Carter, R. A. L.; Zellner, A.

Working Paper

The ARAR error model for univariate time series and distributed lag models

Research Report, No. 2002-5

Provided in Cooperation with:
Department of Economics, University of Western Ontario

Suggested Citation: Carter, R. A. L.; Zellner, A. (2002) : The ARAR error model for univariate time series and distributed lag models, Research Report, No. 2002-5, The University of Western Ontario, Department of Economics, London (Ontario)

This Version is available at:
https://hdl.handle.net/10419/70413

The ARAR Error Model for Univariate Time Series and Distributed Lag Models

R. A. L. Carter*
University of Western Ontario and University of Calgary
A. Zellner†
University of Chicago

Abstract

We show that the use of prior information derived from former empirical findings and/or subject matter theory regarding the lag structure of the observable variables together with an AR process for the error terms can produce univariate and single equation models that are intuitively appealing, simple to implement and work well in practice. JEL Classification C11, C22.

1 Introduction

"None of the previous work should be construed as a demonstration of the inevitability of MA disturbances in econometric models. As Parzen's proverb[1] reminds us the disturbance term is essentially man-made, and it is up to man to decide if some of his creations are more

reasonable than others." Nicholls, Pagan and Terrell (1975) p.117.

Early empirical work in economics, e.g. Burns and Mitchell (1946), O'Leary and Lewis (1955), discovered cycles in many time series which were consistent with the early dynamic models of Samuelson (1939a,b), Metzler (1941) and Goodwin (1947). Thus, economists building time series models often have quite strong prior beliefs, based on this earlier theoretical and applied work, about the lag structure for an observable variable. However, they usually have quite weak prior beliefs about the lag structure for the error in a model. At one time the AR model for the error was popular. Then the methods of Box and Jenkins (1976) became dominant in the field of univariate time series analysis and the MA model for the error was introduced without much theoretical justification. This produced complicated likelihood functions and estimation procedures and implied infinite AR processes for observed variables. We show here how a return to the use of prior information, derived from former empirical findings and/or subject matter theory, about the lag structure of the observable variables and to the AR model for the error is intuitively appealing, yields finite, rather than infinite, AR processes for observed variables and produces models for which inference procedures[2]are simplified.

2 Univariate Models

To explain the behavior of a random variable Y_t, $t = 1, \ldots, T$, researchers often have fairly precise beliefs, from subject matter study, about its lag struc-

ture. For example, Samuelson's (1939) multiplier-accelerator model produces an AR(2) for output, Metzler's (1941) model yields an AR(3) for inventories and Goodwin's (1947) cobweb model produces an AR(2) form for the price. More recently: Garcia-Ferrer *et. al.* (1987) employed an AR(3) model as well as an AR(3) model including leading indicator variables to forecast output growth rates for several countries; Geweke (1988) used an AR(3) model for real, per capita GDP to study its dynamics for 19 OECD countries; Zellner and Hong (1989) and Zellner, Hong and Min (1991) also used an AR(3), together with leading indicators, to forecast turning points in the growth rate of real output in 18 countries. In such cases researchers have assumed that a reasonable starting point for an analysis is to specify that a random variable of interest has been generated by a process with an autoregressive component that is

$$\phi(L)(Y_t - \mu) = U_t, \tag{1}$$

where $\phi(L)$ is a polynomial of degree p in the lag operator L, μ is the origin from which Y_t is measured (the mean if Y_t is stationary) and U_t is a covariance-stationary error with zero mean. Researchers may have quite strong beliefs about the value of p and be willing to specify that Y_t is stationary, perhaps as a result of differencing, which implies restrictions on the roots of $\phi(L)$. In such cases the parameters of interest are the ϕ_i, the coefficients of $\phi(L)$, or some function of them, such as the roots of $\phi(L)$. However, it is rare for researchers to have such strong prior beliefs about the error process and thus a variety of models for U_t involving input variables and/or AR, MA or ARMA error processes may be entertained : see e.g. Fuller and Martin (1961), Zellner *et. al.*

(1965) and Zellner and Geisel (1970).

One possible model for the error is

$$U_t = \varepsilon_t, \tag{2}$$

where ε_t is white noise with zero mean and variance σ^2. However, this is a rather restrictive model: it implies that a shock to the subject-matter portion of the model, $\phi(L)(Y_t - \mu)$, has an impact only in the current period. Of course, if $\phi(L)$ is invertible the MA representation of Y_t,

$$Y_t = \mu + \frac{1}{\phi(L)}\varepsilon_t, \tag{3}$$

has infinite length.

2.1 The ARMA Model

A popular alternative model for U_t is the MA model of Box and Jenkins (1976).

$$U_t = \theta(L)\varepsilon_t, \tag{4}$$

where $\theta(L)$ is an invertible polynomial of degree q in L. This model was introduced "To achieve greater flexibility in fitting actual time series, ..."[3]rather than from any explicit prior knowledge about the behavior of the error. But this model is only slightly less restrictive than the white noise model because it implies that the effect on $\phi(L)(Y_t - \mu)$ of a shock ε_t dies out completely after q periods. Also strong restrictions on the values of the autocorrelations of Y_t are needed for $\theta(L)$ to be invertible: see Box and Jenkins (1976, Chapter 3). In

this case too the MA representation for Y_t,

$$Y_t = \mu + \frac{\theta(L)}{\phi(L)}\varepsilon_t, \tag{5}$$

is infinite in length and allows somewhat richer behavior than does the white noise case. However, the AR representation,

$$\frac{\phi(L)}{\theta(L)}(Y_t - \mu) = \varepsilon_t, \tag{6}$$

is also infinite, a result which may be found undesirable from a subject matter viewpoint and which complicates inference about the parameters of interest, the ϕ_i.

Of course, there are circumstances in which an MA model for U_t is desirable on *a priori* grounds. For example if the observations on Y_t were the result of temporal aggregation an MA model for U_t may be considered. However, we note that in many cases an MA model for U_t arises from an examination of sample autocorrelations and is not justified by any prior theory. In these cases the model described in the next section is worthy of consideration.

2.2 The ARAR Model

We propose an alternative model for the error U_t which is parsimonious, allows rich time series behavior for U_t and simplifies inference procedures for the parameters of interest. Our proposal is to model U_t as an AR process, as in many earlier studies,

$$\omega(L)U_t = \varepsilon_t, \tag{7}$$

where $\omega(L)$ is a polynomial of degree r in L. If $\omega(L)$ is assumed to be invertible the MA representation of U_t is

$$U_t = \frac{1}{\omega(L)}\varepsilon_t, \tag{8}$$

which is infinite in length. Thus, in this model,

$$\phi(L)(Y_t - \mu) = U_t = \frac{1}{\omega(L)}\varepsilon_t, \tag{9}$$

the impact upon the subject-matter portion, $\phi(L)(Y_t - \mu)$, of a shock ε_t may die out slowly over many lags, rather than being cut off abruptly after a small number of lags. We believe that this feature of our model, which we label $ARAR(p, r)$, makes it a useful addition to the family of earlier $ARMA(p, q)$ models discussed above. Our model is no poorer in terms of possible behavior of U_t in that values of r as small as 2 produce very rich behavior in (8). Also, (8) is consistent with Wold's (1938) Decomposition Theorem, in contrast to MA models for U_t which impose a truncation on the Wold representation of U_t. If $\phi(L)$ is invertible and Y_t is covariance-stationary, then the MA representations of all the above models for Y_t are infinite and, therefore, consistent with Wold's theorem. However, (8) has the advantage of being the simplest model that also imposes this consistency on U_t.

To see the parsimony of (8) as compared to (4) consider an example in which $\omega(L) = 1 - \omega_1 L - \omega_2 L^2$. We can factor $\omega(L)$ in terms of its inverse roots as $\omega(L) = (1 - \xi_1 L)(1 - \xi_2 L)$. Assume that $\xi_1 \neq \xi_2$, $|\xi_1| < 1$ and $|\xi_2| < 1$. Then we can rewrite (8) as

$$U_t = \varepsilon_t + (c_1\xi_1 + c_2\xi_2)\varepsilon_{t-1} + (c_1\xi_1^2 + c_2\xi_2^2)\varepsilon_{t-2} + (c_1\xi_1^3 + c_2\xi_2^3)\varepsilon_{t-3} + \ldots, \tag{10}$$

where $c_1 = \xi_1/(\xi_1 - \xi_2)$ and $c_2 = -\xi_2/(\xi_1 - \xi_2)$. Now assume that the roots are real with $\xi_1 = .7$ and $\xi_2 = -.5$. Then (10) becomes

$$U_t = \varepsilon_t + .6167\varepsilon_{t-1} + .1867\varepsilon_{t-2} + .2522\varepsilon_{t-3} + .1140\varepsilon_{t-4} + .1111\varepsilon_{t-5} + .06212\varepsilon_{t-6} + \ldots \tag{11}$$

so it would take an MA process with at least five parameters to approximate the inverse of the AR with two parameters. Next assume that there are complex conjugate roots with $\xi_1 = a + bi$ and $\xi_2 = a - bi$. Now (10) can be written as

$$U_t = \varepsilon_t + 2a\varepsilon_{t-1} + (3a^2 - b^2)\varepsilon_{t-2} + 4a(a^2 - b^2)\varepsilon_{t-3} + \ldots \tag{12}$$

If $a = .7$ and $b = .5$ (12) becomes

$$U_t = \varepsilon_t + 1.400\varepsilon_{t-1} + 1.220\varepsilon_{t-2} + .6720\varepsilon_{t-3} + .2879\varepsilon_{t-4} + .2418\varepsilon_{t-5} + \ldots \tag{13}$$

and so here too an approximating MA would have to have at least five parameters, as compared to two in $\omega(L)$.

Our model for Y_t is, from (1) and (8),

$$\omega(L)\phi(L)(Y_t - \mu) = \varepsilon_t, \tag{14}$$

which can be written as

$$\alpha(L)Y_t = \alpha_0 + \varepsilon_t \tag{15}$$

where $\alpha(L) = \omega(L)\phi(L)$ is a polynomial in L of degree $p + r$ with restricted coefficients and $\alpha_0 = \omega(1)\phi(1)\mu$. If $\phi(L)$ and $\omega(L)$ are invertible there is an MA representation of Y_t as

$$Y_t = \mu + \frac{1}{\omega(L)\phi(L)}\varepsilon_t. \tag{16}$$

Both (1) and (15) are finite AR processes in contrast to (6) which is infinite. Hence the ARAR model is simpler than the ARMA model and, according to the Wrinch-Jeffreys (1921) Simplicity Postulate, should, therefore, be accorded higher prior probability. We note too that almost no scientific theories are framed as infinite AR processes, although many are in the form of finite difference or differential equations.

Note that (3), (5) and (16) are all examples of a linear filter, with the white noise ε_t as input and Y_t as output, which is the basis for the ARMA model of Box and Jenkins (1976). The difference between our model and their's is that we approximate the filter with the inverse of the product of two polynomials while they approximate it with the ratio of two polynomials. The only sort of behavior that is allowed by the ARMA model but ruled out by the ARAR model (14) is an infinite AR in Y_t; although if the degrees of $\phi(L)$ and $\omega(L)$ are high then (14) will be a very long AR. In some cases this restriction will be a notable advantage of our model. Seasonal effects are easily handled in the ARAR model by expressing either $\phi(L)$ or $\omega(L)$, or both, as the product of nonseasonal and seasonal polynomials. Another attractive feature of our model is the facility of separate analyses, using proper priors to achieve identification, of the polynomials $\phi(L)$ and $\omega(L)$, which would be appropriate if subject matter study provided prior information about the coefficients of $\phi(L)$.

Note that some of the coefficients, α_i, of $\alpha(L)$ may be quite small in absolute value, even though no ϕ_i or ω_i is small. This can result in small and imprecise estimates of these α_i if the restrictions implied by the ARAR structure are

ignored and only the unrestricted $AR(p+r)$ (15) is used in estimation with a small sample. Then researchers may be tempted to impose invalid zero restrictions on $\alpha(L)$ in an effort to achieve statistical efficiency. As we demonstrate below, estimation of ϕ and ω is easy so there is no reason to concentrate on (15) exclusively.

Before considering inferences about the parameters of $\phi(L)$ and $\omega(L)$ we note that, without further information, they are not identified. One easy way to see this is to interchange $\phi(L)$ and $\omega(L)$ in (9) which would yield the same unrestricted $AR(p+r)$ as (15) so the likelihood function, which is based on (15), would be unchanged.

Another potential failure of identification is revealed by multiplying both sides of the model (9) by a nonzero scalar υ_0. This would leave (15) unchanged: however, it would result in

$$\upsilon_0\phi(L) = \upsilon_0 - \upsilon_0\phi_1 L - \ldots \upsilon_0\phi_p L^p.$$

We avoid this identification problem by assuming that the model has been normalized to have the first term in $\phi(L)$ equal to 1.0.

A second potential identification failure arises if both sides of (9) are multiplied by the invertible lag polynomial $\upsilon(L)$ to give

$$\upsilon(L)\phi(L)Y_t = \upsilon(1)\mu + \frac{\upsilon(L)}{\omega(L)}\varepsilon_t. \tag{17}$$

Here too (15) and the likelihood function will be unchanged. This is analogous to model multiplicity in the ARMA context and it also arises with ARMA models which contain the products of seasonal and nonseasonal lag polynomials. We

follow Box and Jenkins (1976) and others in assuming that all common factor polynomials like $v(L)$ have been cancelled out of the model.

Even after normalization and common factor cancellations have been performed, there is an identification problem to be solved. Write $\phi(L)$ and $\omega(L)$ in terms of their roots as

$$\phi(L) = (1 - \lambda_1 L)(1 - \lambda_2 L) \dots (1 - \lambda_p L) \tag{18}$$

and

$$\omega(L) = (1 - \xi_1 L)(1 - \xi_2 L) \dots (1 - \xi_r L) \tag{19}$$

Thus

$$\begin{aligned} \alpha(L) &= \phi(L)\omega(L) \\ &= (1 - \lambda_1 L) \dots (1 - \lambda_p L)(1 - \xi_1 L) \dots (1 - \xi_r L). \end{aligned} \tag{20}$$

Now, to focus on the identification problem, assume the λ_i and ξ_i have unknown values. Then write

$$\alpha(L) = (1 - \eta_1 L)(1 - \eta_2 L) \dots (1 - \eta_{p+r} L) \tag{21}$$

and assume we know the values of the roots η_i. The identification problem lies in deciding which terms $(1 - \eta_i L)$ are part of $\phi(L)$ and which are part of $\omega(L)$. The solution is to use the same prior information that was used to specify $\phi(L)$ to make this decision. For example, in many cases $\phi(L)$ will be specified to be of degree 2 displaying damped cyclical behavior with a period in some *a priori* most probable range. This implies that the roots η_i belonging to $\phi(L)$ must be a complex conjugate pair with modulus less than one and a period lying in the

most probable range. Any roots η_i which are real or complex conjugate with a smaller period must belong to $\omega(L)$.

In practice the values of the η_i are unknown. But once we specify p and r we can fit an $AR(p+r)$ by ordinary least squares (OLS) and find the roots of the resulting $\hat{\alpha}(L)$. These estimated roots can be used to find initial values of the $\hat{\phi}_i$ and $\hat{\omega}_i$ for use in nonlinear least squares (NLS) estimation of $\omega(L)\phi(L)$. Indeed, we could use this idea with an AR model produced by someone else. Upon reading their results we may suspect that their unrestricted AR is long enough to be the product of a $\phi(L)$ and an $\omega(L)$. Using their estimated AR coefficients we could obtain estimated roots, moduli and periods and, with large samples, these results could help us to confirm or reject our hypothesis regarding the degree of the $\phi(L)$ polynomial.

Alternatively, someone building an ARAR model could use *a priori* plausible initial values of the ϕ_i and ω_i in NLS to obtain estimates $\hat{\omega}(L)$ and $\hat{\phi}(L)$. Then the roots of $\hat{\omega}(L)$ and $\hat{\phi}(L)$ can be compared to those of $\hat{\alpha}(L)$ from (15) as a check that the identification is correct. We do this in our empirical example on housing starts, below.

An interesting simple case is obtained by assuming that p and r are both one so that (14) becomes

$$\begin{aligned} \omega(L)\phi(L)(Y_t-\mu) &= (1-\omega L)(1-\phi L)(Y_t-\mu) && (22)\\ &= (1-[\omega+\phi]L+\omega\phi L^2)Y_t-(1-\omega)(1-\phi)\mu && (23)\\ &= \varepsilon_t \end{aligned}$$

from which

$$Y_t = (1-\omega)(1-\phi)\mu + (\omega+\phi)Y_{t-1} - \phi\omega Y_{t-2} + \varepsilon_t \quad (24)$$

$$= \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \varepsilon_t. \quad (25)$$

We believe this $ARAR(1,1)$ model is a useful alternative to the popular $ARMA(1,1)$ model with the same number of parameters. Imposing stationarity and invertibility on the $ARMA(1,1)$ model results in very strong restrictions on the range of admissible values for the autocorrelations of Y_t at lags 1 and 2, ρ_1 and ρ_2: see Box and Jenkins(1976) Chapter 3, especially Figure 3.10(b). The analogous restrictions on the $ARAR(1,1)$ model result in restrictions on ρ_1 and ρ_2 which are much weaker. We assume that ϕ and ω are real and that $-1 < \phi < 1$ and $-1 < \omega < 1$, so that Y_t and U_t are both stationary. These restrictions result in the feasible set of α_1 and α_2 values being a subset of that for an unrestricted $AR(2)$ which, since ϕ and ω are real, excludes the subset associated with complex roots. This is shown in Figure 1 which should be compared with Box and Jenkins (1976) Figure 3.9. The restrictions on α_1 and α_2 also result in a restriction of the ρ_1, ρ_2 space which is a subset of that for an unrestricted $AR(2)$. This space is shown in Figure 2a which should be compared with Box and Jenkins (1976) Figures 3.3(b) and, especially, 3.10(b), which is reproduced below as Figure 2b. Note that the admissible parameter space shown in Figure 2a is considerably larger than that shown in Figure 2b and that negative values of ρ_2 are excluded. This last point should aid in model identification as sample autocorrelations which are negative at both lags one and two would be evidence

against an $ARAR(1,1)$ model being appropriate. We note that for many economic time series the first two sample autocorrelations are positive so that the ARAR model is not ruled out.

If Y_t or U_t are assumed to be nonstationary in (24) we can write it as

$$(Y_t - \phi Y_{t-1}) = \alpha_0 + \omega(Y_{t-1} - \phi Y_{t-2}) + \varepsilon_t \tag{26}$$

or as

$$(Y_t - \omega Y_{t-1}) = \alpha_0 + \phi(Y_{t-1} - \omega Y_{t-2}) + \varepsilon_t. \tag{27}$$

Thus, instead of using, say, first differencing to induce stationarity we might assume that the more general "ϕ" differencing or "ω" differencing induces stationarity and these restrictions can be imposed in analyzing (24).

Unrestricted inferences about the α_i in (25) can be easily obtained using OLS, ML, MAD, traditional Bayes and Bayesian Method of Moments (BMOM): see Zellner (1996, 1997), Tobias and Zellner (2001) and Green and Strawderman (1995). The stationarity restrictions imply that $-2 < \alpha_1 < 2$ and $-1 < \alpha_2 < 1$. Note that if, for example, $\phi > 0$ and $\omega < 0$, α_1 may be quite small in absolute value but imposing $\alpha_1 = 0$ would be a specification error.

To achieve identification in the present example we might impose the prior restrictions that $|\phi| > |\omega|$ and that both ϕ and ω are real. Then to obtain the posterior densities for ϕ and ω from the unrestricted posterior distribution for α_1 and α_2, draw a realization of α_1 and α_2, say $\alpha_1(1)$ and $\alpha_2(1)$. If the stationarity restriction is to be imposed check that $-2 < \alpha_1(1) < 2$ and $-1 < \alpha_2(1) < 1$. If $\alpha_1(1)$ and $\alpha_2(1)$ violate this restriction discard them and draw again. If they

satisfy it, check that $\alpha_1^2(1) + 4\alpha_2(1) \geq 0$ so that they correspond to the real roots ϕ and ω. If $\alpha_1(1)$ and $\alpha_2(1)$ violate this condition discard them and draw again. If they satisfy it, calculate $r_1(1) = .5\left(\alpha_1(1) + \sqrt{\alpha_1^2(1) + 4\alpha_2(1)}\right)$ and $r_2(1) = .5\left(\alpha_1(1) - \sqrt{\alpha_1^2(1) + 4\alpha_2(1)}\right)$. Then $\phi(1) = max(|r_1(1)|, |r_2(1)|)$ and $\omega(1) = min(|r_1(1)|, |r_2(1)||)$ are draws from the posterior distributions of ϕ and ω. This simple process can be repeated many times to obtain the posterior densities and moments for ϕ and ω.

Alternatively, one could analyze (22) directly, without first obtaining the posterior for $[\alpha_1, \alpha_2]$. In most cases one will want to specify only that $\omega(L)$ is invertible. A prior that does this is the uniform prior, $p(\omega) = .5$ over the range $-1 < \omega < 1$. To impose the prior belief that Y_t is stationary one could use the prior

$$p(\phi) \propto (1-\phi)^{a-1}(1+\phi)^{b-1} : \tag{28}$$

for $-1 < \phi < 1$: see Zellner (1971 p.190). Finally, adopt a diffuse, uniform prior on $\log(\sigma)$ so that $p(\sigma) \propto \sigma^{-1}$. Then, given a $T \times 1$ vector y of data and assuming ε_t to be independent normal, σ can be easily integrated out analytically. After integrating out σ, we have a joint posterior density for ϕ and ω that can be easily analyzed using bivariate numerical integration techniques. This will be needed since with the prior in (28) the integration over ϕ will be analytically intractable. Also, since bivariate numerical techniques are to be employed, it is possible to introduce a non-diffuse prior on ω.

Now assume a quarterly seasonal model for the error, $\omega(L) = 1 - \omega_4 L^4$ with the nonseasonal, subject-matter structure $\phi(L) = 1 - \phi L$, as before. This would

lead to the restricted $AR(5)$ model

$$(1 - \omega_4 L^4)(1 - \phi L)(Y_t - \mu) = \varepsilon_t \tag{29}$$

containing only three free parameters in contrast to an unrestricted $AR(5)$ model which contains six free parameters. Inferences would be obtained in the same fashion as in the earlier case.

The first step in building an ARAR model is to select a degree p for the polynomial $\phi(L)$ based on subject matter knowledge, previous research, and any other prior information. Data plots, sample autocorrelations and sample partial autocorrelations may also be useful at this stage, e.g. to confirm the presence of cycles or unit roots. Prior belief in the presence of cycles will lead to a specification permitting $\phi(L)$ to have complex roots. A tentative $AR(r)$ model for the error U_t should now be specified where r is believed to be large enough to make ε_t white noise. This will imply that the degree of $\alpha(L)$ is $p + r$. The appropriateness of the choice of r can be gauged from the residual autocorrelations from (14), using NLS, or from (15), using OLS ignoring the restrictions. Note that this procedure differs from the Box and Jenkins (1976) method which bases the lag lengths primarily on the sizes of sample autocorrelations and partial autocorrelations. Their procedure may be suitable if the researcher has no prior beliefs about $\phi(L)$, although it involves considerable pretesting that can have adverse effects upon subsequent inference; see Judge and Bock (1978). Of course, models initially based on features of a sample can in the future be rationalized by theory and confirmed with new samples.

In Bayesian analysis the choice of p and r should be guided by posterior

odds ratios. If the odds analysis leads to several favored models, they can all be carried along and results regarding parameters and predictions averaged over the models. Allowing for model uncertainty in this way often leads to forecasts which are superior to non-combined forecasts. It also avoids problems of pre-test bias and incorrect confidence intervals which arise from using the same data to select the model and to estimate its parameters. Let Θ_i be the vector of all q_i parameters in model i and $l(\Theta_i/y)$ be the likelihood function for model i. If the prior odds ratio is set to one, the posterior odds ratio becomes the Bayes factor $B_{1,2} = p_1(y)/p_2(y)$, where $p_i(y) = \int p(\Theta_i)l(\Theta_i/y)d\Theta$ is the predictive density for model i evaluated at the data y. For small models, like the one in the previous example, $p_i(y)$ can be calculated exactly: see Monahan (1983) for a discussion of $p_i(y)$ for ARMA models. Alternatively, the Schwarz (1978) criteria provides a large sample approximation

$$B_{1,2} \simeq T^{q/2}e^{lr}, \tag{30}$$

where the log-likelihood ratio $lr = \log[l(\hat{\Theta}_1/y)/l(\hat{\Theta}_2/y)]$, $\hat{\Theta}_i$ is the maximum-likelihood estimate of Θ from model i, and $q = q_2 - q_1$.

A common preliminary step in analyzing univariate time series is to test for the presence of unit roots and/or polynomial trends. The most popular model for this purpose is

$$\alpha(L)Y_t = \alpha_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t. \tag{31}$$

which is usually reparameterized to

$$Y_t = \alpha_0 + \beta_1 t + \beta_2 t^2 + \rho Y_{t-1} + \xi(L)\Delta Y_t + \varepsilon_t. \tag{32}$$

where t is a trend variable (a quadratic trend is shown above because it is useful in the example to follow), ρ is the sum of the coefficients of $\alpha(L)$ ($\alpha(1) = 1 - \rho$) and $\xi(L)$ is of degree one less than $\alpha(L)$ with $\xi_0 = 0$: see Hamilton (1994, p. 517). Unit root tests are sensitive to the presence of serial correlation in the error so the degree of $\alpha(L)$, and hence of $\xi(L)$, is set large enough to ensure that ε_t is white noise. In general there could be: several unit roots, roots exceeding 1.0 and conjugate roots with modulus of one or more. However, in many applications Y_t will be the natural logarithm of a variable of interest and the question is whether Y_t is stationary after removal of the polynomial trend as opposed to there being a single unit root in addition to the trend.

Note that (31) and (32) are just (15) with a quadratic trend added: i.e. the test regressions for popular unit root tests are already in a form consistent with an ARAR model. Also if Y_t really were generated by an ARMA(p,q) process it would impossible to find a finite AR model like (32) in which ε_t really was white noise. Thus, the ARAR model is superior to the ARMA model for unit root tests.

In this model the unit root, if it exists, can be in either the subject matter polynomial $\phi(L)$ or in the error polynomial $\omega(L)$. If $\alpha(L)$ has a unit root $\rho = 1$ and (32) becomes

$$\Delta Y_t = \alpha_0 + \beta_1 t + \beta_2 t^2 + \xi(L)\Delta Y_t + \varepsilon_t, \tag{33}$$

so that ΔY_t is stationary after removal of the polynomial trend. An important property of this model is that it allows for the simultaneous existence of a unit root and a polynomial trend.

There is another attractive, but less popular, way to introduce an intercept and trend into the ARAR model: i.e.

$$\alpha(L)(Y_t - \mu_0 - \mu_1 t - \mu_2 t^2) = \varepsilon_t. \tag{34}$$

Now the basic variable in the model is the detrended version of Y_t and the question is whether there is a unit root in the lag structure of this detrended variable. If we rewrite (34) as

$$\alpha(L)Y_t = [\mu_0\alpha(1) + \mu_1\kappa - \mu_2\eta] + [\mu_1\alpha(1) + 2\mu_2\kappa]t + \mu_2\alpha(1)t^2 + \varepsilon_t, \tag{35}$$

where $\kappa = \Sigma_{j=1}^{p+r} j\alpha_j$ and $\eta = \Sigma_{j=1}^{p+r} j^2\alpha_j$, we can see that (35) is also just (15) with a quadratic trend added. Inferences about ρ will be the same as with (32) however, in contrast to (32), the presence of a unit root in $\alpha(L)$, which makes $\alpha(1) = 0$, would exclude a quadratic trend from (35) leaving an intercept and linear trend only if $\mu_1 \neq 0$ and $\mu_2 \neq 0$. However, in the absence of a unit root (35) is a more convenient form for detrending Y_t if an ARMA model is to be entertained.

The ARAR model is also well suited to a more general investigation of the roots. If the unit root is believed to be in $\phi(L)$ which is specified to be of degree 2 or more, repeated draws can be made from the joint posterior of the ϕ coefficients and the roots calculated for each draw. Then the proportion of draws for which the roots are complex versus real and with modulus below 1.0 versus 1.0 or more can be easily calculated and are estimates of the posterior probabilities of these properties of $\phi(L)$. For applications of this technique see Geweke(1988) and Hong(1989).

2.3 Empirical Example: Housing Starts

One of the leading indicators published by the U.S. Department of Commerce is the number of permits issued by local authorities for the building of new private housing units. Pankratz (1983) has modelled this series using 84 seasonally adjusted, quarterly observations, from the first quarter of 1947 to the fourth quarter of 1967. He describes it as "an especially challenging series to model"[4]with Box-Jenkins procedures. We suspect that the absence of exogenous variables, such as real per capita income, the stock of housing and the real price of housing, from the univariate time series model accounts for much of the difficulty in modelling and, especially, forecasting this series. Nevertheless, it is still an attractive series to use in illustrating the ARAR technique.

Our first step in modelling this series as an ARAR was to form a prior belief about the degree of the $\phi(L)$ polynomial. Since this series is thought to lead the business cycle, we believed *a priori* that it should have a cycle. Therefore, we specified a second degree polynomial[5]for $\phi(L)$ and we chose a proper prior for its parameters, below, which placed a modest amount of probability on the region corresponding to conjugate complex roots. Because the data are quarterly, we chose an $AR(4)$ process as our tentative model for U_t. This will allow for imperfections in the seasonal adjustment process and for the presence in the error of omitted variables which are seasonally unadjusted. Thus we began with the model

$$(1 - \phi_1 L - \phi_2 L^2)(Y_t - \mu) = U_t, \tag{36}$$

with

$$(1 - \omega_1 L - \omega_2 L^2 - \omega_3 L^3 - \omega_4 L^4) U_t = \varepsilon_t. \tag{37}$$

From (36) and (37) we obtained the restricted AR model

$$\begin{aligned} Y_t &= \mu\phi(1)\omega(1) + (\phi_1 + \omega_1) Y_{t-1} + (\phi_2 + \omega_2 - \omega_1\phi_1) Y_{t-2} + \\ &\quad (\omega_3 - \omega_1\phi_2 - \omega_2\phi_1) Y_{t-3} + (\omega_4 - \omega_2\phi_2 - \omega_3\phi_1) Y_{t-4} - \\ &\quad (\omega_3\phi_2 + \omega_4\phi_1) Y_{t-5} - \omega_4\phi_2 Y_{t-6} + \varepsilon_t, \end{aligned} \tag{38}$$

which implies an unrestricted AR(6)

$$\begin{aligned} Y_t &= \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \alpha_4 Y_{t-4} + \\ &\quad \alpha_5 Y_{t-5} + \alpha_6 Y_{t-6} + \varepsilon_t. \end{aligned} \tag{39}$$

Note that if $\phi(L)$ has a unit root then $\phi(1) = 0$ and if $\omega(L)$ has a unit root then $\omega(1) = 0$. Thus, given a strong prior belief that $\mu \neq 0$ or a large sample mean for Y_t, an estimate of α_0 close to 0 would lead us to question the stationarity of the both Y_t and U_t.

Our sample of 191 observations was obtained from the U.S. Department of Commerce, Bureau of Economic Analysis, Survey of Current Business, October 1995 and January 1996. These data are seasonally adjusted index numbers (based in 1987) and extend from January 1948 until October 1995. We converted them to quarterly form by averaging over the months of each quarter: they are plotted in Figure 3. Neither Figure 3, nor the sample autocorrelations and partial autocorrelations[6], in Table 1, display the pattern typical of a unit root but they do show the cyclical pattern rather clearly. These results, and

those which follow, used the observations from the third quarter of 1949 until the third quarter of 1990, 165 observations. The earlier observations provided lagged values and the observations from the fourth quarter of 1990 to the third quarter of 1995 were used for post-sample forecasts, which are discussed below.

If our specification in (36) and (37) is adequate we should observe two things about the results from applying OLS to the unrestricted AR(6) (39). First, the residuals should be uncorrelated. Second, the roots of the estimated lag polynomial, $\hat{\alpha}(L)$, should contain a conjugate complex pair whose modulus is less than one and whose period is somewhere between 12 and 24 quarters, giving a damped cycle with a period of about three to six years. There may also be additional conjugate complex pairs of roots with moduli less than one but with shorter periods which model the dynamics of the error U_t.

The estimated lag polynomial, with standard errors in parentheses, is

$$\begin{array}{rcccccccccc} \hat{\alpha}(L) & = & 1 & - & 1.189L & - & .002388L^2 & + & .4343L^3 & - & .01077L^4 \\ & & & & (07685) & & (.1208) & & (.1236) & & (.1228) \\ & & & - & .2700L^5 & + & .2014L^6 & & & & \\ & & & & (.1228) & & (.07871) & & & & \end{array} \tag{40}$$

The sample autocorrelations and partial autocorrelations for the OLS residuals are shown in the second and third columns of Table 1: they indicate a lack of serial correlations. The roots of $\hat{\alpha}(L)$ plus their moduli and periods are given in Table 2. There are three complex conjugate pairs all of which have moduli less than one. The pair in the first line have a period of about 23 quarters which corresponds to our prior belief about the period of the business cycle. The other two pairs have shorter periods which we attribute to dynamics of

the error. Thus on the basis of these results it appears that our specification is adequate.

For purposes of illustration, unit root and trend analysis was carried out using both of the parameterizations (31) and (34). As noted above, inferences about unit roots will be the same for the two forms. An inspection of Figure 3 suggested that a weak quadratic trend might be present so trends of degree 0, 1 and 2 were used. A normal likelihood was assumed. The results are in Table 3 where the means of the posteriors are denoted by $\hat{\alpha}_0$, $\hat{\beta}_i$, $\hat{\rho}$ and $\hat{\mu}_i$ and the standard deviations by "Std Dev". $P_D(\rho \geq 1)$ is the posterior probability that there is a root of one or more when a diffuse prior is used while $P_J(\rho \geq 1)$ is obtained[7]when the Jefferys prior discussed in Phillips (1991) is used. These results lead us to infer $\phi(L)$ does not have a unit root and there is no polynomial trend[8].

If the parameters ϕ_i and ω_i are not of explicit interest and all that is wanted are forecasts of future values of Y_t they can easily be obtained by both Bayesian and frequentist analyses of the unrestricted $AR(6)$ of (39). If uniform, diffuse[9]priors for the α_i and $\log(\sigma^2)$ and a normal likelihood are assumed, OLS can be used, with suitable adjustment (see e.g. Zellner (1971)), to obtain the first two moments of the posterior distributions of the α_i.

Our primary interest here is to draw inferences about ϕ_1 and ϕ_2 with μ, ω_1, ω_2, ω_3, ω_4 and σ being nuisance parameters. For this purpose NLS is a widely used frequentist technique. If we assume diffuse priors and a normal likelihood, NLS also provides the means and standard deviations of the large-

sample normal approximations to the posteriors for the ϕ_i and ω_i. The results are shown in the second column of Table 4 where the means of the approximate posteriors are denoted by $\hat{\phi}_i$ or $\hat{\omega}_i$ and the standard deviations by "Std Dev". The approximate posterior mode for σ is denoted by $\hat{\sigma}$.

The posterior means of ω_2 and ω_3 are small compared to their standard deviations. We therefore imposed $\omega_2 \equiv \omega_3 \equiv 0$ and calculated the results shown in the third column of Table 4. Using the approximation in (30) we obtained a Bayes factor in favor of the restricted model against the unrestricted model of 32.5.

The results summarized in Table 4 can be used to find approximate posterior probabilities for various aspects of the dynamic behavior of the series. Such things as the existence of a cycle, whether a cycle is damped and the period of a cycle all depend on nonlinear functions of ϕ_1 and ϕ_2 . The first two terms in a Taylor series expansion of these functions about $\hat{\phi}_1$ and $\hat{\phi}_2$ provide linear approximations which have asymptotically normal posterior distributions. The roots of $\phi(L)$ are complex, leading to a cycle in housing starts, if $\phi_1^2 + 4\phi_2 < 0$. The parameters of the asymptotic posterior for this quantity are shown in Table 5, in the row labelled "cycle", for both the unrestricted and restricted ARAR(2,4) models: they show that the approximate posterior probability that $\phi(L)$ has complex roots is greater than .99. The cycle in housing starts will be damped if the modulus of the roots, given by $\sqrt{-\phi_2}$, is less than 1.0. The results in the row of Table 5 labelled "modulus" show that the approximate posterior probability that the cycles are damped is over .99 for both versions of the model.

The period of the cycle is given by $2\pi/\arctan(\vartheta)$, where $\vartheta = \sqrt{-\phi_1^2 - 4\phi_2}/\phi_1$. The asymptotic means and standard deviations for the period are shown in the last row of Table 5. Note that the entries in the modulus and period rows of Table 5 are close to those in first row of Table 2, confirming that we have correctly identified $\phi(L)$.

While the ARAR model is best suited to obtaining inferences about the parameters of interest, the ϕ_i, it is also useful to compare its forecasting performance with that of a traditional ARMA model. For the current example an ARMA model might be suggested by the fact that the data were converted from monthly to quarterly form by averaging. An ARMA(2,2) is the model obtained by applying Box-Jenkins model identification techniques to the sample autocorrelations and partial autocorrelations in Table 1. It is one of the models found by Pankratz[10]to be adequate on the grounds of its uncorrelated residuals and small forecast root-mean-square-errors and it retains the form of $\phi(L)$ suggested by our prior beliefs. The NLS results for this model appear in the last two columns of Table 4. The restriction $\theta_1 \equiv 0$, imposed to produce the results in the last column, was also imposed by Pankratz. The ARMA(2,2) estimates were harder to compute than the ARAR(2,4) estimates. Using a starting point derived from the sample autocorrelations and partial autocorrelations, the ARMA routine took 22 iterations to converge. In contrast, the ARAR routine started from 0.0 and took five iterations to converge. Also, as a vehicle for learning about ϕ_1 and ϕ_2 the restricted ARMA(2,2) is less successful than the restricted ARAR(2,4): the Q statistics are larger for the restricted ARMA(2,2),

the absolute ratios of the posterior means to standard deviations are larger for the restricted ARAR model and the approximate Bayes factor is 1.38 in favor of the restricted ARAR. Also, the ARMA(2,2) results have quite different dynamic properties than those of the ARAR(2,4) models: the approximate posterior probability of cycles for the unrestricted ARMA(2,2) is only .276 while that for the restricted ARMA(2,2) is only .401.

We obtained one-step-ahead forecasts by the restricted and unrestricted ARAR(2,4) model and by the restricted and unrestricted ARMA(2,2) for 20 quarters beyond the end of the sample. For the first forecast the coefficients were set at the sample period posterior means. Then for subsequent forecasts they were updated for each period. Since actual data for this period had been left out of the sample used in estimation, we were able to calculate the forecast errors for each of the 20 quarters in the forecast period. Summary statistics for these forecast errors, measured as percentages of the actual future values of Y_t, are shown in Table 6. The main difference between the methods is in the size of the mean percentage forecast errors, which were larger for the ARAR models. This made the RMSE for the unrestricted ARAR(2,4) larger than that for the unrestricted ARMA(2,2), since there standard deviations were the same. However, the standard deviation of the restricted ARAR percentage forecast errors was smaller than that of the restricted ARMA forecast errors so that in this case ARAR has a smaller RMSE. Of course, these forecasts are rather imprecise because, as noted above, obvious exogenous variables have been omitted and no account has been taken of long cycles in building activity. Nevertheless, this

example does illustrate that the ARAR model has the potential to give forecasts more precise than the traditional ARMA model.

Although a sample of size 165 may seem large enough to justify large-sample approximate posteriors, it is still useful to consider the calculation of exact posteriors for the ϕ_i and ω_i. The procedure we followed is an adaptation of Zellner (1971) Chapter IV. We will discuss exact posteriors for ϕ_1, ϕ_2, ω_1 and ω_4 from the restricted ARAR(2,4) model of column 3 in Table 4[11].

Let $\phi' = [\phi_0, \phi_1, \phi_2]$, where $\phi_0 = \mu\phi(1)$, and[12] $\omega' = [\omega_1, \omega_4]$. Our prior on ϕ, ω and σ was

$$p(\phi, \omega, \sigma) = p(\phi/\sigma)p(\omega/\sigma)p(\sigma). \tag{41}$$

An inverted gamma distribution was used for $p(\sigma)$ with parameters $\underline{s}^2 = 1.0$ and $\underline{v} = 1$ giving a mode of .7071. We set $p(\omega/\sigma) \propto const.$, a diffuse prior, in keeping with our lack of prior knowledge regarding the behavior of the error U_t.

We chose a normal form for $p(\phi/\sigma)$ with

$$E(\phi/\sigma)' = \underline{\phi}' = [0, 1.0, -.5]$$

and covariance matrix $V(\phi/\sigma) = \sigma^2 \underline{W}^{-1}$ with

$$\underline{W}^{-1} = \begin{pmatrix} 400 & 0 & 0 \\ 0 & 16 & -8 \\ 0 & -8 & 16 \end{pmatrix}.$$

This prior is centered in the region corresponding to complex roots for $\phi(L)$ but it is still rather uninformative with respect to ϕ_1 and ϕ_2. For example, conditional on $\sigma = .7071$, the prior probability of obtaining complex roots is only about .12.

Let the vector of initial values of Y_t be $\mathbf{y}_\mathbf{0}' = [y_1, \ldots, y_{p+r}]$ and the vector of the T remaining values be $\mathbf{y}' = [y_{p+r+1}, \ldots, y_{p+r+T}]$. For this example $p = 2$, $r = 4$, $T = 165$ and all inferences are conditional on $\mathbf{y_0}$. Define $\mathbf{y}(\omega)$ as the $T \times 1$ vector with elements $y_t(\omega) = \omega(L)y_t$ and $\mathbf{z}(\omega)$ as the $T \times k$ matrix with rows $z_t'(\omega) = [\omega(1), y_{t-1}(\omega), y_{t-2}(\omega)]$, where $k = p + 1$. Assume $\varepsilon_t \sim IN(0, \sigma^2)$. Then the joint posterior, after completing the square on ϕ, can be written as

$$p(\phi, \omega, \sigma/\mathbf{y}, \mathbf{y_0}) \propto \sigma^{-(T+\underline{v}+k+1)} \times \exp\left\{ -\frac{(\tilde{v}\widetilde{s^2(\omega)} + [\phi - \widetilde{\phi(\omega)}]'[\underline{W} + \mathbf{z}'(\omega)\mathbf{z}(\omega)][\phi - \widetilde{\phi(\omega)}])}{2\sigma^2} \right\}$$

where: $\tilde{v} = T + \underline{v}$; $\widetilde{\phi(\omega)} = [\underline{W} + \mathbf{z}'(\omega)\mathbf{z}(\omega)]^{-1}[\underline{W}\underline{\phi} + \mathbf{z}'(\omega)\mathbf{y}(\omega)]$ and

$$\tilde{v}\widetilde{s^2(\omega)} = \underline{v}\underline{s}^2 + \underline{\phi}'\underline{W}\underline{\phi} + \mathbf{y}'(\omega)\mathbf{y}(\omega) - \widetilde{\phi(\omega)}'[\underline{W} + \mathbf{z}'(\omega)\mathbf{z}(\omega)]\widetilde{\phi(\omega)}.$$

Note that the use of the proper prior $p(\phi/\sigma)$ in (41) ensures that $[\underline{W} + \mathbf{z}'(\omega)\mathbf{z}(\omega)]$ will be nonsingular even if $\omega(1) = 0$, thus avoiding the need to confine attention to models with $\mu \equiv 0$ or to impose the restriction $\omega(1) \neq 0$: see Zellner (1971) page 89.

Integrating out σ^2 gives

$$p(\phi, \omega/\mathbf{y}, \mathbf{y_0}) \propto \left[\widetilde{s^2(\omega)}\right]^{-(\tilde{v}+k)/2} \left\{\tilde{v} + [\phi - \widetilde{\phi(\omega)}]'\mathbf{H}(\omega)[\phi - \widetilde{\phi(\omega)}]\right\}^{-(\tilde{v}+k)/2} \tag{42}$$

where $\mathbf{H}(\omega) = \left[\widetilde{s^2(\omega)}\right]^{-1} [\underline{W} + \mathbf{z}'(\omega)\mathbf{z}(\omega)]$.

Integrating (42) over ϕ gives the joint posterior density for ω_1 and ω_4.

$$p(\omega/\mathbf{y}, \mathbf{y_0}) \propto \left[\tilde{v}\widetilde{s^2(\omega)}\right]^{-\tilde{v}/2} |\underline{W} + \mathbf{z}'(\omega)\mathbf{z}(\omega)|^{-1/2}. \tag{43}$$

However, the lack of identification mentioned above means that (43) could also

be the posterior density for ϕ_1 and ϕ_2. This composite posterior, scaled for graphical clarity, is shown in Figure 4. Identification was achieved by imposing our prior belief that the values of ϕ_1 and ϕ_2 are such that $\phi(L)$ has conjugate complex roots leading to cyclical behavior in y_t. Thus we identified the higher of the two "hills" in the background in Figure 4, which lies over a region leading to complex roots, as the joint posterior for ϕ_1 and ϕ_2 and the lower of the "hills" in the foreground, which lies over a region leading to real roots, as the joint posterior for ω_1 and ω_4.

Since (43) is difficult to integrate analytically, the marginal posteriors for ω_1, ω_4, ϕ_1 and ϕ_2 were obtained numerically. To obtain the marginal posterior for either ω_1 or ω_4 we numerically integrated[13]the portion of (43) which we have identified as the posterior of ω over the other ω_i. Marginal posteriors for the ϕ_i were obtained by noting that, conditional on ω, (42) is in the multivariate Student-t form so the conditional posterior densities of the ϕ_i, given ω, are in the univariate Student-t form which can be obtained analytically. Then the products of these conditional densities and the joint density for ω in (43) are the conditional densities of the ϕ_i, given ω. The marginal densities of the ϕ_i were obtained by integrating these products over ω_1 and ω_4. Alternatively, Gibbs sampling procedures could have been employed to compute these integrals[14].

The marginal posteriors for ω_1 and ω_4 are plotted in Figures 5 and 6, those for ϕ_1 and ϕ_2 are in Figures 7 and 8. In each figure the exact posterior is graphed with a solid line while the large-sample normal approximate posterior, moments of which are given in Table 4, is graphed with a dashed line. The exact posterior

moments are given in Table 7. The exact and approximate posterior densities differ because they are based on different priors and because the sample is only moderately large. However, they are quite close in the case of ϕ_1 and ϕ_2, which are the parameters of most interest.

This example has illustrated several advantages of the ARAR model over the ARMA model. Estimates of the ARAR parameters were easier to compute, taking many fewer iterations than ARMA estimates. There was more evidence of serial correlation in the ARMA residuals than in the ARAR residuals and the Bayes factor favored the ARAR model. The NLS point estimates of the ARAR parameters were in closer accord with our prior beliefs than were the estimates or the ARMA parameters. Furthermore, the dynamic properties of the estimated ARAR model were far more satisfactory than those of the estimated ARMA model. Also, the restricted ARAR model had percentage forecast errors with smaller RMSE than did the restricted ARMA model. Exact posterior densities for the ARAR parameters were easily obtained and were close to the asymptotic normal posteriors derived from the NLS estimates. In the next section these techniques will be extended to single equation, distributed lag models.

3 Distributed Lag Models

3.1 The Rational Distributed Lag Model

Jorgenson (1966) considered the general distributed lag model

$$Y_t = \mu + \lambda(L)x_t + V_t \tag{44}$$

where x_t is the value of an exogenous variable X_t and the lag polynomial $\lambda(L)$ is infinitely long. Originally the error was modelled as $V_t = \varepsilon_t$, white noise. Then the infinite lag structure on x_t was approximated as $\lambda(L) = \delta(L)/\phi(L)$ with $\delta(L) = \delta_0 + \delta_1 L + \ldots + \delta_m L^m$ and $\phi(L)$ defined as in (1) having no factors in common with $\delta(L)$. With these assumptions (44) becomes

$$Y_t = \mu + \frac{\delta(L)}{\phi(L)} x_t + \varepsilon_t \tag{45}$$

or

$$\phi(L)Y_t = \mu_0 + \delta(L)x_t + \phi(L)\varepsilon_t \tag{46}$$

where $\mu_0 = \phi(1)\mu$. The parameters of interest in (45) are the coefficients of the lag polynomials $\phi(L)$ and $\delta(L)$, the ϕ_i and δ_i. The disadvantage of this model for V_t are similar to those of the previous section. First, a white noise error in (44) is often too restrictive in that it confines the effect of a shock ε_t on $Y_t - \mu - \lambda(L)x_t$ to only the current period. Secondly, it implies an MA form for the error term in (46) leading to complicated inference procedures. Note that these complications are avoided if the error term $\phi(L)\varepsilon_t$ is white noise; see Zellner and Geisel (1970).

3.2 The Transfer Function Model

Specify the same infinite lag structure on x_t as in (44) with $\lambda(L) = \delta(L)/\phi(L)$, but model V_t as a stationary ARMA;

$$\pi(L)V_t = \theta(L)\varepsilon_t. \tag{47}$$

This leads to the Box and Jenkins(1976) transfer function model

$$\pi(L)\phi(L)Y_t = \mu_0 + \pi(L)\delta(L)x_t + \phi(L)\theta(L)\varepsilon_t. \tag{48}$$

Now the error model is more general but it's even more difficult to draw inferences about the parameters of interest, ϕ_i and δ_i, because of the MA error in (48). Also it is important not to overparameterize the pairs of polynomials $\delta(L)$, $\phi(L)$ and $\pi(L)$, $\theta(L)$ as this can lead to the existence of common factors in one or both of these pairs which implies a lack of identification.

3.3 The ARMAX Model

The ARMAX model can be obtained by either imposing $\pi(L) \equiv \phi(L)$ on the transfer function error model (47) or by adding a distributed lag in x_t to the univariate ARMA model (5) leading to

$$\phi(L)Y_t = \mu_0 + \delta(L)x_t + \theta(L)\varepsilon_t, \tag{49}$$

It is assumed that $\phi(L)$, $\delta(L)$ and $\theta(L)$ contain no common factors. This model is somewhat simpler to analyze than the transfer function model but it retains the inconvenient MA error structure.

3.4 The ARAR Distributed Lag Model (ARDLAR)

Our general distributed lag model is the ARAR model (1) for univariate time series extended by the addition of a distributed lag in x_t to obtain

$$\phi(L)Y_t = \mu_0 + \delta(L)x_t + U_t, \tag{50}$$

We retain the assumption from the ARAR univariate model that the degree of $\phi(L)$ can be specified *a priori*, at least approximately, from subject matter knowledge and we extend that assumption to $\delta(L)$. We also assume that $\phi(L)$ is invertible and has no factors in common with $\delta(L)$. Then we can write the model as

$$\begin{aligned} Y_t &= \frac{\mu_0}{\phi(1)} + \frac{\delta(L)}{\phi(L)}x_t + \frac{1}{\phi(L)}U_t \\ &= \mu + \lambda(L)x_t + V_t. \end{aligned} \tag{51}$$

This model retains the same infinite lag structure, $\lambda(L) = \delta(L)/\phi(L)$, on x_t as (44), (48) and (49). However, in contrast to these models, we specify the same stationary AR process for U_t as in the univariate case:

$$\omega(L)U_t = \varepsilon_t \tag{52}$$

and

$$\phi(L)V_t = U_t \tag{53}$$

Then (44) becomes[15]

$$Y_t = \mu + \frac{\delta(L)}{\phi(L)}x_t + \frac{1}{\phi(L)\omega(L)}\varepsilon_t, \tag{54}$$

which we label the $ARDLAR(p, m, r)$ model, where p is the degree of $\phi(L)$, m is the degree of $\delta(L)$ and r is the degree of $\omega(L)$. We can rewrite (54) as

$$\omega(L)\phi(L)Y_t = \alpha_0 + \omega(L)\delta(L)x_t + \varepsilon_t, \tag{55}$$

where $\alpha_0 = \mu\omega(1)\phi(1)$. This is a simpler model than the transfer function model but it produces nearly the same behavior in Y_t if $\pi(L)$ is approximately

$\theta(L)\omega(L)$. Our ARDLAR model has several desirable features. It permits a very general process for both V_t and U_t as both $\lambda(L)$ and $1/(\omega(L)\phi(L))$ can be infinite in length and there is no need to assume that U_t is white noise. Since distributed lag models are often implemented using monthly or quarterly data, our model is especially convenient because it allows seasonal behavior to be included in $\omega(L)$, which could be the product of seasonal and nonseasonal polynomials.

Because the ARDLAR lag model is an extension of the univariate ARAR model, it is not surprising that it may require the same sort of prior information to identify the ϕ_i and ω_i parameters. To see what identifying restrictions are necessary write (55) as

$$\alpha(L)Y_t = \alpha_0 + \beta(L)x_t + \varepsilon_t \tag{56}$$

where $\alpha(L) = \omega(L)\phi(L)$ and $\beta(L) = \omega(L)\delta(L)$. In contrast to the univariate case, interchanging $\phi(L)$ and $\omega(L)$ in (55) will change the likelihood function, which is based on (56), except in the special case that $\phi(L) = \omega(L)$. Similarily, interchanging $\phi(L)$ and $\delta(L)$ will change the likelihood function, unless $\phi(L) = \delta(L)$, and so will interchanging $\omega(L)$ and $\delta(L)$, unless $\omega(L) = \delta(L)$. Thus, so long as $\delta_0 \neq 1.0$, a sufficient condition for identification is that $\phi(L)$ and $\omega(L)$ be of different degrees. Two other sources of identification failure are removed by making two assumptions. First, as in the univariate case, we will assume that $\phi(L)$ has been normalized to have its first term equal to 1.0. Secondly, as mentioned above, we assume that there is no invertible polynomial $\upsilon(L)$ appearing as a common factor in $\phi(L)$ and $\delta(L)$.

Now assume that $\alpha(L)$ and $\beta(L)$ are known but $\phi(L)$, $\omega(L)$ and $\delta(L)$ are unknown. Write $\alpha(L)$ in terms of its roots, η_i, as in (21). Given the values of η_i we would use prior information in the same way as with the univariate ARAR to decide which $(1 - \eta_i L)$ are part of $\phi(L)$ and which are part of $\omega(L)$. Once $\omega(L)$ had been identified in this way it could be factored out of $\beta(L)$ to leave $\delta(L)$. In practice, the roots η_i and the polynomial $\beta(L)$ are unknown. Given a large enough sample, estimates of them may provide some guidance as to identification.

Note that (56) is the autoregressive-distributed-lag (ARDL) model much favored by Hendry: see *inter alia* Hendry *et al* (1986). However, the two models are built up in different ways. The ARDLAR model begins with fairly precise prior beliefs about the systematic or subject matter dynamics represented by $\phi(L)$ and $\delta(L)$. It then explicitly introduces the common factor $\omega(L)$ as the model for the error U_t. Prior beliefs about $\omega(L)$ are often diffuse. The parameters of interest are the ϕ_i and δ_i while the ω_i and σ are nuisance parameters

In contrast, ARDL model begins with the unrestricted polynomials $\alpha(L)$ and $\beta(L)$ and tries, through COMFAC analysis, to discover if they contain any common factors, like $\omega(L)$. There are no prior beliefs regarding $\phi(L)$ and $\delta(L)$. Any common factor lag polynomials that are discovered are then put into an AR model for the error, resulting in an ARDLAR model. In this procedure the parameters of interest are initially the α_i and β_i but if common factors are discovered interest presumably switches to the ϕ_i and δ_i. Note that since this procedure uses the same data over several rounds of testing, it is subject to the

pre-test problems mentioned above.

Assuming identification, the common factor $\omega(L)$ in (55) imposes a restriction linking the parameters in $\alpha(L)$ and $\beta(L)$. To exploit this restriction, first condition on the coefficients of $\omega(L)$ to obtain

$$\phi(L)[\omega(L)Y_t] = \delta(L)[\omega(L)x_t] + \varepsilon_t. \tag{57}$$

Then condition on the elements of $\phi(L)$ to obtain

$$\omega(L)[\phi(L)Y_t] = \omega(L)[\phi(L)x_t] + \varepsilon_t. \tag{58}$$

Equations (57) and (58) can be used to generate NLS estimates conveniently. They can also provide the basis for the application of the Gibbs sampler in some cases.

An interesting simple example is obtained by assuming that the degrees of $\delta(L)$ and $\phi(L)$ are both one so that

$$Y_t = \mu + \frac{(\delta_0 + \delta_1 L)}{(1 - \phi L)} x_t + V_t \tag{59}$$

with the restriction $\mid \phi \mid < 1$. Assume that observations will be taken quarterly without seasonal adjustment. Then an attractive AR model for U_t is the product of a nonseasonal $AR(1)$ and a seasonal $AR(1)$:

$$\begin{aligned} \omega(L)U_t &= (1 - \omega_1 L)(1 - \omega_4 L^4)U_t \\ &= \varepsilon_t \end{aligned} \tag{60}$$

with the restrictions $\mid \omega_1 \mid < 1$ and $\mid \omega_4 \mid < 1$. Thus our model for U_t is a restricted $AR(5)$ with two free parameters and our model for V_t is a restricted

$AR(6)$ with three free parameters (ϕ, ω_1 and ω_4) plus σ^2, the variance of ε_t. Note that since $\delta(L) \neq \delta_0$ and the degree of $\omega(L)$ is different from the degree of $\phi(L)$, ϕ, ω_1 and ω_4 are identified, as are δ_0 and δ_1. The model for the observable random variable Y_t can be written as

$$(1-\omega_1 L)(1-\omega_4 L^4)(1-\phi L)Y_t = \alpha_0 + (1-\omega_1 L)(1-\omega_4 L^4)(\delta_0 - \delta_1 L)x_t + \varepsilon_t \quad (61)$$

or as the restricted ARDL model

$$Y_t = \alpha_0 + \Sigma_{j=1}^{6}\alpha_j Y_{t-j} + \Sigma_{j=0}^{6}\beta_j x_{t-j} + \varepsilon_t, \quad (62)$$

where the restrictions are

$$\begin{aligned}
\alpha_0 &= \mu\phi(1)\omega(1) & \beta_0 &= \delta_0 \\
\alpha_1 &= \omega_1 + \phi & \beta_1 &= \delta_1 - \omega_1\delta_0 \\
\alpha_2 &= -\omega_1\phi & \beta_2 &= -\omega_1\delta_1 \\
\alpha_3 &= 0 & \beta_3 &= 0 \\
\alpha_4 &= \omega_4 & \beta_4 &= -\omega_4\delta_0 \\
\alpha_5 &= -\omega_4(\omega_1 + \phi) & \beta_5 &= \omega_4(\omega_1\delta_0 + \delta_1) \\
\alpha_6 &= \omega_1\omega_4\phi & \beta_6 &= \omega_1\omega_4\delta_1
\end{aligned} \quad (63)$$

To take advantage of the common factors in (62) first condition on ω_1 and ω_4 to obtain

$$\begin{aligned}
[Y_t - \omega_1 Y_{t-1} - \omega_4 Y_{t-4} + \omega_1\omega_4 Y_{t-5}] &= \phi[Y_{t-1} - \omega_1 Y_{t-2} - \omega_4 Y_{t-5} + \omega_1\omega_4 Y_{t-6}] \\
&+ \delta_0[x_t - \omega_1 x_{t-1} - \omega_4 x_{t-4} + \omega_1\omega_4 x_{t-5}]
\end{aligned}$$

$$+ \quad \delta_1[x_{t-1} - \omega_1 x_{t-2} - \omega_4 x_{t-5} + \omega_1\omega_4 x_{t-6}]$$

$$+ \quad \varepsilon_t. \tag{64}$$

Then condition on ϕ, δ_0 and δ_1 to obtain

$$\begin{aligned}
[Y_t - \phi Y_{t-1} - \delta_0 x_t + \delta_1 x_{t-1}] &= \omega_1[Y_{t-1} - \phi Y_{t-2} - \delta_0 x_{t-1} + \delta_1 x_{t-2}] \\
&+ \omega_4[Y_{t-4} - \phi Y_{t-5} - \delta_0 x_{t-4} + \delta_1 x_{t-5}] \\
&- \omega_1\omega_4[Y_{t-5} - \phi Y_{t-6} - \delta_0 x_{t-5} + \delta_1 x_{t-6}] \\
&+ \varepsilon_t. \qquad (65)
\end{aligned}$$

These two equations can form the basis for application of the Gibbs sampler in some cases.

The first step in building an ARDLAR model is to use subject matter knowledge, previous research, and any other prior information to select p, the degree of $\phi(L)$, and m, the degree of $\delta(L)$. These selections should reflect prior belief about the presence of cycles or trends and about the pattern of the coefficients, λ_i, in the ratio $\lambda(L) = \delta(L)/\phi(L)$. A tentative AR model of degree r for the error U_t should now be chosen. This will imply that $\alpha(L)$ is of degree $p + r$ and $\beta(L)$ is of degree $m + r$. The adequacy of the choice of r can be checked by examining the residuals from the OLS fit of (56). If the initial choice of r was found to be too large it can be reduced, although this leads to the pre-test problems noted above. If the aim of the model is only to produce forecasts one may wish to stop at this point. However, if the aim is to gain knowledge about the parameters of $\phi(L)$ and $\delta(L)$ the procedures discussed above and illustrated below should be followed.

3.5 The Error Correction Model (ECM)

This popular model is simply another form of a restricted ARDL model (62): see Banerjee *et al* (1993) or Hendry *et al* (1986). For the simple example discussed above, where the AR process for the error is the product of a non-seasonal $AR(1)$ and a seasonal $AR(1)$, the ECM can be written as[16]

$$\begin{aligned} \Delta Y_t &= \alpha_0 + \beta_0 \Delta x_t + (\alpha_1 - 1)(Y_{t-1} - x_{t-1}) \\ &+ \Sigma_{j=2}^{6} \alpha_j (Y_{t-j} - x_{t-j}) + (\alpha_1 - 1 + \beta_0 + \beta_1) x_{t-1} \\ &+ \Sigma_{j=2}^{6} (\alpha_j + \beta_j) x_{t-j} + \varepsilon_t, \end{aligned} \tag{66}$$

where terms involving α_3 and $\alpha_3 + \beta_3$ drop out.

If nonseasonal data are used the model can be simplified by removing the seasonal component in U_t so that $\omega_4 = 0$. Then $\alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ and (66) simplifies to

$$\begin{aligned} \Delta Y_t &= \alpha_0 + \beta_0 \Delta x_t + (\alpha_1 - 1)(Y_{t-1} - x_{t-1}) \\ &+ \alpha_2 (Y_{t-2} - x_{t-2}) + (\alpha_1 - 1 + \beta_0 + \beta_1) x_{t-1} \\ &+ (\alpha_2 + \beta_2) x_{t-2} + \varepsilon, \end{aligned} \tag{67}$$

This illustrates the utility of using prior information in specifying the model.

3.6 The Partial Adjustment Model

The partial adjustment model provides the short run dynamic response to a departure from long run equilibrium: e.g. see Harberger(1960), Jorgenson(1965), Lintner(1967), Nerlove(1958), Phillips(1957) and Pagan(1985).

$$Y_t - Y_{t-1} = \mu^* + \phi^*(y_t^* - Y_{t-1}) + U_t, \tag{68}$$

where y_t^* is the long run equilibrium value of the random variable Y_t and $0 \leq \phi^* < 1$ is the speed of the short run dynamic response: i.e. ϕ^* is an error correction parameter. The simplest form of this model sets

$$y_t^* = \psi_0 + \psi_1 x_t. \tag{69}$$

Let $\phi = 1 - \phi^*$, $\delta_0 = \phi^*\psi_1$ and $\mu_0 = \mu^* + \phi^*\psi_0$. Then substitute (69) into (68) to obtain

$$(1 - \phi L)Y_t = \mu_0 + \delta_0 x_t + U_t. \tag{70}$$

and

$$Y_t = \mu + \frac{\delta_0}{(1 - \phi L)} x_t + \frac{1}{(1 - \phi L)} U_t \tag{71}$$

It is common to model U_t as white noise so that $\omega(L) = 1$ and (70) is in the same form as (56); i.e. the partial adjustment model is a simple ARDLAR model. Note that $\delta(L) \neq 0$ and the degree of $\omega(L)$ is 0 while the degree of $\phi(L)$ is 1 so ϕ and δ_0 are identified.

3.7 The Adaptive Expectations Model

This is just a special case of the rational distributed lag model. Consider the simple case

$$Y_t = \mu + \psi x_{t+1}^* + V_t, \tag{72}$$

where x^*_{t+1} is the expectation of the future value of an exogenous variable X_t which is formed according to

$$x^*_{t+1} = \phi x^*_t + (1-\phi)x_t, \tag{73}$$

with $0 \leq \phi \leq 1$. Substituting (73) into (72) gives

$$Y_t = \mu + \frac{\delta_0}{(1-\phi L)}x_t + V_t, \tag{74}$$

where $\delta_0 = \psi(1-\phi)$, which is in the form of (44).

Usually V_t is modelled as white noise which makes (74) the same as (45) with the same disadvantages. Indeed this error specification is the only difference between the adaptive expectations model (74) and the partial adjustment model (71). However if, as in (54), we model the error V_t as

$$\omega(L)(1-\phi L)V_t = \varepsilon_t. \tag{75}$$

then

$$\omega(L)(1-\phi L)Y_t = \delta_0\omega(L)x_t + \varepsilon_t \tag{76}$$

which is identical to the partial adjustment model (71). This model has been analyzed by Zellner and Geisel (1970) and by Zellner (1971, Chapter 7) who give the joint posterior density function for ϕ and ω based on a normal likelihood, diffuse priors and various models for V_t and U_t.

3.8 The Finite Distributed Lag Model

In many cases an infinite lag structure for the observable causal component will *a priori* be unattractive. Such cases are easily handled by restricting the

rational distributed lag model by imposing $\phi(L) \equiv 1.0$, which serves to identify the ω_i. Now (54) becomes

$$Y_t = \delta(L)x_t + \frac{1}{\omega(L)}\varepsilon_t \quad (77)$$

$$= \delta(L)x_t + U_t \quad (78)$$

which is just a linear regression model with AR errors for which inferences procedures are relatively straightforward.

3.9 Empirical Example: Growth of Real GDP

The AR(3) model has often been used in forecasting real GDP or its rate of growth, as in Geweke(1988). However, Garcia-Ferrer *et al*(1987), Zellner and Hong(1989), Hong(1989) and Zellner *et al*(1991) have shown that AR(3) models give very poor forecasts of turning points. Their forecasting performance can be improved by converting them to distributed lag models which include lagged endogenous or exogenous variables as leading indicators. We show in this example how a further extension to an ARDLAR is an attractive model in this context.

Let Y_t be the first difference of the log of real GDP. We begin by building univariate AR and ARAR models for Y_t to further illustrate the ARAR technique and to provide a basis for comparison with the subsequent ARDLAR model. We then introduce two leading indicators and use them as the exogenous variables in an ARDLAR model.

We follow earlier literature in specifying the first of our univariate models as

an AR(3), which is an example of the unrestricted AR model (15). *A priori* we believe that Y_t follows a damped cycle so the first step in building our ARAR model was to specify $\phi(L)$ as a second degree polynomial. An $\omega(L)$ of degree one when combined with the $\phi(L)$ of degree two led to an $\alpha(L)$ of degree three, as in the unrestricted AR(3) above. Thus we specify our ARAR model as

$$(1 - \omega_1 L)(1 - \phi_1 L - \phi_2 L^2)Y_t = \alpha_0 + \varepsilon_t. \tag{79}$$

Y_t was constructed from quarterly, seasonally adjusted, observations on United States GDP in millions of 1987 dollars from the fourth quarter of 1949 to the fourth quarter of 1990; 165 observations. Observations from the first quarter of 1991 until the third quarter of 1995 were held back for use in calculating forecast errors. The results of trend and unit roots analyses appear in the top panel of Table 8 which used the same techniques as were used to obtain the top panel of Table 3. The number of lags used, three, was the same as the number of lags in (79). As in Table 3, $P_D(\rho \geq 1)$ and $P_J(\rho \geq 1)$ are the posterior probabilities of a root of one or more using diffuse and "Jeffreys" priors, respectively. They are all zero to eight significant digits.

OLS results for the unrestricted AR(3) model are reported in the second column of Table 9. We view them as the moments of posterior distributions based on uniform priors and a normal likelihoods. In the top panel of Table 10 we present the roots of $\hat{\alpha}(L)$, their moduli and periods. They support our *a priori* belief in a damped cycle.

NLS results for the ARAR(2,1) model are shown in column three of Table 9. Note that the AR(3) point estimates obey the ARAR(2,1) restrictions exactly

so the ln likelihoods and residual diagnostics in columns two and three are the same. Also the estimated roots, moduli and periods in the second panel of Table 10 are the same as those in the first panel. Note too that the posterior mean for α_3 obtained by OLS, in column two of Table 9, is small compared to its standard deviation but the asymptotic posterior means of both ϕ_2 and ω_1, whose product equals α_3, are large compared to their asymptotic standard deviations. Thus setting α_3 to 0, in an attempt to increase statistical efficiency, would be a specification error. Table 11 uses the same notation as Table 5 to present posterior results for the dynamic properties of the estimated $\phi(L)$ from Table 9. The asymptotic posterior probability of the presence of a cycle is .72.

We now introduce two leading indicator variables which have been used in the past by Garcia-Ferrer *et al*(1987), Zellner and Hong(1989), Hong(1989) and Zellner*et al*(1991); the rate of growth of real M2 and the real rate of return on stocks. Data on the value of M2 in billions of 1987 dollars and on an index of the prices of 500 common stocks were both taken from U.S. Department of Commerce, Bureau of Economic Analysis, Survey of Current Business, October 1995 and January 1996. These series are both monthly while the GDP series are quarterly. Several methods are available to transform the monthly series to quarterly series: averaging over the quarter, taking the median over the quarter, or taking one of the months to represent the quarter. We used the value for the third month in each quarter as the value for that quarter. Then the rate of growth of real M2, $x_{1,t}$, was obtained as the first difference of the log of M2. The index of common stock prices was deflated by the consumer price

index and the real rate of return on stocks, $x_{2,t}$ was the obtained as the first difference of the log of the deflated series. The results of trend and unit roots analyses for $x_{1,t}$ and $x_{2,t}$ appear in the second and third panels of Table 8. We refrained from using prior information as to the appropriate lag length for $x_{1,t}$ and $x_{2,t}$; instead we chose the lag length which gave the minimum BIC for an unrestricted AR model. This was a lag of one in both cases. Here too, all the posterior probabilities of stochastic nonstationarity were zero to eight significant digits.

To obtain an ARDLAR model we must add distributed lags in $x_{1,t}$ and $x_{2,t}$ to the ARAR model (79). In making these additions we specified the lag polynomials so as to: reflect the character of $x_{1,t}$ and $x_{2,t}$ as leading indicators; take account of the fact that the reporting lag before data on the growth of M2 are publicly available is much longer than the reporting lag for real stock returns; take advantage of the way in which monthly values of $x_{1,t}$ and $x_{2,t}$ were converted to quarterly values. This produced a generalization of (55) which we label ARDLAR(p, m_1, m_2, r)

$$\omega(L)\phi(L)Y_t = \alpha_0 + \omega(L)\delta_1(L)x_{1,t} + \omega(L)\delta_2(L)x_{2,t} + \varepsilon_t \tag{80}$$

where: $\omega(L)$ and $\phi(L)$ are specified as in the ARAR model (79); $\delta_1(L) = \delta_{1,2}L^2$ (so $m_1 = 2$ with $\delta_{1,0} = \delta_{1,1} = 0$) and $\delta_2(L) = \delta_{2,1}L$ (so $m_2 = 1$ with $\delta_{2,0} = 0$). We can also write (80) in the form of the restricted ARDL(3,3,2) model

$$Y_t \quad = \quad \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} +$$

$$\beta_{1,2}x_{1,t-2} + \beta_{1,3}x_{1,t-3} + \beta_{2,1}x_{2,t-1} + \beta_{2,2}x_{2,t-2} + \varepsilon_t. \qquad (81)$$

where the restrictions are

$$\alpha_0 = \mu\phi(1)\omega(1) \qquad \alpha_1 = \omega_1 + \phi_1$$

$$\alpha_2 = \phi_2 - \omega_1\phi_1 \qquad \alpha_3 = -\omega_1\phi_2$$

$$\beta_{1,2} = \delta_{1,2} \qquad \beta_{1,3} = -\omega_1\delta_{1,2}$$

$$\beta_{2,1} = \delta_{2,1} \qquad \beta_{2,2} = -\omega_1\delta_{2,1}.$$

If, for example, Y_t pertains to the fourth quarter of 1970 then $x_{1,t-2}$ pertains to June 1970 and $x_{2,t-1}$ pertains to September 1970: i.e. there is a delay of three months before $x_{1,t}$ has an impact on Y_t and a delay of only one month before $x_{2,t}$ has an impact. Note that there is no identification problem with this model.

NLS results for this model are in the fourth column of Tables 9. They show that the addition of the distributed lags in $x_{1,t}$ and $x_{2,t}$ was useful because the posterior means $\hat{\delta}_{1,2}$ and $\hat{\delta}_{2,1}$ are large compared to the posterior standard deviations and because the residual variance was reduced. Also the approximate Bayes factor is 36100 in favour of the ARDLAR(2,2,1,1) over the ARAR(2,1). The estimated roots, plus their modulus and period, of $\hat{\alpha}(L)$ implied by this model are given in the third panel of Table 10: they agree closely with those from the AR(3) and ARAR(2,1) models. The size of $\hat{\phi}_2$ compared to its standard deviation is smaller for the ARDLAR(2,2,1,1) model than for the ARAR(2,1). This results in less precise inferences about the dynamic properties $\phi(L)$, especially the period of the cycle, as shown in Table 11. Never-the-less,

the approximate posterior probability of the presence of a cycle is .64 for the ARDLAR(2,2,1,1) model.

The fifth column of Table 9 shows the results for the ARDL(3,3,2) implied by imposing the restrictions shown below equation (81), while those in the sixth column were obtained by applying OLS to (81) without regard for these restrictions. A researcher who focussed on these OLS results might be tempted to set α_3 and $\beta_{1,3}$ to zero, to gain statistical efficiency. But the results in columns four and five of Table 9 suggest that these restrictions would be inappropriate. Also the approximate Bayes factor in favor of the ADLAR(2,2,1,1) model over the unrestricted ARDL(3,3,2) is 40.69. Finally, we note that the roots of $\hat{\alpha}(L)$ from the unrestricted ARDL(3,3,2), shown in the last panel of Table 10, do not suggest the presence of a cycle.

Next we obtained forecasts of the level of GDP for 16 periods beyond the end of the sample used in estimation. For this purpose the ARAR(2,1) model was written as a restricted AR(3) (which in this case is the same as the unrestricted AR(3)) and the ARDLAR(2,2,1,1) as a restricted ARDL(3,3,2). Then forecasts were calculated for these two models and for the unrestricted ARDL(3,3,2) for the 16 post-sample periods and converted to forecasts of the level of GDP. Summary statistics for the percentage forecast errors are shown in Table 12. The forecasting performance of the ARAR model was the best of the three models, in spite of its lack of exogenous variable. On RMSE grounds the ARDLAR and ARDL models were nearly equal.

To illustrate the calculation of exact posterior results we used a procedure

similar to that used for the housing starts model above. Write $\mathbf{x_t'} = [1, x_{1,t}, x_{2,t}]$ and $\gamma' = [\alpha_0, \phi_1, \phi_2, \delta_{1,2}, \delta_{2,1}]$. In contrast to the housing model we adopted a uniform, diffuse prior for ω_1, γ' and $\log\sigma$. Write the initial values of Y_t and $\mathbf{x_t}$ as $\mathbf{y_0'} = [y_1, \ldots, y_s]$ and $\mathbf{x_0'} = [\mathbf{x_1'}, \ldots, \mathbf{x_s'}]$, where $s = max(p+r, m+r)$, and let $\mathbf{z_0'} = [\mathbf{y_0'}, \mathbf{x_0'}]$. Write the T remaining observations as $\mathbf{y}' = [y_{s+1}, \ldots, y_{s+T}]$ and $\mathbf{x}' = [\mathbf{x_{s+1}}, \ldots, \mathbf{x_{s+T}}]$ and write $\mathbf{z}' = [\mathbf{y}', \mathbf{x}']$. For this example $s = 3$ and T = 165. All inferences are conditional on $\mathbf{z_0'}$. Define $\mathbf{y}(\omega)$ as the $T \times 1$ vector with elements $y_t(\omega) = y_t - \omega_1 y_{t-1}$ and $\mathbf{x}(\omega)$ as the $T \times 2$ matrix with rows $\mathbf{x_t} - \omega_1 \mathbf{x_{t-1}}$. Then let $\mathbf{z}(\omega)$ be the $T \times k$ matrix with rows $\mathbf{z_t'}(\omega) = [\mathbf{y_{t-1}}(\omega), \mathbf{x_t}(\omega), \mathbf{x_{t-1}}(\omega)]$, where $k = m + p + 1 = 5$. Assume $\varepsilon_t \sim IN(0, \sigma^2)$. Then the joint posterior, after completing the square on γ, can be written as

$$p(\gamma, \omega, \sigma / \mathbf{y}, \mathbf{z}, \mathbf{y_0}, \mathbf{z_0}) \propto \sigma^{-(T+2)} \exp\left\{ -\frac{(v\widetilde{s^2(\omega)} + [\gamma - \widetilde{\gamma(\omega)}]'[\mathbf{z}'(\omega)\mathbf{z}(\omega)][\gamma - \widetilde{\gamma(\omega)}])}{2\sigma^2} \right\}$$

where: $v = T - k$;

$$\widetilde{\gamma(\omega)} = [\mathbf{z}'(\omega)\mathbf{z}(\omega)]^{-1}\mathbf{z}'(\omega)\mathbf{y}(\omega)$$

and

$$v\widetilde{s^2(\omega)} = \mathbf{y}'(\omega)\mathbf{y}(\omega) - \widetilde{\gamma(\omega)}'[\mathbf{z}'(\omega)\mathbf{z}(\omega)]\widetilde{\gamma(\omega)}.$$

Integrating out σ^2 gives

$$p(\gamma, \omega / \mathbf{y}, \mathbf{z}, \mathbf{y_0}, \mathbf{z_0}) \propto \left[\widetilde{s^2(\omega)}\right]^{-(v+k)/2} \left\{ v + [\gamma - \widetilde{\gamma(\omega)}]'\mathbf{H}(\omega)[\gamma - \widetilde{\gamma(\omega)}] \right\}^{-(v+k)/2} \tag{82}$$

where $\mathbf{H}(\omega) = \left[\widetilde{s^2(\omega)}\right]^{-1} [\mathbf{z}'(\omega)\mathbf{z}(\omega)]$.

Integrating (82) over γ gives the marginal posterior density for ω_1.

$$p(\omega_1 / \mathbf{y}, \mathbf{z}, \mathbf{y_0}, \mathbf{z_0}) \propto \left[v\widetilde{s^2(\omega)}\right]^{-v/2} |\mathbf{z}'(\omega)\mathbf{z}(\omega)|^{-1/2}. \tag{83}$$

This density is plotted, together with the large-sample approximation, in Figure 9. Posterior moments for ω_1 are in Table 13. Since a diffuse prior was used, the difference between the two densities is due entirely to the approximation error resulting from the small sample size. As in the housing starts example, conditional on ω_1, (82) is a multivariate Student-t density so the conditional posterior densities of the γ_i, given ω_1, are in the univariate Student-t form. To obtain the marginal posterior density for an individual γ_i integrate ω_1 out of the product of the conditional Student-t density and the marginal density for ω_1 given in (83). This integration was done using Gaussian quadrature over one dimension which is quick and accurate. The resulting posterior densities are plotted in Figures 10, 11, 12 and 13 and their moments are in Table 13. Clearly, the asymptotic approximations are not nearly so close to the exact densities as was the case in the housing example above.

4 Conclusions

We have shown in this paper that a useful model for the error, in univariate and single equation distributed lag models, is a finite, stationary AR. In the past researchers have often used a white noise or MA model for the error without much prior or other information to support their choice. The ARAR and ARDLAR models allows for very rich behavior of the error process and yet are usually easier to implement empirically than models with MA errors. Other appealing features of our ARAR model are its parsimony and the fact that all

its components obey the Wold decomposition. We have shown theoretically and empirically how a researcher's prior beliefs about the autoregressive structure of the observable variable can be used to solve the identification problem inherent in the univariate version of the model. There is no such identification problem in versions of the model, such as distributed lag models, which contain exogenous variables.

The extension of these ideas to vector autoregression and simultaneous equations models is the subject of our current research.

Acknowledgements

* I am grateful to the Graduate School of Business, The University of Chicago for the generous hospitality extended to me during the conduct of this research.

† Research financed in part by the National Science Foundation and by the H.G.B. Alexander Endowment Fund, Graduate School of Business, University of Chicago.

Notes

1. "God made X (the data), man made all the rest (especially, the error term)." quoted in Nicholls *et al* (1975). We add that in many instances humans, not God, construct the data.

2. For work on inference procedures for time series models see, *inter alia*: Box and Jenkins (1976), Chib (1993), Harvey (1981), and Hamilton (1994).

3. Box and Jenkins (1976)page 11.

4. Pankratz (1983) p.369.

5. In this example we are abstracting from the problem of analyzing multiple cycles, such as long cycles in construction activity.

6. For all the calculations reported in Tables 1 to 4 we used TSP Version 4.3A for OS/2.

7. These values were calculated using the COINT routines for GAUSS which scale the trend value to range from 1 to 1/T.

8. We drew the same inference from the results of augmented Dickey-Fuller (Said and Dickey (1984)), Phillips-Perron (1988) and weighted symmetric (Pantula *et.al.* (1994)) tests.

9. While a uniform, diffuse prior is uninformative for the α_i in (39), it is not uninformative for the roots of $\alpha(L)$. See Hong(1989) and Phillips (1991) for a discussion of this point.

10. The other model which Pankratz found adequate was an ARMA(3,2) with ϕ_2 set to 0. For our data that model has slightly higher residual autocorrelation, was less parsimonious and had larger forecast errors than did the ARAR(2,2).

11. For these calculations the data was measured as deviations from the estimated sample mean divided by the standard deviation of these deviations. This had no effect on the $\hat{\phi}_1$, $\hat{\phi}_2$, $\hat{\omega}_1$ or $\hat{\omega}_4$ but it served to eliminate overflow and underflow errors in the numerical evaluations of the integrals discussed below.

12. There is no $\omega_0 = 0$, analogous to ϕ_0, because $E(U_t) = 0$ by assumption.

13. All the numerical integrations reported here were done by Gaussian quadrature using GAUSS 386 version 3.2.13.

14. Marginal posteriors for ϕ_1 or ϕ_2 can also be obtained by integrating the portion of (43) which we have identified as the joint posterior of ϕ_1 and ϕ_2 over the other ϕ_i.

15. This produces the same results as the technique introduced by Fuller and Martin (1961).

16. This formulation assumes that Y_t and x_t have the same unit of measurement.

References

Banerjee, Anindya, Juan Dolado, John W. Galbraith and David F. Hendry, Co-Integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data (Oxford, Oxford University Press, 1993).

Box, George E. P. and Gwilym Jenkins, Time Series Analysis: Forecasting and Control, 2nd ed. (San Francisco, Holden Day, 1976).

Chib, Siddhartha, Bayes Regression with Autoregressive Errors: A Gibbs Sampling Approach, Journal of Econometrics 20 (August, 1993) 275-294.

Fuller, Wayne. A. and James E. Martin, The Effects of Autocorrelated Errors on the Statistical Estimation of Distributed Lag Models, Journal of Farm Economics 63 (February, 1961). 71-82.

Garcia-Ferrer, Antonio, R. A. Highfield, Franz Palm and Arnold Zellner, Macro-economic Forecasting Using Pooled International Data, Journal of Business and Economic Statistics 5 (January, 1961). 53-67.

Geweke, John K., The Secular and Cyclical Behavior of Real GDP in 19 OECD Countries,1957-1983, Journal of Business and Economic Statistics 6 (October, 1988) 479-486.

Goodwin, Richard M., Dynamical Coupling With Especial Reference to Markets Having Production Lags, Econometrica 15 (July, 1947) 181-204.

Hamilton, James D., Time Series Analysis (Princeton, Princeton University Press, 1994).

Harberger, Arnold C., The Demand for Durable Goods, (Chicago, University of Chicago Press, 1960).

Harvey, Andrew C., Time Series Models (Oxford, Philip Allan,1981).

Hendry, David F., Adrian R. Pagan and J. Denis Sargan, Dynamic Specification, in Zvi Griliches and Michael D. Intriligator (Eds), Handbook of Econometrics, Volume 3 (Amsterdam, North-Holland, 1986).

Hong, Chansik, Forecasting Real Output Rates and Cyclical Properties of Models, A Bayesian Approach, PhD Thesis (Chicago, Department of Economics, University of Chicago, 1989).

Jorgenson, Dale, Anticipations and Investment Behavior, in James S. Dusenberry , Gary Fromm, Lawrence R. Klein and Edward Kuh (Eds), The Brookings Quarterly Econometric Model of the United States (Chicago, Rand McNally, 1965).

Jorgenson, Dale, Rational Distributed Lag Functions, Econometrica 34 (January, 1966) 135-139.

Judge, George G. and M. E. Bock, The Statistical Implications of Pre-Testing and Stein-Rule Estimators in Econometrics (Amsterdam, North-Holland, 1978).

Koyck, Leendert M., Distributed Lags and Investment Analysis (Amsterdam, North-Holland, 1954).

Lintner, John., Corporation Finance: Risk and Investment, in J. Ferber (Ed.) Determinants of Investment Behavior. A Conference of the Universities–National Bureau Committee for Economic Research (New York, National Bureau of Economic Research, 1967).

Ljung, Greta M. and George E. P. Box, On a Measure of Lack of Fit in Time Series Models, Biometrika 66 (April, 1978) 297-303.

Metzler, Lloyd A., The Nature and Stability of Inventory Cycles, The Review of Economics and Statistics, 23 (August, 1941) 113-129.

Monahan, John F., Fully Bayesian Analysis of ARMA Time Series Models, Journal of Econometrics 21 (1983) 307-331.

Nerlove, Marc, The Dynamics of Supply: Estimation of Farmers' Response to Price (Baltimore, John Hopkins Press, 1958).

Nicholls, Des F., Adrian R. Pagan and R. Deane Terrell, The Estimation and Use of Models With Moving Average Disturbance Terms: A Survey, International Economic Review 16 (February, 1958) 112-134.

Pagan, Adrian, Time Series Behavior and Dynamic Specification, Oxford Bulletin of Economics and Statistics 47 (August, 1985) 199-211.

Pankratz, Alan, Forecasting With Univariate Box-Jenkins Models (New York, Wiley, 1985).

Pantula, Sastry G., Graciela Gonzalez-Farias and Wayne A. Fuller, A Comparison of Unit-Root Test Criteria, Journal of Business and Economic Statistics 12 (October, 1994) 449-459.

Phillips, A. W., Stabilization Policy and the Time Form of Lagged Responses, Economic Journal 67 (June, 1957) 265-267.

Phillips, Peter C. B. and Pierre Perron, Testing for a Unit Root in Time Series Regression, Biometrika 76 (June, 1988) 335-346.

Phillips, Peter C. B., To Criticize the Critics: an Objective Bayesian Analysis of Stochastic Trends, Journal of Applied Econometrics 6 (October-December, 1991) 333-364.

Samuelson, Paul A., Interactions Between the Multiplier Analysis and the Principal of Acceleration, Review of Economics and Statistics 21 (May, 1939) 75-78.

Samuelson, Paul A., A Synthesis of the Principle of Acceleration and the Multiplier, Journal of Political Economy 47 (December, 1939) 786-797.

Said, E. Said and David. E. Dickey, Testing for Unit Roots in Autoregessive-Moving Average Models of Unknown Order, Biometrika 71 (December, 1984) 599-607.

Schotman, Peter C. and Herma K. Van Dijk, On Bayesian Routes to Unit Roots, Journal of Applied Econometrics 6 (October-December, 1991) 387-401.

Schwarz, Gideon, Estimating the Dimension of a Model, Annals of Statistics 6 (March, 1978) 461-464.

Stock, James H., Unit Roots, Structural Breaks and Trends, in Robert F. Engle and Daniel L. Mcfadden (Eds.) Handbook of Econometrics Volume 4, Chapter 46 (Amsterdam, North-Holland, 1994).

Tobias, Justin and Arnold Zellner, Further Results on Bayesian Method of Moments Analysis of the Multiple Regression Model, International Economic Review 42 (February, 2001) 121-132.

Wrinch, Dorothy H. and Harold Jeffreys, On Certain Fundamental Principles of Scientific Inquiry, Philosophical Magazine Series 6 42 (1921) 369-390.

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

Zellner, Arnold, Statistical Analysis of Econometric Models, Journal of the American Statistical Association 74 (September, 1979) 628-651.

Zellner, Arnold and Chansik Hong, Forecasting International Growth Rates Using Bayesian Shrinkage and Other Procedures, Journal of Econometrics 40 (July-August, 1989) 183-202.

Zellner, Arnold, Chansik Hong and Chung-Ki Min, Forecasting Turning Points in International Growth Rates Using Bayesian Exponentially Weighted Autoregression, Time-Varying Parameters and Pooling Techniques, Journal of Econometrics 49 (July, 1991) 275-304.

Zellner, Arnold, Bayesian Method of Moments/Instrumental Variables (BMOM) Analysis of Mean and Regression Models, in J. C. Lee, A. Zellner and W. O. Johnson (Eds.) Modelling and Prediction Honoring Seymour Geisser (New York, Springer-Verlag, 1996) and in Zellner Arnold, Bayesian

Analysis in Econometrics and Statistics (Cheltenham, U.K., Edward Elgar, 1997) 291-304.

Zellner, Arnold, The Bayesian Method of Moments (BMOM): Theory and Applications, in T. Fomby and R. C. Hill (Eds.) Advances in Econometrics, Vol 12 (New York, Springer-Verlag, 1997) 85-105.

Zellner, Arnold and Martin S. Geisel, Analysis of Distributed Lag Models with Applications to Consumption Function Estimation, Econometrica 38 (November, 1970) 865-888.

Zellner, Arnold, D. S. Huang and L. C. Chau, Further Analysis of the Short-Run Consumption Function with Emphasis on the Role of Liquid Assets, Econometrica 33 (July, 1965) 571-581.

Table 1
Housing Starts Model
Autocorrelations and Partial Autocorrelations

	Data		AR(6) Residuals	
	Autocorr	Partial	Autocorr	Partial
Lag	(S.E.)	(S.E.)	(S.E.)	(S.E.)
1	.911	.911	.00213	.00213
	(.0765)	(.0765)	(.0778)	(.0778)
2	.765	-.376	-.0273	-.0273
	(.125)	(.0765)	(.0779)	(.0778)
3	.581	-.242	-.0238	-.0237
	(.150)	(.0765)	(.0779)	(.0778)
4	.397	-.0111	.0146	.0139
	(.162)	(.0765)	(.0780)	(.0778)
5	.230	-.0207	.0481	.0468
	(.168)	(.0765)	(.780)	(.0778))
6	.0751	-.126	-.0628	-.0630
	(.144)	(.0765)	(.0781)	(.0778))
7	-.0459	.0382	.00562	.00920
	(.170)	(.0765)	(.0785)	(.0778)
8	-.141	-.0471	.0232	.0221
	(.170)	(.0765)	(.0785)	(.0778)
9	-.210	-.0548	.0322	.0283
	(.171)	(.0765)	(.0785)	(.0778)
10	-.267	-.106	-.0387	-.0384
	(.172)	(.0765)	(.0786)	(.0778))
Box-Ljung (1978) portmanteau statistics at lags 5 & 10				
Q_5	344		19.3	
Q_{10}	370		30.3	

Data source: U.S. Department of Commerce,
Survey of Current Business

Table 2
Housing Starts Model
Roots of $\hat{\alpha}(L)$ by OLS

Roots	Modulus	Period
$.860 \pm .242i$	.893	23.0
$-.624 \pm .382i$	.731	2.42
$.358 \pm .586i$	.687	6.14

Table 3
Housing Starts Model
Posteriors for Unit Roots and Trends

	No Trend	Linear Trend	Quadratic Trend
	Using Equation (31)		
$\hat{\alpha}_0$	15.4	15.2	14.3
Std Dev	3.39	3.43	3.67
$\hat{\beta}_1$	—-	.00568	.0403
Std Dev	—-	.0123	.0494
$\hat{\beta}_2$	—-	—-	-.000206
Std Dev	—-	—-	.000285
$\hat{\rho}$	.836	.833	.833
Std Dev	.0352	.0359	.0359
$P_D(\rho \geq 1)$	.000	.000	.000
$P_J(\rho \geq 1)$	.000	.000	.000
	Using Equation (34)		
$\hat{\mu}_0$	94.2	91.2	84.1
Std Dev	3.51	7.36	12.2
$\hat{\mu}_1$	—-	.0341	.254
Std Dev	—-	.0729	.315
$\hat{\mu}_2$	—-	—-	-.00123
Std Dev	—-	—-	.172

Table 4
Housing Starts Models
Asymptotic Posterior Moments

	Unrestricted ARAR(2,4)	Restricted ARAR(2,4)	Unrestricted ARMA(2,2)	Restricted ARMA(2,2)
$\hat{\phi}_1$	1.72	1.63	1.04	1.22
Std Dev	.0727	.0628	.189	.0771
$\hat{\phi}_2$	-.798	-.722	-.217	-.371
Std Dev	.0699	.0623	.176	.0741
$\hat{\omega}_1$	-.531	-.461	—-	—-
Std Dev	.0985	.0805	—-	—-
$\hat{\omega}_2$	-.114	—-	—-	—-
Std Dev	.118	—-	—-	—-
$\hat{\omega}_3$	-.206	—-	—-	—-
Std Dev	.110	—-	—-	—-
$\hat{\omega}_4$	-.252	-.149	—-	—-
Std Dev	.0882	.0743	—-	—-
$\hat{\theta}_1$	—-	—-	-.175	—-
Std Dev	—-	—-	.177	—-
$\hat{\theta}_2$	—-	—-	-.398	-.329
Std Dev	—-	—-	.0910	.0845
$\hat{\sigma}$	7.38	7.41	7.52	7.48
ln likelihood	-560	-562	-565	-565
Q_5	.658	4.80	2.37	3.96
Q_{10}	1.89	5.79	6.96	7.39

Table 5
Housing Starts Models
Asymptotic Posterior Moments
For Functions of ϕ_1 and ϕ_2

	Unrestricted ARAR(2,4)		Restricted ARAR(2,4)	
Dynamic Property	Mean	Standard Deviation	Mean	Standard Deviation
Cycle	-.233	.0631	-.232	.0815
Modulus	.893	.0391	.850	.366
Period	23.0	2.75	21.9	3.41

Table 6
Housing Starts Models
Percentage Forecast Error Summary Statistics

	Unrestricted ARAR(2,4)	Restricted ARAR(2,4)	Unrestricted ARMA(2,2)	Restricted ARMA(2,2)
Mean	2.92	2.52	1.85	1.86
Std Dev	7.25	6.68	7.25	6.98
RMSE	7.82	7.14	7.49	7.22

Table 7
Housing Starts Model
Exact Posterior Moments
Restricted ARAR(2,4)

Coeff.	Mean	Std. Dev.	Skewness
ϕ_1	1.62	.0680	-.0000186
ϕ_2	-.706	.0671	-.000210
ω_1	-.440	.0860	-.00176
ω_4	-.137	.0768	-.00848

Table 8
Posteriors for Unit Roots and Trends in Rates of Growth
for Real GDP, M2 and Stock Returns Using Equation (31)

	No Trend	Linear Trend	Quadratic Trend
Rate of Growth of Real GDP			
$\hat{\alpha}_0$	.04604	.006394	.006502
Std Dev	.001045	.001776	.002524
$\hat{\beta}_1$	—-	-.003146	-.003737
Std Dev	—-	.002529	.01013
$\hat{\beta}_2$	—-	—-	.0005825
Std Dev	—-	—-	.009668
$\hat{\rho}$	.3894	.36375	.36299
Std Dev	.09417	.09600	.09695
$P_D(\rho \geq 1)$	.00000	.00000	.00000
$P_J(\rho \geq 1)$	.00000	.00000	.00000
Rate of Growth of Real M2			
$\hat{\alpha}_0$	.002594	.0032332	.0008665
Std Dev	.0008632	.001585	.002280
$\hat{\beta}_1$	—-	-.001253	.01345
Std Dev	—-	.002611	.01057
$\hat{\beta}_2$	—-	—-	-.01464
Std Dev	—-	—-	.01020
$\hat{\rho}$	.5922	.5910	.5750
Std Dev	.06351	.06352	.06410
$P_D(\rho \geq 1)$	.00000	.00000	.00000
$P_J(\rho \geq 1)$	.00000	.00000	.00000
Rate of Growth of Real Stock Returns			
$\hat{\alpha}_0$	.006503	.01626	.03827
Std Dev	.005782	.01163	.01764
$\hat{\beta}_1$	—-	-.01928	-.1479
Std Dev	—-	.01996	.08041
$\hat{\beta}_2$	—-	—-	.1273
Std Dev	—-	—-	.07716
$\hat{\rho}$	.1447	.1370	.1180
Std Dev	.07708	.07726	.07750
$P_D(\rho \geq 1)$	.00000	.00000	.00000
$P_J(\rho \geq 1)$	.00000	.00000	.00000

Table 9
Real GDP Growth Models
Asymptotic Posterior Moments

	AR(3) Eqn (15)	ARAR(2,1) Eqn (79)	ARDLAR(2,2,1,1) Eqn (80)	Restricted ARDL(3,3,2) Eqn (81)	Unrestricted ARDL(3,3,2) Eqn(81)
$\hat{\alpha}_0$	.004884	.004884	.004049	.004049	.004023
Std Dev	.001103	.001103	.001056	.001056	.001056
$\hat{\alpha}_1$	.3403	—-	—-	.2091	.1993
Std Dev	.07883	—-	—-	.07766	.07826
$\hat{\alpha}_2$	.1257	—-	—-	.1357	.1158
Std Dev	.08267	—-	—-	.07206	.07693
$\hat{\alpha}_3$	-.08642	—-	—-	-.05417	-.01090
Std Dev	.07841	—-	—-	.06628	.07346
$\hat{\beta}_{1,2}$	—-	—-	—-	.1100	.1328
Std Dev	—-	—-	—-	.05283	.07869
$\hat{\beta}_{1,3}$	—-	—-	—-	.04623	.003707
Std Dev	—-	—-	—-	.02255	.07698
$\hat{\beta}_{2,1}$	—-	—-	—-	.04137	.03639
Std Dev	—-	—-	—-	.008879	.009529
$\hat{\beta}_{2,2}$	—-	—-	—-	.01739	.02818
Std Dev	—-	—-	—-	.005981	.01047
$\hat{\phi}_1$	—-	.7676	.6295	—-	—-
Std Dev	—-	.1443	.1450	—-	—-
$\hat{\phi}_2$	—-	-.2023	-.1289	—-	—-
Std Dev	—-	.1279	.1208	—-	—-
$\hat{\omega}_1$	—-	-.4273	-.4203	—-	—-
Std Dev	—-	.1364	.1410	—-	—-
$\hat{\delta}_{1,2}$	—-	—-	.1100	—-	—-
Std Dev	—-	—-	.05283	—-	—-
$\hat{\delta}_{2,1}$	—-	—-	.04137	—-	—-
Std Dev	—-	—-	.008879	—-	—-
$\hat{\sigma}$	.09761	.009761	.008921	.008921	.08919
ln(L)	531.7	531.7	547.3	547.3	548.7
Q_5	.459	.459	.803	.803	.537
Q_{10}	6.78	6.78	9.41	9.41	7.96

Table 10
Real GDP Growth Models
Roots of Estimated $\hat{\alpha}(L)$

Roots	Modulus	Period
Unrestricted AR(3) by OLS		
$.3838 \pm .2344i$	.4497	11.46
-.4273	.4273	
$\hat{\alpha}(L)$ From ARAR(2,1) by NLS		
$.3838 \pm .2343i$	.4497	11.46
-.4273	.4273	
$\hat{\alpha}(L)$ From ARDLAR(2,2,1,1) by NLS		
$.3147 \pm .1727i$	.3590	12.52
-.4203	.4203	
$\hat{\alpha}(L)$ From ARDL(3,3,2) by OLS		
.4151	.4151	
-.3026	.3026	
.08681	.08681	

Table 11
Real GDP Growth Models
Asymptotic Posterior Moments
For Functions of ϕ_1 and ϕ_2

	ARAR(2,1)		ARDLAR(2,2,1,1)	
Dynamic Property	Mean	Standard Deviation	Mean	Standard Deviation
Cycle	-.2195	.3345	-.1193	.3339
Modulus	.4497	.1422	.3590	.1682
Period	11.46	6.024	12.52	13.07

Table 12
Real GDP Growth Models
Percentage Forecast Error Summary Statistics

	ARAR(2,1)	ARDLAR(2,2,1,1)	Unrestricted ARDL(3,3,2)
Mean	-.0434	-.180	-.170
Std Dev	1.65	1.74	1.74
RMSE	1.66	1.75	1.75

Table 13
ARDLAR(2,2,1,1) Model for Real GDP Growth
Exact Posterior Moments

Coeff.	Mean	Std. Dev.	Skewness
ϕ_1	.4678	.1842	-.4976
ϕ_2	-.02723	.1190	-.02367
ω_1	-.4203	.1898	.5911
$\delta_{1,2}$	.1472	.06145	.3166
$\delta_{2,1}$	.03836	.008749	-.02477

Figure 1: Admissible AR Coefficients; Stationary ARAR(1,1)

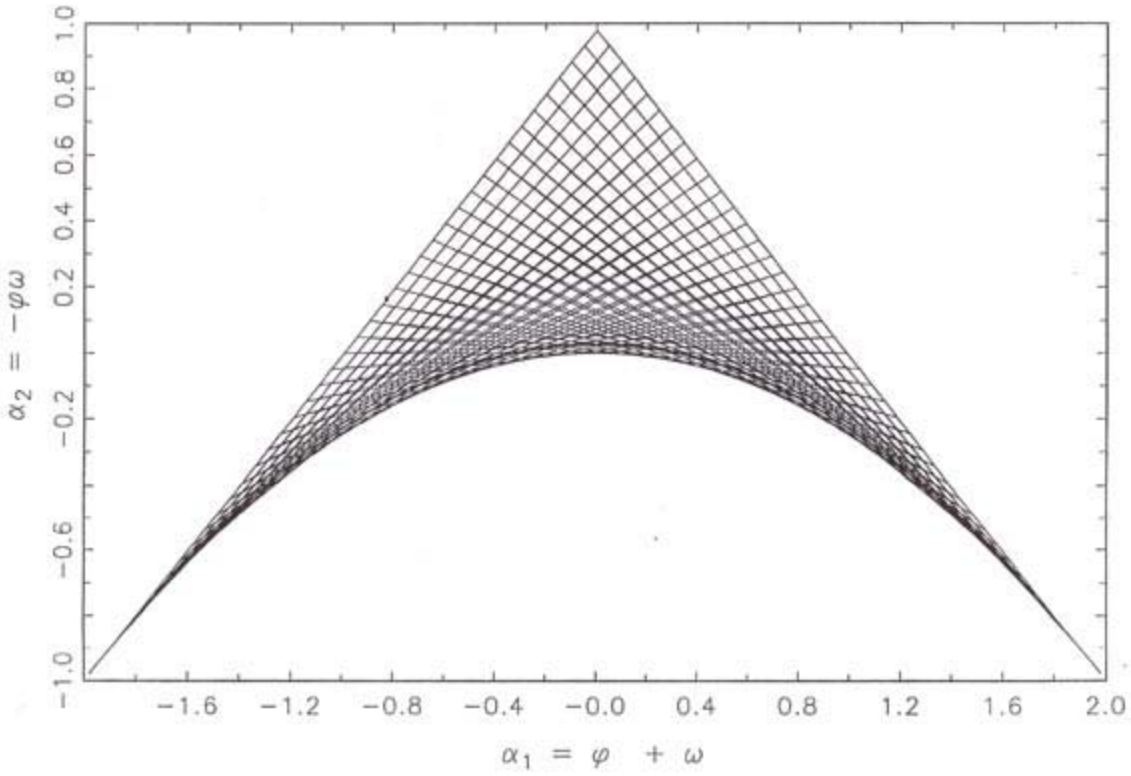

Figure 2a: Admissible Autocorrelations; Stationary ARAR(1,1)

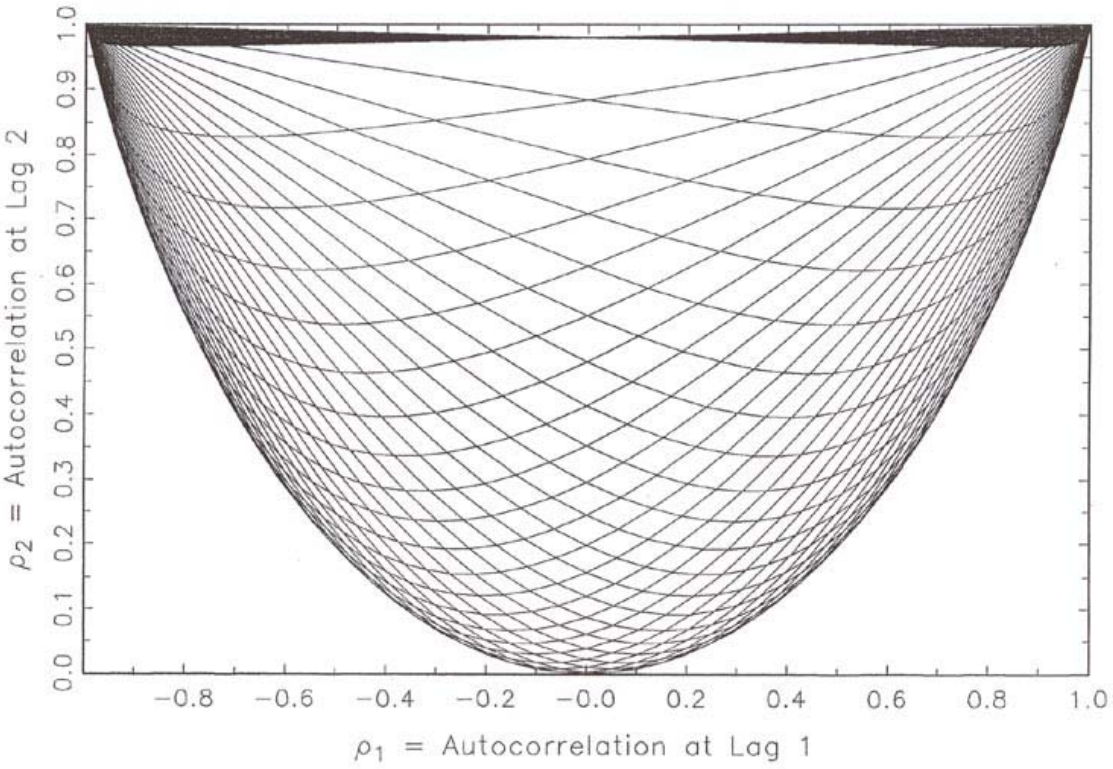

Figure 2b: Admissible Autocorrelations; Stationary ARMA(1,1)

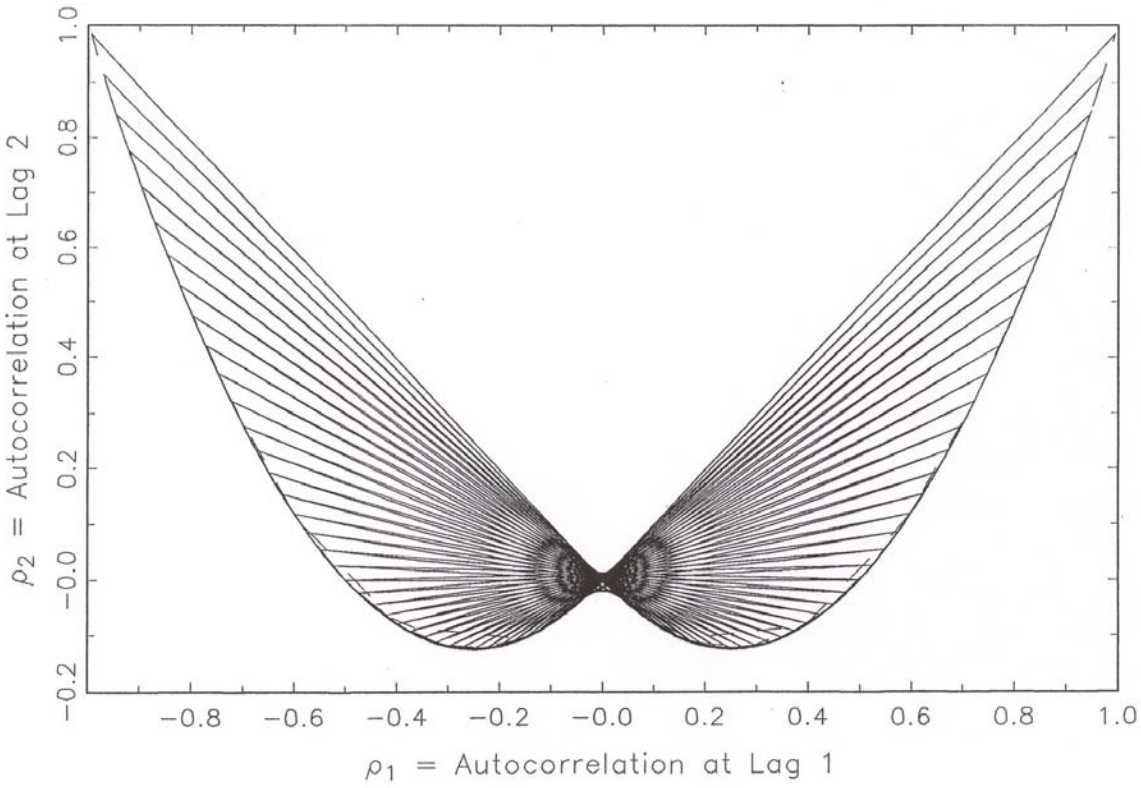

Figure 3: Housing Starts

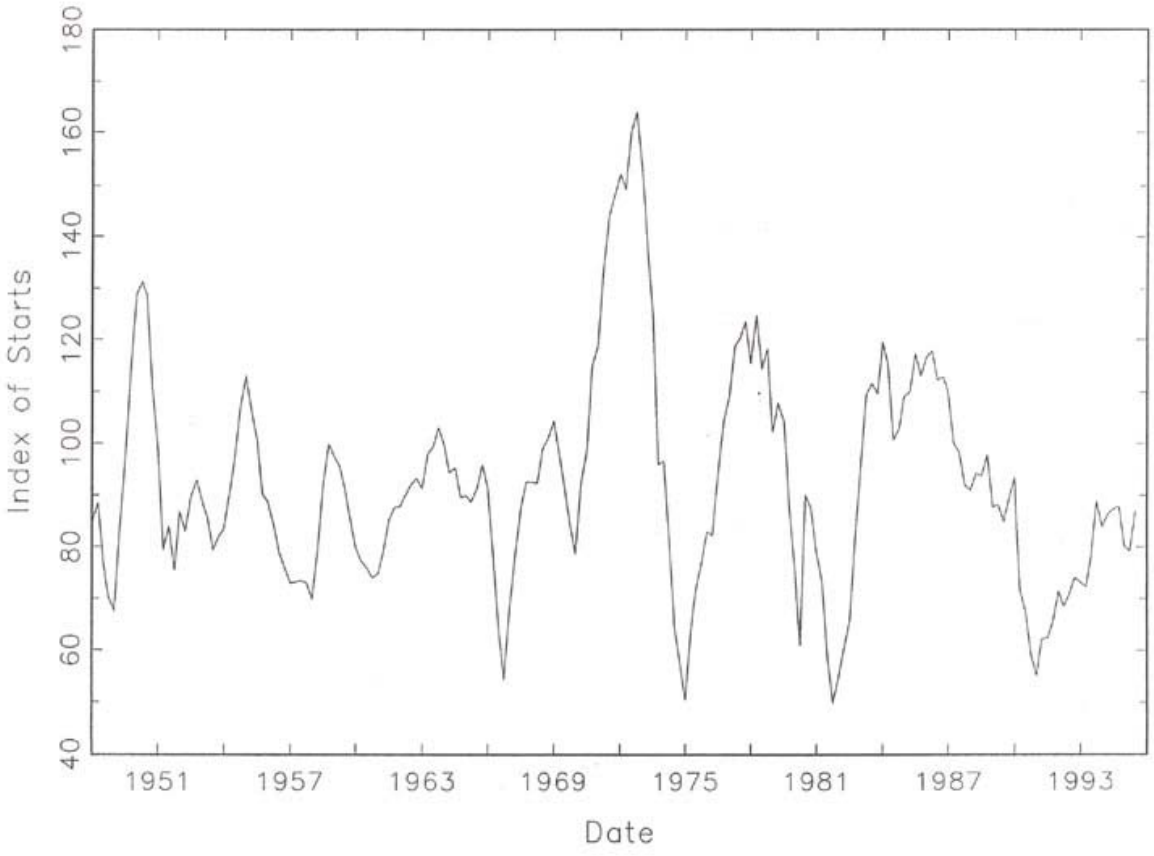

Figure 4: Housing Starts Model Joint Posterior

For φ_1 and φ_2 or ω_1 and ω_4

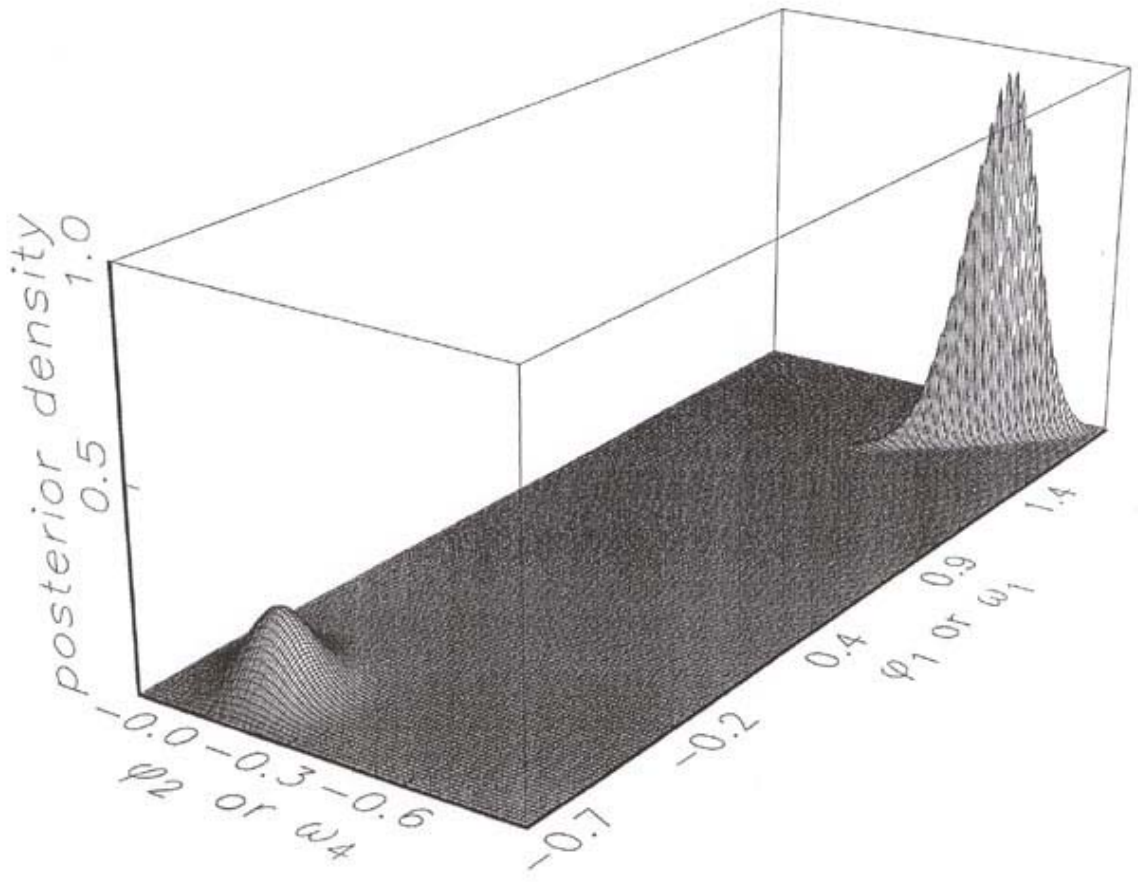

Figure 5: Housing Starts
Posterior for ω_1

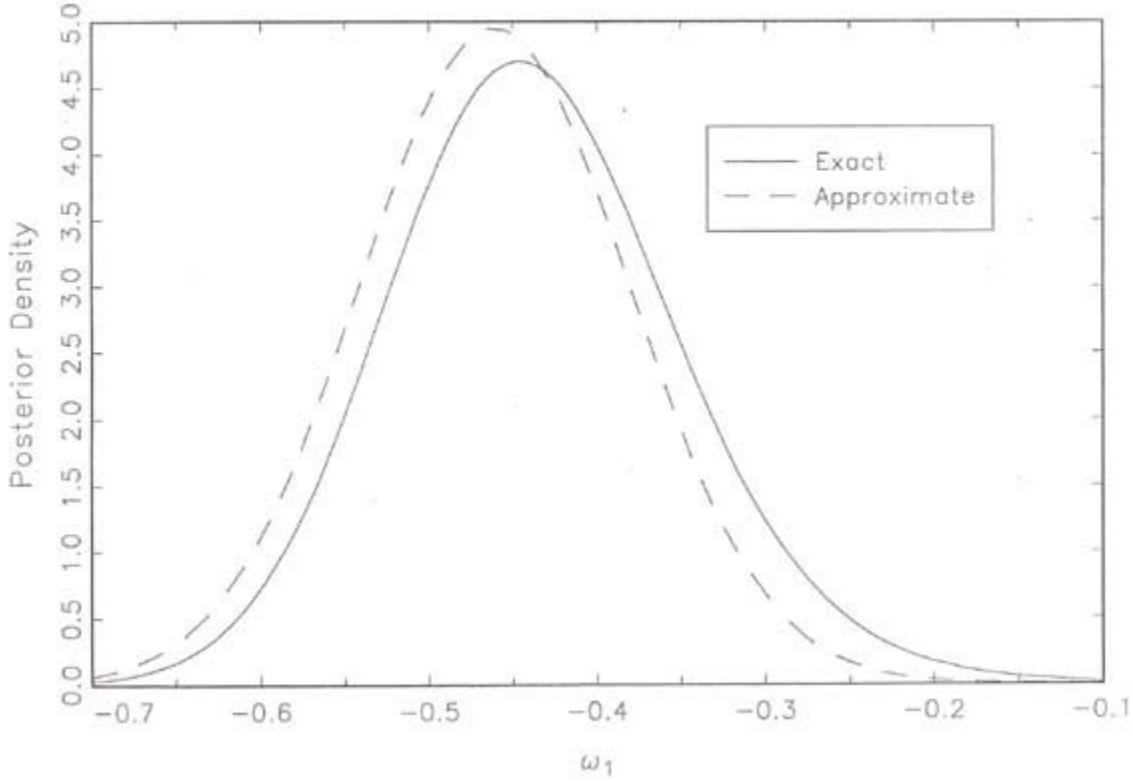

Figure 6: Housing Starts
Posterior for ω_4

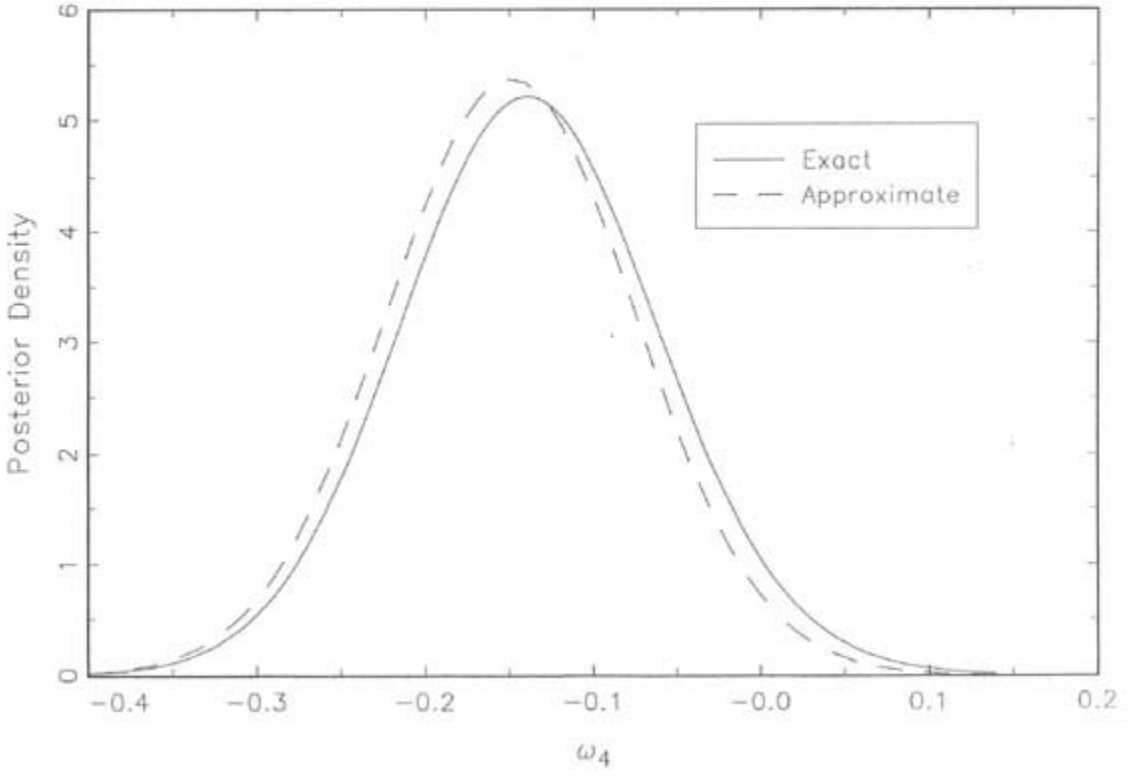

Figure 7: Housing Starts
Posterior for φ_1

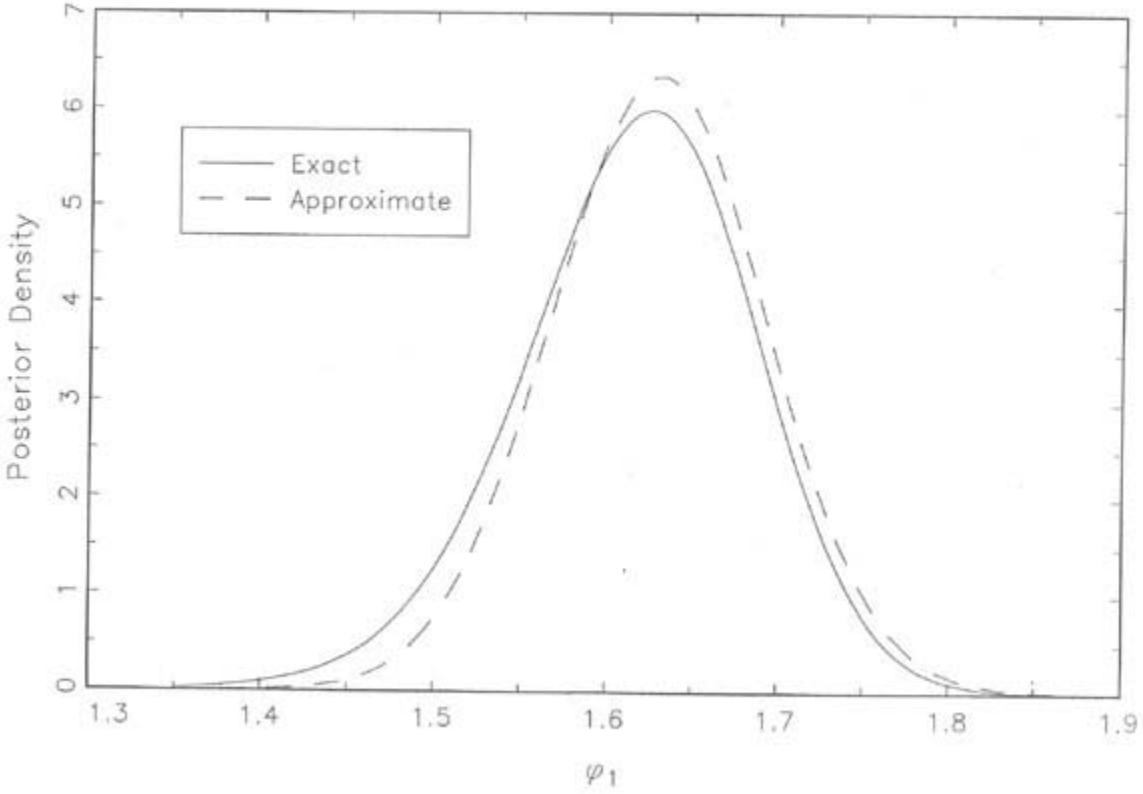

Figure 8: Housing Starts
Posterior for φ_2

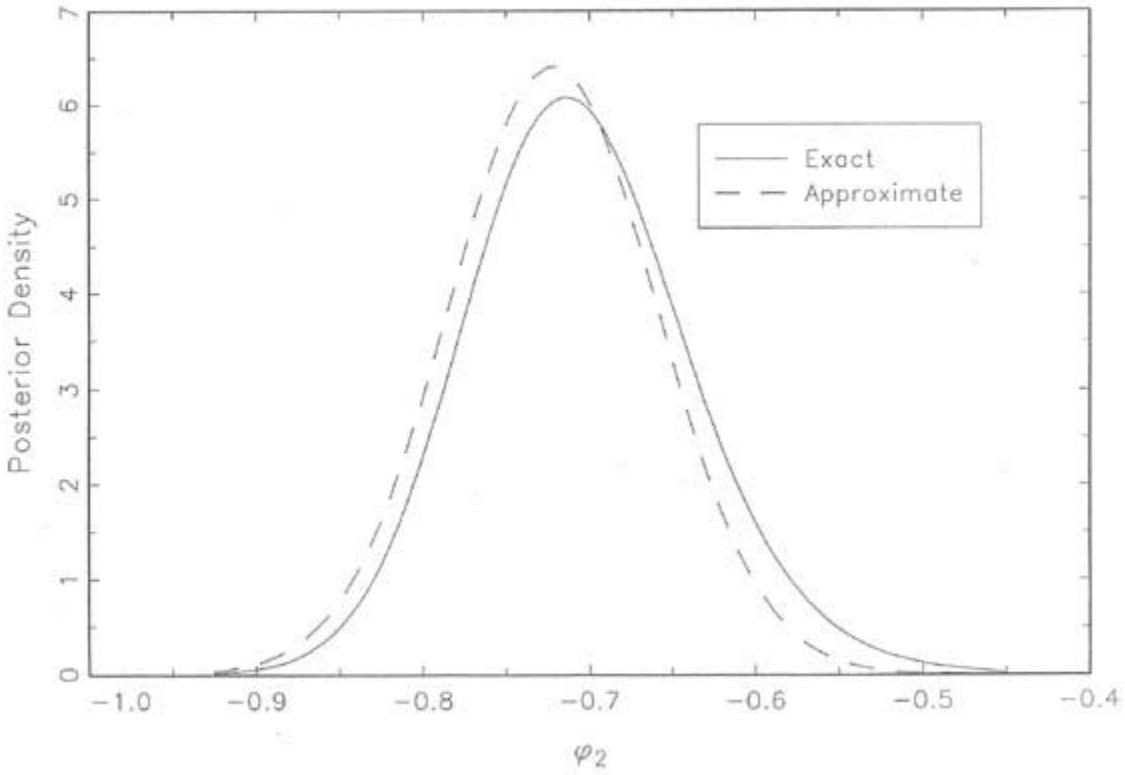

Figure 9: GDP Distributed Lag Model
Posterior for ω_1

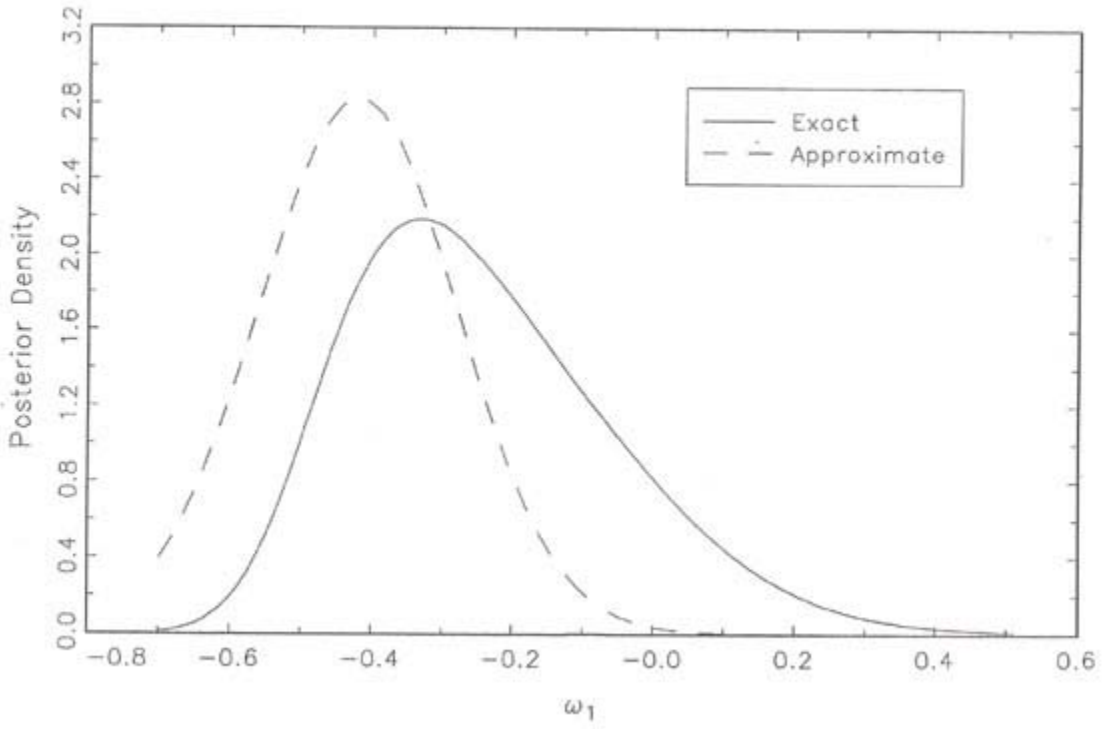

Figure 10: GDP Distributed Lag Model
Posterior for φ_1

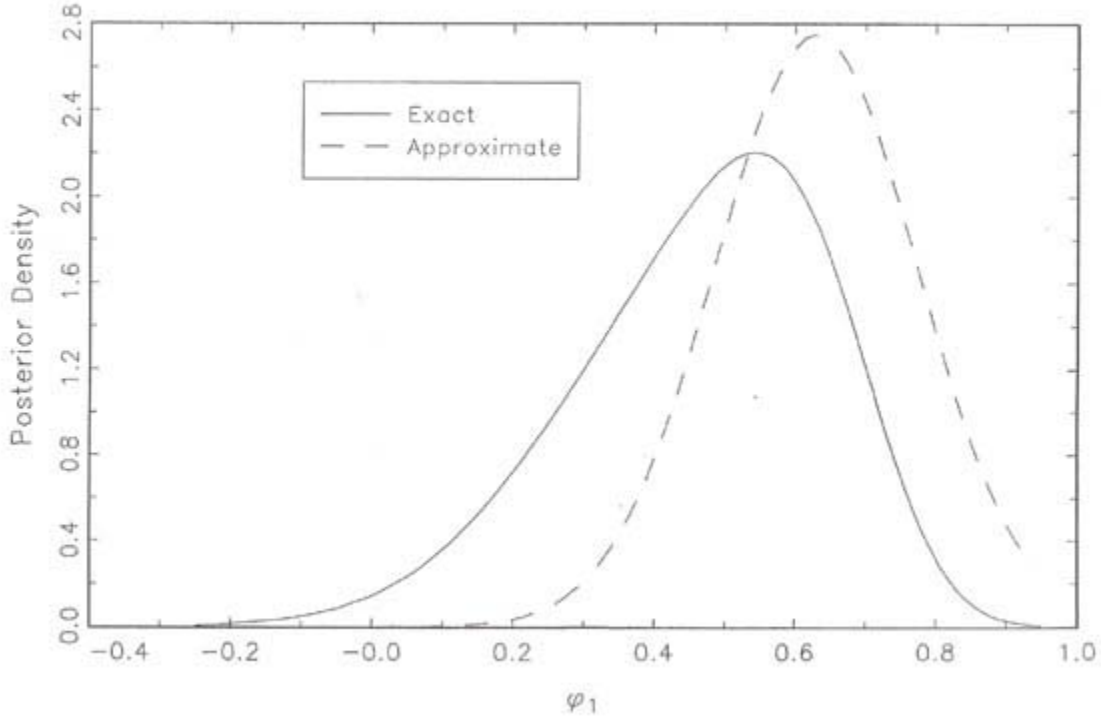

Figure 11: GDP Distributed Lag Model
Posterior for φ_2

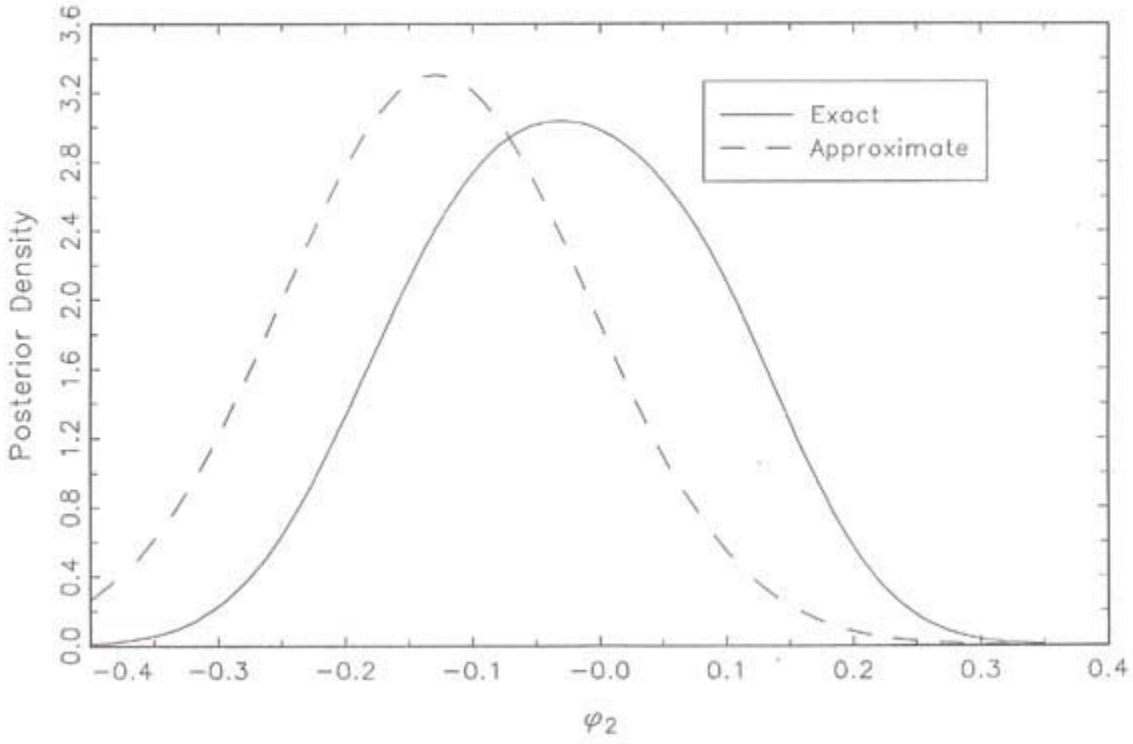

Figure 12: GDP Distributed Lag Model
Posterior for $\delta_{1,2}$

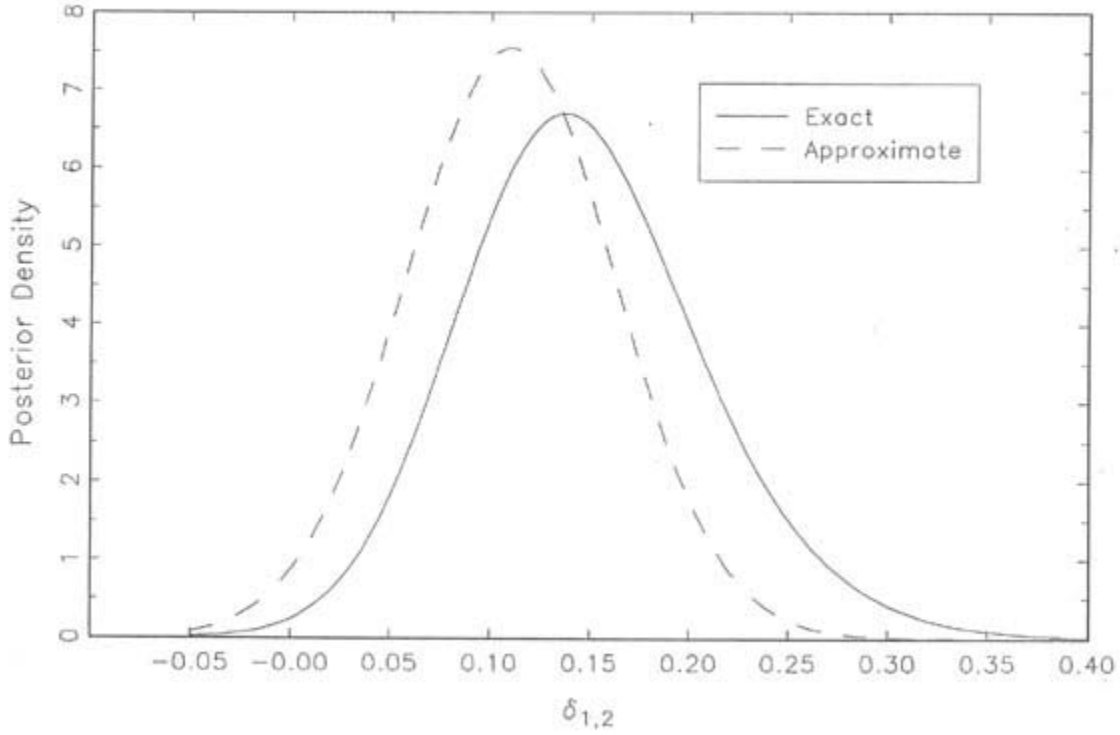

Figure 13: GDP Distributed Lag Model
Posterior for $\delta_{2,1}$