

KALMAN FILTERING ESTIMATION OF UNOBSERVED RATIONAL EXPECTATIONS WITH AN APPLICATION TO THE GERMAN HYPERINFLATION*

Edwin BURMEISTER

*University of Illinois, Champaign, IL 61820, USA
University of Virginia, Charlottesville, VA 22901, USA*

Kent D. WALL

University of Virginia, Charlottesville, VA 22901, USA

Received September 1980, final version received April 1982

The assumption that rational expectations always lie on a convergent path is subject to an empirical test using the German hyperinflation data. The estimation technique employs a Kalman filtering algorithm. After presenting a brief background for the convergent expectations problem and a derivation of the various model specifications, a generalized expectations model and its attendant Kalman filtering estimation technique are discussed. Additional estimation details and empirical results are then presented. Based on an assumption of normally distributed errors, the null hypothesis of convergent paths is rejected in all situations involving a deterministic specification of the evolution of the unobserved parameter which characterizes the convergent path. The same null hypothesis is rejected in four of the six cases corresponding to a stochastic specification of the evolution of the unobserved parameter which characterizes the convergent path. A discussion of these findings, their economic significance, and suggestions for further research concludes the paper.

1. Introduction

Elsewhere Burmeister (1980, 1982) has summarized the conceptual problems which arise in rational expectations modelling due to the common *assumption* that rationally formed expectations always lie on convergent paths. When this assumption is not made — and sometimes even when it is — the model is not determinant in the sense that there exist many stochastic paths for the actual variables, all of which are consistent both with equilibrium in every time period and with the rational expectations hypothesis. Accordingly, since the postulate of convergent expectations

*The authors thank the National Science Foundation (SOC-76-03608-1) and the Federal Trade Commission (L0638) for financial support. Earlier comments from Robert P. Flood, Peter M. Garber, Bennett T. McCallum, and especially Kenneth J. Singleton are gratefully acknowledged, and suggestions from two referees on an earlier version have provided many improvements. E. Burmeister is also grateful for research support from Duke University where some of this work was completed.

carries with it important economic implications, and because the postulate is testable, a careful empirical investigation is merited.

A general framework within which this stability issue can be analyzed empirically has been developed by Wall (1980). A brief background for the convergent expectations question is provided in section 2, along with a derivation of the various model specifications we shall investigate. In section 3 we turn to the generalized expectation model and discuss how it is applied to obtain econometric estimates for data from the German hyperinflation. Additional details regarding estimation are presented in section 4 along with the empirical results.

2. A brief background and derivation of the basic model

The common assumption that rationally formed expectations always converge is crucial for at least three reasons. First, in many models this assumption is needed to determine a unique monetary equilibrium at each instant; without this convergence assumption, the future path of the economy modeled might be indeterminate, even without uncertainty. Second, the assumption of convergent expectations is commonly used to provide cross-equation restrictions which facilitate identification and econometric estimation; thus, most econometric estimates of rational expectations models are conditional upon this assumption. Third, if markets are not always in equilibrium, then there exist many cases in which the assumption of convergent expectations is untenable because it implies a contradiction.^{1,2}

¹See Burmeister, Flood and Turnovsky (1979). Burmeister (1980, 1982) provides a survey of some conceptual issues in rational expectations modelling, including the question of convergent expectations. Burmeister, Flood and Garber (1983) have shown that there is only one type of indeterminacy in rational expectation models; in the model considered here, there is indeterminacy whenever expectations are not convergent because there exist an infinity of equilibrium rational expectations paths with divergent expectations.

²The fundamental problem of non-convergent price expectations which in turn may cause the problem of divergent actual prices, has been recognized for over twenty years. Writing in 1957, Samuelson discussed the issue of non-convergent paths:

So much for the avoidable difficulties introduced by infinite time. Now to return to the intrinsic difficulty. I shall call it the 'tulip-mania phenomenon.' Let the market maximize over any finite time, adding in at the end into the thing to be maximized a value for the terminal amount of grain left. At what level should this terminal grain be valued? We could extend the period in order to find out how much it is really worth in the remaining time left; but this obviously leads us back into our infinite regression, since there is always time left beyond any extended time. We are back into maximizing over infinite time.

But suppose we do what the market itself does in evaluating any stock $Q(t)$ at any given date; suppose we simply evaluate it at the then ruling market price $P_0(t)$. Then we immediately run into the paradox that any speculative bidding up of prices at a rate equal to carrying costs can last *forever*. This is precisely what happens in a tulip mania or newer bull stock market. The market literally lives on its own dreams, and each individual at every moment of time is perfectly rational to be doing what he is doing. [Samuelson (1957, pp. 215-216)].

In order to make the above ideas concrete, we shall base our discussion on the following simple stochastic monetary model:

$$m^d(t) - p(t) = b - a[p^*(t+1, t) - p^*(t, t)], \quad a > 0, \quad (2.1a)$$

$$p^*(t+1, t) \equiv E_t p(t+1), \quad (2.1b)$$

$$m(t) = \alpha_0 + \alpha_1 m(t-1) + \varepsilon(t), \quad (2.1c)$$

$$m^d(t) = m(t) + \mu(t), \quad (2.1d)$$

where

$p(t)$ = logarithm of the price level at time t ;

$E_t p(t+h)$ = conditional expectation of $p(t+h)$, $h=0, 1, \dots$, formed at time t based upon all the information available at time t , which is assumed to be $I(t) = \{a, b, \alpha_0, \alpha_1; p(t-1), p(t-2), \dots; m(t-1), m(t-2), \dots\}$;

$m(t)$ = logarithm of the nominal stock of money at time t ;

$m^d(t)$ = logarithm of the demand for the nominal stock of money at time t ;

$\varepsilon(t)$ = serially uncorrelated stochastic disturbance in the money supply over the period $(t-1, t)$ with $E[\varepsilon(t) | I(t)] = 0$;

$\mu(t)$ = serially uncorrelated stochastic disturbance in the price adjustment equation over the period $(t-1, t)$ with $E[\mu(t) | I(t)] = 0$;

a = a positive constant related to elasticity of the demand for real balances with respect to the expected rate of inflation;

b = a constant reflecting other variables influencing the demand for money which are held constant;

α_0, α_1 = parameters of the money supply equation.

The system described by these four equations can be viewed as a simple extension of the models used by Sargent and Wallace (1973) and Black (1974), which now include stochastic disturbances in the money supply. Eq. (2.1a) follows these previous authors in specifying the demand for real money balances to be a function of the anticipated rate of inflation over the period $(t, t+1)$. These expectations are formed rationally in the sense of Muth (1961), and accordingly the expected price level is specified to equal the conditional expectations of actual prices, as described by (2.1b). Eq. (2.1c) determines the evolution of the stochastic money supply. Finally, eq. (2.1d) describes stochastic equilibrium in the money market. The stochastic

disturbances, $\varepsilon(t)$ and $\mu(t)$, are assumed to be sequentially independent with zero means.³

The model reduces to

$$m(t) - p(t) = b - a[p^*(t+1, t) - p^*(t, t)] + \mu(t), \quad (2.2)$$

and

$$m(t) = \alpha_0 + \alpha_1 m(t-1) + \varepsilon(t). \quad (2.3)$$

Assuming that economic agents know the parameters a , b , α_0 , and α_1 with certainty, taking rational expectations of (2.2) and (2.3) at any time $t+h \geq t$, yields

$$m^*(t+h, t) - p^*(t+h, t) = b - a[p^*(t+h+1, t) - p^*(t+h, t)], \quad (2.2')$$

and

$$m^*(t+h, t) = \alpha_0 + \alpha_1 m^*(t+h-1, t), \quad (2.3')$$

where $h=0, 1, 2, \dots$.

The solution to the difference eq. (2.3') with initial condition $m^*(t-1, t) = m(t-1)$ is

$$m^*(t+h, t) = \left[\frac{\alpha_0}{\alpha_1 - 1} + m(t-1) \right] \alpha_1^{h+1} - \frac{\alpha_0}{\alpha_1 - 1}, \quad h=0, 1, 2, \dots \quad (2.4)$$

Substitution of (2.4) into (2.2') gives

$$p^*(t+h+1, t) = \frac{b}{a} + \left(1 + \frac{1}{a} \right) p^*(t+h, t) - \frac{1}{a} \left\{ \left[\frac{\alpha_0}{\alpha_1 - 1} + m(t-1) \right] \alpha_1^{h+1} - \frac{\alpha_0}{\alpha_1 - 1} \right\},$$

³We assume that the joint process, $\varepsilon(t)$, defined by

$$\varepsilon(t) = [\varepsilon(t), \mu(t)]',$$

constitutes a zero mean, sequentially independent vector stochastic disturbance. Specifically,

$$E_t \varepsilon(t) = \theta, \quad (a)$$

$$E_t \{ \varepsilon(t) \varepsilon'(\tau) \} = \theta, \quad (b)$$

for all $\tau \neq t$, and

$$E \{ \varepsilon(t) \varepsilon'(\tau) \} = R < \infty, \quad (c)$$

for $t = \tau$.

or

$$p^*(t+h+1, t) = \frac{b\alpha_1 - b + \alpha_0}{a(\alpha_1 - 1)} + \left(1 + \frac{1}{a}\right) p^*(t+h, t) - \frac{1}{a} \left[\frac{\alpha_0}{\alpha_1 - 1} + m(t-1) \right] \alpha_1^{h+1}, \quad h=0, 1, 2, \dots \quad (2.5)$$

Since $(1 + 1/a) > 1$ as the parameter a is positive, the only possibility for convergent price expectations arises from the ‘forward-looking’ solution to (2.5),

$$p^*(t+h, t) = \frac{-\alpha_0 + b - b\alpha_1}{\alpha_1 - 1} + \frac{(1 + 1/a)^h}{a} \left[m(t-1) + \frac{\alpha_0}{\alpha_1 - 1} \right] \sum_{\theta=h+1}^{\infty} \left(\frac{\alpha_1}{1 + 1/a} \right)^\theta. \quad (2.6)$$

Accordingly, the initial condition consistent with convergent expectations is uniquely given by setting $h=0$ in (2.6); assuming that $-1 < (\alpha_1/(1 + 1/a)) < +1$, this yields the special initial condition

$$p^*(t, t) = \frac{-\alpha_0 + b - b\alpha_1}{\alpha_1 - 1} + \frac{\alpha_1}{1 + a - a\alpha_1} \left[m(t-1) + \frac{\alpha_0}{\alpha_1 - 1} \right]. \quad (2.7)$$

The crucial economic significance of (2.7) — which itself was derived by imposing the *assumption* that price expectations are convergent — stems from the fact that it determines actual prices which are then given by

$$p(t) = p^*(t, t) - \mu(t) + \varepsilon(t), \quad (2.8)$$

where $p^*(t, t)$ is determined from (2.7). To derive (2.8), subtract (2.2') with $h=0$ from (2.2) and subtract (2.3') with $h=0$ from (2.3), giving

$$[m(t) - p(t)] - [m^*(t, t) - p^*(t, t)] = \mu(t),$$

and

$$m(t) - m^*(t, t) = \varepsilon(t).$$

Substituting the latter into the former yields (2.8).

If alternatively one assumes that

$$p(t) = \pi_0 + \pi_1 m(t) + \pi_2 m(t - 1) + \pi_3 \mu(t),$$

for some non-zero π_0, π_1, π_2 and π_3 , and then imposes rational expectations, the solution obtained is identical to that given by (2.7) and (2.8). This latter procedure (sometimes called the ‘undetermined coefficient’ algorithm) may be *equivalent*, as it is in this example, to assuming convergent expectations.⁴

However, (2.6) is *not* the unique ‘forward-looking’ solution to (2.5) if the assumption of convergent expectations is dropped. There are an infinity of divergent ‘forward-looking’ rational expectations solutions obtained by adding the unstable term

$$c_t(1 + 1/a)^t \tag{2.9}$$

to (2.6), where c_t is an arbitrary constant. This gives

$$p^*(t, t) = c_t + \frac{-\alpha_0 + b - b\alpha_1}{\alpha_1 - 1} + \frac{\alpha_1}{1 + a - a\alpha_1} \left[m(t - 1) + \frac{\alpha_0}{\alpha_1 - 1} \right]. \tag{2.10}$$

In this model the *assumption* that expectations are always convergent is equivalent to imposing $c_t = 0$ for all t .

Our approach will be to treat c_t as an unobserved variable which is estimated jointly with the other unobserved variables — $p^*(t, t), p^*(t + 1, t)$, and $m^*(t, t)$ — and with the other parameters of the model. The details of this technique are spelled out in section 3 and 4 below.

The c_t ’s are restricted by the rational expectations hypothesis to satisfy

$$E[c_{t+h} | I(t)] = c_t(1 + 1/a)^h, \tag{2.11}$$

where the information set $I(t)$ is now expanded to include $c_t, c_{t-1}, c_{t-2}, \dots$.⁵ A stochastic process consistent with the restriction (2.9) is

$$c_{t+1} = \left(1 + \frac{1}{a} \right) c_t + \eta(t), \tag{2.12}$$

where $E[\eta(t + h) | I(t)] = 0$ for $h = 0, 1, 2, \dots$.⁶ Thus the actual c_{t+1} differs from

⁴See Burmeister (1982) for a complete discussion of this issue.

⁵This restriction follows from the fact that rationality necessitates

$$E[p^*(t + h + 1, t + h) | I(t)] = p^*(t + h + 1, t),$$

i.e.: ‘What I now think I will expect in the future is the same as what I do expect now.’

⁶For estimation purposes we shall require the stronger assumption that the joint process $[\varepsilon_1(t), \mu_1(t), \eta(t)]'$ is a zero mean, sequentially independent vector stochastic disturbance, where $\varepsilon_1(t) \equiv \varepsilon(t) - \varepsilon(t - 1)$ and $\mu_1(t) \equiv \mu(t) - \mu(t - 1)$; see footnote 3. We furthermore assume that $E[\varepsilon_1(t)\eta(t)] = E[\mu_1(t)\eta(t)] = 0$ but $\varepsilon_1(t)$ and $\mu_1(t)$ may be correlated.

the expectation $E[c_{t+1} | I(t)]$ by the random error term $\eta(t)$ which we assume has finite variance.

Our complete model consists of eqs. (2.3), (2.8) and (2.12).

Derivation in difference form

Estimation difficulties were encountered which were circumvented by respecifying the model in first differences. Thus we define the *expected inflation rate*

$$\pi^*(t+h, t+h) \equiv p^*(t+h+1, t+h) - p^*(t+h, t+h), \quad (2.13)$$

and the corresponding *actual inflation rate*

$$\pi(t+h) \equiv p(t+h+1) - p(t+h), \quad h=0, 1, 2, \dots \quad (2.14)$$

Likewise, the *expected* and *actual money growth rates* are defined as

$$g^*(t+h, t+h) \equiv m^*(t+h+1, t+h) - m^*(t+h, t+h), \quad (2.15)$$

and

$$g(t+h) \equiv m(t+h+1) - m(t+h), \quad h=0, 1, 2, \dots, \quad (2.16)$$

respectively. Moreover, for all $h=0, 1, 2, \dots$ we have that

$$E[\pi^*(t+h, t+h) | I(t)] = \pi^*(t+h, t), \quad (2.17)$$

and

$$E[g^*(t+h, t+h) | I(t)] = g^*(t+h, t). \quad (2.18)$$

From eq. (2.2),

$$\begin{aligned} g(t+h) - \pi(t+h) = & -a[\pi^*(t+h+1, t+h+1) - \pi^*(t+h, t+h)] \\ & + \mu(t+h+1) - \mu(t+h), \end{aligned} \quad (2.19)$$

for all $h=0, 1, 2, \dots$, while (2.3) implies that

$$g(t+h) = \alpha_1 g(t+h-1) + \varepsilon(t+h+1) - \varepsilon(t+h), \quad h=0, 1, 2, \dots \quad (2.20)$$

We make the following alternative assumption about the stochastic processes governing (2.2) and (2.3); it implies that $\varepsilon(t)$ and $\mu(t)$ are random walks:

For all $h=0, 1, 2, \dots$,

$$E[\mu(t+h+1) - \mu(t+h) | I(t)] = 0, \quad (2.21)$$

and

$$E[\varepsilon(t+h+1) - \varepsilon(t+h) | I(t)] = 0, \quad (2.22)$$

where

$$I(t) = \{a, \alpha_0, \alpha_1; p(t-1), p(t-2), \dots; m(t-1), m(t-2), \dots; c_t, c_{t-1}, \dots\}.$$

In view of assumption (2.21), taking conditional expectations of (2.19) at time t , using the information set $I(t)$ known at that time, yields

$$g^*(t+h, t) - \pi^*(t+h, t) = -a[\pi^*(t+h+1, t) - \pi^*(t+h, t)] + 0, \quad (2.23)$$

$$h=0, 1, 2, \dots$$

Analogously, (2.20) and (2.22) imply that

$$g^*(t+h, t) = \alpha_1 g^*(t+h-1, t) + 0, \quad h=0, 1, 2, \dots \quad (2.24)$$

The solution to the difference eq. (2.24) is

$$g^*(t+h, t) = \alpha_1^{h+1} g^*(t-1, t), \quad h=0, 1, 2, \dots, \quad (2.25)$$

where the initial condition is calculated as

$$g^*(t-1, t) = E[g(t-1) | I(t)] = \alpha_1 g(t-2). \quad (2.26)$$

Then, substituting (2.25) and (2.26) in (2.23) results in the difference equation

$$\pi^*(t+h+1, t) = (1+1/a)\pi^*(t+h, t) - (1/a)[\alpha_1^{h+2} g(t-2)]. \quad (2.27)$$

The 'forward-looking' stable solution to (2.27), which is analogous to (2.6), is

$$\begin{aligned} \pi^*(t+h, t) &= (1/a)(1+1/a)^h \alpha_1 g(t-2) \sum_{\theta=h+1}^{\infty} \left[\frac{\alpha_1}{1+1/a} \right]^\theta \\ &= \frac{\alpha_1^{h+2}}{1+a-a\alpha_1} g(t-2), \end{aligned} \quad (2.28)$$

provided $-1 < |\alpha_1/(1+1/a)| < +1$ as we shall assume. Thus setting $h=0$ and $h=1$ in (2.28) we may calculate

$$\pi^*(t, t) = \frac{\alpha_1^2}{1+a-a\alpha_1} g(t-2), \quad (2.29)$$

and

$$\pi^*(t+1, t) = \frac{\alpha_1^3}{1+a-a\alpha_1} g(t-2). \quad (2.30)$$

It is easily verified that (2.29) and (2.30) satisfy (2.27) with $h=0$; they represent rationally formed expected inflation rates which are consistent with the assumption of convergent expectations.

In general, however, the solutions to (2.27) which are consistent with 'forward-looking' rational expectations are of the form

$$\pi^*(t+h, t) = c_t(1+1/a)^h + \frac{\alpha_1^{h+2}}{1+a-a\alpha_1} g(t-2), \quad h=0, 1, 2, \dots, \quad (2.31)$$

where the c_t 's satisfy (2.12). Note that (2.31) and (2.28) are equivalent if and only if $c_t=0$ in (2.31). However, for non-zero c_t , (2.31) is not convergent since $\lim_{h \rightarrow \infty} (1+1/a)^h = +\infty$. In view of (2.12), we see that

$$\pi^*(t, t) = c_t + \frac{\alpha_1^2}{1+a-a\alpha_1} g(t-2), \quad (2.32)$$

and

$$\pi^*(t+1, t+1) = c_{t+1} + \frac{\alpha_1^2}{1+a-a\alpha_1} g(t-1). \quad (2.33)$$

Substituting (2.32) and (2.33) into (2.19) with $h=0$ gives

$$\begin{aligned} g(t) - \pi(t) &= -a(c_{t+1} - c_t) - \frac{a\alpha_1^2}{1+a-a\alpha_1} [g(t-1) - g(t-2)] \\ &\quad + \mu(t+1) - \mu(t). \end{aligned} \quad (2.34)$$

We now take conditional expectations of (2.34), using the fact that $E[c_{t+1} | I(t)] = c_t(1+1/a)$ from assumption (2.12),

$$g^*(t, t) - \pi^*(t, t) = -c_t - \frac{a\alpha_1^2}{1+a-a\alpha_1} [g(t-1) - g(t-2)]. \quad (2.35)$$

Subtracting (2.34) from (2.35) yields

$$\pi(t) - \pi^*(t) = ac_{t+1} - (1+a)c_t + \varepsilon(t+1) - \varepsilon(t) - \mu(t+1) + \mu(t), \quad (2.36)$$

since $g(t) - g^*(t, t) = \varepsilon(t+1) - \varepsilon(t)$. The first two terms on the right-hand side of (2.36) sum to $a\eta(t)$, and therefore

$$E[\pi(t) - \pi^*(t, t) | I(t)] = 0,$$

confirming that expectations have been formed rationally.

As noted below, we found it necessary to consider a second-order money growth rate process. Thus following Flood and Garber (1980), we also have estimated another version of the model for which the model supply process is respecified as

$$g(t) = \alpha_0 + (1 + \alpha_1)g(t-1) - \alpha_1 g(t-2) + \omega(t), \quad (2.37)$$

where $\omega(t)$ is a white noise process. Eq. (2.37) implies

$$\begin{aligned} g^*(t+h, t) &= g(t-2) + \frac{\alpha_1}{1-\alpha_1} (1-\alpha_1^{h+2}) [g(t-2) - g(t-3)] \\ &\quad + \frac{\alpha_0}{1-\alpha_1} (h+2) + \frac{\alpha_0 \alpha_1}{(1-\alpha_1)^2} (\alpha_1^{h+2} - 1). \end{aligned} \quad (2.38)$$

These expectations for the rate of growth of the money supply then yield the following specification for the evolution of the inflation expectations:⁷

$$\begin{aligned} \pi^*(t+1, t+1) - \pi^*(t, t) &= [c_{t+1} - c_t] + [g(t-1) - g(t-2)] \\ &\quad + \frac{\alpha_1}{1-\alpha_1} \left(1 - \frac{\alpha_1^2}{1+a-a\alpha_1} \right) \\ &\quad \times [g(t-1) - 2g(t-2) + g(t-3)]. \end{aligned} \quad (2.39)$$

⁷Eq. (2.39) is derived from

$$\begin{aligned} &\pi^*(t+h, t)c_t(1+1/a)^h + (1/a)(1+1/a)^h \sum_{\theta=h+1}^{\infty} (1+1/a)^{-\theta} g^*(t+\theta-1, t) \\ &= c_t(1+1/a)^h + g(t-2) + \frac{\alpha_0}{1-\alpha_1} (h+2+a) \\ &\quad + \frac{\alpha_0 \alpha_1}{(1-\alpha_1)^2} \left[\frac{\alpha_1^{h+1}}{1+a-a\alpha_1} - 1 \right] - \frac{\alpha_1}{1-\alpha_1} \left[\frac{\alpha_1^{h+1}}{1+a-a\alpha_1} - 1 \right] [g(t-2) - g(t-3)], \end{aligned}$$

which is the general solution to the difference equation

$$\pi^*(t+h+1, t) = (1+1/a)\pi^*(t+h, t) - (1/a)g^*(t+h, t),$$

i.e., it is the solution to (2.23).

Finally, the actual rate of inflation is given by

$$\pi(t) = \pi^*(t, t) + a\eta(t) + \omega(t) - \mu(t+1) + \mu(t), \quad (2.40)$$

with $E[\pi(t) | I(t)] = \pi^*(t, t)$, again confirming rational expectations.

3. A state space model of expectation formation

The empirical investigation of the convergent expectations question can be approached in a number of ways. Perhaps the most attractive from a traditional econometric viewpoint is to solve the system analytically and substitute these solutions into the behavioural system (2.2)–(2.3). The result is a system of equations involving interequation restrictions that are amenable, at least in principle, to standard regression techniques. While such an approach eliminates the bothersome appearance of unobserved variables like $p^*(t+h, t)$ and c_t , it may not be the least complicated since it necessitates the formation of convolutions on $m(t-1)$ and $\mu(t)$.

An alternative approach, and the one taken here, is to retain the explicit reference to the unobserved variables. To wit, a state space form representation is employed wherein the unobserved variables become state variables. Drawing upon a well-developed theory from control engineering, it is then possible to obtain simultaneously estimates of both the model parameters and the unobserved expectations. This permits a much easier test of certain hypotheses than would otherwise be the case. Since such an approach is relatively unfamiliar, we first present the state space interpretation of models introduced in section 2, and describe its estimation.

3.1. The state space model

Let $x(t)$ be an n -vector of state variables, $u(t)$ an m -vector of 'inputs', and $y(t)$ an l -vector of 'outputs'. Then a state space model relating these variables to one another is given by

$$x(t+1) = Fx(t) + Gu(t) + \Gamma\eta(t), \quad (3.1)$$

$$y(t) = Hx(t) + Du(t) + \varepsilon(t). \quad (3.2)$$

Both $\eta(t)$ and $\varepsilon(t)$ denote vectors of random variables such that

$$E\{\eta(t)\} = 0, \quad E\{\varepsilon(t)\} = 0,$$

and

$$E\{\eta(t), \eta'(s)\} = Q\delta_{ts}, \quad E\{\varepsilon(t), \varepsilon'(s)\} = R\delta_{ts},$$

F , G , Γ , H , D , Q , and R are constant real matrices of dimensions compatible with the vector-matrix operations above, and δ_{i_s} denotes the Kronecker delta function. The $l \times l$ variance-covariance matrix R is always assumed positive definite, while the $n \times n$ variance-covariance matrix Q need be only non-negative definite (i.e., positive semi-definitive).

In the context of section 2, three state space representations of the rational expectations hypothesis can be obtained, depending upon the use of $p(t)$ and $m(t)$, or $\pi(t)$ and $g(t)$. In the first instance, a state space model arises once we make the following definitions:

$$\left. \begin{aligned} \mathbf{x}'(t) &= [c_t] \\ \mathbf{u}'(t) &= [m(t), p(t), m(t-1), 1] \\ \boldsymbol{\eta}'(t) &= [\eta(t)] \\ \mathbf{y}'(t) &= [m(t), p(t)] \\ \boldsymbol{\varepsilon}'(t) &= [\varepsilon(t), \varepsilon(t) - \mu(t)] \end{aligned} \right\} \quad (3.3)$$

The coefficient matrices appearing in (3.1) are then

$$F = [1 + 1/a], \quad G = [0 \ 0 \ 0 \ 0], \quad \Gamma = [1],$$

while those appearing in (3.2) become

$$H = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & 0 & \alpha_1 & \alpha_0 \\ 0 & 0 & \frac{\alpha_1}{1+a-a\alpha_1} & \frac{1}{(\alpha_1-1)} \left[b(1-\alpha_1) - \alpha_0 + \frac{\alpha_0\alpha_1}{1+a-a\alpha_1} \right] \end{bmatrix}.$$

There is only a single state equation since imposing rational expectations produces a closed form expression for $p^*(t, t)$. Thus, given (2.10) and $m(t-1)$ one can compute $p^*(t, t) - c_t$ exactly. Price expectations under the rational expectations hypothesis only become 'unobserved variables' insofar as their c_t component is concerned, and thus only c_t need be a state variable. The first equation of (3.2) is merely (2.3), while the second is a combination of (2.8) and (2.10).

A second state space model representation is required to describe the differenced version of section 2. This is obtained by making the following

definitions once the α_0 term has been reintroduced in (2.20),

$$\left. \begin{aligned} \mathbf{x}'(t) &= [c_t, c_{t+1} - c_t] \\ \mathbf{u}'(t) &= [g(t), \pi(t), g(t-1), g(t-2), 1] \\ \boldsymbol{\eta}'(t) &= [\eta(t), \eta(t+1)] \\ \mathbf{y}'(t) &= [g(t), \pi(t)] \\ \boldsymbol{\varepsilon}'(t) &= [\varepsilon(t+1) - \varepsilon(t), \varepsilon(t+1) - \varepsilon(t) - \mu(t+1) + \mu(t)] \end{aligned} \right\} \quad (3.4)$$

The coefficient matrices of the state equations for this version of the model become

$$F = \begin{bmatrix} 1 + 1/a & 0 \\ (1/a)(1 + 1/a) & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 1 & 0 \\ 1/a & 1 \end{bmatrix},$$

while those of the output equation, (3.2), become

$$H = \begin{bmatrix} 0 & 0 \\ 0 & a \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 & \alpha_1 & 0 & \alpha_0 \\ 0 & 0 & \frac{(1+a)\alpha_1}{1+a-a\alpha_1} & \frac{-a\alpha^2}{1+a-a\alpha_1} & \alpha_0 \end{bmatrix}.$$

The first state equation is just (2.12). The second represents the evolution of $c_{t+1} - c_t$, which can be obtained by noting that

$$\begin{aligned} c_{t+2} - c_{t+1} &= (1/a)c_{t+1} + \eta(t+1) \\ &= (1/a)(1 + 1/a)c_t + (1/a)\eta(t) + \eta(t+1). \end{aligned}$$

The two equations in (3.2) are derived from (2.20) and (2.34), respectively.

As described in the next section, it was found necessary to respecify a second-order money growth rate process in place of (2.20). Following Flood and Garber (1980), $g(t)$ was assumed to obey a first-order autoregression in its second differences, i.e.,

$$[g(t) - g(t-1)] = \alpha_0 + \alpha_1 [g(t-1) - g(t-2)] + \omega(t),$$

or

$$g(t) = \alpha_0 + (1 + \alpha_1)g(t-1) - \alpha_1 g(t-2) + \omega(t).$$

Here $\omega(t)$, as usual, is assumed to be a zero mean white noise process. With $g(t)$ obeying the above, a third state space model representation is obtained by employing the following definitions:

$$\left. \begin{aligned} x'(t) &= [c_t, c_{t+1} - c_t] \\ u'(t) &= [g(t), \pi(t), g(t-1), g(t-2), g(t-3), 1] \\ \eta'(t) &= [\eta(t), \eta(t-1)] \\ y'(t) &= [g(t), \pi(t)] \\ \varepsilon'(t) &= [\omega(t), \omega(t) - \mu(t+1) + \mu(t)] \end{aligned} \right\} \quad (3.5)$$

The coefficient matrices of the state space representation for this version of the model become

$$F = \begin{bmatrix} 1 + 1/a & 0 \\ (1/a)(1 + 1/a) & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 1 & 0 \\ 1/a & 1 \end{bmatrix},$$

while those of the output equation, (3.2), become

$$H = \begin{bmatrix} 0 & 0 \\ 0 & a \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 & 1 + \alpha_1 & -\alpha_1 & 0 & \alpha_0 \\ 0 & 0 & 1 + \alpha_1 + aA & -\alpha_1 - aB & aC & \alpha_0 \end{bmatrix},$$

where

$$A = \left[1 + \frac{\alpha_1}{1 - \alpha_1} - \frac{\alpha_1^3}{1 - \alpha_1} \cdot \frac{1}{1 + a - a\alpha_1} \right],$$

$$B = \left[1 + \frac{2\alpha_1}{1 - \alpha_1} - \frac{2\alpha_1^3}{1 - \alpha_1} \cdot \frac{1}{1 + a - a\alpha_1} \right],$$

$$C = \left[\frac{\alpha_1}{1 - \alpha_1} - \frac{\alpha_1^3}{1 - \alpha_1} \cdot \frac{1}{1 + a - a\alpha_1} \right].$$

The state equations appear exactly as in the model associated with (3.4). The only differences appear in the output equations.

Conversion of the model into (3.1)–(3.2) is desirable primarily to permit the use of Kalman filtering. Estimates of the elements of $x(t)$ can be directly obtained, and these estimates are minimum-mean-square error. Furthermore, the Kalman filter can be employed to produce innovations sequences (residuals) which may be used to iteratively compute estimates of the model

parameters along with estimates of $x(t)$. Moreover, both state and parameters estimation can be effected in the absence of stability and stationarity assumptions. This makes the state space representation and its attendant estimation via Kalman filtering particularly well suited to address the issues raised in section 1, and is the prime reason for representing the model in the form (3.1)–(3.2). In estimation algorithm is briefly described below.

3.2. State estimation

For expositional convenience first consider the problem of estimating $x(t)$ given the parameters of F , G , Γ , H , and D , together with the first two moments of $y(t)$ and $u(t)$. If $\hat{x}(t, \tau)$ denotes the minimum-mean-square error estimate of $x(t)$ given the model and all observed data up through time τ ,

$$Y^\tau = \{y(1), y(2), \dots, y(\tau)\},$$

$$U^\tau = \{u(1), u(2), \dots, u(\tau)\},$$

then $\hat{x}(t, t)$ is produced by the following recursive computation:

$$\hat{x}(t+1, t) = F\hat{x}(t, t) + Gu(t), \quad (3.6)$$

$$P(t+1, t) = FP(y, t)F' + \Gamma Q \Gamma', \quad (3.7)$$

$$B(t+1, t) = HP(t+1, t)H' + R, \quad (3.8)$$

$$\hat{e}(t+1, t) = y(t+1) - H\hat{x}(t+1, t) - Du(t+1), \quad (3.9)$$

$$K(t+1) = P(t+1, t)H' B^{-1}(t+1, t), \quad (3.10)$$

$$\hat{x}(t+1, t+1) = \hat{x}(t+1, t) + K(t+1)\hat{e}(t+1, t), \quad (3.11)$$

$$P(t+1, t+1) = [I - K(t+1)H]P(t+1, t), \quad (3.12)$$

for $t_0 \leq t \leq T$. $P(t+1, t)$ is the variance-covariance matrix of the estimation matrix error in $\hat{x}(t+1, t)$, i.e.,

$$P(t+1, t) = E\{[x(t+1) - \hat{x}(t+1, t)][x(t+1) - \hat{x}(t+1, t)]'\}.$$

$B(t+1, t)$ is the variance-covariance matrix of the innovation, i.e.,

$$B(t+1, t) = E\{\hat{e}(t+1, t)\hat{e}(t+1, t)'\}.$$

The initial values for $\hat{x}(t, t)$ and $P(t, t)$ are assumed known and given by

$$\hat{x}(t_0, t_0) = \hat{x}(0) = E\{x(t_0) \text{ given all information at time } t_0\},$$

$$P(t_0, t_0) = P(0) = E\{[x(t_0) - \hat{x}(0)][x(t_0) - \hat{x}(0)]'\}.$$

Thus $P(t, \tau)$ is the variance-covariance matrix of the error in estimating $x(t)$ given all observations up through time ($\tau \leq t$). The vector $\hat{\epsilon}(t+1, t)$ represents the innovations process and is analogous to the model residuals used in econometric estimation. Eqs. (3.6)–(3.12) constitute the Kalman filter.

More efficient estimates of the states can be obtained by utilizing all the sample information available; i.e., $\hat{x}(t, T)$. This is referred to as the *smoothed* estimate. It is derived from the filtered estimate, $\hat{x}(t, t)$, by means of a reverse 'sweep' over the data from T back to $t+1$. Broadly speaking, computation is as follows: the recursive Kalman filter is employed in reverse time 'beginning' at time T using a diffuse 'prior' for $\hat{x}(T, T+1)$, i.e., $P(T, T+1) = \infty$. For any time t in the closed interval $[0, T]$, this reverse time filter produces an estimate, $\hat{x}(t, t+1)$, along with its corresponding variance-covariance matrix, $P(t, t+1)$. This represents our best estimate of $x(t)$ using data only over the interval $[t+1, T]$. Combining this with our forward time estimate, $\hat{x}(t, t)$, using only data over the interval $[0, t]$, gives us the desired result, $\hat{x}(t, T)$. The method of combination follows from a classical result in probability and statistics; namely, the optimal combination of two independent estimates $\hat{x}(t, t)$ [with precision matrix $P^{-1}(t, t)$] and $\hat{x}(t, t+1)$ [with precision matrix $P^{-1}(t, t+1)$] is

$$\hat{x}(t, T) = P(t, T)[P^{-1}(t, t)\hat{x}(t, t) + P^{-1}(t, t+1)\hat{x}(t, t+1)],$$

with corresponding precision matrix

$$P^{-1}(t, T) = P^{-1}(t, t) + P^{-1}(t, t+1).$$

Details of the smoothing algorithm employed below are given in Cooley, Rosenberg and Wall (1977). Thus, once filtered estimates are obtained, they can be revised by the smoothing algorithm to produce the most efficient estimates of $x(t)$.⁸

3.3. Parameter estimation

Using the Kalman filter to generate model residuals enables the formation of a loss function that can be used in parameter estimation. The parameters

⁸It can be shown that $P(t, t) \geq P(t, T)$ for $t_0 \leq t \leq T$. This should be intuitively clear since, by definition, $\hat{x}(t, T)$ uses more information than $\hat{x}(t, t)$. See Jazwinski (1970, ch. 7) or Bryson and Ho (1969, ch. 13).

to be estimated may not only include the unknown elements of H , D , and R (the parameters of the behavioral equations), but more importantly those of F , G , Γ , and Q . The algorithms for estimation of the unknowns in this manner are called ‘prediction error’ methods and, like the Kalman filter, are thoroughly treated in the control literature [see Caines (1976), Ljung (1979), and Ljung and Caines (1979)]. The algorithm employed in the present study is outlined by the following steps:

Step 1. Collect the unknown parameters into a vector θ of dimension $N \times 1$. Denote an initial guess at its true value by θ^0 and insert this into the Kalman filter eqs. (3.6)–(3.12). Set $i=0$.

Step 2. Using the Kalman filter eqs. (3.6)–(3.12), compute the model innovations sequence $\{\hat{\varepsilon}(t+1, t); t_0 \leq t \leq T-1\}$ where $\hat{\varepsilon}(t+1, t) = \hat{\varepsilon}(t+1, t, \theta^i)$ is an implicit function of θ^i .

Step 3. Form the loss function $J(\theta^i)$ where

$$J(\theta^i) = \frac{1}{2} \sum_{t_0}^{T-1} [\hat{\varepsilon}(t+1, t)' A_{t+1|t}^{-1} \hat{\varepsilon}(t+1, t) + \ln(\det A_{t+1|t})], \tag{3.13}$$

and $A_{t+1|t}$ is some positive definite weighting matrix.

Step 4. Compute an improved estimate of θ , denoted θ^{i+1} , such that $J(\theta^{i+1}) \leq J(\theta^i)$. Use

$$\theta^{i+1} = \theta^i - \rho^i M^{-1} \partial J(\theta^i) / \partial \theta,$$

where ρ^i is a (scalar) step size parameter and M_i^{-1} is a positive definite $N \times N$ matrix such that in the limit (as $i \rightarrow \infty$) it tends to the inverse Hessian of J . (See discussion after Step 5.)

Step 5. Check to see if $\|\theta^{i+1} - \theta^i\| < \delta_1$ and/or $\|\partial J(\theta^{i+1}) / \partial \theta\| \leq \delta_2$. If so, stop; θ^{i+1} is accepted as the ‘best’ estimate of θ . Otherwise, set θ^i to θ^{i+1} , $i=i+1$, and return to Step 2. If it is assumed that $\varepsilon(t)$ is normally distributed for each t , and $A_{t+1|t}$ is set equal to $B(t+1, t)$, then approximate maximum likelihood estimates are obtained; see Anderson and Moore (1979).

The iterative algorithm given above requires an initial estimate, θ^0 ; a convergence criterion, δ_1 and/or δ_2 ; and expressions for the components of the gradient vector $\partial J(\theta) / \partial \theta$. The gradients may be computed numerically using simple finite first differences of $J(\theta)$ or analytically using a straightforward application of differential calculus to (3.13). The method by

which ρ^i and M_i computed depends on the particular function minimization algorithm employed. A Davidson–Fletcher–Powell Variable Metric algorithm is used here since then M_i^{-1} is computed automatically, with $\partial J(\theta)/\partial \theta$ being the only user supplied information. Upon convergence M_i^{-1} is the inverse Hessian of $J(\theta)$ (i.e., the information matrix) and yields valuable information concerning estimated parameter standard errors, correlations (covariances), and indentifiability. In particular, once the algorithm converges, a simple scaling of M_i^{-1} produces an estimate of the parameter variance–covariance matrix. Not only is this estimate useful in hypothesis tests on elements of θ , but also in examining identifiability. Since local identifiability and asymptotic non-singularity of the Hessian matrix are equivalent, a nearly singular M_i^{-1} indicates identification problems. In practice, this is most easily tested by converting the parameter variance–covariance matrix to a correlation matrix and examining the off-diagonal elements. Interparameter correlations near unity, say ± 0.996 , lead to a singular condition suggesting an over-parameterized specification and lack of complete identification. Moreover, a singular Hessian for $J(\theta)$ results in non-convergence of the numerical optimization algorithm so that lack of convergence and lack of identification are highly related. These facts are extremely helpful in deciphering the estimation results given in section 4.

4. Estimation details and results

The state space model associated with the definitions in (3.3), (3.4), and (3.5) each contain six unknown parameters: a , α_0 , α_1 , σ_η^2 , and the two distinct elements of the symmetric 2×2 matrix R , σ_ε^2 and σ_μ^2 . In order to guarantee that the estimates of R always be positive definite during estimation, the lower triangular Cholesky factorization was used to define the unknowns related to R ; namely, R was represented as

$$R = \begin{bmatrix} \sigma_\varepsilon & 0 \\ \sigma_\varepsilon & -\sigma_\mu \end{bmatrix} \begin{bmatrix} \sigma_\varepsilon & \sigma_\varepsilon \\ 0 & -\sigma_\mu \end{bmatrix}.$$

A likewise representation was employed for the Q matrix, i.e.,

$$Q = [\sigma_\eta][\sigma_\eta]$$

for the levels model, while

$$Q = \begin{bmatrix} \sigma_\eta & 0 \\ \sigma_\eta/a & \sigma_\eta \end{bmatrix} \begin{bmatrix} \sigma_\eta & \sigma_\eta/a \\ 0 & \sigma_\eta \end{bmatrix}$$

for the differenced models.

The initial state estimate, $\hat{x}(0)$, and its error variance-covariance matrix, $P(0)$, constitute another five unknowns. These were not estimated but fixed at a specified value since, asymptotically they do not affect the results.⁹ Thus the θ vector is defined as

$$\theta' = [a, b, \alpha_0, \alpha_1, \sigma_\eta, \sigma_\varepsilon, \sigma_\mu].$$

Upon convergence of the iterative algorithm for θ , the Kalman filter is called to produce 'filtered' estimates $\hat{x}(t, t)$ for $x(t)$, i.e.,

$$\hat{x}_1(t, t) = E\{c_t | Y^t, U^t\},$$

for the levels models, and

$$\hat{x}_1(t, t) = E\{c_t | Y^t, U^t\},$$

$$\hat{x}_2(t, t) = E\{c_{t+1} - c_t | Y^t, U^t\},$$

for the differences models. Finally, given θ and $\hat{x}(t, t)$, the smoothing algorithm is called to produce $\hat{x}(t, T)$ which constitute our best estimates of the state variables given all the sample information. These are to be interpreted as our conditional expectation (conditioned on the observed data sample) of the variables employed by a representative economic agent in forming the rational expectations consistent with the given model.

4.1. Estimation results

Estimation of (3.1)–(3.2) was carried out using price data taken from *Zahlen zur Geldentwertung in Deutschland 1914 bis 1923*, Reimar Hobbing, Berlin, 1925. Money supply figures were taken from table VII of Flood and Garber (1980). The same period for both price and money covered January 1919 through June 1923. Six different bubble phenomena were considered, all lasting until June 1923 but beginning in different months. Beginning dates for the bubbles were: January 1919, October 1919, July 1920, June 1921, June 1922, and January 1923. The third, fifth, and sixth bubbles correspond exactly with those treated by Flood and Garber (1980).

Initial attempts to fit the levels model, (3.3), failed. No convergence in the iterative procedure was obtained due to severe negative correlation between

⁹Given a completely observable model, it can be shown that the filter is uniformly asymptotically stable with asymptotic solutions independent of the initial data; see Jazwinski (1970), Bucy and Joseph (1968), for example.

α_0 and α_1 (estimated at -0.994 after approximately sixty iterations). This problem was also evident in the serial correlation properties of the residuals in the money supply equation. For these reasons the levels model was abandoned in favor of the differenced version (3.4).

Two versions of (3.4) were estimated. In the first α_1 was estimated along with the other model parameters, while in the second it was constrained to unity (reflecting the belief that the money growth rate was best modeled as a perfect second difference). In both cases the residuals for the money growth rate failed to satisfy tests for no serial correlation, significant serial correlation being present at lag one. Hence, this model was dismissed from further consideration and a slightly more elaborate money supply process specified.

Estimation of the more elaborate model, (3.5), produced none of the problems associated with the previous cases. Convergence of the iterative estimation scheme was achieved in each instance and all estimated residual series met the serial correlation tests. Results for this case are presented in tables 1, 2, and 3^{10,11} and in fig. 1. The first table presents the parameters estimates. In each of the six bubbles a significant non-zero σ_η is obtained. Moreover, a slight negative correlation between α and σ_η is revealed. The second table presents the smoothed estimates for the c_t trajectories and their associated estimation error variances. Significant c_t 's appear in the last two months of the sample for all bubbles except those bubbles starting in to June 1922 and January 1923.¹² This behavior is due to a combination of the

¹⁰All standard errors reported in table 1 require a special interpretation similar to that given in Flood and Garber (1980, fn. 18) because the state space representations associated with (3.3) and (3.4) are unstable (i.e., each has an F matrix with one eigenvalue equal to $1 + 1/a$ which lies outside the unit circle), and result in a situation analogous to the 'exploding regressor' case in econometrics. Thus, we must view our data sample as one drawing from a cross-section of repeated hyperinflations, all with the same pre-1920 events and behavioral parameters. In this sense the length of the data sample is fixed at 40 observations (March 1920 through June 1923), while the number of repetitions, N , of this sample tends to infinity. The estimates obtained here then are both asymptotically consistent and normal [see Goodrich and Caines (1979)].

¹¹The degrees of freedom for the $g(t)$ equation Q statistic are obtained by interpreting this equation as a standard ARMA (2,0) model for $g(t)$. For the $\pi(t)$ equation the interpretation is not so clear since a state variable appears on the right-hand side; we no longer have an equivalent ARMA (p, q) model. What has been done is to assume the estimated residuals themselves comprise a given time series and that an ARMA (0,0) model has been fit. A Q -statistic test for 'model' adequacy then translates into an approximate test for 'whiteness'. Alternatively, both residuals could be subjected to a test of significance using as the 95% confidence band $\pm 2/\sqrt{51} \approx \pm 0.28$ [see Box and Jenkins (1973, pp. 177, 178, 290)].

¹²The confidence regions are computed from the diagonal elements of the $P(t, t)$ matrix in the filtered case, and $P(t, T)$ matrix in the smoothed case. Each of these matrices implicitly depends on the θ vector through the $F, H, Q,$ and R matrices [see (3.6)–(3.12)]. Thus in computing these variances we are treating the parameters as known exactly, i.e., that the representative economic agent takes these parameter estimates as given. Incorporation of the uncertainty in θ in generating both $P(t, t)$ and $P(t, T)$ requires solution of a nonlinear filtering problem for which only approximate solutions exist [see Jazwinski (1970, chs. 7 and 8)].

Table 1
 Estimation results for stochastic case.^a

Date of 'bubble'	a	α_0	α_1	σ_η	σ_ϵ	σ_μ	Q -statistic ^b	L.M. test ($1 \leq l \leq 4$) ^c
2/19-6/23	2.1918 (± 0.4308)	0.008012 (± 0.00937)	-0.40958 (± 0.1261)	0.07911 (± 0.02959)	0.06806 (± 0.00670)	0.1318 (± 0.0817)	9.275(22) 14.31 (24)	0.3051 1.0997 1.2432 1.9519
10/19-6/23	2.0217 (± 0.4108)	0.007951 (± 0.009345)	-0.39318 (± 0.12619)	0.089965 (± 0.02131)	0.067944 (± 0.00668)	-0.13311 (± 0.03237)	9.143(22) 15.238(24)	0.2860 1.0602 1.1704 1.9290
7/20-6/23	1.9881 (± 0.4425)	0.007893 (± 0.009323)	-0.37768 (± 0.12802)	0.09447 (± 0.02266)	0.067824 (± 0.00665)	0.15029 (± 0.02526)	9.805(22) 13.118(24)	0.2647 1.0151 1.0931 1.9042
6/21-6/23	1.8635 (± 0.4672)	0.007821 (± 0.009316)	-0.35826 (± 0.13276)	0.13687 (± 0.032509)	0.067694 (± 0.006721)	0.11911 (± 0.01575)	6.228(22) 14.05 (24)	0.2387 0.9580 1.0037 1.8746
6/22-6/23	1.5286 (± 0.4912)	0.007654 (± 0.00925)	-0.3136 (± 0.13471)	0.2206 (± 0.06925)	0.06748 (± 0.00637)	0.1252 (± 0.0138)	9.275(22) 17.755(24)	0.1821 0.8257 0.8322 1.8110
1/23-6/23	1.3726 (± 0.53879)	0.0075525 (± 0.009269)	-0.2853 (± 0.13941)	0.3304 (± 0.1353)	0.067402 (± 0.006553)	0.15587 (± 0.01594)	9.540(22) 7.95 (24)	0.1491 0.7427 0.7487 1.7726

^aAsymptotic estimates of standard errors are in parentheses.

^bBox-Pierce Q -statistic with degrees of freedom in parentheses [see Box and Jenkins (1973, p. 291)]. The first line gives the Q value for the $g(t)$ equation while the second line indicates the Q value associated with the $\pi(t)$ equation.

^cLagrange Multiplier test applied to the $g(t)$ equation residuals where l gives the assumed degree of autoregression or moving averaging [see Godfrey (1978)].

Table 2
Smoothed estimates for c_t and its error variance (stochastic case).

Date	\hat{c}_t	var(\hat{c}_t)	\hat{c}_t	var(\hat{c}_t)
	-0.96978E-02	0.67962E-02		
	-0.14638E-01	0.60995E-02		
	-0.35449E-02	0.58531E-02		
	0.62814E-02	0.57661E-02		
	0.12753E-01	0.57353E-02		
	0.25363E-01	0.57244E-02		
	0.42444E-01	0.57205E-02		
1919	0.65732E-01	0.57192E-02		
	0.74394E-01	0.57187E-02		
	0.65182E-01	0.57185E-02		
	0.60193E-01	0.57185E-02	0.18349E-01	0.52747E-02
	0.62803E-01	0.57184E-02	0.28148E-01	0.51242E-02
	0.71523E-01	0.57184E-02	0.38369E-01	0.50705E-02
	0.51461E-01	0.57184E-02	0.36600E-01	0.50512E-02
	0.13465E-01	0.57184E-02	0.84662E-01	0.50443E-02
	-0.14556E-01	0.57184E-02	-0.12408E-01	0.50419E-02
	-0.20715E-01	0.57184E-02	-0.16966E-01	0.50410E-02
	-0.20090E-01	0.57184E-02	-0.16530E-01	0.50407E-02
1920	-0.13201E-01	0.57184E-02	-0.11154E-01	0.50406E-02
	-0.85290E-02	0.57184E-02	-0.72770E-02	0.50405E-02
	-0.77643E-02	0.57184E-02	-0.65281E-02	0.50405E-02
	-0.10285E-01	0.57184E-02	-0.84264E-02	0.50405E-02
	-0.77554E-02	0.57184E-02	-0.64241E-02	0.50405E-02
	-0.68827E-02	0.57184E-02	-0.59086E-02	0.50405E-02
	-0.49274E-02	0.57184E-02	-0.44136E-02	0.50405E-02
	-0.73732E-02	0.57184E-02	-0.60923E-02	0.50405E-02
	-0.59325E-02	0.57184E-02	-0.48097E-02	0.50405E-02
	-0.25324E-02	0.57184E-02	-0.19065E-02	0.50405E-02
	0.38328E-02	0.57184E-02	0.31695E-02	0.50405E-02
	0.13562E-01	0.57184E-02	0.10986E-01	0.50405E-02
1921	0.25212E-02	0.57184E-02	0.20197E-01	0.50405E-02
	0.42244E-01	0.57184E-02	0.33643E-01	0.50405E-02
	0.38616E-01	0.57184E-02	0.30889E-01	0.50405E-02
	0.39410E-01	0.57184E-02	0.31698E-01	0.50405E-02
	0.39058E-01	0.57184E-02	0.31950E-01	0.50405E-02
	0.22489E-01	0.57184E-02	0.18688E-01	0.50405E-02
	0.25583E-01	0.57184E-02	0.20876E-01	0.50405E-02
	0.40303E-02	0.57184E-02	0.33426E-01	0.50405E-02
	0.55538E-01	0.57184E-02	0.43451E-01	0.50405E-02
	0.40235E-01	0.57184E-02	0.32238E-01	0.50405E-02
	0.29562E-01	0.57185E-02	0.24695E-01	0.50406E-02
	0.48818E-01	0.57185E-02	0.36697E-01	0.50406E-02
	0.68459E-01	0.57187E-02	0.55665E-01	0.50408E-02
1922	0.79930E-01	0.57191E-02	0.65666E-01	0.50414E-02
	0.56356E-01	0.57202E-02	0.48586E-01	0.50429E-02
	0.40878E-01	0.57235E-02	0.38051E-01	0.50473E-02
	0.13801E-01	0.57328E-02	0.18189E-01	0.50595E-02
	-0.30168E-01	0.57590E-02	-0.14239E-01	0.50937E-02
	-0.17545E-01	0.58332E-02	-0.24994E-02	0.51892E-02
	0.19843E-01	0.60432E-02	0.29375E-01	0.54563E-02
	0.51674E-02	0.66370E-02	0.25994E-01	0.62034E-02
1923	0.11964E+00	0.83165E-02	0.12541E+00	0.82927E-01
	0.28851E+00	0.13065E-01	0.27570E+00	0.14136E-01
	0.46800E+00	0.26503E-01	0.45073E+00	0.30477E-01

Table 2 (continued)

Date	\hat{c}_t	var(\hat{c}_t)	\hat{c}_t	var(\hat{c}_t)
1920	-0.57448E-03	0.61546E-02		
	-0.89755E-03	0.57464E-02		
	-0.26551E-02	0.56006E-02		
	-0.59627E-02	0.55485E-02		
	-0.49082E-02	0.55299E-02		
	-0.49542E-02	0.55233E-02		
	-0.38566E-02	0.55209E-02		
	-0.56433E-02	0.55201E-02		
1921	-0.44986E-02	0.55198E-02		
	-0.17578E-02	0.55197E-02		
	0.31062E-02	0.55196E-02		
	0.10604E-02	0.55196E-02	0.15446E-01	0.11119E-01
	0.19477E-01	0.55196E-02	0.25618E-01	0.10364E-01
	0.32412E-01	0.55196E-02	0.42682E-01	0.10126E-01
	0.29773E-01	0.55196E-02	0.37539E-01	0.10051E-01
	0.30542E-01	0.55196E-02	0.38207E-01	0.10027E-01
	0.30825E-01	0.55196E-02	0.38662E-01	0.10020E-01
	0.18055E-01	0.55196E-02	0.19781E-01	0.10017E-01
	0.20125E-01	0.55196E-02	0.22281E-01	0.10017E-01
	0.32150E-01	0.55196E-02	0.38858E-01	0.10016E-01
	0.41736E-01	0.55196E-02	0.52068E-01	0.10016E-01
	0.30978E-02	0.55196E-02	0.36295E-01	0.10016E-01
	0.23840E-01	0.55197E-02	0.25795E-01	0.10016E-01
	0.35442E-01	0.55197E-02	0.41825E-01	0.10016E-01
1922	0.53742E-01	0.55200E-02	0.67814E-01	0.10016E-01
	0.63540E-02	0.55206E-02	0.82394E-01	0.10017E-01
	0.47308E-01	0.55224E-02	0.59642E-01	0.10018E-01
	0.37418E-01	0.55274E-02	0.46840E-01	0.10020E-01
	0.18588E-01	0.55414E-02	0.19865E-01	0.10029E-01
	-0.12353E-03	0.55805E-02	-0.27379E-01	0.10056E-01
	-0.63553E-01	0.56900E-02	-0.16650E-01	0.10143E-01
	0.30585E-01	0.59968E-02	0.16144E-01	0.10417E-01
	0.28463E-01	0.68557E-02	-0.71530E-02	0.11285E-01
	0.12628E+00	0.92610E-02	0.10344E+00	0.14032E-01
1923	0.27423E+00	0.15996E-01	0.27320E+00	0.22729E-01
	0.44914E+00	0.34855E-01	0.46822E+00	0.50258E-01
	0.49062E-01	0.25527E-01		
	0.81298E-01	0.21350E-01		
	0.10192E+00	0.20286E-01		
	0.72849E-01	0.20015E-01		
	0.61082E-01	0.19947E-01		
	0.27987E-01	0.19933E-01		
1922	-0.36545E-01	0.19943E-01		
	-0.26214E-01	0.19998E-01	0.13397E-01	0.50322E-01
	0.55866E-02	0.20217E-01	0.23228E-01	0.41713E-01
	-0.47027E-01	0.21079E-01	-0.56428E-01	0.40855E-01
	0.69991E-01	0.24465E-01	0.58425E-01	0.45566E-01
	0.25782E+00	0.37753E-01	0.25225E+00	0.67808E-01
	0.48235E+00	0.89916E-01	0.49226E+00	0.16728E-00

Table 3
Sensitivity check on $P(t_0, t_0)$ ^a

$P(t_0, t_0)$	a	α_0	α_1	σ_η	σ_ε	σ_μ
$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	1.3591 (± 0.5367)	0.00756 (± 0.00929)	-0.2891 (± 0.1390)	0.2897 (± 0.1275)	0.06741 (± 0.00657)	0.15600 (± 0.01597)
$\begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$	1.3702 (± 0.5388)	0.00755 (± 0.00927)	-0.2865 (± 0.1394)	0.3162 (± 0.1353)	0.0674 (± 0.00655)	0.1559 (± 0.01594)
$\begin{bmatrix} 1.0 & 0 \\ 0 & 1.0 \end{bmatrix}$	1.3740 (± 0.5399)	0.00755 (± 0.00927)	-0.2855 (± 0.1396)	0.3284 (± 0.1422)	0.0674 (± 0.00656)	0.1558 (± 0.0159)
$\begin{bmatrix} 10.0 & 0 \\ 0 & 10.0 \end{bmatrix}$	1.3746 (± 0.5408)	0.00755 (± 0.00925)	-0.2853 (± 0.1394)	0.3304 (± 0.1431)	0.06740 (± 0.00653)	0.1558 (± 0.0159)

^aAsymptotic estimates of standard errors are in parentheses.

stochastic specification for the evolution of c_t [i.e., (2.33)] and the increasing value of the σ_η estimates.¹³

Table 3 presents the results of a sensitivity check. Referring back to the description of the estimation algorithm, it is clear that the values of all state variables estimates and variances [see (3.6)–(3.12)] depend to an extent on the assumed prior. In turn, from the manner in which the loss function (3.13) is constructed, the influence of the assumed prior may carry through all the parameter estimates. If the sensitivity to the prior is high, then all our estimates must be viewed as highly conditional. Furthermore, this potential problem would be expected to be great for short duration bubbles, i.e., when $T-t_0$ is small, since then the solution to the difference equations for $\hat{x}(t, t)$ and $P(t, t)$ will be most influenced by their initial conditions. Therefore, a 'worse case' situation was studied: estimation of the January 1923–June 1923 bubble. Four different prior distributions were considered, ranging from a certain prior centered on $\hat{x}(t_0, t_0) = \theta$ to one approximating a diffuse prior centered on the same state value. As can be seen from the table, the only perceptible variations in the estimated parameters are found in a and σ_η . The total percent change in a amount to about 1.1%, while that for σ_η is about 14%. At the same time, the range in $P(t_0, t_0)$ is over several orders of magnitude. Further increases in the diagonal of $P(t_0, t_0)$ produced negligible

¹³In the stochastic model for the c_t process, (2.12), it is possible to have a significant c_t for some $t \geq t_0$ while c_t is insignificant for $t = t_0$ because of the presence of $\eta(t)$. This is not the case, of course, in a deterministic model for c_t , as in (2.12) when the $\eta(t)$ term is removed; in this instance a significant c_t for some $t > t_0$ must be accompanied by a statistically significant estimate of c_t at $t = t_0$.

variations in the parameter estimates. Sensitivity to the assumed prior does not appear to be a problem.

Fig. 1 depicts the trajectories of c_t produced by application of the smoothing algorithm to the model with the parameters given in table 1. Table 2 presents the c_t estimates plotted in this figure, together with the corresponding estimated error variances.

The presence of significant c_t estimates in the stochastic case, prompted a further estimation of the model under a restriction corresponding to a deterministic evolution for the c_t 's. More specifically, the model was re-

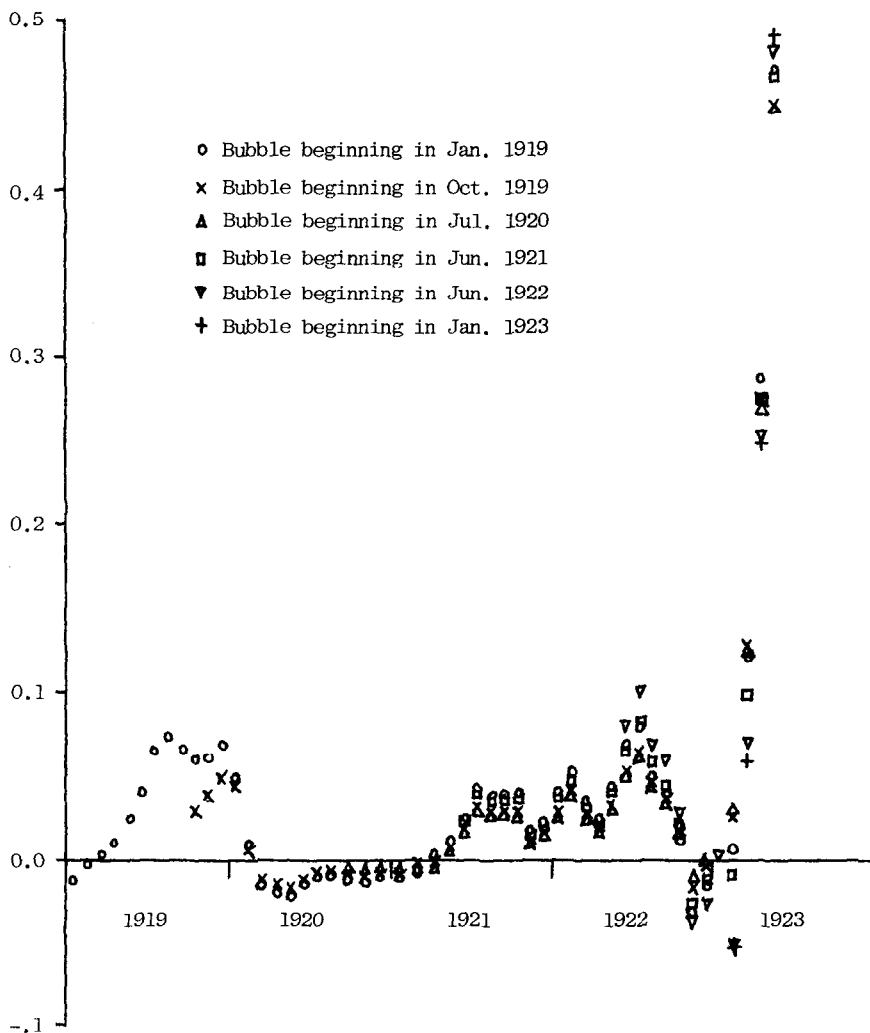


Fig. 1. Smoothed estimates for c_t (stochastic).

Table 4
 Estimation results for deterministic case.

'Bubble' dates	a	α_0	α_1	σ_ε	σ_μ	Q -statistic ^a	L.M. test ($1 \leq i \leq 4$)
1/19-6/23	2.3123 (± 0.4238)	0.007965 (± 0.009326)	-0.3970 (± 0.12411)	0.06798 (± 0.006666)	0.2325 (± 0.0262)	9.143(22) 16.298(24)	0.2912 1.0713 1.1903 1.9353
10/19-6/23	2.0794 (± 0.4157)	0.007919 (± 0.00942)	-0.3845 (± 0.1253)	0.06788 (± 0.00671)	0.2259 (± 0.0242)	9.805(22) 15.635(24)	0.2741 1.0352 1.1267 1.9151
7/20-6/23	1.8253 (± 0.3914)	0.007864 (± 0.009404)	-0.3699 (± 0.1286)	0.06778 (± 0.00681)	0.2189 (± 0.0237)	9.408(22) 14.575(24)	0.2543 0.9925 1.0566 1.8923
6/21-6/23	1.4710 (± 0.3568)	0.00778 (± 0.00967)	-0.3480 (± 0.1329)	0.06763 (± 0.00679)	0.2098 (± 0.0229)	9.275(22) 13.780(24)	0.2252 0.9276 0.9599 1.8594
6/22-6/23	0.9848 (± 0.2800)	0.00766 (± 0.00942)	-0.3153 (± 0.1361)	0.06749 (± 0.00669)	0.1984 (± 0.0208)	9.275(22) 18.550(24)	0.1841 0.8306 0.8377 1.8127
6/23-6/23	0.5988 (± 0.2570)	0.00756 (± 0.00928)	-0.2888 (± 0.1398)	0.06742 (± 0.00656)	0.1905 (± 0.0193)	9.540 (22) 6.095(24)	0.1530 0.7526 0.7577 1.7768

^aSee note b to table 1.

Table 5
Smoothed estimates for c_t and its error variance (deterministic case).

'Bubble' dates	c_0	$\text{var}(c_0)$	c_T	$\text{var}(c_T)$
1/19-6/23	0.204×10^{-8}	0.790×10^{-18}	0.3816	0.0277
10/19-6/23	0.123×10^{-7}	0.273×10^{-16}	0.3903	0.0278
7/20-6/23	0.916×10^{-7}	0.146×10^{-14}	0.4006	0.0279
6/21-6/23	0.164×10^{-5}	0.438×10^{-12}	0.4170	0.0284
6/22-6/23	0.991×10^{-4}	0.147×10^{-8}	0.4453	0.0297
1/23-6/23	0.349×10^{-2}	0.171×10^{-5}	0.4714	0.0312

estimated with σ_η constrained to zero. If large estimates for σ_η led to insignificance of the c_t 's, as was the case for the last two bubbles, then restriction of σ_η to zero should produce significant c_t in all instances.

Tables 4 and 5 and fig. 2 summarize the results of this re-estimation. As in the previous estimation using the second-order money growth rate model, (3.5), convergence of the iterative estimation procedure was achieved in each instance. In addition, all residual series passed their whiteness tests. Parameter estimates are given in table 4 and reveal an interesting variation in the estimates for a , α_1 , and σ_μ . Table 5 presents the smoothed estimates for the terminal and initial values of c_t 's and its estimation error variance.¹⁴ In each case the c_t 's are significant under the usual normal distribution assumption.

5. Conclusions

In summary, our empirical evidence strongly suggests that for this model it is impossible to maintain the common assumption that rational expectations are *always* convergent. This conclusion is most disturbing in large part because it means that the actual price at each t may be indeterminate without some additional assumption to determine the c_t 's. One solution to this difficulty having considerable intuitive appeal to many economists involves a slow price adjustment mechanism which might determine the initial condition by $p^*(0, 0) = p(0)$, where $p(0)$ is historically given. However, the issue of stochastic determinacy must be left as unresolved for now.

¹⁴Under the restriction of $\sigma_\eta = 0$, the smoothed estimates of c_t and $P_{11}(t, T)$ become mere backward extrapolations of $\hat{x}_1(T, T)$ and $P_{11}(T, T)$. Thus, the smoothed estimate of c_t ($t < T$) is obtained from $\hat{x}_1(t, T) = (1 + 1/a)^{t-T} \hat{x}_1(T, T)$ while the smoothed estimation error variance is obtained from $P_{11}(t, T) = (1 + 1/a)^{2(t-T)} P_{11}(T, T)$. See Anderson and Moore (1979, pp. 187-190) or Jazwinski (1970, pp. 215-218) for details.

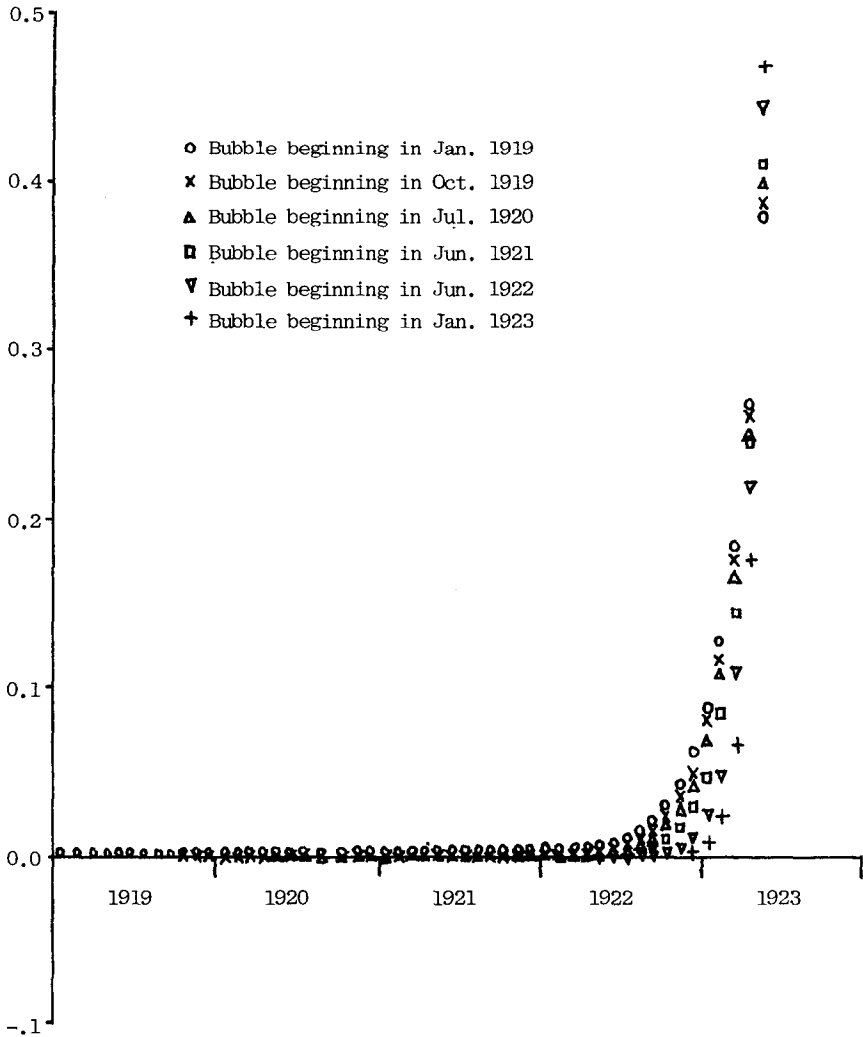


Fig. 2. Smoothed estimates for c , (deterministic).

It is important to recognize that the issues we have raised in this paper are not restricted to the simple monetary model of inflation studied here. The fact that identical conceptual issues arise in more complex rational expectations models of the Lucas-Sargent type is evident from the analysis of Burmeister (1980, 1982) and Burmeister, Flood and Turnovsky (1981).

Our preliminary work has left us with several unanswered questions. One should investigate whether or not the assumption of known and constant parameters is justified, and it is important to test alternative specifications for

the money supply function which include, for example, the variable $p^*(t+1, t)$. Likewise one might try alternative specifications for the stochastic c_t process.

Our primary conclusion is that estimation techniques which impose restrictions implied by the assumption of convergent expectations, and which therefore are conditional upon this convergent expectations assumption, are suspect without additional verification of the underlying stability hypothesis. We have demonstrated the feasibility of an alternative estimation methodology which does not preclude the possibility that rationally formed expectations are unstable. We have shown how to test an important hypothesis and to obtain parameter estimates which are not conditional upon a perhaps invalid stability assumption.

References

- Anderson, B.D.O. and J.B. Moore, 1979, *Optimal filtering* (Prentice-Hall, Englewood Cliffs, NJ).
- Black, Fischer, 1974, Uniqueness of the price level in monetary growth models with rational expectations, *Journal of Economic Theory*, Jan.
- Box, George E.P. and Gwilym M. Jenkins, 1973, *Time series analysis: Forecasting and control*, rev. ed. (Holden Day, San Francisco, CA).
- Bryson, A.E. and Y.-C. Ho, 1969, *Applied optimal control* (Blaisdell, Waltham, MA).
- Bucy, R.S. and P.D. Joseph, 1968, *Filtering for stochastic processes with applications* (Wiley, New York).
- Burmeister, Edwin, 1980, On some conceptual issues in rational expectations modelling, *Journal of Money, Credit and Banking*, Nov.
- Burmeister, Edwin, 1981, Indeterminacy and stability in both rational expectations and perfect foresight models, Paper presented at Summer Econometric Society Meetings, San Diego, June (revised 1982).
- Burmeister, Edwin, Robert P. Flood and Stephen J. Turnovsky, 1979, Rational expectations and stability in a stochastic monetary model of inflation, Working paper (University of Virginia, Charlottesville, VA).
- Burmeister, Edwin, Robert P. Flood and Peter M. Garber, 1983, On the equivalence of solutions in rational expectations mode, *Journal of Economic Dynamics and Control*, forthcoming.
- Caines, P.E., 1976, Prediction error methods for stationary stochastic processes, *IEEE Transactions on Automatic Control* AC-21, no. 4, Aug., 500-505.
- Cooley, Thomas F., Barr Rosenberg and Kent D. Wall, 1977, A note on optimal smoothing for time varying coefficient problems, *Annals of Eco. Soc. Meas.* 6, no. 4.
- Flood, Robert P. and Peter M. Garber, 1980, Market fundamentals vs. price level bubbles: The first tests, *Journal of Political Economy* 88, no. 4, Aug., 745-770.
- Godfrey, L.G., 1978, Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables, *Econometrica* 46, no. 6, Nov.
- Goodrich, R.L. and Peter E. Caines, 1979, Linear system identification from nonstationary cross-sectional data, *IEEE Transactions on Automatic Control* AC-24, no. 3, June, 403-410.
- Green, Jerry R. and Seppo Honkapohja, 1979, Variance minimizing monetary policies with lagged price adjustment and rational expectations, Discussion paper no. 721, Sept. (Harvard Institute of Economic Research, Cambridge, MA).
- Hansen, Lars P. and Thomas J. Sargent, 1980, Formulating and estimating dynamic linear rational expectations models, *Journal of Economic Dynamics and Control* 2, no. 1, 7-46.
- Jazwinski, A.H., 1970, *Stochastic processes and filtering theory* (Academic Press, New York).
- Ljung, Lennart, 1978, Convergence analysis of parametric identification methods, *IEEE Transactions on Automatic Control* AC-23, 770-783.

- Ljung, Lennart and Peter E. Caines, 1979, Asymptotic normality of prediction error estimators for approximate systems models, *Stochastics* 3, 29–46.
- McCallum, Bennett T., 1980, Rational expectations and macroeconomic stabilization policy: An overview, *Journal of Money, Credit and Banking*, Nov.
- Muth, J.F., 1967, Rational expectations and the theory of price movements, *Econometrica*, July.
- Samuelson, Paul A., 1957, Intertemporal price equilibrium: A prologue to the theory of speculation, *Weltwirtschaftliches Archiv* 79, no. 2. Reprinted in: Joseph E. Stiglitz, ed., 1966, *The collected scientific papers of Paul A. Samuelson*, Vol. 2 (M.I.T. Press, Cambridge, MA).
- Samuelson, Paul A., 1967, Indeterminacy of developments in a heterogeneous-capital model with constant saving propensity, in: Karl Shell, ed., *Essays on the theory optimal economic growth* (M.I.T. Press, Cambridge, MA). Reprinted in: Robert C. Merton, ed., 1972, *The collected scientific papers of Paul A. Samuelson*, Vol. 3 (M.I.T. Press, Cambridge, MA).
- Sargent, Thomas J. and Neil Wallace, 1973, Rational expectations and the dynamics of hyperinflation, *International Economic Review* 4, June.
- Wall, Kent D., 1980, Generalized expectations modeling in econometrics, *Journal of Economic Dynamics and Control* 2, no. 2, May.