Evidences of Unit Roots and Co-integration in the Time Series for US GNP, MI and CPI

by

Masanao Aoki
University of California, Los Angeles
USA

UCLA Working Paper #439
March 1987

Evidences of Unit Roots and Co-integration

in the Time Series for

US GNP, MI and CPI

by
Masanao Aoki *
University of California, Los Angeles
USA

Abstract

This paper describes the dynamic factor (state space) models constructed for some US time series and test if the MI and CPI series are co-integrated in the sense of Granger and measure how much random walk component is present in the GNP series.

Dynamic factor (state space) models are constructed for the US MI and CPI monthly time series with a common dynamic factor that explains significant movements in those two series. Discovery of such a common factor is equivalent to the notion of co-integration advanced by Granger. To this end, a recently developed method of Aoki is used to (i) first construct a low order dynamics for trend components and (ii) then the residuals are treated as weakly stationary to which another model is fitted. This procedure results in a recursive dynamic model in which short-run dynamics affects but is not affected by the longer-run trend dynamics. This decomposition differs from the random detrending advanced by Kitagawn or Harvey because the latters produce block-diagonal dynamic matrices rather than block-triangular, i.e., the trend dynamics and shorter-run cyclical dynamics are not allowed to interact.

* 4731 Boelter Hall, University of California, Los Angeles, California 90024, USA

† Research supported in part by an NSF grant.

When the procedure is applied to the bivariate monthly series of the US MI and CPI from January 1975 on, consisting of 117 data points, one dynamic factor is discovered common to both series, so that 1.13 CPI - MI is co-integrated. The US GNP quarterly series exhibit the strongest evidence for the unit roots among the three series examined. Here the issue is how to define the random walk components. The paper proposes a measure different from that used by Nelson and Plosser and Cochrane. The real GNP series from 1974 seems to contain less than 6% of the random walk components.

1. Introduction

There now exists a growing body of literature on finding and measuring unit roots or random walk components in nonstationary time series and on testing if some of them are co-integrated. for example, a recent special issue of *Oxford Bulletin of Economics and Statistics*, (1987), is devoted entirely to econometric modeling with cointegrated variables.

In particular, several studies are available on the quarterly US real GNP time series data. These studies aim at testing the presence of random walk components or measuring their proportion. Perron-Phillips (1986) report mixed results depending on the series they used. They attribute their divergent findings possibly to the low power of the unit root test when the number of samples is small. Nelson-Plosser series contain only 62 observations. Walsh (1986) and King, Plosser, Stock and Watson (1987) examined other macroeconomic series.

The findings on the percentages of random walks in the US GNP fall into two classes: Nelson and Plosser (1982) and Campbell and Mankiw (1986) found large components, while Watson (1986) and Clark (1986) and Cochrane (1986) found small percentages of random walk components. All of them used slightly differently specified models. Cochrane attributes these divergent results to certain errors in neglecting sums of large numbers of small correlation coefficients. Some studies fit a pure random walk model $y_t = \rho y_{t-1} + u_t$, where the disturbance term is mean zero and uncorrelated across time and the hypotheses that $\rho = 1$ is tested. Others allow for more structured residuals such as u_t being AR (p) or rather rarely as ARMA (p, q). According to Cochrane those using low values of p and q found larger random walk component than those with higher values of p.

Stock and Watson (1986) arbitrarily detrend the series by extracting a 1.5% annual trend growth before testing for the presence of unit roots. This paper employs a two-step state space modeling procedure which conforms to a canonical or standard representation of time series with unit roots, recently developed by Aoki (1987) to re-examine the issue of percentages of random walks in the US real GNP and the existence of co-integration in the US M1 and CPI series, by first determining dynamics for random trends and then modeling the residuals by another state space model. This procedure produces a recursive dynamic structure and is superior to prior detrending and separate modeling of the residuals which produce block diagonal dynamic structure with no interactions from the cyclical component dynamics to the trend dynamics.

Use of state space models is also advantageous because it uses a finite number of parameters when data processes have rational spectral density functions, thus avoiding one source of error of neglecting an infinite number of small correlations which Cochrane attributes as a source of discrepancy of the findings in the literature. The dichotomous classification of Nelson and Plosser is to posit either 1) y_t has no determinstic trend, more specifically that $y_t - y_{t-1}$ is weakly stationary, or 2) y_t contains a linear deterministic trend bt. This dichotomy is by no means the only possible one. Rather than the dichotomy of deterministic v.s. random trends, it is more useful to decompose y_t into $y_{1t} + y_{2t}$ where y_{1t} has unit roots while y_{2t} is weakly stationary.

2. Two-Step Modeling Procedure

The singular value decomposition of the Hankel matrix, which results from evaluating the covariance matrix of a finite sequence of future realizations and past data stacked in a certain way, is the basis of the two-step modeling procedure. Let d_{t+1}^{+} be $Kp \times 1$ composed of $d_{t+1}, \ldots, d_{t+K}$ stacked in that order and let d_t^- be $Kp \times 1$ constructed by stacking d_t, d_{t-1}, $d_{t-2}, \ldots, d_{t-K+1}$ in that order, where p is the dimension of the data vector. Let $H_K = Ed_{t+1}^{+}d_t^{-\prime}$.

When $\{d_t\}$ is nonstationary containing unit roots, then the singular values of H_k, $\sigma_1 \geq \sigma_2 \geq \ldots$ are such that σ_2 / σ_1 and/or σ_3 / σ_1 are very small. (The reasons for this are discussed later in the Appendix.) This fact allows one to construct low (usually one- or two-) dimensional random trend dynamic model by extending the procedure described in Aoki (1983, 1987) and applied in Aoki and Havenner (1986) to directly construct a model of this nonstationary series

$$\tau_{t+1} = \Lambda\tau_t + Gu_t,$$

$$d_t = H\tau_t + u_t,$$

where u_t is the (usually highly correlated) residual vector sequence.

When τ_t is scalar-valued, Λ is one. When τ_t is vector-valued, the eigenvalues of Λ have magnitude one. To simplify exposition suppose that τ_t is scalar and rewrite the model as

$$\tau_{t+1} = \rho\tau_t + g'u_t, \quad \text{and} \quad d_t = h\tau_t + u_t \tag{1}$$

where $|\rho| = 1$.

Next, apply the modeling procedure to $\{u_t\}$ to produce

$$\begin{cases} x_{t+1} = Ax_\tau + Be_\tau \\ u_\tau = Cx_\tau + e_t \end{cases} \tag{2}$$

where (2) is minimal dimensional in the sense that (2) is obervable and A is invertable. See Lindquist and Picci (1979) or Aoki (1987, Chapter 5) for more detail.

Jointly $\{d_t\}$ is modeled as

$$\begin{bmatrix}\tau_{t+1} \\ x_{t+1}\end{bmatrix} = \Phi \begin{bmatrix}\tau_t \\ x_t\end{bmatrix} + \begin{bmatrix}g' \\ B\end{bmatrix} e_t,$$

$$d_t = (h \quad c) \begin{bmatrix}\tau_t \\ x_t\end{bmatrix} + e_t, \tag{3}$$

where

$$\Phi = \begin{bmatrix} \rho & g'c \\ 0 & A \end{bmatrix}.$$

Here $e_t = d_t - d_{t|t-1}$ is the weak innovation vector of d_t. Choose the initial condition τ_0 and x_0 to be independent of e_0. In the Appendix we show that a unique symmetric, positive definite $\pi = cov\, x$ exists if and only if the eigenvalues of F all have magnitude less than one. Similarly from (1) and (2)

$$T_{t+1} = aT_t + g'D_t g.$$

where

$$a = (\rho - g'h)^2, \quad T_t = E\tau_t^2, \quad \text{and} \quad D_t = Ed_t d_t.'$$

Even when d_t is a pure random walk (then D_t linearly grows with time), the above equations converge if $a < 1$.

Model representation (3) is used in empirical works because it conforms to one of the standard representation of time series with unit roots. See Aoki (1987, p. 39).

3. Percentage of Random Walk

Cochrane (1986) proposed the ratio of the limit of $var(y_t - y_{t-k}) / k$ as k goes to infinity over $var(y_t - y_{t-1})$ as a measure for random walk components in $\{y_t\}$. We point out an undesirable feature of this measure, and propose an alternative. To simplify exposition we assume y_t to be scalar-valued, even though vector-valued y_t can be analyzed almost as easily as the scalar-valued ones. His measure can over or understate the content of random walks as we show below. We approach the time series modeling from the state space viewpoint assuming that y_t has a rational spectral density function. This allows us to use a finite parametrization of the model and avoid approximation of infinite sums by finite sums which Cochrane attributes as a source of disagreement of empirical findings.

In the pure random walk model

$$y_t = y_{t-1} + d_t$$

where d_t is mean zero and uncorrelated with $Ed_t^2 = \mu_0$, the ratio of $var(y_t - y_{t-k}) / k$ over $var(y_t - y_{t-1})$ is one.

Consider next a model

$$y_t = y_{t-1} + d_t$$

where d_t is a covariance stationary process,

$$d_t = A(L)e_t, \tag{4}$$

$$A(L) = \sum_0^\infty a_j L^j, \qquad a_0 = 1, \qquad \Sigma a_j^2 < \infty, \tag{5}$$

and e_t is an i.i.d mean zero white noise with variance Δ. When d_t has a rational spectral density, it has a finite-dimensional state space representation

$$x_{t+1} = Fx_t + ge_t,$$

$$d_t = h'x_t + e_t, \tag{6}$$

with some finite-dimensional vector x_t. The parameter a_j in (5) is the impulse response of (6), i.e., a_j can be represented as

$$a_j = h'F^{j-1}g, \qquad j \geq 1, \qquad a_0 = 1. \tag{7}$$

This finite parameter representation is more useful than (5) because no infinite sequence need be used. For example

$$\sum_{j=1}^\infty a_j = h'(I - F)^{-1}g \ , \qquad \text{and} \qquad \Sigma_{j=0}^\infty a_j^2 = 1 + h'X\ h$$

where

$$X = F\ X\ F' + gg' . \tag{8}$$

The constraint on a's imply that the eigenvalues of F all lie inside the unit disk. The Lyapunov equation (8) has a positive definite solution if and only if F is asymptotically stable because X is symmetric and positive definite. (Bellman (1960, Chapter 12) or Aoki (1987, p. 251)). Since

$$var(y_t - y_{t-k}) = var(d_t + d_{t-1} + \ldots + d_{t-k+1}) \ .$$

$$= k\mu_0 + 2\sum_1^{k-1}(k - l)\,\mu_l$$

where

$$\mu_0 = var\ d_t \ ,$$

and

$$\mu_l = E\ d_{t+l}d_t,$$

the limit of $var(y_t - y_{t-k})\ /\ k$ is $\mu_0 + 2\Sigma\,\mu_l$. Clearly

$$\mu_0 = (\sum_0^\infty a_j^2)\ \sigma_e^2 \ ,$$

and the limit of $var(y_t - y_{t-k})\ /\ k$ is given by $(\Sigma\, a_j)^2\ \sigma_e^2$. A simple way of seeing this is to use the spectral decomposition theorem. (See Whittle (1963) or Aoki (1987, p. 69) for example.) The sequence $\{d_t\}$ has the spectral density function

$$S(z) = \sum_{-\infty}^\infty \mu_e z^{-l}$$

$$= W(z)\, W'(1/z)\, \sigma_e^2$$

where $W(z)$ is the transfer function of the system (6), i.e., $W(z) = I + h'(zI - F)^{-1}g$. Therefore, noting (7), we have

$$S(1) = [1 + h'(I - F)^{-1}g]^2\sigma_e^2$$

$$= (\Sigma_{j=0}^{\infty} a_j)^2\, \sigma_e^2 .$$

Cochrane's ratio is thus equal to

$$\left[\Sigma_{j=0}^{\infty} a_j\right]^2 / \Sigma_{j=0}^{\infty} a_j^2 .$$

This ratio can be greater or less than one. For example, suppose $a_j = 0$ for $j \geq 2$. (This corresponds to $F = 0$ in (3).) Then $(1 + a_1)^2 / (1 + a_1^2)$ is greater than 1 if $a_1 > 0$ but less than 1 if $a_1 < 0$. This defect calls for an alternative measure. The reason that the ratio may become greater than one is that the denominator becomes less than the numerator if the correlation between the pure random components and weakly stationary components are negative. For example,

$$y_t - y_{t-1} = e_t + a_1 e_{t-1}$$

$$= (1 + a_1)e_t - a_1(e_t - e_{t-1}) .$$

Here the random component and the weakly stationary component is negatively correlated when $a_1 > 0$.

We propose a measure θ as

$$\theta = \frac{var\ \Delta y_{1t}}{var\ \Delta y_{1t} + var\ \Delta \tilde{y}_{2t}}$$

after y_t is decomposed into $y_{1t} + y_{2t}$ where y_{1t} is the pure random component. The notation Δ stands for the first difference and $\Delta \tilde{y}_{2t}$ is defined to be $\Delta y_{2t} - \hat{E}(\Delta y_{2t} \mid \Delta y_{1t})$ where $\hat{E}$ denotes orthogonal projection.

The value of $var\ \Delta y_{1t}$ is the same as the limit of $var\ (y_t - y_{t-k}) / k$ as k goes to infinity and can be identified as $A(1)^2\sigma_e^2$ when d_t is given by (1). To see this, rewrite the model (4) as

$$d_t = e_t + h'(zI - F)^{-1}ge_t$$

$$= A(1)e_t - (z - 1)\, h'(zI - F)^{-1}(I - F)^{-1}\, ge_t$$

Then

$$y_t = \frac{A(1)}{z - 1}\, e_t - h'(zI - F)^{-1}(I - F)^{-1}\, ge_t$$

In the above equation, we can identify the pure random walk component,

$$\Delta y_{1t} = A(1)\, e_t$$

and a weakly stationary component as

$$y_{2t} = -h'(zI - F)^{-1}(I - F)^{-1} g e_t .$$

var $\Delta \tilde{y}_{2t}$ can be calculated as follows. First, note that the component y_{2t} has a representation

$$y_{2t} = h' s_t$$

$$s_{t+1} = F s_t - \psi e_t$$

where

$$\psi = (I - F)^{-1} g.$$

Second

$$\Delta y_{2t} = h'(s_t - s_{t-1})$$

$$= h' F s_{t-1} - h' \psi e_{t-1}.$$

Since s_t and e_t are uncorrelated for all $t \geq 0$ by construction (s_0 is chosen to be independent of e_o),

$$var\ \Delta \tilde{y}_{2t} = var\ \Delta y_{2t} = h'FSF'h + (h'\psi)^2 \sigma_e^2$$

where

$$S = FSF' + \psi\psi'\sigma_e^2,$$

or

$$S = \sigma_e^2 Z$$

where

$$Z = FZF' + \psi\psi'.$$

Thus,

$$\theta = \frac{[1 + h'(I - F)^{-1} g]^2}{[1 + h'(I - F)^{-1} g]^2 + (h'\psi)^2 + h'FSF'h}$$

where

$$h'FZF'h = (h'F \otimes h'F)\,(I - F \otimes F)^{-1}\,(\psi \otimes \psi).$$

Random walk components of a more general model can be similarly examined. Before we proceed to it, we remark that the measure suggested by Nelson and Plosser (1982) also have some undesirable features. To demonstrate, consider their model

$$y_t = u_t + v_t$$

$$(1 - L)u_t = A(L)\, e_t$$

and

$$v_t = B(L)\, \delta_t .$$

In the above, we set the deterministic drift term to zero since it contributes nothing to the variances. Note that the term $(1 - L)\, u_t$ is the same as d_t analyzed above, and hence

$$var\ \Delta y_{1t} = A(1)^2\, \sigma_e^2,$$

because v_t has no random walk component. Here, $\Delta y_{2t} = (1 - L)\, B(L)\, \delta_t$. The term v_t has its own state space representation

$$\begin{cases} s_{t+1} = Fs_t + g\,\delta_t, \\ v_t = h's_t + \delta_t . \end{cases}$$

Noting that

$$v_t - v_{t-1} = h'Fs_{t-1} + \delta_t + (h'g - 1)\,\delta_{t-1},$$

we can write

$$var\ \Delta \tilde{y}_{2t} = [(h'g - 1)^2 + 1]\,\sigma_\delta^2 + h'F\ X\ F'h\ \sigma_\delta^2,$$

where

$$S = X\sigma_\delta^2,$$

and

$$X = FXF' + yg',$$

assuming that δ_t and e_t are uncorrelated for simplicity. Then

$$\theta = \frac{A(1)^2}{A(1)^2 + \{[1 + (h'g - 1)^2] + h'FXF'h\}\mu}$$

where

$$\mu = \sigma_\delta^2 / \sigma_e^2.$$

This shows that μ alone is not a good measure because its effect on $var\ \Delta\bar{y}_{2t}$ is affected by the system parameters h, g and F, and μ does not uniquely specify the random walk components. Only in the extreme cases, such as $F = 0$, $h'g = 1$, μ alone suffices to specify the variance ratios of components of $y\,(\theta = 4\,/\,(4+\mu)$ then).

4. Example: Logarithm of the real US GNP

When a mean zero time series is modeled as

$$y_{t+1} = y_t + u_t \tag{9}$$

where u_t is a mean-zero independent process, then (9) describes a pure random walk process. More often, however, the series $\{u_t\}$ is correlated. As an example, the quarterly US GNP series from the first quarter 1947 to the second quarter in 1986 is fitted by a state space model.

State space models may be thought of as a generalized and flexible way for handling serially correlated residuals after a major portion of the trend effects are captured by a univariate model of the form

$$\begin{aligned} x_{t+1} &= \rho x_t + b u_t, \\ y_t &= c x_t + u_t, \end{aligned} \tag{10}$$

where the residuals are highly correlated.

Here $Ey_t^2 = .129$ and $Eu_t^2 = .5 \times 10^{-2}$. With $K = 2$, $\sigma_2 / \sigma_1 = .15 \times 10^{-4}$ The system parameters are: $\rho = .979$, $b = 2.87$ and $c = .356$. Figure 1 plots the u_t series. The first order autocorrelation coefficient of u_t is about .60.

The series $\{u_t\}$ is then further modeled by

$$\begin{aligned} x_{t+1} &= Ax_t + Be_t, \\ u_t &= Hx_t + e_t. \end{aligned}$$

For example, with $dim\, x_t = 4$, the 4×4 Hankel matrices constructed from $\{u_t\}$ with $K = 6$ and 7 are used. With $K = 7$, the matrix A has eigenvalues $\lambda_1 = .968$, $\lambda_2 = .472$ and $\lambda_{3,4} = .018 \pm j\ .803$. The residuals $\{e_t\}$ has covariance $.11 \times 10^{-3}$ with excellent residual characteristics. For example, the first order autocorrelation coefficient is 1.9×10^{-4} and the second is 1.9×10^{-2}.

For this model $\theta = .0554$ is obtained. With $K = 6$ and $n = 4$ the ratio is $\theta = .044$. After several sensitivity analysis, the models with $K = 6$ or 7 and $n = 4$ seem to have the smallest Δ. From these, we conclude that the random walk component is about 6% or less.

5. Are M1 and CPI Cointegrated?

When *ln* CPI is modeled as a univariate series by the state space method, the model is

$$\begin{cases} x_{t+1} = .978x_t + 4.69u_t \\ y_t = .228x_t + u_t \end{cases}$$

where $\{u_t\}$ is correlated. For example, $\rho_1 = .78$

The univariate M1 is modeled by itself as

$$x_{t+1} = .972x_t + 5.31u_t$$
$$y_t = .198x_t + u_t$$

where

$$\rho_1 = .71.$$

When jointly modeled as a bivariate series by a first order model, the residuals are shown by Figure 2. A second order state model, however, produces much smaller residuals shown in Figure 3. The plot also makes clear that M1 and CPI are closely related, co-integrated, as it turns out. The second-order random trend is modeled as

$$f_{t+1} = \begin{bmatrix} .977 & -.030 \\ -.004 & .953 \end{bmatrix} f_t + \begin{bmatrix} 3.05 & 1.56 \\ 39.77 & -35.96 \end{bmatrix} u_t$$

$$\begin{bmatrix} m \\ cpi \end{bmatrix}_t = \begin{bmatrix} .197 & .020 \\ .227 & -.017 \end{bmatrix} f_t + u_t$$

where the residuals are then modeled by

$$x_{t+1} = \begin{bmatrix} .943 & -.125 \\ .066 & -.075 \end{bmatrix} x_t + \begin{bmatrix} 42.30 & 34.29 \\ -917.07 & 49.55 \end{bmatrix} e_t$$

$$u_t = \begin{bmatrix} .0069 & .00203 \\ .0073 & .00191 \end{bmatrix} x_t + e_t$$

The joint model has the dynamic matrix

$$\begin{bmatrix} .977 & -.030 & .032 & .009 \\ -.004 & .953 & .012 & .012 \\ 0 & 0 & .943 & -.125 \\ 0 & 0 & .066 & -.075 \end{bmatrix}$$

A single factor model for the trend is

$$x_{t+1} = .977x_t + (1.418 \quad 3.174)\, u_t,$$

$$\begin{bmatrix} m \\ cpi \end{bmatrix}_t = \begin{bmatrix} .197 \\ .227 \end{bmatrix} x_t + u_t.$$

Thus, the series are co-integrated with the ratio .227 / .197 = 1.15 because $m_t - 1.15\, cpi_t$ eliminates x_t leaving only the linear combination of u_t which is weakly stationary. The two factor models has two factors

$$f_{1t+1} = .981 f_{1t} + (-33.07 \quad 33.67)\, u_t$$

and

$$f_{2t+1} = .948 f_{2t} + (36.13 \quad 32.11)\, u_t.$$

and

$$\begin{bmatrix} m \\ cpi \end{bmatrix}_t = \begin{bmatrix} .195 \\ .229 \end{bmatrix} f_{1t} + \begin{bmatrix} .262 \\ .210 \end{bmatrix} f_{2t} + u_t.$$

Here the first factor has the ratio 1.17 and the second .80. These two factors are uncorrelated except through the u_t disturbances.

The two factors in the two-factor model are uncorrelated with the same magnitude;

$$\sqrt{.9788} \approx \sqrt{.9769.}$$

Since the contribution of f_{2t} to the trends is about 10% or less of that of f_{1t}, one could ignore f_{2t}. If f_{2t} is retained, then the two-dimensional vector $y_t = (m_t\ , cpi_t)'$ is co-intgrated if lagged variables are allowed, i.e., $y_{t+2} - 1.93y_{t+1} - .931_{yt}$ is a functions of u_{t+2}, u_{t+1} and u_t alone and thus weakly stationary. Where 1.93 is the trace of the dynamic matrix and .931 its determinant.

6. Concluding Remarks

This paper shows that the state space modeling procedure in Aoki (1987) can be extended to a two-step procedure to deal with nonstationary time series which is superior to prior detrending of individual series as used by Kitagawa in Akaike (1985). This two-step procedure has been applied first to the US real GNP series after a new measure of the random walk components in nonstationary series is proposed. This new measure corrects some undesirable features of some previously proposed measures. The two-step procedure is then applied to the US MI and CPI to obtain the vector of co-integration. These two series are shown to have two-dimensional random trends rather than a scalar trend.

References

Akaike, H., et al., *TIMSAC-84, Part 1,* Inst. Stat. Math., Tokyo, 1985

Aoki, M., *State Space Modeling of Time Series,* Springer-Verlag, Heidelberg, 1987

Aoki, M., "On Unit Roots in Time Series Modeling", unpublished ms

Aoki, M. and Havenner, *Approximate State Space Models of Some Vector-Valued Macroeconomic Time Series for Cross-Country Comparisons J.E.D.C., 10* p. 149-155, 1986

Bellman, R., *Introduction to Matrix Analysis,* McGraw Hill, 1960

Granger, C.W.J., "Co-Integrated Variables and Error-Correcting Models", UCSD Discussion Paper, 83-13a, 1983

Cochrane, J.H., "How Big is the Random Walk in GNP", *Working Paper,* University of Chicago, April 1986

King, R., C. Plosser, J. Stock and M. Watson, "Stochastic Trends and Economic Fluctuations", unpublished ms, 1987

Lindquist, A., and G. Picci, "On the Stochastic Realization Problem" *Siam Journal of Control and Optimization, 17,* p. 365-389, 1979

Nelson, C.R., and C.I. Plosser, "Trends and Random Walks in Economic Time Series", *Journal of Monetary Economics 10,* (1982), pp. 135-162

Ouliaris, S., "Testing for Unit Roots and Co-integration in Multiple Time Series", Ph.D. Thesis, Yale University, 1987

Oxford Bulletin of Economics and Statistics, 48, No.3, August, 1986

Perror, P., and P.C.B. Phillips, "Does GNP have a Unit Root? A Revelation" *Cahier 2686, C.R.D.E.,* 1986

Phillips, P.C.B., "Time Series Regression with a Unit Root", *Cowles Foundation Discussion Paper 740R,* 1986

Said, S.E., and D.A. Dickey, "Testing for Unit Roots in Autoregressive-moving Average Models of Unknown Order", *Biometrika 71 3,* pp. 599-607, 1984

Stock, J.H., and M.W. Watson, "Does GNP Have a Unit Root?" *Economics Letters,* (forthcoming)

Walsh, C., "The Impact of Monetary Targeting in the United States: 1976-1984", unpublished ms 1986

Watson, M. W., "Univariate Detrending Methods with Stochastic Trends", *Journal of Monetary Economics, 18,* p.45-75, 1986

Whittle, P., *Prediction and Regulation,* Van Nostrand, 1963

Appendix 1

Convergence of the Iterative Procedure

The matrix Riccati equation

$$\pi = A\pi A' + (M - A\pi C')(\Lambda_0 - C\pi C')^{-1}(M - A\pi C')' \tag{1}$$

may be solved iteratively

$$\pi_{n+1} = A\pi_n A' + (M - A\pi_n C')(\Lambda_0 - C\pi_n C')^{-1}(M - A\pi_n C')'.$$

Let

$$Z_n = \pi_{n+1} - \pi.$$

Then $\{Z_n\}$ is generated according to

$$\begin{aligned} Z_{n+1} &= AZ_n A' - AZ_n C'(\Lambda_0 - C\pi C')^{-1}(M - A\pi C')' \\ &\quad - (M - A\pi C')(\Lambda_0 - C\pi C')^{-1} CZ_n A' \\ &\quad + (M - A\pi C')(\Lambda_0 - C\pi C')^{-1} CZ_n C'(\Lambda_0 - C\pi C')^{-1}(M - A\pi C')' + \dots \end{aligned} \tag{2}$$

to the first order of smallness in some norm measure of Z_n. Here we used an approximate expression

$$(X - Y)^{-1} = X^{-1} + X^{-1}YX^{-1} + \dots$$

when the indicated inverses exist. Taking the *vec* of (2) and noting that

$$(A + B) \otimes (A + B) = A \otimes A + B \otimes A + A \otimes B + B \otimes B$$

we can rewrite (2) as

$$vecZ_{n+1} = (F \otimes F)\, vecZ_n \tag{3}$$

where

$$F = A - BC$$

because $(M - A\pi C')(\Lambda_0 - C\pi C')^{-1}$ is equal to B.

Eq (3) shows that $Z_n \to 0$ as $n \to \infty$ if and only if all the eigenvalues of the matrix F are less than one in magnitude, i.e., if F is asymptotically stable. This stability condition fails if $|\lambda| = 1$ is an uncontrollable eigenvalue for the pair (A, M).

To see this, it is more convenient to examine the eigenvalues of $F' = A' - C'B'$. Suppose that $A'z = \lambda z$ and $M'z = 0$ for $|\lambda| = 1$. Then

$$F' = A' - C'\Delta^{-1}(M' - C\pi A')$$

$$= (I + C'\Delta^{-1}C\pi)A' - C'\Delta^{-1}M^{-1}.$$

Rewrite the right side of the Riccati equation as

$$A\pi A' + (M - A\pi C')'(\Lambda_0 - C\pi C')^{-1}(M - A\pi C')'$$

$$= \Psi'\pi\Psi + D + \Psi'\pi C'\Delta^{-1}C\pi\Psi$$

where

$$\Psi' = A - M\Lambda_0^{-1}C, \qquad Q = C'\Lambda_0^{-1}C,$$

and

$$D = M\Lambda_0^{-1}M^{-1}.$$

From which we note that

$$\Psi'^{-1}(\pi - D)\Psi^{-1} = (\pi^{-1}Q)^{-1} = \pi(I - Q\pi)^{-1}$$

i.e.,

$$I + C'\Delta^{-1}C\pi = (I - Q\pi)^{-1}.$$

Express the symplectic matrix associated with the Riccati equation as $\Phi_1^{-1}\Phi_2$ where

$$\Phi_1 = \begin{bmatrix} I & -Q \\ 0 & \Phi' \end{bmatrix} \quad and \quad \Phi_2 = \begin{bmatrix} \Psi & 0 \\ -D & I \end{bmatrix}.$$

Then π satisfies

$$\Phi_2 \begin{bmatrix} I \\ \pi \end{bmatrix} = \Phi_1 \begin{bmatrix} I \\ \pi \end{bmatrix} K.$$

Then K is related to Γ of the Shur decomposition of the symplectic matrix

$$\Phi_2 \begin{bmatrix} U \\ V \end{bmatrix} = \Phi_1 \begin{bmatrix} U \\ B \end{bmatrix} \Gamma$$

by $KU = U\Gamma$ and $\pi = VU^{-1}$. From the above follows

$$(I - Q\pi)^{-1} = K\Psi^{-1},$$

i.e.,

$$I + C'\Delta^{-1}C\pi = K\Psi^{-1}.$$

Therefore

$$F' = K\Psi^{-1}A - C'\Delta^{-1}M'$$

$$= K + K\Psi^{-1}C'M' - C'\Delta^{-1}M^1.$$

Let z be as assumed above. Then

$$F'z = Kz = \lambda z,$$

i.e., the uncontrollable eigenvalue for the pair (A, M) is inherited by the Kalman filter dynamic matrix F.

Otherwise, $|\lambda(F)| < 1$ and the iterative procedure converges.

Appendix 2

State space model of a mixed data generating process.

This appendix shows analytically how the proposed modeling procedure works when applied directly to a time series which is a mixture of random walk and a weakly stationary series.

Suppose that y_t has two components

$$y_t = y_{1t} + y_{2t}$$

where y_{1t} is a random walk

$$y_{1t+1} = y_{1t} + n_t$$

and y_{2t} is weakly stationary. The weakly stationary component is modeled then by

$$x_{t+1} = \rho x_t + b w_t, \qquad |\rho| < 1,$$

$$y_{2t} = c x_t + w_t .$$

For simplicity assume that n_t and w_t are uncorrelated and that x_t is scalar. From the model dynamics

$$y_{2t+1} - y_{2t} = c(\rho - 1)x_t + (bc - 1)w_t + w_{t+1},$$

we see that

$$var(y_{2t+1} - y_{2t}) = c^2(\rho - 1)^2 \pi + [(bc - 1)^2 + 1]\, \sigma_w^2$$

where

$$\pi = b^2 \sigma_w^2 / (1 - \rho^2)$$

and

$$\overline{y_{2t}^2} = \sigma_w^2 + c^2 \pi.$$

The covariances have the structure

$$\gamma_k = \overline{y_{t+k} y_t} = \overline{y_{1t}^2} + c^2 \rho^{k-1}, \qquad k \geq 1,$$

where

$$\sigma_w^2 = \overline{y_{2t}^2} / [1 + b^2 c^2 / (1 - \rho^2)],$$

and where

$$\overline{y_{1t-k}^2} = \overline{y_{1t}^2} - k\sigma^2, \qquad \sigma^2 = var\ n_t .$$

The Hankel matrix

$$H = E \begin{bmatrix} y_{t+1} \\ \vdots \\ y_{t+K} \end{bmatrix} [y_t \; y_{t-1} \cdots y_{t-K+1}]$$

can be written as a sum of two rank one matrices

$$H = \sigma_1 u_1 u'_1 + \xi u_2 v'$$

where

$$\sigma_1 = c^2(1-\rho^{2K}) / (1-\rho^2)$$

$$u_1' = \sqrt{\frac{1-\rho^2}{1-\rho^{2K}}} \, (1 \; \rho \; \rho^2 \ldots \rho^{K-1}),$$

$$\xi = \sqrt{K}\, \theta_K$$

$$u'_2 = \frac{1}{\sqrt{K}} (1 \ldots 1)$$

$$v' = \frac{1}{\theta_K} (\eta, \eta - \sigma^2, \eta - 2\sigma^2, \ldots, \eta - (K-1)\,\sigma^2)$$

where

$$\eta = var\, y_{1t}$$

and

$$\theta_K^2 = K\,[\eta^2 - \eta\sigma^2 (K-1) + \sigma^4 \phi(K)]$$

where

$$K\phi(K) = \sum_{i=1}^{K} i^2 = K(K+1)(2K+1)/6.$$

To illustrate the procedure for building a state space model for this mixed process, we assume that the random walk component is small in the sense that ξ/σ_1 is a small parameter μ, which is of the order η and we drop terms ξ^2/σ_1^2 and higher in our later developments where

$$\xi / \sigma_1 = \frac{K(1-\rho^2)}{c^2(1-\rho^{2K})} \sqrt{\eta^2 \{1-\rho^2(K-1)\} + \sigma^4 \Phi(K)}$$

For later use also define two constants

$$a = u_1' v = \sqrt{\frac{1-\rho^2}{1-\rho^{2K}}} \left\{ \frac{1-\rho^K}{1-\rho} \eta - \frac{\sigma^2 \rho K\,(K-1)}{2} \right\}$$

and

$$b = u_1'u_2 = \sqrt{\frac{1-\rho^2}{(1-\rho^{2K})K}} \left[\frac{1-\rho^K}{1-\rho}\right] = \sqrt{\frac{1+\rho}{(1+\rho^K)K}} > 0.$$

We assume that $|a|$ and $|b|$ are less than one since a is of the order η, a is expected to be much smaller than b.

First, we obtain the singular value decomposition of the Hankel matrix

$$H = U\Sigma V'$$

by calculating the eigenvalues and the eigenvectors for HH' and $H'H$. We have

$$HH'[u_1u_2] = [u_1u_2]\,\Xi$$

where

$$\Xi = \begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix},$$

$$\xi_{11} = \sigma_1^2 + \sigma_1 ab\,\xi, \quad \xi_{12} = \sigma_1^2 b + \sigma_1 a\,\xi,$$

$$\xi_{21} = \sigma_1 a\,\xi + b\,\xi^2, \quad \xi_{22} = \sigma_1 ab\,\xi + \xi^2.$$

The matrix U is given by

$$U = [u_1, u_2]X$$

where

$$\Xi X = X\Sigma^2$$

$$\Sigma^2 = diag.\,(\lambda_1, \lambda_2),$$

and the eigenvalues are

$$\lambda_1 = \sigma_1^2\,(1 + 2ab\,\mu) + 0(\xi)$$

and

$$\lambda_2 = \sigma_1^2\mu^2(1-a^2)\,(1-b^2) + 0(\xi^3).$$

The matrix X is of the form

$$X = \begin{bmatrix} x_1 & l_2x_2 \\ l_1x_1 & x_2 \end{bmatrix}$$

where

$$l_1 = \frac{\xi_{21}}{\lambda_1 - \xi_{22}} \approx a\,\mu$$

and

$$l_2 = \frac{\lambda_2 - \xi_{22}}{\xi_{21}} \approx -b + 0\,(\mu).$$

The matrix U is orthogonal, i.e.,

$$I = X' \begin{bmatrix} 1 & b \\ b & 1 \end{bmatrix} X.$$

The matrix X is given as

$$X \approx \begin{bmatrix} 1 & -b \,/\, \sqrt{1-b^2} \\ 0 & 1 \,/\, \sqrt{1-b^2} \end{bmatrix}.$$

Proceeding analogously, define Θ by

$$H'H\,[u_1 \quad v] = [u_1 \quad v]\,\Theta$$

where

$$\Theta = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix},$$

$$\theta_{11} = \sigma_1^2 + \sigma_1 ab\,\xi, \quad \theta_{12} = \sigma_1^2\,a + \sigma_1 b\,\xi,$$
$$\theta_{21} = \sigma_1 b\,\xi + a\,\xi^2, \quad \theta_{22} = \sigma_1 ab\,\xi + \xi^2.$$

The matrix V is then given by

$$V = [u_1 \quad v]\,Y$$

where

$$\Theta\,Y = Y\Sigma^2,$$

The normalization $V'V = I$ gives Y to be

$$Y \approx \begin{bmatrix} 1 & -a \,/\, \sqrt{1-a^2} \\ 0 & 1 \,/\, \sqrt{1-a^2} \end{bmatrix}.$$

Since

$$H^A = \rho\sigma_1\,u_1 u_1' + \xi u_2 v',$$

we obtain

$$\hat{A} = \Sigma^{-1/2} U' H^A V \Sigma^{-1/2},$$

and from

$$H^c = c\sqrt{\sigma_1}\,u_1' + \frac{\xi}{\sqrt{K}}\,v'$$

we compute

$$\hat{c} = H^c V \Sigma^{-1/2}.$$

The model has two factors, x_{1t} and x_{2t}, which evolve with time according to

$$\begin{cases} x_{1t+1} \approx \rho x_{1t} + \sqrt{\mu}\, b \left[\dfrac{1-a^2}{1-b^2}\right]^{1/4} x_{2t} + g_1 e_t \\ x_{2t+1} \approx a\sqrt{\mu} \left[\dfrac{1-b^2}{1-a^2}\right]^{1/4} x_{1t} + x_{2t} + g_2 e_t \approx x_{2t} + g_2 e_t \end{cases}$$

and because a is of the order η which is typically smaller than b.

$$y_t = (c + a\mu\sqrt{\sigma_1 / K})\, x_{1t} + (\sigma_1 \mu / K)^{1/2} \left[\frac{1-a^2}{1-b^2}\right]^{1/4} x_{2t} + e_t$$

$$\approx c x_{1t} + (\sigma_1 \mu / K)^{1/2} [(1-a^2)/(1-b^2)]^{1/4} x_{2t} + e_t \qquad (1)$$

where

$$\frac{\sigma_1 \mu}{K} = \eta \sqrt{1 - \delta^2 (K-1) + \frac{\sigma^4}{\eta^2} \Phi(K).}$$

At one extreme point $\mu = 0$, the procedure recovers the state space model for y_{1t} and y_{2t}. Note that if the y_{2t} component is absent, then $H = \xi u_2 v'$, $H^A = \xi u_2 v'$ and $H^c = \theta_K v'$ $U = u_2$, and $V = v_2$, and $\Sigma = \xi$. From these, one calculates $\xi^{-1/2} U' H^A V \xi^{-1/2} = 1$ and $H^c V \Sigma^{-1/2} = \sqrt{\xi / K}$ as a and c, i.e., we recover a random walk model

$$x_{t+1} = x_t + b e_t$$

$$y_t = \sqrt{\xi / K}\, x_t + e_t,$$

or

$$y_{t+1} = y_t + \sqrt{\frac{\xi}{K}}\, b e_t$$

$$= y_t + u_t.$$

For small μ, the model has the triangular dynamics of the canonical representation in Aoki (1987, p. 39)

$$\begin{bmatrix} x_{2t+1} \\ x_{1t+1} \end{bmatrix} = \begin{bmatrix} 1 & h \\ 0 & \rho \end{bmatrix} \begin{bmatrix} x_{2t} \\ x_{1t} \end{bmatrix} + e_t, \quad \text{and} \quad y_t = c_1 x_{1t} + c_2 x_{2t} + e_t$$

where c_1 and c_2 are as in (1).

The ratio of the singular value is

$$\sqrt{\lambda_2 / \lambda_1} = \mu \sqrt{(1 - a^2)(1 - b^2)}$$

which is typically of the order 10^{-3} or less for small K. The ratio σ_2 / σ_1 of the real GNP series is 2.4×10^{-4} when $K = 2$. This indicates that η is of the order 10^{-4}.

residuals of ln rgnp by m2.1 for 1982 constant dollar data

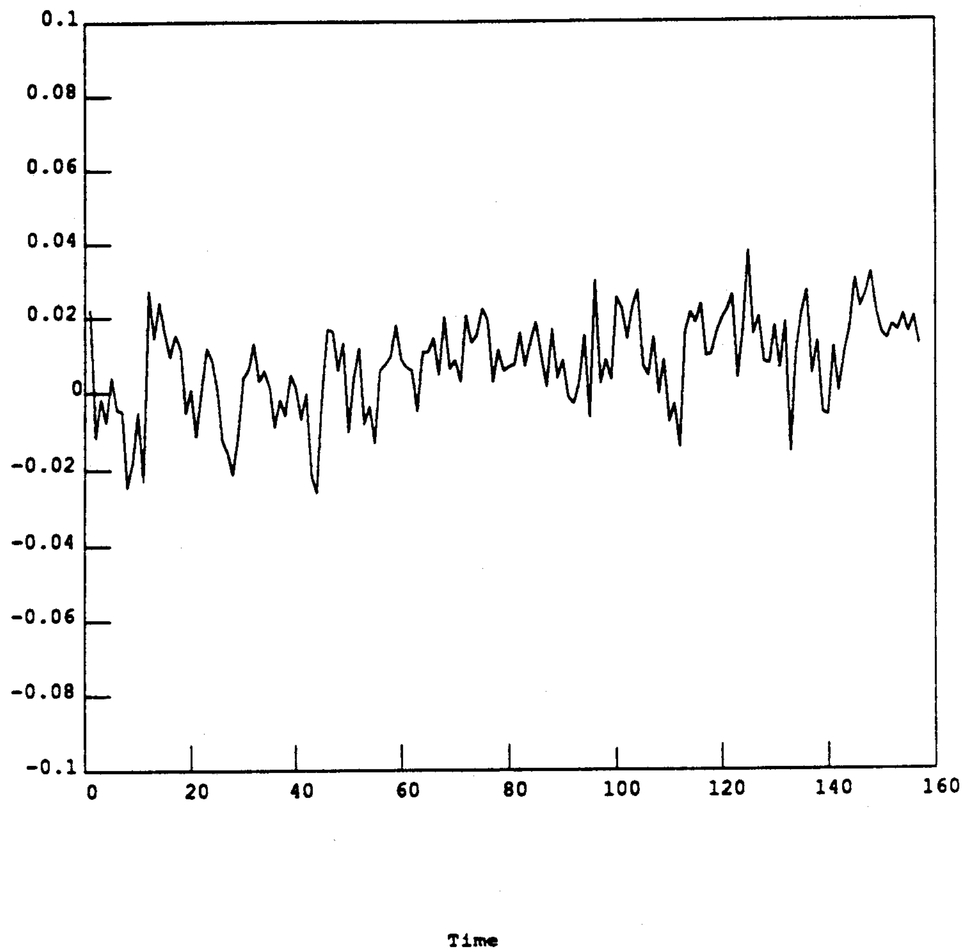

Figure 1

Residuals of ln GNP at 1982 constant dollar by a scalar dynamics (K = 2, n = 1)

residuals of ln m1 cpi by mcpi1-2

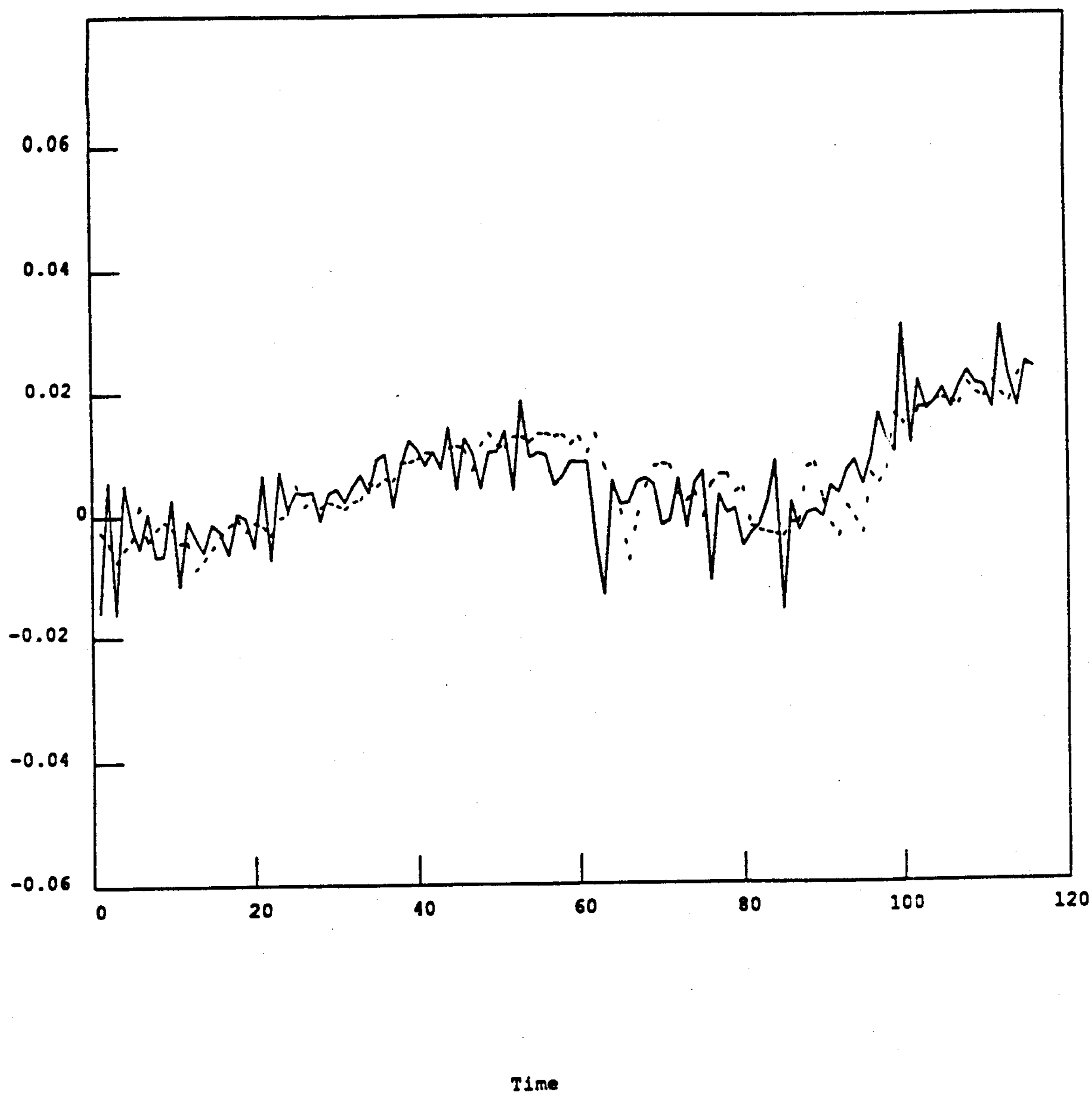

Figure 2

Residuals of ln M1 and ln CPI by 2-dim trend dynamics. (K = 1, n = 2)

residuals of ln ml cpi by mcpil.1

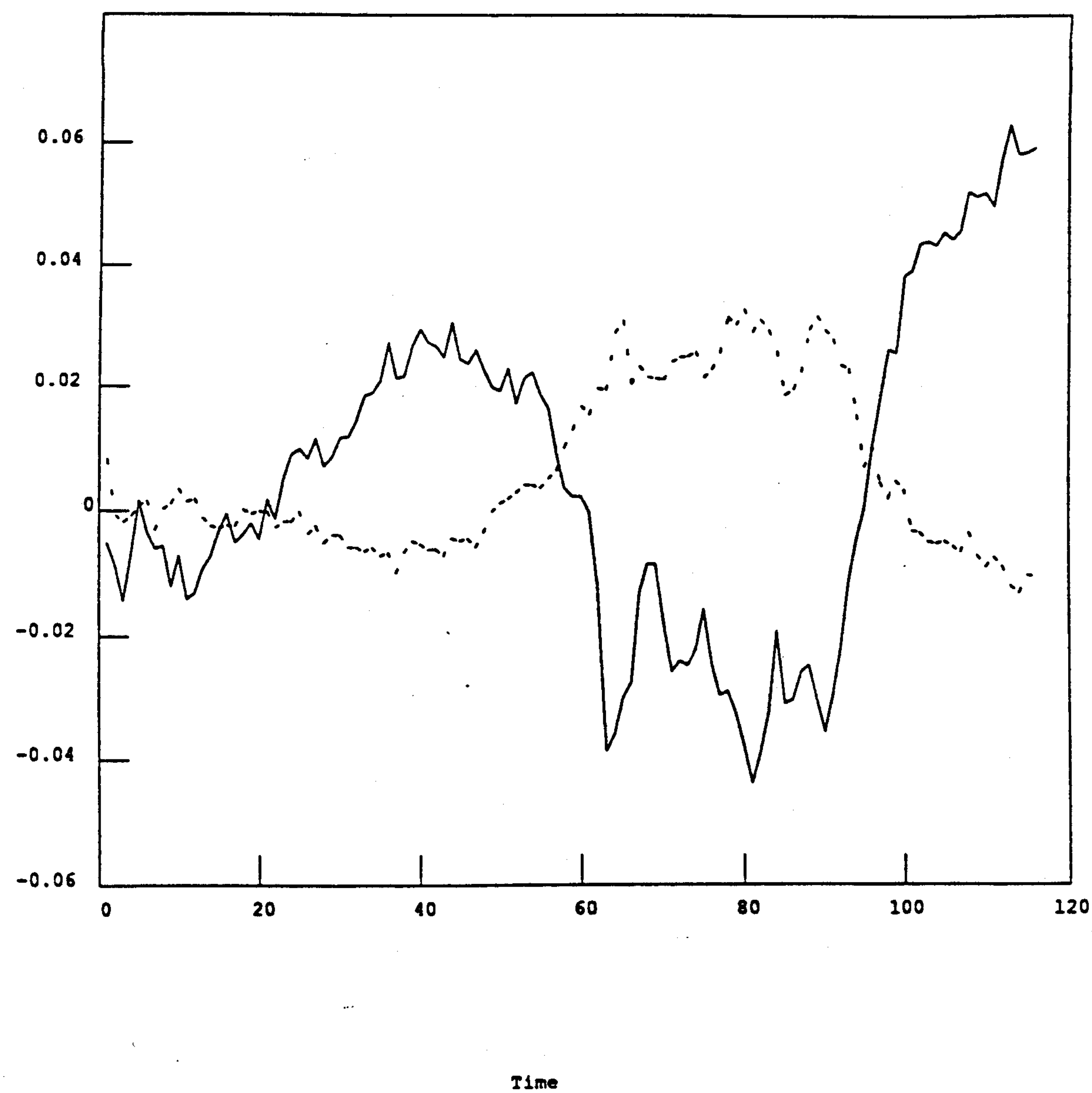

Figure 3

Residuals of ln of M1 and CPI by a scalar model (k = 1, n = 1).