) + c\text{Inf}_{-4} + \text{error}, \tag{32}$$

or, using the variable RCL

$$\ln(C) = a + b \ln(YHD) + cRCL + \text{error}, \tag{33}$$

$$\ln(\$C) = a + b \ln(\$YHD) + cRCL + \text{error}. \tag{34}$$

In all cases, the anticipated sign of the coefficient c is negative.

Since data on M3 and hence RCL are available to me only from 1965 whereas the Inf_{-4} series is computed from 1960, it is preferable to use relations (31) and (32) with the longer series. The empirical results we discuss below therefore concentrate on these augmented regression equations for consumption. However, we will report some results also for (33) and (34) based on the shorter series.

Table 11.8 gives the results of the posterior-odds test applied to the residuals of (31) – (34). In every case the results unambiguously confirm the presence of a cointegrating relationship among these variables.

Observe that the nominal equations receive more support than the real equations but in both cases the evidence is strongly in favour of cointegration. The results also seem to be more decisive for the specifications (31) and (32) that employ the Inf_{-4} variable, although the sample size differential makes this comparison more tentative.

Figure 11.5(a) plots the OLS residuals from the two regressions (31) and (32). As in the case of Figure 11.4(a) the behaviour of the residuals from the real and nominal regressions is very similar. However, in this case there appears to be no persistent overprediction of consumption

Table 11.8: Results of Phillips-Ploberger posterior odds tests for cointegration on the alternative forms; values of the BLR criterion

Consumption Function	Lag Length in Residual Regression 1	2	3	4	5	6
(31) real	1803.570	18.818	73.030	188.439	13.008	6.890
(32) nominal	21099.576	77.078	409.129	2267.961	99.859	55.193
(33) real	118.318	14.988	40.784	100.006	7.154	4.076
(34) nominal	643.022	40.371	167.375	1129.116	54.803	24.442
(36) real	40.477	5.467	11.805	28.383	3.848	2.690

Legend: reject unit root in residuals if 'BLR > 1.0'
(i.e. accept the presence of cointegration if BLR > 1.0)

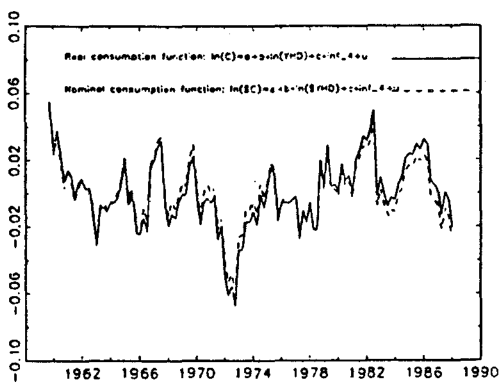
expenditure in the 1974 – 78 period, in contrast to that of the regressions (28) and (29). In this respect the residual graphs in Figure 11.5(a) are quite different from those of Figure 11.4(a). The visual evidence therefore corroborates the statistical finding from Table 11.8 that (31) and (32) are cointegrating.

Bayesian posteriors for ρ in model (9) fitted with these regression residuals are shown in Figures 11.5(b) and (c). Both figures show that virtually all of the mass of the posterior distributions is located in the stationary region of ρ. The calculations given in the lower panel of Table 11.6 confirm this directly. For every prior and for each model the posterior probability that ρ lies in the stationary region $\rho < 1$ is greater than 0.99. The Bayesian evidence therefore overwhelmingly supports the hypothesis that (31) and (32) are cointegrating relations.

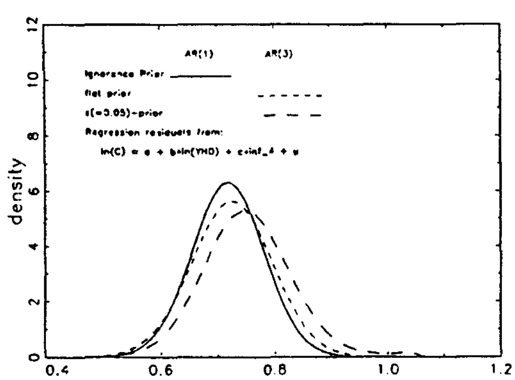
Table 11.7 (5th vertical panel) reports results of classical tests applied to the residuals from (31) and (32). The $Z(a)$ and $Z(t)$ tests reject the presence of a unit root in these residuals. The ADF rejects in the case of residuals from (32) for a model with a lag length $k = 2$ but otherwise fails to reject. Thus, the outcome is somewhat mixed. But the $Z(a)$ and $Z(t)$ tests both rather convincingly reject the presence of a unit root, and thereby corroborate the Bayesian evidence in support of cointegration for both (31) and (32).

Figure 11.5: Graphs of the residuals for regressions (31) and (32) and their posterior densities for ρ.

(a) Cointegrating residuals for (31) and (32)



(b) Posterior densities for ρ on equation (31)'s residuals



(c) Posterior densities for ρ on equation (32)'s residuals

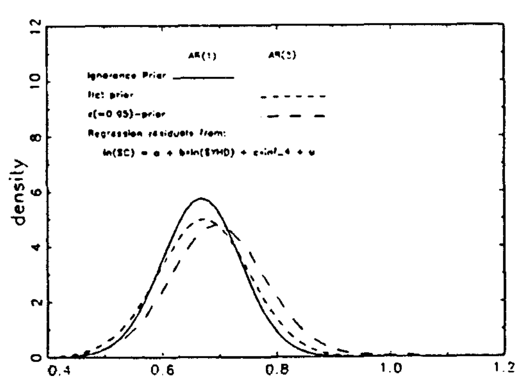


Table 11.9: Long-run parameter estimates of equations (31) and (32)

		Equation (31)		Equation (32)	
Method	Lag	b	c	b	c
Spectral Regression	2	1.010	−1.680	1.014	−1.888
		(137.229)	(−6.843)	(461.751)	(−8.658)
	6	0.993	−1.079	1.008	−1.120
		(159.422)	(−5.418)	(531.968)	(−6.140)
Fully Modified OLS	2	1.048	−2.498	1.018	−2.263
		(113.905)	(−8.649)	(404.436)	(−9.396)
	6	1.054	−2.714	1.019	−2.463
		(76.431)	(−6.841)	(268.595)	(−7.120)
Canonical Cointegrating Regression	2	1.055	−2.669	1.020	−2.510
		(88.678)	(−7.297)	(317.791)	(−7.924)
	6	1.063	−2.922	1.021	−2.644
		(70.914)	(−6.387)	(267.091)	(−7.460)
OLS		1.046	−2.306	1.019	−2.345
		(120.808)	(−8.484)	(449.602)	(−10.825)

Legend: t-ratios in parentheses

Equations (31) and (32) were estimated by several asymptotically-efficient regression techniques. The methods included were spectral regression (Phillips, 1991b), fully modified OLS (Phillips and Hansen, 1990), canonical cointegrating regression (Park, 1991) and simple OLS. The resulting estimates for the long-run coefficients in (31) and (32) are presented in Table 11.9. The results appear highly consistent both across estimation methods and across models. The estimates of b and c obtained from the real variable regression (31) and the nominal variable regression (32) are certainly very close, confirming this aspect of the theory by which they were derived. In both models the estimates are highly significant and b, the long-run propensity to consume, is particularly well determined. Estimates of c are all significantly negative, confirming the theory underlying (31) and (32). The spectral regression estimates differ by the greatest amount from the other estimates. In fact, the spectral regression estimates of b in model (31) are the only estimates of this parameter that are not significantly different from unity.

3.6 Some Final Empirical Formulations

With the shorter data set available for M3 and the capital loss variable RCL we performed some additional empirical analysis. Two formulations were studied:

$$\ln(\mathrm{C}) = a + b \ln(\mathrm{YHD}) + c\,\mathrm{RCL} + \text{error} \tag{35}$$

as suggested in (33) above; and

$$\ln(\mathrm{C}) = a + b \ln(\mathrm{YHD}) + c\,\mathrm{RCL} + d \ln(\mathrm{M3}) + \text{error} \tag{36}$$

According to our earlier decomposition (30) we may write

$$\ln(\mathrm{YHD}^*) = \ln(\mathrm{YHD}) - \beta\,\mathrm{RCL} \tag{37}$$

Using this representation of YHD* in (28) yields (35) directly. Hence, the specification (35) should be an alternative to (31). Again, the anticipated sign of c is negative but the coefficient can be expected to be different from the coefficient of Inf_{-4} in (31).

Equation (36) is based on a steady-state solution that arises in the work of HVS. HVS employ an integral correction mechanism in their empirical study of U.K. consumption behaviour and use a liquid assets/income variable to capture this effect. The presence of this additional variable in the regression alters the steady-state solution. In our notation the steady state would be of the form

$$\begin{aligned}\ln(\mathrm{C}) &= a + b \ln(\mathrm{YHD}^*) + c \ln(\mathrm{M3}/\mathrm{YHD}^*)\\ &= a + (b - c)\ln(\mathrm{YHD}^*) + c \ln(\mathrm{M3})\end{aligned}$$

As remarked earlier, this steady-state formulation would only be possible in a larger model that explained M3 demand if the long-run income elasticity of demand for M3 were unity. Using (30), the expression above can be approximated as

$$\ln(\mathrm{C}) = a + (b - c)\ln(\mathrm{YHD}) + d\,\mathrm{RCL} + c \ln(\mathrm{M3}),$$

which yields the specification (36) above. According to this formulation, the long-run propensity to consume is given by the sum of the coefficients of ln(YHD) and ln(M3).

Bayes model posterior-odds tests of the long-run specifications (35) (i.e.(33)) and (36) are given in Table 11.8. Clearly the data support equation (35) as an alternative to (31) according to this test but give much less support to (36). This outcome is interesting because it shows that the inclusion of additional variables in a cointegrating relation does carry some cost. Especially in small samples (and recall here that we are using series of reduced length due to the availability of the M3 measure), there will be additional uncertainty in the regression residuals from (36) arising from the presence of the additional variable in the regression and imprecision in the estimation of its coefficient. Posterior-odds criteria typically involve a penalty for such additional regressors and thereby inherently operate a form of model selection. In the present case, we find that the posterior-odds criterion selects the more parsimonious representation (35).

Bayesian posteriors provide additional evidence on this point. Figures 11.6(a) and (b) give graphs of the posteriors for ρ in model (9) fitted with the regression residuals from equations (35) and (36). These posteriors show an appreciable probability of misspecification in equation (36) (i.e. no cointegration amongst the four variables). The posteriors in Figure 11.6(a) provide more support for the specification of equation (35). However, the l-posterior indicates some uncertainty about this specification. The posterior probabilities of nonstationarity in the residuals of these two regressions are given in Table 11.10. According to these outcomes, there is strong evidence in both the l-posterior and the ε-posterior against equation (36). Equation (35) is more satisfactory, although the l-posterior probability of $\{\rho \geq 1\}$ is non-negligible. In this case the uncertainty may be the result of the shorter data series from which these regression equations were estimated.

Thus, Bayesian evidence rejects the long-run specification (36) but is tentatively supportive of (35). Given that (35) and (31) are close alternatives, given that (31) is estimated with longer time series and given the unambiguous support for (31) and its nominal version (32) in the evidence, we conclude that specifications (31) and (32) of the long-run consumption equation are to be preferred.

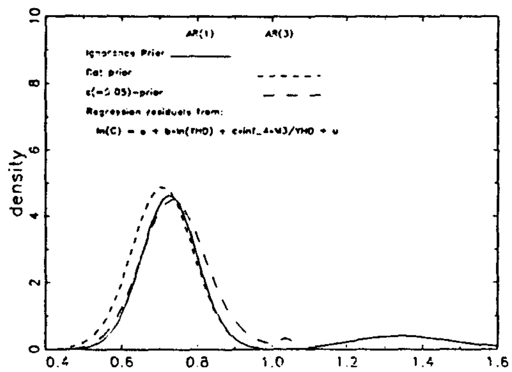
4. CONCLUSION

This paper is an empirical exercise in new Bayesian methodology. The outcome of this exercise can be summarized as follows.

1. Objective Bayesian methods seem to provide helpful empirical evidence that complements classical tests. The Bayesian methods we have

Figure 11.6: Graphs of the posterior densities of ρ for the residuals of regressions (35) and (36)

(a) Posterior densities for ρ on equation (35)'s residuals



(b) Posterior densities for ρ on equation (36)'s residuals

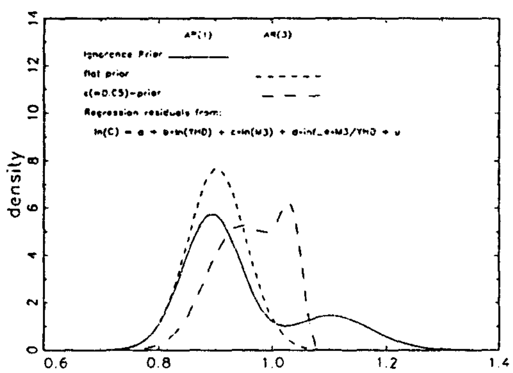


Table 11.10: Posterior probabilities of stochastic nonstationarity in regression residuals of equations (35) and (36)

Regression Equation	AR(1) + trend model		AR(3) + trend model	
	$P_I(\rho \geq 1)$	$P_F(\rho \geq 1)$	$P_\varepsilon(\rho \geq 1)$	$P_F(\rho \geq 1)$
(35)	0.140	0.000	0.018	0.000
(36)	0.255	0.011	0.330	0.032

Legend: P_I = ignorance posterior; P_ε = ε-posterior (ε=0.05); P_F=flat posterior used elicit information from the data in ways that are quite distinct from classical procedures. Posterior graphics exhibit in a rather clear way the plausibility of the underlying hypotheses and show sensitivities, when they exist, to prior distributions and model choice. Uncertainty in model specifications tend to show up in widely dispersed and bimodal posterior distributions. Hypotheses like cointegration may be examined in a straightforward way by Bayesian posterior analysis of the residuals and by direct tests using the Phillips-Ploberger 'Bayes model' posterior-odds test.

2. The methods used in this paper are easy to apply in practice. Computer programs in GAUSS 2.0 have now been written to implement the procedures used here and are quick and convenient to use. On modern desktop PCs the programs are executed in a few seconds and graphics are especially easy to produce in GAUSS software.

3. Bayesian methods are especially useful when classical tests give conflicting results. It is known that ADF tests, for instance, have low power especially as the lag length increases, whereas this is much less true of the $Z(a)$ and $Z(t)$ tests. Our empirical results in Section 3 often show divergent conclusions in the application of these classical tests. This type of mixed outcome is common in empirical practice and hard to interpret. In such cases Bayesian methods are particularly valuable because they present an additional form of empirical evidence that often points to new possibilities. For instance, as seen in the posterior analysis of Section 3.6, a full posterior distribution (with appreciable stationary and nonstationary set probabilities) can rather clearly indicate an unsatisfactory empirical specification, especially when it is compared with the posteriors of other competing specifications.

4. Our empirical results on the long-run Australian consumption function appear conclusive. An inflation measure should certainly figure in the consumption equation and this empirical finding fits well with underlying theory of consumption behaviour which suggests the use of an appropriate 'perceived income' measure that compensates for capital losses in periods of inflation. Our results show that instantaneous inflation measures are inappropriate for this purpose but that annual rates like Inf_{-4} work well and correspond better with longer-term perceptions of capital losses in inflationary periods. We find Inf_{-4} to be nonstationary and we find strong support for a cointegrating relationship between C, YHD and Inf_{-4} from a variety of Bayesian and classical tests. Furthermore, our posited long-run relationship between these variables holds for nominal variables ($C, $YHD) and real variables (C, YHD). Empirical estimation of equations (31) and (32) by several asymptotically efficient methods leads to broadly similar results. Correspondence between the real equation (31) estimates and those of the nominal equation (32) is close and

again fits well with underlying theory. We conclude that the presence of a suitable measure of inflation or relative capital loss is supported empirically, at least for Australian data on private consumption and household disposable income.

NOTE

1. All of the computations and graphics reported herein were carried out by the author using programs written in GAUSS 2.0 on a 386 – 20 Mhz desktop PC. Thanks go to Rob Trevor and Sam Ouliaris for supplying the data used in Section 3, to Glena Ames for outstanding wordprocessing and to the NSF for research support under grant no. SES 8821180.

REFERENCES

Anstie, R.K., M.R. Gray and A.R. Pagan (1983), 'Inflation and the consumption ratio,' in A.R. Pagan and P.K. Trivedi (eds), *The Effects of Inflation: Theoretical Investigations and Australian Evidence*, Canberra: Centre for Economic Policy Research, ANU.

Box, G.E.P. and G.M. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden Day.

Campbell, J.Y. and P. Perron (1991), 'Pitfalls and opportunities: What macroeconomists should know about unit roots', *NBER Working Paper* No. 100.

DeJong, D.N. and C.H. Whiteman (1989), 'Trends and cycles as unobserved components in U.S. real GNP: A Bayesian perspective', *Proceedings of the American Statistical Association*.

DeJong, D.N. and C.H. Whiteman (1991), 'The temporal stability of dividends and stock prices: Evidence from the likelihood function', *American Economic Review*, Vol. 81, pp. 600 – 17.

Diebold, F.X. and M. Nerlove (1990), 'Unit roots in economic time series', *Advances in Econometrics*, Vol. 8, pp. 3 – 70.

Dolado, J.J., T. Jenkinson and S. Sosvilla-Rivero (1990), 'Cointegration and unit roots', *Journal of Economic Surveys*, Vol. 4, pp. 249 – 73.

Engle, R.F. and C.W.J. Granger (1987), 'Cointegration and Error-correction: Representation, Estimation and Testing', *Econometrica*, Vol. 55, pp. 251 – 76.

Geweke, J. (1988), 'Comment on Poirier: Operational Bayesian methods in econometrics', *Journal of Economic Perspectives*, Vol. 2, pp. 159 – 66.

Hall, V.B. and R.G. Trevor (1991), 'Long-run equilibrium estimation and inference: A non-parametric approach', to appear in P.C.B. Phillips (ed.), *Models, Methods and Applications of Econometrics: Essays in Honour of Rex Bergstrom*, Oxford: Basil Blackwell (forthcoming).

Hendry, D.F. and T. von Ungern-Sternberg (1981), 'Liquidity and inflation effects on consumers' expenditure', in A. Deaton (ed.), *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, Cambridge: Cambridge University Press, pp. 237 – 60.

Jeffreys, H. (1946), 'An invariant form for the prior probability in estimation problems', *Proceedings of the Royal Society of London*, Series A, Vol. 186, pp. 453 – 61.

Johansen, S. (1988), 'Statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control*, Vol. 12, pp. 231 – 54.

Johansen, S. (1989/1991), 'Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models', University of Copenhagen, to appear in *Econometrica*, 1991.

Leamer, E.E. (1978), *Specification Searches: Ad hoc Inferences with Nonexperimental Data*, New York: John Wiley & Sons.

Leamer, E.E. (1983), 'Let's take the con out of econometrics', *American Economic Review*, Vol. 73, pp. 31 – 44.

Leamer, E.E. (1988), 'Things that bother me', *Economic Record*, Vol. 64, pp. 344 – 59.

Malinvaud, E. (1980), *Statistical Methods of Econometrics*, 3rd Edition, Amsterdam: North Holland.

Park, J.Y. (1990), 'Testing for unit roots and cointegration by variable addition', *Advances in Econometrics*, Vol. 8, pp. 107 – 33.

Park, J.Y. (1991), 'Canonical cointegrating regressions', *Econometrica* (forthcoming).

Park, J.Y. and B. Choi (1988), 'A new approach to testing for a unit root', Working Paper No. 88 – 23, Center for Analytical Economics, Cornell University.

Park, J.Y. and P.C.B. Phillips (1988), 'Statistical inference in regressions with integrated processes: Part 1', *Econometric Theory*, Vol. 4, pp. 468 – 97.

Park, J.Y. and P.C.B. Phillips (1989), 'Statistical inference in regressions with integrated processes: Part 2', *Econometric Theory*, Vol. 5, pp. 95 – 131.

Phillips, P.C.B. (1991a), 'Optimal inference in cointegrated systems', *Econometrica*, Vol. 59, pp. 283 – 306.

Phillips, P.C.B. (1991b), 'Spectral regression for cointegrated time series', in W. Barnett, J. Powell and G. Tauchen (eds), *Nonparametric and Semiparametric Methods in Economics and Statistics*, New York: Cambridge University Press.

Phillips, P.C.B. (1991c), 'Unit roots', Essay for the *New Palgrave Dictionary of Money and Finance*.

Phillips, P.C.B. (1991d), 'To criticize the critics: An objective Bayesian analysis of stochastic trends', *Journal of Applied Econometrics* (forthcoming).

Phillips, P.C.B. (1991e), 'Bayesian routes and unit roots: de rebus prioribus semper est disputandum', *Journal of Applied Econometrics* (forthcoming).

Phillips, P.C.B. and B. E. Hansen (1990), 'Statistical inference in instrumental variables regression with I(1) processes', *Review of Economic Studies*, Vol. 57, pp. 99 – 125.

Phillips, P.C.B. and M. Loretan (1991), 'Estimating long-run economic equilibria', *Review of Economic Studies*, Vol. 59, pp. 407 – 36.

Phillips, P.C.B. and W. Ploberger (1991a), 'Time series modeling with a Bayesian frame of reference: I. Concepts and illustrations', Cowles Foundation Discussion Paper No. 980.

Phillips, P.C.B. and W. Ploberger (1991b), 'Time series modeling with a Bayesian frame of reference: II. General theory and applications', in preparation.

Poirier, D.J. (1988), 'Frequentist and subjectivist perspectives on the problems of model building in economics' (with discussion), *Journal of Economic Perspective*, Vol. 2, pp. 121 – 44.

Saikkonen, P. (1991), 'Asymptotically efficient estimation of cointegration regressions', *Econometric Theory*, Vol. 7, pp. 1 – 21.

Sims, C.A. (1988), 'Bayesian skepticism on unit root econometrics', *Journal of Economic Dynamics and Control*, Vol. 12, pp. 463 – 74.

Sims, C.A., J.H. Stock and M.W. Watson (1990), 'Inference in linear time series models with some unit roots', *Econometrica*, Vol. 58, pp. 113 – 44.

Sims, C.A. and H. Uhlig (1988/1991), 'Understanding unit rooters: A helicopter tour', *Econometrica*, Vol 59(6), pp. 1591 – 9.

Stock, J.H. and M.W. Watson (1991), 'A simple MLE of cointegrating vectors in higher order integrated systems', mimeographed, Department of Economics, Northwestern University.

Toda, H. and P.C.B. Phillips (1991a), 'Vector autoregressions and causality', Cowles Foundation Discussion Paper No. 977.

Toda, H. and P.C.B. Phillips (1991b), 'The spurious effect of unit roots on exogeneity tests in vector autoregressions: An analytical study', Cowles Foundation Discussion Paper No. 978.

Toda, H. and P.C.B. Phillips (1991c), 'Vector autoregression and causality: A theoretical overview and simulation study', *Econometric Reviews*, forthcoming.

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

Zivot, E. and P.C.B. Phillips (1991), 'A Bayesian analysis of trend determination in economic time series', Cowles Foundation Discussion Paper, No. 1002, Yale University.