ESSAYS IN MACROECONOMICS AND FINANCE

Essay I: Money Demand, Seigniorage and the Welfare Cost of Inflation: Evidence from an Intertemporal Model of Money and Consumption for the U.S. Economy

Essay II: Implementation of the Heath-Jarrow-Morton and the Black-Derman-Toy Interest Rate Models for Pricing Options on Eurodollar Futures

by

TURAN GOKCEN BALI

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1998

UMI Number: 9908293

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the Degree of Doctor of Philosophy.

July 24, 98
Date

Alvin L. Marty
Chair of Examining Committee

July 24, 1998
Date

Linda Edwards
Executive Officer

Alvin L. Marty

Thom B. Thurston

Salih N. Neftci

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

ESSAYS IN MACROECONOMICS AND FINANCE

by

Turan Gokcen Bali

Advisor: Professor Alvin L. Marty

Essay I: The first essay emphasizes that it is crucial to identify the proper specification of money demand as well as the appropriate definition of monetary aggregate and scale variable to find the exact welfare cost of inflation. The econometric test results obtained from the nonlinear form of money demand with Box-Cox restriction indicate that not the semi-logarithmic form but the double-log form with constant elasticity of less than one is a more accurate characterization of the actual data. Furthermore, the empirical results suggest that consumption spending produces more stable measures of monetary velocity and outperforms GDP in estimated money demand equations for the United States. This paper also shows the estimated welfare cost of inflation is proportional to the money stock and since M1 is about three times the monetary base in the United States, identifying M1 without modeling the distinctive roles of currency and deposits as the relevant definition of money overestimates the true welfare cost of inflation. I estimate the welfare cost of inflation for the U.S. economy using Bailey's (1956) consumer's surplus and Lucas's (1993) compensating variation approaches. The welfare cost estimates imply that for each monetary model (currency-deposit, single-asset) and each scale variable (income, consumption) the double-log function, compared to the constant semi-elasticity Cagan-

type demand for money, yields substantial welfare gains in moving from zero inflation to the Friedman optimal deflation rate needed to bring nominal interest rates to zero.

Essay II: The second essay explains arbitrage-free term structure models used for pricing options on Eurodollar Futures with particular emphasis on the Black-Derman-Toy (BDT) and the Heath-Jarrow-Morton (HJM) interest rate models and their applications. In this essay, I concentrate on valuing options on Eurodollar futures using the BDT and the HJM models *with different volatility structures*. I compare the estimated option prices of Eurodollar futures with the actual values to determine which of these two models [specified with different dynamics of interest rate volatility: historical volatility, exponentially smoothed volatility, implied volatility, GARCH, GARCH-X] perform better in pricing interest rate sensitive derivative securities.

ACKNOWLEDGEMENTS

I owe an enormous debt to my dissertation supervisor Prof. Alvin L. Marty who provided the intellectual inspiration and guidance at every stage of the work, and taught me how to ask vital economic questions and which questions are worth attacking. His concern with the issues of seigniorage and the welfare cost of inflation had a profound influence on my intellectual development in macro and monetary theory. I am grateful to Prof. Thom B. Thurston who was my teacher, has been my co-author, has been my friend and now my colleague at Queens College. He is the person who got me interested in macroeconomics, patiently listened to my initial findings and made editorial and substantive suggestions on numerous drafts of my thesis. He is really a critical resource in developing my ideas in monetary theory. I am also thankful to Prof. Salih N. Neftci for his comments and suggestions in the course of preparation of the finance section of this dissertation. I appreciate Prof. Linda Edwards for her helping me in administrative issues in my Ph.D. studies and especially for her crucial help to my wife. I would like to express my deepest gratitude to Prof. Michael Grossman [one of the best people I have ever seen in my life] who provided invaluable assistance, supportive and lively atmosphere at every stage of my Ph.D. studies at the Graduate Center. Finally, I would like to thank my wife Mehtap for her continuous support and encouragement. Without her emotional and intellectual support, this dissertation would not have been completed.

TABLE OF CONTENTS

ESSAY I:

ESSAY II:

LIST OF TABLES

LIST OF FIGURES

ESSAY I:

ESSAY II:

Money Demand, Seigniorage and the Welfare Cost of Inflation: Evidence from an Intertemporal Model of Money and Consumption for the U.S. Economy

I. INTRODUCTION

This paper proposes new formulas for evaluating seigniorage revenue and the welfare cost of inflation as a fraction of GDP in the U.S. economy by providing new estimates of U.S. money demand. In the theoretical literature, it is a widely accepted idea that seigniorage is a certain amount of revenue a government collects, with the aid of its central bank, from the issue of monetary base (reserves plus currency). With the base money stock denoted M_t and the price level P_t, seigniorage revenue S_t can be defined as

$$S_t = \frac{M_t}{P_t} = \mu_t m_t = m_t + \pi_t m_t \qquad \text{(I.1)}$$

where $\mu_t = \frac{M_t}{M_t}$ denotes the rate of growth of the monetary base and m_t real money balances.

The first expression in Equation (I.1) defines seigniorage as the change in the nominal money stock divided by the price level. The second expression defines total seigniorage as the product of the rate of nominal money growth and real money balances held by the public. By analogy with the public finance literature, μ_t is often referred to as the tax rate and m_t, which is equal to the demand for cash balances under the assumption of money market equilibrium, as the tax base. The third expression in Equation (I.1) expresses the value of the government revenue as the sum of the increase in the real stock of money and the product of the inflation rate and the real money balances, $\pi_t m_t$. The last expression represents the inflation tax, I_{tax}:

$$I_{tax} = \pi_t m_t \tag{I.2}$$

so that

$$S_t = I_{tax} + m_t \tag{I.1'}$$

which implies that in a stationary state (with $m_t = 0$), seigniorage is equal to the inflation tax.

The traditional definition of seigniorage implies that when the nominal money stock is not growing, no revenue accrues to the government. Alternatively, the Phelps (1973) - Auernheimer (1974) definition uses the nominal rate of interest (rather than the growth rate of the monetary base) times stock of real balances as the revenue accruing from the authorities. The rationale for this definition is as follows. Less revenue is spent by the authorities because governments have monopoly power to hold part of their debt to the private sector in the form of non-interest bearing money rather than as interest bearing bonds. As long as the nominal rate of interest is greater than zero, this monopoly power implies that on the government debt in the form of money, no interest is paid. This saving of interest payment is measured by the stock of non-interest bearing money times the nominal rate of interest.

Seigniorage can be viewed as a tax on private agents' domestic currency holdings since money creation causes inflation, thereby lowering the real value of nominal assets. In other words, inflation imposes a tax on money holdings because it is the rate at which individuals lose the purchasing power of a dollar. Therefore, individuals change their holdings and their use of money to lower the total cost of holding money when inflation rises. Their efforts to do so, however, reduce total services from real money balances, thereby lowering individuals' welfare. This loss is the welfare cost of inflation.

In practice, estimating welfare costs involves using a theoretical model of monetary economy to interpret available data on money, income, and interest rates. Following Bailey (1956) and Marty (1967,1976), the welfare loss to society is traditionally measured by integrating the area under the liquidity preference schedule from the stock of real money balances held at a zero nominal rate of interest to that held at the positive nominal rate of interest. As a alternative to Bailey's consumer's surplus argument, Lucas (1994), using the compensating variation approach, defines the welfare cost of a nominal rate of interest rate $i>0$ to be the percentage income compensation needed to leave the household indifferent between $i>0$ and $i=0$.

Bailey's (1956) original study says that the welfare cost of inflationary finance increase with the *square* of the interest rate, assigning large costs to very high rates of inflation but trivial costs to moderate inflation's. According to this approach, the objective of a zero inflation rate has negligible benefit, and the additional gain in moving from zero inflation to the steady deflation of around 3 percent that would attain the Friedman optimum is assigned an even smaller value.

In this paper we argue that the traditional quadratic approximation to the welfare cost of inflation may not be suitable for this question, and to propose a new estimates of the welfare cost of inflation. By providing the microfoundations of the money demand function which suggests a *square root* formula for welfare costs, we show that the new estimates assign much greater benefit to the reduction of moderate inflations. Deriving both the double-log money demand function and the square root formula in the context of a simplified version of Sidrauski's (1967) model, we emphasize the quantitative analysis of welfare cost instead of qualitative issues in the welfare analysis of inflation.

We test the sensitivity of estimates of the revenue maximizing inflation rate, the maximum seigniorage (as a ratio to income), and the marginal welfare cost of seigniorage to the specification of money demand. The seigniorage-maximizing inflation rate and the maximum revenue from money creation are overestimated with the semi-logarithmic form compared to the nonlinear form of money demand. For each monetary aggregate (monetary base, M1) and each scale variable (income, consumption), the double-log function yields substantial welfare cost estimates compared to the semi-log function.

We show that in estimating the welfare cost of inflation the use of a unitary income (or consumption) elasticity assumption is not only improper but also leads to overestimate the true welfare cost of inflation since the unitary income and consumption elasticity hypotheses are strongly rejected and the elasticities are estimated well below one. We also find that defining M1 as the relevant definition of money overstates the actual welfare cost of inflation since the estimated welfare cost of inflation is proportional to the money stock and M1 is about three times the monetary base in the United States.

This paper is organized as follows. Section II summarizes previous discussions regarding the procedure of introducing money into the utility function. Section III first describes our macro model - the intertemporal optimizing framework in which the representative agent is assumed to have an infinite planning horizon, to face perfect capital markets, and to have perfect foresight - then presents the microfoundations for the double-log form of money demand. Section IV discusses why the specification of the money demand function has important implications for a number of macroeconomic issues.

Section V exposits the econometric analysis of the choice of a scale variable (consumption or income) in money demand. Section VI sets out the estimation techniques employed in U.S. money demand regression equations. Section VII provides four different econometric tests to determine whether a semi-log or a double-log money demand function is the proper specification for the U.S. economy. Section VIII presents the time series analysis of the variables employed in estimated money demand equations. Section IX studies the implications of different money demand specifications for seigniorage revenue and the welfare cost of inflation. Section X first develops a currency-deposit model that provides some insight into the implications of the analysis for the optimal rate of inflation and required reserve ratio. Section XI derives both the quadratic approximation and the square root formula for welfare costs in the context of a simplified version of Sidrauski's model. Section XII presents the estimates of income and consumption elasticities of demand for money and tests the unitary income and consumption elasticity hypotheses. This section also demonstrates the degree of overestimation of the welfare cost of inflation due to the use of M1 and/or the unitary income (or consumption) elasticity assumption. Section XIII measures the exact welfare cost of inflation for the U.S. economy in a currency-deposit model and estimates the costs of deviating from a zero inflation and from an optimal deflation policy. Section XIV concludes the paper.

II. MONEY IN THE UTILITY FUNCTION

Several approaches can be found in the literature for introducing the role of money into the intertemporal optimizing framework[1]. These are intended to capture the three key roles of money in the economy: (*i*) its role as a store of wealth; (*ii*) its role as the medium of exchange; and (*iii*) its role as a unit of account. Within the infinite horizon model two approaches have been adopted to incorporate the role of money. The first is to incorporate its role as a medium of exchange through the so-called cash-in-advance constraint, originally proposed by Clower (1967). The basic idea here is to formulate the role that money plays in carrying out transactions by the explicit introduction of a "transactions technology". Following McCallum (1987), a similar way of expressing the idea of introducing money into the utility function is to assume that, while households derive utility only from consumption and leisure, the acquisition of consumption goods requires "shopping" which reduces the time available for leisure or employment. In a monetary economy, however, the amount of shopping time required for a given amount of consumption depends negatively on the quantity of real money balances held by the household. If the required shopping time is functionally related to c_t and m_t, substitution into the basic utility function will yield an indirect utility function in which m_t appears.[2] The second, originally due to Sidrauski (1967), is to introduce money directly into the utility function. By facilitating transactions, money is assumed to yield a direct utility to the representative agent that is not associated with other assets such as bonds, which yield only an indirect utility through the income they generate and the consumption goods they enable the agent to purchase. Therefore, we should think about the money stock as "making life easier" since it allows people to get consumption goods without having to go to the bank and transform bonds into consumption goods all the time.

The introduction of money into the utility function has often been the subject of severe criticism by monetary economists, who have argued that one should model the process of transactions explicitly. In the words of Turnovsky (1995), "however, this criticism appears to have been muted because of an important paper by Feenstra (1986), who studied the relationship between the two approaches. He showed that under certain regulatory conditions, the maximization problem with money, modeled by means of a cash-in-advance constraint, may be equivalent to a maximization problem with money in the utility function. Thus the procedure of introducing money directly into the utility function seems to be generally viewed as being an acceptable approximation". This is the approach we will use in the following section.

III. SETUP OF THE MODEL

Following Sidrauski (1967), we will assume that the economy is populated by infinitely lived consumers or dynasties who derive utility from the only consumption good and from real money stock. The households in the economy provide labor services in exchange for wages, receive interest income on assets, purchase goods for consumption, and save by accumulating additional monetary and non-monetary assets. We will further assume that the utility function is time separable, that is

$$U = \int_0^\infty e^{-\rho t} N(t)\, u[c(t), m(t)]\, dt, \qquad \text{(III.1)}$$

where ρ is the subjective rate of time preference (or the personal discount rate), $N(t)$ is the total amount of people alive at time t and $N(t)$ will be assumed to grow at an exogenous rate n. By normalizing initial population to 1, we have $N(t) = e^{nt}$. It is also assumed that the rate of time preference is greater than the growth rate of population $(\rho > n)$, which implies that the overall utility is bounded if c and m are constant over time.

The instantaneous per capita utility, $u(\,)$, is a function of per capita consumption, $c(t)$, and per capita real money balances, $m(t)$, both of which are assumed to be normal goods. In order to estimate the model and derive its implications for seigniorage and the welfare cost of inflation, we use the utility function:

$$u[c(t),\, m(t)] = \left(\frac{1}{1-s}\right)\left[\left(\frac{1}{1-a}\right)(c_t^{\,1-a} - 1) + \left(\frac{d}{1-b}\right)(m_t^{\,1-b} - 1)\right]^{1-s}, \qquad \text{(III.2)}$$

where *a, b, d* and *s* are positive constants so that *u( )* is a strictly concave utility function[3] with continuous first and second derivatives. To find a steady state solution to the model in which the real rate of interest and the growth rate of various quantities are constant, the elasticity of marginal utility of consumption and money must be constant asymptotically. We therefore use this form which reduces to *constant intertemporal elasticity of substitution* (CIES) utility function when $s = 0$. This utility function, which is similar to MaCurdy's (1981), has, as special cases, an additively separable utility function in per capita consumption and per capita real money balances, ($s=0$)[4]; a logarithmic utility function ($a=1,\ b=1,\ s=0$)[5]. This functional form also provides for the possibility of differential degrees of intertemporal substitution in consumption and money. This is the easiest to see when $s=0$, so that $1/a$ represents the elasticity of intertemporal substitution of consumption and $1/b$ represents the corresponding elasticity for money[6].

Individuals will be assumed to maximize utility subject to the budget constraint

$$\frac{F}{N} + \frac{M}{PN} = r(1-\tau)\frac{F}{N} + \frac{H}{N} + \frac{W}{N}(1-\tau) - c, \qquad \text{(III.3)}$$

where F is non-monetary assets such as capital and bonds, M is nominal money, τ is an income tax rate, r is the real interest rate on real assets F, W is the real wage rate and H is some real lump-sum transfer from the government. Equation (III.3) says that per capita savings (non-consumed resources) are equal to the per capita investment plus money accumulation. If we define lower case variables as the real per capita versions of their capital letter counterparts $(m = M / PN, f = F / N, h = H / N, w = W / N)$ and we denote per capita real assets by v_t ($v_t = m_t + f_t$), we can rewrite the household's budget constraint as

$$\dot{v}_t = r(1-\tau)v_t + h_t + w_t(1-\tau) - c_t - nv_t - i_t m_t{}^{7}, \tag{III.4}$$

where i_t is the after-tax nominal interest rate[8] $i_t = r(1-\tau)+\pi_t$, where π_t is the (expected) and the actual inflation rate. Equation (III.4) implies that assets (monetary and non-monetary) per person rise with per capita income, $r(1-\tau)v_t + w_t(1-\tau)$, and the lump-sum transfer (assuming $h_t > 0$) from the government, fall with per capita consumption, c_t, and the opportunity cost of holding per capita real balances, $i_t\, m_t$, as well as the expansion of the population in accordance with the term nv_t.

We can now set up the present-value Hamiltonian[9] and get the following set of first-order conditions:

$$J = e^{-(\rho-n)t}\left(\frac{1}{1-s}\right)\left[\left(\frac{1}{1-a}\right)(c_t{}^{1-a}-1)+\left(\frac{d}{1-b}\right)(m_t{}^{1-b}-1)\right]^{1-s} + \tag{III.5a}$$

$$\Omega_t\,[\, r(1-\tau)v_t + h_t + w_t(1-\tau) - c_t - nv_t - i_t m_t]$$

$$e^{-(\rho-n)t}\left[\left(\frac{1}{1-a}\right)(c_t{}^{1-a}-1)+\left(\frac{d}{1-b}\right)(m_t{}^{1-b}-1)\right]^{-s} c_t{}^{-a} = \Omega_t\,, \tag{III.5b}$$

$$e^{-(\rho-n)t}\left[\left(\frac{1}{1-a}\right)(c_t{}^{1-a}-1)+\left(\frac{d}{1-b}\right)(m_t{}^{1-b}-1)\right]^{-s} d\, m_t{}^{-b} = \Omega_t\, i_t\,, \tag{III.5c}$$

$$\Omega_t\,(r(1-\tau)-n) = -\,\dot{\Omega}_t\,, \qquad \text{(III.5d)}$$

$$\lim_{t\to\infty}\,\Omega_t\,v_t = 0. \qquad \text{(III.5e)}$$

The transversality condition[10] in Equation (III.5e) says that the value of the household's assets (both monetary and non-monetary) - which equals the quantity v_t times the shadow price of income Ω_t - must approach 0 as time approaches infinity. That is, optimizing agents do not want to have any valuable assets left over at the end of the planning horizon.

The shadow price Ω_t evolves over time in accordance with Equation (III.5d). Integration of this equation with respect to time yields

$$\Omega_t = \Omega_0\, e^{-\int_0^t \{[r(1-\tau)](\Omega)-n\}d\Omega} \qquad \text{(III.6)}$$

The term Ω_0 equals $U_c\,[c(0),\ m(0)]$, which is positive. If we substitute the result for Ω_t into Equation (III.5e), then the transversality condition becomes

$$\lim_{t\to\infty}\left[v_t e^{-\int_0^t \{[r(1-\tau)](\Omega)-n\}d\Omega}\right] = 0 \qquad \text{(III.7)}$$

This equation implies that the quantity of assets per person, v_t, does not grow asymptotically at a rate as high as $r(1-\tau) - n$ or, equivalently, that the level of assets does not grow at a rate as high as $r(1-\tau)$. It would be sub-optimal for households to accumulate positive assets forever at the rate $r(1-\tau)$ or higher, because utility would increase if these assets were instead consumed in finite time.

Dividing (III.5c) by (III.5b), we can get the double-log money demand function,

$$m_t = d^{1/b}\, i_t^{-1/b} c_t^{a/b} \qquad \text{(III.8)}$$

where *-1/b* and *a/b* are the interest rate and the per capita consumption elasticity of demand for real money balances, respectively. Equation (III.8) indicates that per capita real demand for money is a positive function of real consumption per capita and a negative function of the opportunity cost of holding it, the nominal interest rate.

In order to find closed form solutions for the growth rate, we will now consider the special case of $s=0$, reducing Equation (III.5b) to

$$e^{-(\rho-n)t} c_t^{-a} = \Omega_t. \qquad \text{(III.5a')}$$

If we take logarithmic differentiation of Equation (III.5a') with respect to time, we get the basic condition for choosing consumption over time[11]:

$$\frac{\dot{c}_t}{c_t} = \frac{r(1-\tau)-\rho}{a}. \qquad \text{(III.9)}$$

where we used the fact that $\frac{\dot{\Omega}_t}{\Omega_t} = n - r(1-\tau)$ from Equation (III.5d). Therefore, the relation between $r(1-\tau)$ and ρ determines whether households choose a pattern of per capita real consumption that rises over time, stays constant, or falls over time. Households would select a flat consumption profile, with $\frac{\dot{c}_t}{c_t} = 0$, if $r(1-\tau) = \rho$. Households would be willing to depart from this flat pattern and sacrifice some consumption today for more consumption tomorrow - that is $\frac{\dot{c}_t}{c_t} > 0$, only if they are compensated by the after-tax real interest rate, $r(1-\tau)$, that is sufficiently above ρ. A lower willingness to substitute intertemporally (a higher value of "*a*") implies a smaller responsiveness of $\frac{\dot{c}_t}{c_t}$ to the gap between $r(1-\tau)$ and ρ. Therefore, consumption growth is a positive function of the difference between the after-tax real interest rate and the discount rate, and it is a positive function of the rate of accumulation of per capita real money balances. If the nominal interest rate is constant we can take logs and derivatives of per capita real money demand equation (III.8) and find that

$$\frac{m_t}{m_t} = \frac{a}{b}\frac{c_t}{c_t}, \qquad \text{(III.10)}$$

which can be rewritten as

$$\frac{m_t}{m_t} = \frac{r(1-\tau)-\rho}{b}, \qquad \text{(III.11)}$$

where b is the reciprocal of the elasticity of intertemporal substitution in money.

The next step is to specify the behavior of firms. Firms produce goods, pay wages for labor input, and make rental payments for capital input. We will model the production side of the economy with the following function,

$$Y_t = F\,(K_t\,,\, N_t), \qquad \text{(III.12)}$$

where Y_t is the flow of output, K_t is the capital input (in units of consumables), N_t is labor input (in person-hours per year). The production function $F(\,)$ is *neoclassical* if it satisfies the following properties. First, for all $K_t > 0$ and $N_t > 0$, $F(\,)$ exhibits positive and diminishing marginal products with respect to each input:

$$\partial/\partial K > 0, \quad \partial^2 F/\partial K^2 < 0\,, \quad \partial F/\partial N > 0, \quad \partial^2 F/\partial N^2 < 0. \qquad \text{(III.13a)}$$

Second, $F(\,)$ exhibits constant returns to scale[12]:

$$F(\phi K, \phi N) = \phi F(K,N) \text{ for all } \phi > 0. \qquad \text{(III.13b)}$$

Third, the marginal product of capital (or labor) approaches infinity as capital (or labor) goes to zero and approaches zero as capital (or labor) goes to infinity[13]:

$$\lim_{K \to 0} (F_K) = \lim_{L \to 0} (F_L) = \infty,$$

$$\lim_{K \to \infty} (F_K) = \lim_{L \to \infty} (F_L) = 0. \qquad \text{(III.13c)}$$

The condition of constant returns to scale implies that output can be written as

$$Y = F(K, N) = N\, F(K/N, 1) = N\, (f(k), 1)$$

where $k \equiv K/N$ is the capital-labor ratio, $y \equiv Y/N$ is per capita output, and the function $f(k)$, is defined to equal $F(k, 1)$. This result means that the production function can be expressed in *intensive form*[14] as

$$y = f(k), \qquad \text{(III.14)}$$

where the marginal products of the factor inputs are given by

$$\partial Y / \partial K = f'(k), \quad \partial Y / \partial N = f(k) - k f'(k). \qquad \text{(III.15)}$$

Let R be the rental price for a unit of capital services, and assume that capital stocks depreciate at the constant rate $\delta \geq 0$. The net rate of return to a household that owns a unit of capital is then $R-\delta$. Recall that households can also receive the interest rate r on bonds, since capital and bonds are perfect substitutes as stores of value, we must have $r = R-\delta$ or, equivalently, $R = r + \delta$. The representative firm's flow of net receipts or profit at any point in time is given by

$$Profit = F(K, N) - (r + \delta)\, K - w\, N, \qquad \text{(III.16)}$$

that is gross receipts from the sale of output, $F(K, N)$, less the factor payments, which are rentals to capital, $(r + \delta)\, K$, and wages to workers, $w\, N$. We assume that the firm seeks to maximize the profits and the competitive firm, which takes r and w as given, maximizes profit by setting

$$r = f'(k) - \delta,$$

$$w = f(k) - kf'(k). \qquad \text{(III.17)}$$

That is, in order for profit to be zero, the competitive firm chooses capital-labor ratio to equate the marginal product of capital to the rental price and the marginal product of labor to the wage rate.

Since, in our model, we have introduced some taxes fall on all forms of income - including interest income and returns to capital as well as wage earnings, we can see how the presence of an income tax alters the theoretical model. As pointed out in footnote 8, the marginal tax rate, τ, in this study, applies to all forms of income - including interest income and returns to capital, as well as labor earnings. Then, if the tax rate is constant over time, the after-tax real rate is $r_t\,(1-\tau)$, and the after-tax return to capital is $(f'(k_t) - \delta)(1-\tau)$ where δ is the constant depreciation rate. Investors now equate their return after tax, $(f'(k_t) - \delta)(1-\tau)$, to the after-tax interest rate, $r_t\,(1-\tau)$ which yields $f'(k_t) - \delta = r_t$. A steady state requires equality between $r(1-\tau)$ and the rate of time preference, ρ, thus implying that

$$r = \rho/(1-\tau). \qquad \text{(III.18)}$$

Since an increase in τ raises the steady-state real interest rate, r, in Equation (III.18). It follows that the steady-state capital-labor ratio, k, declines. Thus, in the words of Barro (1989), "when taxes apply to the earnings from capital, a higher tax rate leads in the long run to a lower intensity (and in the short run to less investment)".

In order to close the model we need to specify the behavior of the government. In general, the government faces the following budget constraint.

$$G + H + rB = M / P + B + T, \tag{III.19}$$

where G is real government spending, H is the total lump-sum transfers that appear in the private budget constraint, B is total real public debt, B is the real budget deficit, M / P is the collection of seigniorage and T is tax collection. The budget constraint, therefore, says that the government spends resources in purchasing goods, G, transferring income to households, H, and paying interest in past debt, rB. The resources of the government are taxes, T, and seigniorage, M / P. The difference between expenditure and revenue is the budget deficit B.

We will make some further simplifying assumptions so as to isolate the important effects. First, for simplicity, we will not allow the government to issue debt so we will set

$$B = B = 0. \tag{III.20}$$

While this is a restrictive hypothesis, we concentrate in this paper on the steady state of this economy in which the government is satisfying its intertemporal budget constraint and (domestic and external) debt accumulation is equal to zero[15].

Second, we need to make an assumption on how the government sets G. Since the economy will be growing we would like to assume that the government can spend increasing amounts of resources. We will assume therefore that G is a constant proportion, ε, of private consumption:

$$G = \varepsilon C, \tag{III.21}$$

where ε will be the policy parameter that tells how large government spending is relative to the economy. We express G as a proportion of C but it would be equally easy to assume that G is a constant fraction of total output or capital stock, given that in steady state they all grow at the same rate.

Third, we have already assumed when we have discussed the consumers and the producers side that the tax revenue is based on some taxes fall on income, i.e., the constant average and marginal tax rate τ applies to the returns to capital as well as labor earnings[16]. Therefore, the per capita real tax revenue is

$$\frac{T}{N} = \tau rk + w\tau. \qquad \text{(III.22)}$$

And finally, we will assume that the government sets the nominal growth rate of money at a constant level μ so

$$M/ M = \mu. \qquad \text{(III.23)}$$

The resulting budget constraint in per capita terms is the following:

$$m = g + h - r\tau k - w\tau - m(\pi + n). \qquad \text{(III.24)}$$

Since we assume that there is no international borrowing and lending, private bonds are in zero net supply so we can identify f (non-liquid assets) with k (real productive capital stock). Then, substituting (III.24) into (III.4), we can get the social budget constraint[17].

$$\dot{k} = f(k) - (n + \delta)k - (1+\varepsilon)c\,, \qquad \text{(III.25)}$$

where the firms' first-order condition (III.17) has been used and where $g = \varepsilon c$. This resource constraint says that the increase in the capital-labor ratio is the difference between total output per capita, $f(k)$, and the effective depreciation, $(n + \delta)k$, plus private consumer expenditures and public spending, $g = \varepsilon c$. Finally, we can rewrite the government budget constraint (III.19) as

$$g = m\mu + \tau (f(k) - \delta k) \qquad \text{(III.26)}$$

where we neglect transfers and we used the firms' first-order condition (III.17) and the fact that $m/ m = \mu - \pi - n$. Equation (III.26) says that total government spending must equal the total tax revenue. The tax revenue, in turn is equal to the sum of revenue from money creation and income taxes.

IV. THE SPECIFICATION OF THE MONEY DEMAND FUNCTION

The specification of the money demand function has important implications for a number of macroeconomic issues. First, if policymakers are to be responsible for achieving price stability they need reliable quantitative estimates of money demand. In particular, if the money demand function is stable, the income elasticity yields the rate of money growth that is consistent with long run price stability. Second, macroeconomic theorists need quantitative estimates of the money demand function in order to determine the exact predictions of their models. In Keynesian models, for instance, the relative ability of monetary and fiscal policy to affect the real economy depends on the elasticities of the demand for money. For a given interest elasticity, a larger income elasticity or for a given income elasticity, a smaller interest elasticity implies a more vertical LM curve; as a result, monetary policy is relatively more potent than fiscal policy. Furthermore, in many macroeconomic models, money demand elasticities are important figures for the equilibrium conditions since both the income and the interest rate elasticity of money demand matter in determining the aggregate price level and the inflation rate, given the growth rate of money. Finally, the specification of the money demand function has important implications for seigniorage and the welfare cost of inflation. In this paper we reexamine two aspects of this specification, the choice of the scale variable and the functional form of the money demand function. While GDP is the standard scale variable for money demand in macroeconomic models, we argue that both theoretical and empirical considerations suggests that it may not be the right choice. Furthermore, we find that consumer expenditure is a more empirically successful scale variable in estimated money demand equations.

V. THE CHOICE OF A SCALE VARIABLE: CONSUMPTION OR INCOME ?

The conclusion that money is an important determinant of aggregate demand is widely accepted among macroeconomists due to the empirical work of Friedman and Schwartz (1963,1982) and others which shows that there is a close empirical connection between monetary fluctuations and the business cycle. These studies emphasize that over long periods of time and under a variety of institutional arrangements, nominal GDP move in accordance with the money supply. In other words, the velocity of money appears relatively stable.

Even though velocity measured with respect to nominal GDP is stable this fact does not imply that GDP is the correct quantity variable in the quantity equation. Gross domestic product is only a proxy for total transactions, and another variable may be a better proxy. Empirically, velocity measured with respect to another variable may be even more stable than velocity traditionally defined.

Mankiw and Summers (1986) examine the standard deviation of velocity by detrending the log of velocity and then regressing the detrended log of velocity on time and the square of time. Furthermore, they compute the standard deviation of the change in the log of velocity for the U.S. economy between 1960:I and 1984:IV. The empirical results that they obtain indicate that velocity measured using consumer expenditures is more stable than velocity measured using income. They also argue that consumption is a better scale variable than income because it more accurately reflects permanent income.

Another study by Faig (1989) also finds that consumption is better scale variable than income in the money demand equation. According to his study, in some countries, such as Germany, consumption and income have quite different seasonal patterns. Faig shows that the seasonal patterns of money balances more closely matches that of consumption.

In this section, following Mankiw and Summers (1986), we compare alternative velocity measures. In particular, we compare velocity measured with respect to gross domestic product (GDP/M) to velocity measured with respect to nominal consumer expenditure (C/M). Our objective is to find the nominal aggregate that yields the more stable measure of velocity.

To measure the standard deviation of velocity, since the velocity of money, which has historically trended upward, is a non stationary process we should either detrend or first-difference the time series data. We compute the standard deviation both ways. We detrend by regressing the log of velocity on time and the square of time. The residuals from this regression are always highly correlated, implying that velocity is possibly not stationary. The first difference of log velocity appears close to white noise, indicating that velocity is a random walk process.

Table 1 presents the standard deviation of the velocity of different monetary aggregates for the period 1957:I and 1994:I using quarterly data. These figures do not support the traditional formulation using nominal GDP. Velocity measured using consumer expenditure (C/M) is unambiguously more stable than velocity measured using total production (Y/M). For the seven monetary aggregates presented in Table 1, consumer expenditure yields the more stable measure of velocity.

Therefore, the empirical results presented in this paper suggest that consumption is a better scale variable in the money demand function than GDP since consumer spending produces more stable measures of monetary velocity and outperforms GDP in estimated money demand equations. The reader may wish to consult Tables 3-8 to see whether the money demand function scaled with GDP or consumption is a more accurate characterization of the actual U.S. data.

VI. ESTIMATION TECHNIQUES OF THE MONEY DEMAND FUNCTION

In estimating the regression equation of the demand for per capita real money balances in U.S., the crucial step is to define the specific form of money demand and to find an efficient mechanism to estimate short-run money demand equations which are consistent with imposed long-run relationships. A standard assumption in money demand models is that short-run actual real balances divert from their long-run desired levels. Transaction costs involved in adjusting money holdings, sluggish adjustment of prices, interest rates or income can all be factors leading to this divergence. Estimation of long-run demand for money balances requires a short-run dynamic specification, and the *partial adjustment model* has been widely used for this purpose. In this section, we will first focus on the derivation of the partial adjustment model and then use an estimation technique that allows us to test whether a semi-log or a double-log money demand function gives better fit for the U.S. economy.

Partial adjustment is typically motivated by cost-minimizing behavior where the costs of disequilibrium are balanced against adjustment costs. Following Hwang (1985), consider a quadratic cost function of the form:

$$C = \gamma_1[\ln M_t^d - \ln M_t]^2 + \gamma_2 [(\ln M_t - \ln M_{t-1}) - \delta(\ln P_t - \ln P_{t-1})]^2, \qquad \text{(VI.1)}$$

where $M_t = m_t P_t$, $M_t^d = m_t^d P_t$, m_t^d is the "desired" stock of real balances, and P_t is the price level. The first and second terms of Equation (VI.1) correspond to the disequilibrium and adjustment costs, respectively. For $\delta = 0$, Equation (VI.1) posits that the adjustment term is only a function of nominal magnitudes, i.e., Equation (VI.1) becomes

$$C = \gamma_1[\ln M_t^d - \ln M_t]^2 + \gamma_2 [\ln M_t - \ln M_{t-1}]^2. \qquad \text{(VI.2)}$$

For $\delta = 1$, both the first and the second term of Equation (VI.1) turns out to be a function of real magnitudes, i.e., Equation (VI.1) reduces to

$$C = \gamma_1[\ln m_t^d - \ln m_t]^2 + \gamma_2 [\ln m_t - \ln m_{t-1}]^2. \qquad \text{(VI.3)}$$

Minimizing costs presented in Equation (VI.1) with respect to M_t yields:

$$\ln M_t - \ln M_{t-1} = \theta (\ln M_t^d - \ln M_{t-1}) + \delta(1 - \theta)(\ln P_t - \ln P_{t-1}), \qquad \text{(VI.4)}$$

where $\theta = \gamma_1/(\gamma_1+\gamma_2)$. Thus, when $\delta = 1$, Equation (VI.4) reduces to the *real partial adjustment model* in which real balances are adjusted. Alternatively, when $\delta = 0$, Equation (VI.4) reduces to the *nominal partial adjustment model* in which nominal balances are adjusted. To complete the story, a standard specification for $m_t^d = M_t/P_t$ is given by Equation (VI.5)

$$\ln m_t^d = \beta_0 + \beta_1 i_t + \beta_2 \ln y_t, \qquad \text{(VI.5)}$$

where y_t is a transactions variable such as real GDP, i_t represents interest rate. Combining (VI.4) and (VI.5) and rearranging terms yields

$$ln\ m_t = \theta\beta_0 + \theta\beta_1 i_t + \theta\beta_2 ln\ y_t + (1-\theta) ln\ m_{t-1}, \tag{VI.6}$$

where we use the real partial adjustment model:

$$ln\ m_t - ln\ m_{t-1} = \theta (ln\ m_t^d - ln\ m_{t-1}). \tag{VI.7}$$

The econometric tests performed in this study aim at finding whether the money demand function is logarithmic or semi-logarithmic in i_t, assuming that it is logarithmic in y_t. To determine what form the money demand will take, the *Box-Cox transformation*[18],

$$g^{(\lambda)}(i_t) = \frac{i_t^{\lambda} - 1}{\lambda}, \tag{VI.8}$$

is applied to the i_t variable. Note that in the following regression equation,

$$ln\ m_t^d = \beta_o + \beta_1 \left(\frac{i_t^{\lambda} - 1}{\lambda} \right) + \beta_2 ln\ y_t + e_t, \tag{VI.9}$$

when λ equals one, Equation (VI.9) reduces to

$$ln\ m_t^d = (\beta_o - \beta_1) + \beta_1 i_t + \beta_2 ln\ y_t + e_t, \tag{VI.10a}$$

and the demand for real money balances per capita becomes semi-logarithmic in i_t. When λ equals zero, the transformation is, by *L'Hopital's rule*,

$$lim_{\lambda \to 0} \left(\frac{i_t^{\lambda} - 1}{\lambda} \right) = lim_{\lambda \to 0} \frac{d(i_t^{\lambda} - 1)/d\lambda}{1}$$

$$= lim_{\lambda \to 0} (i_t^{\lambda} ln\ i_t) = ln\ i_t.$$

Hence, $\lambda = 0$ in $g^{(\lambda)}(i_t)$ implies that the demand for per capita real money balances is logarithmic in i_t,

$$ln\ m_t^d = \beta_o + \beta_1 ln\ i_t + \beta_2 ln\ y_t + e_t. \tag{VI.10b}$$

Substituting (VI.9) into (VI.7) yields

$$\ln m_t^d = (\theta\beta_o) + (\theta\beta_1)\left(\frac{i_t^\lambda - 1}{\lambda}\right) + (\theta\beta_2) \ln y_t + (1-\theta) \ln m_{t-1} + \theta e_t. \qquad \text{(VI.11)}$$

Since our main purpose is to set out a more accurate procedure to use in estimating seigniorage revenue and the welfare cost of inflation -both as a share of real income- in steady state, the per capita income elasticity of demand for real money balances is assumed to be unity so that the equilibrium level of velocity is independent of the growth of output. In other words, seigniorage and the welfare cost analysis is compatible with a steady state only if the velocity of money remains constant. In turn, constant velocity requires a unitary per capita income elasticity of demand for real balances which reduces (VI.11) to

$$\ln (m/y)_t^d = (\theta\beta_o) + (\theta\beta_1)\left(\frac{i_t^\lambda - 1}{\lambda}\right) + (1-\theta) \ln (m/y)_{t-1} + \theta e_t. \qquad \text{(VI.12)}$$

In estimating (VI.12) which is nonlinear in its parameters, a simple one dimensional grid search method[19] and a nonlinear least square estimation technique are used. If λ in (VI.12) is taken to be an unknown parameter, the regression becomes nonlinear in the parameters. Although no transformation will reduce it to linearity, nonlinear least squares is straightforward with the development of easy-to-use software. Since our goal is to find whether the demand for real balances is logarithmic or semi-logarithmic in i_t, the value of λ is expected to be between zero and one. Then, λ is estimated by scanning this range in increments of 0.1. If a minimum of the sum of squares is found and greater precision is desired, the area to the right and left of the current optimum can be searched in increments of 0.01, and so on.

There are two kinds of nonlinearity in economic models. A model can be nonlinear in its variables or nonlinear in its parameters, or both. An equation that is nonlinear in its variables but linear in its parameters can be estimated with least squares or two-stage least squares. But, if the equation is nonlinear in its parameters there is no way to set up variables in advance so that ordinary least squares can be used for estimation. Instead, we will have to use nonlinear least squares.

A nonlinear regression model is one for which the first-order conditions for least squares estimation of the parameters are nonlinear functions of the parameters. Nonlinear least squares methods[20] estimate the maximum likelihood parameters by minimizing the sum of squared residuals. Since the first-order conditions for estimating the parameters will be a set of nonlinear equations that do not have an explicit solution this typically requires an iterative procedure for solution. We can now present the simplified version of this procedure discussed in Greene (1997) and Amemiya (1985).

Let the nonlinear regression model be

$$y = h(x, b) + e. \tag{VI.13}$$

Many of the results that have been obtained for nonlinear regression models are based on a linear Taylor series approximation to $h(x, b)$ at a particular initial value for the parameter vector b^0 :

$$h(x, b) \cong h(x, b^0) + \sum_{k=1}^{K} \left(\frac{\partial h(x, b^0)}{\partial b_k^0} \right) (b_k - b_k^0). \tag{VI.14a}$$

This is called the *linearized regression model*. By collecting terms, we obtain

$$h(x, b) \cong \left[h(x, b^0) - \sum_{k=1}^{K} b_k^0 \left(\frac{\partial h(x, b^0)}{\partial b_k^0} \right) \right] + \sum_{k=1}^{K} b_k \left(\frac{\partial h(x, b^0)}{\partial b_k^0} \right). \tag{VI.14b}$$

Let x_k^0 equal the kth derivative, $\partial h(x, \beta^0)/\partial\beta_k^0$. For a given value of β^0, this is a function only of the data, not of the unknown parameters. We now have

$$h(x, b) \cong \left[h^0 - \sum_{k=1}^{K} x_k^0 b_k^0\right] + \sum_{k=1}^{K} x_k^0 b_k \qquad \text{(VI.14c)}$$

$$= h^0 - (x^0)'b^0 + (x^0)'b$$

or

$$y \cong h^0 - (x^0)'b^0 + (x^0)'b + e.$$

By placing the known terms on the left-hand side of the equation, we obtain a regression model:

$$y^0 = y - h^0 + (x^0)'b^0 \qquad \text{(VI.14d)}$$

$$= (x^0)'b + e^0.$$

With a value of b^0 in hand, we could compute y^0 and x^0 then estimate the parameters of (VI.14d) by linear least squares.

To present the method of nonlinear least squares more explicitly, consider our nonlinear regression model,

$$ln\ m_t^d = \beta_0 + \beta_1 \left(\frac{i_t^\lambda - 1}{\lambda}\right) + \beta_2\, ln\ y_t + \beta_3\, ln\ m_{t-1} + e_t,$$

and note that, in the context of the regression analysis, nonlinear least squares estimation is the method used to estimate the value of parameters that minimize (one half of) the sum of squares deviations,

$$S(\beta) = \frac{1}{2}\sum_{t=1}^{n} e_t^2 = \frac{1}{2}\sum_{t=1}^{n}\left[\ln m_t - \beta_0 - \beta_1\left(\frac{i_t^\lambda - 1}{\lambda}\right) - \beta_2 \ln y_t - \beta_3 \ln m_{t-1}\right]^2. \qquad \text{(VI.15)}$$

The first-order conditions for estimating the parameters of (VI.15) by least squares are

$$\partial S(\beta)/\partial\beta_0 = -\sum_{t=1}^{n}\left[\ln m_t - \beta_0 - \beta_1\left(\frac{i_t^{\lambda}-1}{\lambda}\right) - \beta_2 \ln y_t - \beta_3 \ln m_{t-1}\right] = 0, \qquad \text{(VI.16a)}$$

$$\partial S(\beta)/\partial\beta_1 = -\sum_{t=1}^{n}\left[\ln m_t - \beta_0 - \beta_1\left(\frac{i_t^{\lambda}-1}{\lambda}\right) - \beta_2 \ln y_t - \beta_3 \ln m_{t-1}\right]\left(\frac{i_t^{\lambda}-1}{\lambda}\right) = 0, \qquad \text{(VI.16b)}$$

$$\partial S(\beta)/\partial\beta_2 = -\sum_{t=1}^{n}\left[\ln m_t - \beta_0 - \beta_1\left(\frac{i_t^{\lambda}-1}{\lambda}\right) - \beta_2 \ln y_t - \beta_3 \ln m_{t-1}\right](\ln y_t) = 0, \qquad \text{(VI.16c)}$$

$$\partial S(\beta)/\partial\beta_3 = -\sum_{t=1}^{n}\left[\ln m_t - \beta_0 - \beta_1\left(\frac{i_t^{\lambda}-1}{\lambda}\right) - \beta_2 \ln y_t - \beta_3 \ln m_{t-1}\right](\ln m_{t-1}) = 0, \qquad \text{(VI.16d)}$$

$$\partial S(\beta)/\partial\lambda = -\sum_{t=1}^{n}\left[\ln m_t - \beta_0 - \beta_1\left(\frac{i_t^{\lambda}-1}{\lambda}\right) - \beta_2 \ln y_t - \beta_3 \ln m_{t-1}\right]\left[\beta_1\left(\frac{i^{\lambda}(\lambda \ln i - 1)+1}{\lambda^2}\right)\right] = 0. \qquad \text{(VI.16e)}$$

To get the maximum likelihood or the nonlinear least squares estimator, since we cannot solve the set of nonlinear equations (VI.16) explicitly for β we must solve it iteratively: Start from an initial estimate of β (say β^0) and obtain a sequence of estimates by iteration, which will converge to the values of the parameters that minimize (VI.15). Numerous iterative methods[21] have been proposed and used.

In this paper, for nonlinear estimation of the money demand function, Marquardt algorithm is used in the following way. At each iteration, it computes approximate derivatives with respect to each parameter, by making a small change in the parameters, one at a time, and seeing how much the equation changes as a result. Then it regresses the dependent variable on the derivatives. When necessary, it uses ridge regression in order to avoid problems of collinearity, but the final results do not involve ridge regression. The regression gives a vector of proposed changes in the parameters. It then evaluates whether the proposed changes actually decrease the sum of squared residuals. If they do, it adopts the changes and starts a new iteration. If not, it tries a fraction of the changes. In principle, there should be some fraction such that improvement is obtained. This process is continued until the proposed changes are very small compared to the parameters themselves.

VII. ECONMETRIC TESTS FOR THE SPECIFICATION OF THE MONEY DEMAND FUNCTION

To determine the proper specification of the demand for real money balances per capita, we have first estimated the most general version (nonlinear form) of the money demand equation given in (43). Money demand equations have been estimated with two different monetary aggregates M1 and monetary base (MB). Currency, together with reserves, constitutes the monetary base. M1 is the sum of currency, traveler's checks, demand deposits and other checkable deposits. Furthermore, all these estimates have been carried out by using both a short term (3-month t-bill rate) and a long-term interest rate (10-year government bond yield) to test the idea that time-series estimates of money demand are somewhat sensitive to the choice of an interest rate. The quantitative outcomes presented in Table 5 indicate that the magnitude of the elasticities in money demand equations estimated with the long-term interest rate is slightly bigger than the estimates with the short-term interest rate.

To find out precisely which scale variable is more appropriate in money demand equation, two different quantity variables (consumption and income) have been utilized in our estimations. The empirical results presented in Table 1 suggest that consumption is a better scale variable in the money demand function than GDP since consumer spending produces more stable measures of monetary velocity and outperforms GDP in estimated money demand equations.

We also compute the sum of squared residuals and the log likelihood of each regression to see whether a semi-log or a double-log money demand function with consumption or income gives better fit for the U.S. economy. The statistical figures in Table 3-4 and 6-7 point out that the double-log specification with consumption gives the best fit for the U.S. economy supporting our derivation of money demand from the utility function in the intertemporal optimizing framework presented in Section III.

An important assumption of the classical linear model utilized in this paper is that there is no autocorrelation or serial correlation among the disturbances entering into the population regression function. In the presence of autocorrelation, the ordinary least squares estimators are still linear-unbiased as well as consistent, but they are no longer efficient (i.e., minimum variance). Therefore, if one detects autocorrelation in his OLS estimates, he should use generalized least squares (GLS) instead of ordinary least squares (OLS) even though the estimators derived from OLS are unbiased and consistent. The major problems of using OLS in the presence of autocorrelation are as follows:

The residual variance is likely to underestimate the true σ^2, i.e., the usual residual variance will be downward biased and as a result we are likely to overestimate R^2. Even if σ^2 is not underestimated, the usual OLS variance of the parameters may underestimate its variance under autocorrelation. Therefore, if we use the OLS variance, we will inflate the precision or accuracy (i.e., underestimate the standard error) of the estimators. As a result, in computing the *t* ratio, we will be overestimating the *t* value and hence the statistical significance of the estimators.

Since, in our model, one of the regressors is the lagged value of the dependent variable, the Durbin-Watson statistics is not used to detect serial correlation in the disturbance term. To test autocorrelation in autoregressive models, the Durbin-h statistic[22] is used,

$$h = \hat{\rho}\sqrt{\frac{n}{1-n[Var(\beta_{lag})]}}\ , \qquad \text{(VII.1)}$$

where n = sample size, $Var(\beta_{lag})$ = variance of the coefficient of the lagged dependent variable, and $\hat{\rho}$ = estimate of the first-order serial correlation, which is given by

$$\hat{\rho} = 1 - \frac{D.W.}{2} \qquad \text{(VII.2)}$$

where $D.W.$ is the usual Durbin-Watson statistic. Therefore, (VII.1) can be written as

$$h = \left[1 - \frac{D.W.}{2}\right]\sqrt{\frac{n}{1-n[Var(\beta_{lag})]}} \qquad \text{(VII.3)}$$

where $h \sim AN(0, 1)$ is asymptotically normally distributed with zero mean and unit variance. From the normal distribution we know that

$$P(-1.96 \leq h \leq 1.96) = 0.95, \qquad \text{(VII.4)}$$

that is, the probability that h lying between -1.96 and 1.96 is about 95 percent. Therefore, the decision rule is

(i) if $h > 1.96$ reject the null hypothesis that there is no positive first-order autocorrelation, and

(ii) if $h < -1.96$ reject the null hypothesis that there is no negative first-order autocorrelation, but

(iii) if h lies between *-1.96* and *1.96* do not reject the null hypothesis that there is no first-order (positive or negative) autocorrelation.

As presented in our tables, the statistical results regarding autocorrelation (the Durbin-Watson and the h statistics) indicate that in our money demand estimates we do not have autocorrelation problem since h statistic of each regression is less than 1.96 or greater than -1.96.

As pointed out in the previous discussions, a semi-logarithmic Cagan-type money demand function would be proper specification if the value of λ were found to be statistically not different from one whereas if λ is statistically equal to zero then the demand for real money balances per capita is logarithmic in i_t. To find out precisely what form the money demand function will take, the formal tests have been applied to Equation (VI.12) which is the most relevant money demand equation for steady state seigniorage and welfare cost analysis. Maximum likelihood estimates of all the parameters of (VI.12) are obtained by scanning the range of λ from 0 to 0.1 in increments of 0.01 and the range of λ from 0.1 to 1.0 in increments of 0.1. The sum of squared residuals, which we denote SSR, for a range of values of λ are shown in Table 2. The results in Table 2 that we get from the nonlinear least squares estimation of the model using Marquardt algorithm indicates that over the range of λ from 0 to 1, a minimum of the sum of squares is found at the value of zero suggesting the double-log form of the money demand function.

Having computed the estimate of λ, the next step is to test whether the true λ is statistically different from 1 or different from 0. To serve the purpose, three different tests - the Likelihood ratio test, the Lagrange multiplier test, and the Wald test - have been employed.

<u>*(i) The Likelihood Ratio Test*[23] :</u> The likelihood for a sample of n observations, assuming normally distributed disturbances, is

$$ln\ L = -(n/2)\ ln(2\pi\sigma^2) - (\varepsilon'\varepsilon / 2\sigma^2). \qquad \text{(VII.5)}$$

Letting $ln\ L^*$ the log-likelihood evaluated at the restricted estimates, the likelihood statistic for testing the restrictions is

$$LR = -2(ln\ L^* - ln\ L). \qquad \text{(VII.6)}$$

This statistic is asymptotically distributed as chi-squared with J degrees of freedom where J is the number of restrictions. Since the maximum likelihood estimate of σ^2 is $e'e/n$ and $\varepsilon \sim N(0, \sigma^2 I)$ where I is the identity matrix, the log-likelihood computed at the least squares estimates is

$$ln\ L = -(n/2)\ [1 + ln\ (2\pi) + ln\ (e'e/n)], \qquad \text{(VII.7)}$$

and similarly the log-likelihood for restricted estimates,

$$ln\ L^* = -(n/2)\ [1 + ln\ (2\pi) + ln\ (e^{*\prime}e^*/n)], \qquad \text{(VII.8)}$$

and plugging (VII.7) and (VII.8) into (VII.6), the likelihood ratio statistic for a classical regression model (linear or nonlinear) is

$$LR = n\ (\,ln\ \sigma^2_* - ln\ \sigma^2\,). \qquad \text{(VII.9)}$$

To find out the proper specification of our fundamental money demand equation (VI.12),

$$ln\ (m/y)_t^d = (\theta\beta_o) + (\theta\beta_1)\left(\frac{i_t^\lambda - 1}{\lambda}\right) + (1-\theta)\ ln\ (m/y)_{t-1} + \theta e_t,$$

the nonlinear least squares estimation using Marquardt algorithm yielded the following estimated values for the parameters in (VI.12) with t-ratios in parentheses and the statistical figures,

β_o = -3.138990 (-4.423661)

β_1 = -0.034032 (-0.743663)

λ = -0.820392 (-2.095169)

θ = 0.145334 (27.61332)

$ln\ L$ = 206.8527

σ = 0.060634

R^2 = 0.919609

$D.W.$ = 2.209793

To determine the proper functional form of the money demand equation by using the Likelihood ratio test, the restricted regression models have been estimated with $\lambda=0$ (for testing the double-log form) and (with $\lambda=1$ for testing the semi-log form). The log-likelihood values of the restricted regressions have been found to be

$ln\ L^* = 204.6030$ (for double-log, $\lambda=0$),

$ln\ L^* = 200.5028$ (for semi-log, $\lambda=1$).

The test is carried out by using (VII.6) and the two chi-squared test statistics are

$LR^{log\text{-}log} = -2[ln\ L^*(\lambda=0) - ln\ L(\lambda=\text{MLE})] = -2(204.6030-206.8527) = 4.4994,$

$LR^{semi\text{-}log} = -2[ln\ L^*(\lambda=1) - ln\ L(\lambda=\text{MLE})] = -2(200.5028-206.8527) = 12.6998.$

Semi-logarithmic form of the money demand is rejected, but the double-log form is not rejected at the 1 % significance level since the critical value from the table is $\chi_{(0.01)} = 6.63$.

We can support our previous findings by using (VII.9). To perform the alternative Likelihood ratio test, we utilize the following forecast error variances of the restricted regressions,

$\sigma_{\bullet}^{2}$ = 0.0037639 (log-log),

$\sigma_{\bullet}^{2}$ = 0.003978 (semi-log).

The two chi-squared test statistics for the alternative LR test are

$$LR^{log\text{-}log} = n*ln(\sigma_{\bullet}^{2}/\sigma^{2}) = 148*ln(0.0037639/0.003676) = 148*(0.023499) = 3.4779,$$

$$LR^{semi\text{-}log} = n*ln(\sigma_{\bullet}^{2}/\sigma^{2}) = 148*ln(0.003978/0.003676) = 148*(0.078906) = 11.678.$$

As presented above, the chi-squared test statistics with the alternative LR test are very close to the values obtained by using (VII.6) which implies the same conclusion that the double-log specification of the money demand function is better than the semi-log form since $LR^{log\text{-}log}$ is less but $LR^{semi\text{-}log}$ is greater than the critical value from the table, $\chi_{(0.01)}$ = 6.63.

(ii) The Lagrange Multiplier Test[24] : The Lagrange multiplier test is based on the decrease in the sum of squared residuals that would result if the restrictions in the restricted model were released. For the nonlinear regression model, the test has a particularly appealing form. Let $e_{\bullet}$ be the vector of residuals $y_i - h(x_i, b_{\bullet})$ computed using the restricted estimates. Recall that we defined X^0 as an $n \times K$ matrix of derivatives computed at a particular parameter vector in (VI.14). Let $X_{\bullet}^{0}$ be this matrix computed at the restricted estimates. Then the Lagrange multiplier statistic for the nonlinear regression model is

$$LM = \frac{e_{\bullet}' X_{\bullet}^{0} [X_{\bullet}^{0'} X_{\bullet}^{0}]^{-1} X_{\bullet}^{0'} e_{\bullet}}{e_{\bullet}' e_{\bullet} / n} .$$

This is asymptotically distributed as a chi-squared statistic with degrees of freedom equal to the number of restrictions. Note also that the Lagrange multiplier statistic is n times the uncentered R^2 in the regression of $e_{\bullet}$ on $X_{\bullet}^{0}$, $LM = nR^2$.

(iii) Lagrange Multiplier Test for the Box-Cox Model[25] : To support our previous findings using the one-dimensional grid-search method and the Likelihood ratio test that the double-log specification of the money demand equation is better than the semi-log form, we will now apply the Lagrange multiplier test to our fundamental money demand equation given by

$$ln\ (m/y)_t^d = \alpha_o + \alpha_1 \left(\frac{i_t^\lambda - 1}{\lambda} \right) + \alpha_2\, ln\ (m/y)_{t-1} + e_t, \qquad \text{(VII.10)}$$

$$= f(i_t\ , ln\ (m/y)_{t-1}\ , \alpha_o\ , \ \alpha_1\ , \alpha_2 . \lambda) + e_t.$$

We first consider a Lagrange multiplier test of the hypothesis that λ equals zero. The partial derivative regressors are

$$x_1^\bullet = \partial f(.)/\partial\alpha_o = 1, \qquad \text{(VII.11a)}$$

$$x_2^\bullet = \partial f(.)/\partial\alpha_1 = \left(\frac{i_t^\lambda - 1}{\lambda} \right), \qquad \text{(VII.11b)}$$

$$x_3^\bullet = \partial f(.)/\partial\alpha_2 = ln\ (m/y)_{t-1}\ , \qquad \text{(VII.11c)}$$

$$x_4^\bullet = \partial f(.)/\partial\lambda \ = \alpha_1 \left[\frac{\lambda \cdot i_t^\lambda \cdot \ln i_t - i_t^\lambda + 1}{\lambda^2} \right]. \qquad \text{(VII.11d)}$$

The test is carried out by first regressing $ln\ (m/y)_t$ on a *constant*, $ln\ i_t$, and $ln\ (m/y)_{t-1}$ (i.e., the regressors in (VII.10) evaluated at $\lambda=0$) and then computing $nR_\bullet^2$ in the regression of the residuals from this first regression on $x_1^\bullet$, $x_2^\bullet$, $x_3^\bullet$, and $x_4^\bullet$, also evaluated at $\lambda=0$. As presented above $x_1^\bullet = 1$ and $x_3^\bullet = ln\ (m/y)_{t-1}$. The next step is to find $x_2^\bullet|_{\lambda=0}$ and $x_4^\bullet|_{\lambda=0}$. We first compute $x_2^\bullet$ at $\lambda=0$ as follows,

$$x_2^\bullet\, |_{\lambda=0} = lim_{\lambda\to 0} \left(\frac{i_t^\lambda - 1}{\lambda} \right)$$

$$= lim_{\lambda \to 0} \frac{d(i_t^{\lambda} - 1)/d\lambda}{1}$$

$$= lim_{\lambda \to 0} (i_t^{\lambda} \ ln\ i_t) = ln\ i_t.$$

Second, we can calculate $x_4^{\bullet}$ at $\lambda=0$,

$$x_4^{\bullet}|_{\lambda=0} = \alpha_1 \lim_{\lambda \to 0} \left[\frac{\lambda \cdot i_t^{\lambda} \cdot \ln i_t - i_t^{\lambda} + 1}{\lambda^2} \right],$$

Now applying L'Hopital's rule to the right hand-side, differentiate numerator and denominator with respect to λ:

$$x_4^{\bullet}|_{\lambda=0} = \alpha_1 \lim_{\lambda \to 0} \left[\frac{\ln i_t [i_t^{\lambda} + i_t^{\lambda} \lambda \ln i_t] - i_t^{\lambda} \ln i_t}{2\lambda} \right],$$

$$= \alpha_1 \lim_{\lambda \to 0} \left[\frac{i_t^{\lambda} (\ln i_t)^2 \lambda}{2\lambda} \right],$$

$$= \alpha_1 \left[\frac{(\ln i_t)^2}{2} \right].$$

Therefore, to determine whether the double-log is the proper specification or not, the Lagrange multiplier test, in this study, is carried out in two steps. First, we regress *ln* $(m/y)_t$ on a *constant*, *ln* i_t, and *ln* $(m/y)_{t-1}$ and compute the *residuals*. Second, we regress these *residuals* on a *constant*, *ln* i_t, *ln* $(m/y)_{t-1}$, $\alpha_1 \left[\frac{(\ln i_t)^2}{2} \right]$, where α_1 is the coefficient on *ln* i_t in the first regression. The Lagrange multiplier statistic is $nR_{\bullet}^2$ from the second regression.

As presented in Table 3, we have already computed the first regression. Regressing the residuals of the first regression on $x_1^{\bullet}$, $x_2^{\bullet}|_{\lambda=0}$, $x_3^{\bullet}$, and $x_4^{\bullet}|_{\lambda=0}$, we have estimated the second regression equation,

$$e_t^{log\text{-}log} = \delta_0 + \delta_1 \ln i_t + \delta_2 \ln (m/y)_{t-1} + \delta_3 \{\alpha_1 [(\ln i_t)^2/2]\} + \eta_t \,. \qquad \text{(VII.12)}$$

The estimated values for the parameters in (VII.12) with t-ratios in parentheses and the statistical results are,

$\delta_0 =$ 0.282192 (1.658834)

$\delta_1 =$ 0.200982 (2.189806)

$\delta_2 =$ -0.003331 (-0.107918)

$\delta_3 =$ -1.112625 (-2.213789)

$\ln L$ = 207.0796

σ = 0.060542

R^2 = 0.032914

$D.W.$ = 2.180528

n = 148

The Lagrange multiplier statistic for testing $\lambda=0$ is

$$LM^{log\text{-}log} = n{*}R^2 = 148{*}(0.032914) = 4.8712,$$

which implies that the double-log specification of the money demand function is not rejected at the 1 % significance level since the critical value from the table is 6.63.

We will now consider the Lagrange multiplier test of the hypothesis that λ equals 1. The test is carried out by first regressing $\ln (m/y)_t$ on a *constant*, $(i_t - 1)$, and $\ln (m/y)_{t-1}$ and then computing $n{*}R^2$ in the regression of the residuals from this first regression on z_1^{*}, z_2^{*}, z_3^{*}, and z_4^{*}, also evaluated at $\lambda = 1$. Given the unrestricted model in (VII.10), the same partial derivative regressors presented in (VII.11) are obtained. To test whether the semi-logarithmic form is the proper specification of the money demand function or not, the partial derivative regressors evaluated at $\lambda=1$ are defined as follows,

$$z_1^{\bullet} = \partial f(.)/\partial\alpha_0|_{\lambda=1} = 1, \qquad \text{(VII.13a)}$$

$$z_2^{\bullet} = \partial f(.)/\partial\alpha_1|_{\lambda=1} = i_t - 1, \qquad \text{(VII.13b)}$$

$$z_3^{\bullet} = \partial f(.)/\partial\alpha_2|_{\lambda=1} = ln\ (m/y)_{t-1}, \qquad \text{(VII.13c)}$$

$$z_4^{\bullet} = \partial f(.)/\partial\lambda|_{\lambda=1} = \alpha_1\ [i_t\ (ln\ i_t - 1) + 1]. \qquad \text{(VII.13d)}$$

Thus, to find out whether or not the semi-log is the proper specification, the Lagrange multiplier test, in this study, is carried out in two steps. First, we regress $ln\ (m/y)_t$ on a *constant*, $(i_t - 1)$, and $ln\ (m/y)_{t-1}$ and compute the *residuals*. Second, we regress these *residuals* on a *constant*, $(i_t - 1)$, $ln\ (m/y)_{t-1}$, $\alpha_1\ [i_t\ (ln\ i_t - 1) + 1]$, where α_1 is the coefficient on $(i_t - 1)$ in the first regression. The Lagrange multiplier statistic is $nR_{\bullet}^{2}$ from the second regression.

Having estimated the first regression, regressing the residuals of the first regression on $z_1^{\bullet}$, $z_2^{\bullet}|_{\lambda=1}$, $z_3^{\bullet}$, and $z_4^{\bullet}|_{\lambda=1}$, we have estimated the second regression equation,

$$e_t^{semi\text{-}log} = \gamma_0 + \gamma_1\ (i_t - 1) + \gamma_2\ ln\ (m/y)_{t-1} + \gamma_3\ \{\alpha_1\ [i_t\ (ln\ i_t - 1) + 1]\} + \xi_t. \qquad \text{(VII.14)}$$

The estimated values for the parameters in (VII.14) with t-ratios in parentheses and the statistical results are,

γ_0 = 3.200544 (2.886252)

γ_1 = 5.100607 (2.947646)

γ_2 = -0.030205 (-0.958634)

γ_3 = -2.589748 (-2.972171)

$ln\ L$ = 204.9086

σ = 0.061436

R^2 = 0.057800

$D.W.$ = 2.036683

n = 148

The Lagrange multiplier statistic for testing $\lambda=1$ is

$$LM^{semi\text{-}log} = n*R^2 = 148*(0.057800) = 8.5544,$$

which implies that the semi-log specification of the money demand function is rejected at the 1 % significance level since the critical value from the table is 6.63.

<u>*(iv) The Wald Test*[26] :</u> A practical shortcoming of the likelihood ratio test is that it usually requires estimation of both the restricted and unrestricted parameter vectors. In complex models, one or the other of these estimates may be very difficult to compute. Fortunately, there are two alternative testing procedures, the Wald test and the Lagrange multiplier test which was discussed in the previous section, that circumvent this problem. Both tests are based on an estimator that is asymptotically normally distributed. In this section, we will first introduce the distribution of the Wald test statistic then apply this test to our fundamental regression equation (VI.12).

The Wald statistic is

$$W = [c(\hat{\theta}) - q]'(Var\,[c(\hat{\theta}) - q])^{-1}\,[c(\hat{\theta}) - q]. \qquad \text{(VII.15)}$$

Under $H_0 : c(\theta) = q$, in large samples, W has a chi-squared distribution with degrees of freedom equal to the number of restrictions [*i.e., the number of equations in* $c(\hat{\theta}) - q = 0$]. For testing a set of linear restrictions $R\theta = q$, the Wald test would be based on

$$H_0 : c(\theta) - q = R\theta - q = 0,$$

$$C = \left[\frac{\partial c(\hat{\theta})}{\partial \theta} \right] = R,$$

$$Var\ [c(\hat{\theta}) - q] = R\ Var\ [\hat{\theta}]\ R',$$

and

$$W = [R\hat{\theta} - q]'\ [R\ Var\ [\hat{\theta}]\ R']^{-1} [R\hat{\theta} - q].$$

The degrees of freedom is the number of rows in R.

To determine the proper functional form of the money demand equation, the test is $H_0 : \lambda = 0$ for testing the double-log form and $H_0 : \lambda = 1$ for testing the Cagan-form. The Wald test will be based on

$$W = [(\hat{\lambda} - \lambda_0) - 0](Var[(\hat{\lambda} - \lambda_0) - 0])^{-1}[(\hat{\lambda} - \lambda_0) - 0]$$

$$= \frac{(\hat{\lambda} - \lambda_0)^2}{Var[\hat{\lambda}]}$$

where $\lambda_0 = 1$ for testing the Cagan-form and $\lambda_0 = 0$ for testing the double-log form of the money demand function. Here W has a chi-squared distribution with one degree of freedom. As pointed out on page 29, the estimated value of λ and its t-ratio for our fundamental money demand equation are,

$\hat{\lambda}$ = -0.820392

t_λ = -2.095169

which implies that $Var(\hat{\lambda}) = 0.153322$. Hence, to determine the proper specification of U.S. money demand we can compute the Wald test statistic for testing the semi-log and the double-log form.

$$W^{log\text{-}log} = \frac{(\hat{\lambda}-0)^2}{Var[\hat{\lambda}]} = \frac{(-0.820392-0)^2}{0.153322} = 4.389736$$

$$W^{semi\text{-}log} = \frac{(\hat{\lambda}-1)^2}{Var[\hat{\lambda}]} = \frac{(-0.820392-1)^2}{0.153322} = 21.613513$$

It is not difficult to estimate the Wald test statistic even in the nonlinear regression models with the development of easy-to-use econometrics software packages. By using one of these softwares, we obtained more or less the same Wald test statistics.

Wald test for testing double-log form :

Null Hypothesis: λ=0

F-statistic	4.389713	Probability	0.037907
Chi-square	4.389713	Probability	0.036156

Wald test for testing semi-log (Cagan) form:

Null Hypothesis: λ=1

F-statistic	21.50437	Probability	0.000008
Chi-square	21.50437	Probability	0.000004

The semi-logarithmic form of the money demand function is rejected, but the double-log form is not rejected at the 1% significance level since the critical value from the table is 6.63.

As a result of the one-dimensional grid-search method, the Likelihood ratio test, the Lagrange multiplier test, and the Wald test, we end up with the same conclusion that the demand for real money balances per capita is logarithmic in i_t, approving the double-log form of the money demand function as the proper specification for the U.S. economy.

VIII. TIME SERIES ANALYSIS OF THE VARIABLES

In this section, we develop and illustrate a general procedure to determine whether or not a series contains a unit root by using the augmented Dickey-Fuller test. Unit root tests are sensitive to the presence of deterministic regressors, such as an intercept term or a deterministic time trend. We, therefore, present the formal tests to identify the specification of the variables for testing unit root. After testing variables for trends and unit root, we will introduce the basic concept of cointegration and show how it applies to macroeconomic models. Finally, we study the Engle-Granger(1987) methodology for cointegration. Now, let's introduce such concepts as stationarity, nonstationarity and unit root then we will concentrate on cointegration.

Unit Root Processes: There are important differences between stationary and nonstationary time series. Shocks to a stationary time series are necessarily temporary; over time, the effects of the shocks will dissipate and the series will return to its long-run mean level. In other words, long-term forecasts of a stationary series will converge to the unconditional mean of the series. To identify whether a series is stationary or not we will introduce the major discrepancies between stationary and nonstationary series. We know that a covariance of a stationary series:

(a) exhibits mean reversion in that it fluctuates around a constant long-run mean.

(b) has a finite variance that is time-invariant.

(c) has a theoretical correlogram that diminishes as lag length increases.

On the other hand, a nonstationary series necessarily has permanent components. The mean and/or variance of a nonstationary series are time-dependent. To aid in identification of a nonstationary series, we know that:

(a) there is no long-run mean to which the series returns.

(b) the variance is time-dependent and goes to infinity as time approaches infinity.

(c) theoretical autocorrelations do not decay but, in finite samples, the sample correlogram dies out slowly.

Although the properties of a sample correlogram are useful tools for detecting the possible presence of unit roots, the method is necessarily imprecise. Therefore, we consider the formal test procedures for the presence of unit roots.

Dickey-Fuller Tests: In general, the procedure is to determine whether $a_1 = 1$ in the model

$$y_t = a_1 y_{t-1} + \varepsilon_t.$$

Substracting y_{t-1} from each side of the equation in order to write the equivalent form:

$$\Delta y_t = \mu y_{t-1} + \varepsilon_t , \qquad \text{(VIII.1)}$$

where $\mu = a_1 - 1$. Of course, testing the hypothesis $a_1 = 1$ is equivalent to testing the hypothesis $\mu = 0$. Dickey and Fuller (1979) actually considers three different regression equations that can be used to test for the presence of a unit root:

$$\Delta y_t = \mu y_{t-i} + \varepsilon_t, \qquad \text{(VIII.2a)}$$

$$\Delta y_t = a_0 + \mu y_{t-1} + \varepsilon_t, \qquad \text{(VIII.2b)}$$

$$\Delta y_t = a_0 + \mu y_{t-1} + a_2 t + \varepsilon_t. \qquad \text{(VIII.2c)}$$

The difference between the three regressions concerns the presence of the deterministic elements a_0 and $a_2 t$. The first is a pure random walk model, the second adds an intercept or drift term, and the third includes both a drift and linear time trend.

The parameter of interest in all the regression equations is μ; if $\mu = 0$, the $\{ y_t \}$ sequence contains a unit root. The test involves estimating one (or more) of the equations above using OLS in order to obtain the estimated value of μ and its standard error.

Comparing the resulting t-statistic with the appropriate value reported in the Dickey-Fuller tables allows us to determine whether to accept or reject the null hypothesis $\mu = 0$.

The methodology is precisely the same, regardless of which of the three forms of the equations is estimated. However, be aware that the critical values of the t-statistics do depend on whether an intercept and/or time trend is included in the regression equation. In their Monte Carlo study, Dickey and Fuller (1979) found that the critical values for $\mu = 0$ depend on the form of the regression and the sample size. Therefore, they constructed different statistical tables to use for equations (VIII.2).

As discussed in the next section, Dickey and Fuller (1981) provide three additional F-statistics (called ϕ_1, ϕ_2, and ϕ_3) to test *joint* hypotheses on the coefficients. With (VIII.2b), the null hypothesis $\mu = a_0 = 0$ is tested using the ϕ_1 statistic. Including a time trend in the regression -so that (VIII.2c) is estimated- the joint hypothesis $\mu = a_0 = a_2 = 0$ is tested using the ϕ_2 statistic and the joint hypothesis $\mu = a_2 = 0$ is tested using the ϕ_3 statistic.

The ϕ_1, ϕ_2, and ϕ_3 statistics are constructed in exactly the same way as ordinary F-tests are:

$$\phi_1 = \frac{[RSS(restricted) - RSS(unrestricted)]/q}{RSS(unrestricted)/(n-k)}, \qquad \text{(VIII.3)}$$

where *RSS(restricted)* and *RSS(unrestricted)* = the sum of the squared residuals from the restricted and unrestricted models,

q = number of restrictions,

n = number of usable observations,

k = number of parameters estimated in the unrestricted model.

Comparing the calculated value of ϕ_i to the appropriate value reported in Dickey and Fuller (1981) allows us to determine whether we can reject or accept the null hypothesis that the data are generated by the restricted model. If the calculated value of ϕ_i is smaller than that reported by Dickey and Fuller, we can accept the restricted model. If the calculated value of ϕ_i is larger than that reported by Dickey and Fuller, we can reject the null hypothesis and conclude that the data are generated by the unrestricted model.

Extensions of the Dickey-Fuller Test (Augmented Dickey-Fuller Test) : Not all time-series processes can be represented by the first-order autoregressive process

$$\Delta y_t = a_0 + \mu y_{t-1} + a_2 t + \varepsilon_t.$$

It is possible to use the Dickey-Fuller tests in higher-order equations such as

$$\Delta y_t = \mu y_{t-1} + \sum_{i=2}^{p} \psi_i \Delta y_{t-i+1}\ \varepsilon_t, \qquad \text{(VIII.4a)}$$

$$\Delta y_t = a_0 + \mu y_{t-1} + \sum_{i=2}^{p} \psi_i \Delta y_{t-i+1}\ \varepsilon_t, \qquad \text{(VIII.4b)}$$

$$\Delta y_t = a_0 + \mu y_{t-1} + a_2 t + \sum_{i=2}^{p} \psi_i \Delta y_{t-i+1} + \varepsilon_t. \qquad \text{(VIII.4c)}$$

In (VIII.4), the coefficients of interest is μ ; if $\mu = 0$, the equation is entirely in first differences and so has a unit root. We can test for the presence of a unit root using the same Dickey-Fuller statistics discussed above. Again, the appropriate statistic to use depend on the deterministic components included in the regression equation.

Now the only concern with the augmented Dickey-Fuller test is to determine the appropriate lag length. Including too many lags reduces the power of the test to reject the null of a unit root since the increased number of lags necessitates the estimation of additional parameters and a loss of degrees of freedom. The degrees of freedom decrease since the number of parameters estimated has increased and because the number of usable observations has decreased. (We lose one observation for each additional lag included in the autoregression.) On the other hand, too few lags will not appropriately capture the actual error process, so that μ and its standard error will not be well estimated.

How do we select the appropriate lag length in such circumstances?

Model Selection Criteria: One natural question to ask of any estimated model is: How well does it fit the data? Adding additional lags will necessarily reduce the sum of squares of the estimated residuals. However, adding such lags brings the estimation of additional coefficients and an associated loss of degrees of freedom as a consequence. Moreover, the inclusion of extraneous coefficients will reduce the forecasting performance of the fitted model. There exist various model selection criteria that help in selecting the most appropriate model. The two most commonly used model selection criteria are the Akaike information criterion (AIC) and Schwartz Bayesian criterion (SBC), calculated as

$$\text{AIC} = n \ln (RSS) + 2k$$

$$\text{SBC} = n \ln (RSS) + k \ln (n)$$

where k = number of parameters estimated (including constant term)

n = number of observations.

Ideally, the AIC and SBC will be as small as possible (note that both can be negative); model A is said to fit better than model B if the AIC (or SBC) for A is smaller than that for model B. For the time series analysis of the variables, we use these criteria to select the optimum lag length for unit root and cointegration tests.

In this study, we first used the ordinary F-test statistics to test for the presence of deterministic regressors, such as an intercept term or a deterministic time trend. Using our quarterly data set (1957:I-1994:I) with 149 observations, we fail to reject the null hypothesis at the 5 % and the 1 % level of significance that the data for log(MB/GDP) and log(MB/Consumption) contain both an intercept and a deterministic time trend and the data for log(3-month t-bill rate) contain only an intercept term. Then, we utilized the augmented Dickey-Fuller test to determine whether or not the variables employed in our steady state money demand estimates contain a unit root after selecting the most appropriate lag length using the Akaike information criterion (AIC) and Schwartz Bayesian criterion (SBC). As presented in Table 8a and 8b, the variables in the present paper; log(Monetary Base/GDP), log(Monetary Base /Consumption), log(M1/GDP), log(M1/Consumption), short-term interest rate (3-month t-bill rate), and long-term interest rate (10-year government bond yield) are all characterized as nonstationary I(1) variables since we found that there is a unit root in the level of the variables but there is none in their first differences. The reader may wish to consult Table 8a and 8b to understand our model selection criteria (i.e., selection of optimum lag length) and to see the augmented Dickey-Fuller test results for the level and the first difference of the variables.

Cointegration: In the macroeconometric literature, it is a widely accepted idea that equilibrium theories involving nonstationary variables require that the existance of a combination of the variables that is stationary. Any equilibrium relationship among a set of nonstationary variables implies that their stochastic trends must be linked. After all, the equilibrium relationship means that the variables cannot move independently of each other. This linkage among the stochastic trends necessitates that the variables be cointegrated.

The problem confronting us is that the variables employed in our steady state money demand estimates; log(Monetary Base/GDP), log(Monetary Base /Consumption), log(M1/GDP), log(M1/Consumption), short-term interest rate (3-month t-bill rate), and long-term interest rate (10-year government bond yield) can all be characterized as nonstationary I(1) variables. That is to say, each variable can meander without any tendency to return to a long-run level. In the last section, we showed that all of the variables used in this study are integrated of order 1 so that we can use the Engle-Granger cointegration test. The theory asserts that there exists a linear combination of these nonstationary variables that is stationary. Therefore, the residual term from the regression equation $\{\varepsilon_t\}$ must be stationary.

The concept of cointegration was first introduced by Engle and Granger (1987). Their formal analysis begins by considering a set of economic variables in long-run equilibrium when

$$\omega_1 x_{1t} + \omega_2 x_{2t} + \dots\dots\dots + \omega_n x_{nt} = 0 \qquad \text{(VIII.5)}$$

If we let ω and x_t denote the vectors $(\omega_1, \omega_2, \dots\dots\dots, \omega_n)$ and $(x_{1t}, x_{2t} \dots\dots\dots x_{nt})'$, the system is in long-run equilibrium when $\omega \mathbf{x}_t = 0$. The deviation from long-run equilibrium - called the equilibrium error- is ε_t, so that

$$\omega \mathbf{x}_t = \varepsilon_t. \qquad \text{(VIII.6)}$$

If the equilibrium is meaningful, it must be the case that the equilibrium error process is stationary. Engle and Granger (1987) provide the following definition of cointegration.

The components of the vector $x_t = (x_{1t}, x_{2t} \dots\dots\dots x_{nt})'$ are said to be *cointegrated of order d, b*, denoted by $x_t \sim CI(d, b)$ if

(a) all components of x_t are integrated of order d.

(b) there exists a vector $\omega = (\omega_1, \omega_2, \dots\dots\dots, \omega_n)$ such that linear combination $\omega x_t = \omega_1 x_{1t} + \omega_2 x_{2t} + \dots\dots\dots + \omega_n x_{nt}$ is integrated of order $(d\text{-}b)$, where $b>0$.

The vector ω is called the cointegrating vector. Note that the cointegrating vector is not unique. If $(\omega_1, \omega_2, \dots\dots\dots, \omega_n)$ is a cointegrating vector, then for any nonzero value of Γ, $(\Gamma\omega_1, \Gamma\omega_2, \dots\dots\dots, \Gamma\omega_n)$ is also a cointegrating vector. Typically, one of the variables is used to normalize the cointegrating vector by fixing its coefficients at unity. To normalize the cointegrating vector with respect to x_{1t}, simply select $\Gamma = 1/\omega_1$. Another important point about cointegration is that all variables must be integrated of the same order. Of course, this does not imply that all similarly integrated variables are cointegrated; usually, a set of I(d) variables is not cointegrated. Such a lack of cointegration implies no long-run equilibrium among the variables, so that they can wander arbitrarily far from each other. If the variables are integrated of different orders, they cannot be cointegrated.

Testing for Cointegration: The Engle-Granger Methodology To explain the Engle-Granger testing procedure, let us begin with the type of problem to be encountered in applied studies. Suppose that two variables -say, y_t and z_t- are believed to be integrated of order 1 and we want to determine whether there exists an equilibrium relationship between the two. Engle-Granger (1987) propose a straightforward test whether two I(1) variables are cointegrated of order $CI(1,1)$.

Step **1**: Pretest the variables for their order of integration. By definition, cointegration necessitates that the variables be integrated of the same order. Thus, the first step in the analysis is to pretest each variable to determine its order of integration. If both variables are stationary, it is not necessary to proceed since standard time-series methods apply to stationary variables. If the variables are integrated of different orders, it is possible to conclude that they are not cointegrated.

Step 2: Estimate the long-run equilibrium relationship. If the results of Step 1 indicate that both $\{y_t\}$ and $\{z_t\}$ are I(1), the next step is to estimate the long-run equilibrium relationship in the form:

$$y_t = \beta_0 + \beta_1 z_t + \varepsilon_t. \qquad \text{(VIII.7)}$$

If the variables are cointegrated, an OLS regression yields a "super-consistent" estimator of the cointegrating parameters β_0 and β_1. Stock (1987) proves that the OLS estimates of β_0 and β_1 converge faster than in OLS models using stationary variables.

In order to determine if the variables are actually cointegrated, denote the residual sequence be $\{e_t\}$. Thus, $\{e_t\}$ is the series of the estimated residuals of the long-run relationship. If these deviations from long-run equilibrium are found to be stationary, the $\{y_t\}$ and $\{z_t\}$ sequences are cointegrated of order (1, 1). It would be convenient if we could perform a Dickey-Fuller test on these residuals to determine their order of integration. Consider the autoregression of the residuals:

$$\Delta e_t = a_1 e_{t-1} + \zeta_t. \qquad \text{(VIII.8)}$$

Since the $\{e_t\}$ sequence is a residual from a regression equation, there is no need to include an intercept term; the parameter of interest in (VIII.8) is a_1. If we cannot reject the null hypothesis $a_1 = 0$, we can conclude that the residual series contains a unit root. Hence, we conclude that the $\{y_t\}$ and $\{z_t\}$ sequences are not cointegrated. That is, the rejection of the null hypothesis implies that the residual sequence is stationary. Given that both $\{y_t\}$ and $\{z_t\}$ were found to be I(1) and the residuals are stationary, we can conclude that the series are cointegrated of order (1, 1).

In most applied studies, it is not possible to use the Dickey-Fuller tables for cointegration. Fortunately, Engle and Granger provide test statistics that can be used to test the hypothesis $a_1 = 0$. When more than two variables appear in the equilibrium relationship, the appropriate tables are provided by Engle and Yoo (1987).

If the residuals of (VIII.8) do not appear to be white-noise, an augmented Dickey-Fuller test can be used instead of (VIII.8). Estimate the autoregression:

$$\Delta e_t = a_1 e_{t-1} + \sum_{i=1}^{n} a_{i+1} \Delta e_{t-i}\ \zeta_t. \qquad \text{(VIII.9)}$$

Again, if -2< a_1 <0, we can conclude that the residual sequence is stationary and $\{y_t\}$ and $\{z_t\}$ are CI (1, 1).

In order to determine if the variables used in this study are actually cointegrated, we first obtain the residual sequence from each money demand regression equation and then perform a Dickey-Fuller test on these residuals to determine whether or not the residual series contain a unit root. The test involves estimating Equation (VIII.8) presented above using OLS in order to obtain the estimated value of a_1 and its standard error. Comparing the resulting t-statistic with the appropriate value reported in Table 9 (MacKinnon critical values) allows us to determine whether to accept or reject the null hypothesis $a_1 = 0$. As presented in Table 8a, 8b and 9, given that the variables utilized in U.S. money demand estimates were found to be I(1) and the residuals are stationary, we can conclude that the series are cointegrated of order (1,1). Although the residuals of (VIII.8) appeared to be white-noise, an augmented Dickey-Fuller test was used after the selection of the most appropriate model with optimum lag length. Having estimated the autoregression given in Equation (VIII.9), we again obtained the residual sequence to be stationary.

IX. SEIGNIORAGE REVENUE AND THE WELFARE COST OF INFLATION

Conventional estimates of the seigniorage-maximizing inflation rate generally make use of the Cagan (1956) function, where the log of real money balances is a linear function of the inflation rate. It is very common to use the Cagan form because of its simplicity and its attractive algebraic property: revenue maximizing inflation rate (π_{max}) is given by the reciprocal of coefficient on inflation (the inflation semi-elasticity of demand for real money balances). However, this simplicity is achieved at the expense of a restrictive functional form, which assumes a constant semi-elasticity of money demand with respect to inflation. Our aim, in this section, is to test the sensitivity of estimates of π_{max}, maximum seigniorage revenue as a fraction of GDP, and the marginal as well as the average welfare cost of revenue from inflation to the specification of the money demand function.

To see the implications of different money demand specifications for steady-state seigniorage, we will begin with the most general (nonlinear) form of the money demand function which has already been introduced in section VI,

$$ln\ (m/y)_t^d = \alpha_0 - \alpha_1 \left(\frac{i_t^{\lambda} - 1}{\lambda} \right) \qquad \text{(IX.1)}$$

where the per capita income elasticity of demand for real money balances is assumed to be unity so that the equilibrium level of velocity is independent of the growth of output, which is compatible with steady state seigniorage[27]. As we have already discussed in the previous sections, if the Box-Cox transformation is applied only to the independent variable, i_t, we obtain this general form of money demand equation presented in Equation (74) which gives the Cagan form (when $\lambda = 1$) or the double-log form (when $\lambda = 0$).

Equation (IX.1) implies that the interest semi-elasticity of money demand varies with the nominal rate of interest since

$$\eta^{\text{semi-elasticity}} = d\, ln(m/y)_t \,/\, di_t = -\ \alpha_1\ i_t^{\lambda-1}, \qquad \text{(IX.2a)}$$

and unless $\lambda = 0$ the interest elasticity of money demand also varies with i_t since

$$\eta^{\text{elasticity}} = d\, ln(m/y)_t \,/\, d\, ln i_t = -\ \alpha_1\ i_t^{\lambda}. \qquad \text{(IX.2b)}$$

The nonlinear form of money demand implies that the semi-elasticity and the elasticity of money demand increase or decrease with the change in the nominal rate of interest depending on the value of λ assuming $\alpha_1 > 0$. If $\lambda > 1$ ($\lambda = 1$, $\lambda < 1$), the absolute value of the semi-elasticity rises (does not change, declines) with the increase in the nominal rate of interest since

$$\frac{d\eta^{semi-elasticity}}{di_t} = d^2\, ln(m/y)_t \,/\, di_t^2 = -\alpha_1\ (\lambda - 1)\ i_t^{\lambda-2}. \qquad \text{(IX.2c)}$$

Analogously, if $\lambda > 0$ ($\lambda = 0$, $\lambda < 0$), the absolute value of the elasticity rises (does not change, declines) with the increase in the nominal rate of interest since

$$\frac{d\eta^{elasticity}}{di_t} = \frac{d[d\ln(m/y)_t \,/\, d\ln i_t]}{di_t} = -\ \alpha_1\, \lambda\, i_t^{\lambda-1}. \qquad \text{(IX.2d)}$$

Traditionally, the revenue from money creation is defined as the stock of real money balances held times the growth rate of the nominal money supply. This implies that when the nominal money stock is not growing, no revenue accrues to the government. The Phelps (1973) - Auernheimer (1974) definition uses the nominal rate of interest (rather than the growth rate of the nominal money supply) times stock of real balances as the revenue accruing from the authorities. The rationale for this definition is as follows. Less revenue is spent by the authorities because governments have monopoly power to hold part of their debt to the private sector in the form of non-interest bearing money rather than as interest bearing bonds. As long as the nominal rate of interest is greater than zero, this monopoly power implies that on the government debt in the form of money, no interest is paid. This saving of interest payment is measured by the stock of non-interest bearing money times the nominal rate of interest.

We can now derive the seigniorage maximizing inflation rate using the Phelps-Auernheimer definition of the revenue, S. Since we intend to find the seigniorage revenue and the welfare cost of revenue from inflation as a ratio to real GDP, we can define seigniorage to real income ratio as

$$(S/y)_t^{nonlinear} = (m/y)_t\, i_t = e^{\alpha_0 - \alpha_1 \left(\frac{i_t^{\lambda} - 1}{\lambda} \right)} i_t, \qquad \text{(IX.3)}$$

since $(m/y)_t^{nonlinear} = e^{\alpha_0 - \alpha_1 \left(\frac{i_t^{\lambda} - 1}{\lambda} \right)}$ from Equation (IX.1). Revenue maximizing interest rate, i_{max}, is obtained from the first-order condition for seigniorage maximum,

$$d(S/y)/di = e^{\alpha_0 - \alpha_1 \left(\frac{i_t^{\lambda} - 1}{\lambda} \right)} (1 - \alpha_1\, i_t^{\lambda}) = 0 \qquad \text{(IX.4)}$$

which yields

$$i_{max} = (1/\alpha_1)^{1/\lambda}. \qquad \text{(IX.5)}$$

With the real interest rate, r, the revenue maximizing inflation rate can be obtained by using the fact that $i = r + \pi$. Hence,

$$\pi_{max} = (1/\alpha_1)^{1/\lambda} - r. \qquad \text{(IX.6)}$$

Substituting i_{max} from Equation (IX.5) into $(S/y)^{\text{nonlinear}}$, the maximum value of seigniorage as a share of real income is found to be

$$(S/y)_{max} = e^{\alpha_0 - \left(\frac{1-\alpha_1}{\lambda}\right)} (1/\alpha_1)^{1\,\lambda}. \qquad \text{(IX.7)}$$

When we consider the special case of λ = 1, the nonlinear money demand equation reduces to simple constant semi-elasticity form.

$$(m/y)_t^{\text{Cagan}} = e^{(\alpha_0+\alpha_1)-\alpha_1 i_t}, \qquad \text{(IX.8a)}$$

where

$$i_{max} = (1/\alpha_1) \quad \text{and} \quad (S/y)_{max} = e^{(\alpha_0+\alpha_1-1)}(1/\alpha_1). \qquad \text{(IX.8b)}$$

When λ = 0, we get the double-log money demand function with a constant elasticity of $-\alpha_1$. Unlike the semi-log function used by Cagan and Bailey, the double-log function with a constant elasticity less than unity does not generate a Laffer curve. Along a Laffer curve as the tax rate (the nominal rate of interest) rises, the tax revenue first increases, reaches a maximum and then declines. To see this, note that seigniorage to real income ratio is

$$(S/y)_t^{log\text{-}log} = e^{\alpha 0}\, i_t^{1-\alpha 1} \qquad \text{(IX.9a)}$$

since $(m/y)_t^{log\text{-}log} = e^{\alpha 0}\, i_t^{-\alpha 1}$ when λ = 0. Clearly, the revenue from money creation continually increases as the nominal rate of interest rises[28] since

$$d(S/y)_t/di_t = (1-\alpha_1)\, e^{\alpha 0}\, i_t^{-\alpha 1} > 0, \qquad \text{(IX.9b)}$$

assuming $0 < \alpha_1 < 1$. Hence, there is no revenue maximizing interest rate with the double-log function.

To set out the implications of different money demand specifications for seigniorage revenue and the welfare cost of inflation, we can compare the revenue maximizing nominal interest rate, the maximum value of seigniorage to real income ratio, and the marginal as well as the average welfare cost of revenue from money creation obtained from the Cagan, the double-log, and the nonlinear form of money demand.

As presented on page 29, the nonlinear least squares estimation of Equation (VI.12) yielded the following values for the parameters

$\alpha_0 = -3.138990$

$\alpha_1 = 0.034032$

$\lambda = -0.820392.$

Hence, the revenue maximizing nominal rate of interest obtained from the estimated nonlinear money demand equation is $i_{max} = 0.01624 = 1.62$ % per annum. Assuming 3 % annual real rate of interest, the revenue maximizing inflation rate is found to be $\pi_{max} = -1.38$ % (or deflation rate of 1.38 %). Using these estimated values, we can compute the maximum value of seigniorage to real income ratio from Equation (IX.8b),

$$(S/y)_{max} = e^{\alpha_0 - \left(\frac{1-\alpha_1}{\lambda}\right)} (1/\alpha_1)^{1/\lambda} = 0.23\ \%.$$

To test the sensitivity of estimates of the marginal as well as the average welfare cost of revenue from the creation of money, we will now introduce the quantitative analysis of the welfare cost of inflation[29]. We will follow Bailey's (1956) original study to measure the welfare cost of inflationary finance. The welfare loss to society in his study is measured by integrating the area under the money demand function from the stock of real money balances held at a zero money rate of interest to that held at positive money rate of interest. Hence, the welfare cost of inflation as a share of income is defined as follows,

$$\left(\frac{WC}{y}\right)^{nonlinear} = \int_0^i e^{\alpha_0-\alpha_1\left(\frac{x^\lambda-1}{\lambda}\right)}dx - e^{\alpha_0-\alpha_1\left(\frac{i^\lambda-1}{\lambda}\right)} i, \tag{IX.10}$$

which does not have a closed form solution. However, when λ = 1, Equation (IX.10) reduces to

$$\left(\frac{WC}{y}\right)^{Cagan} = e^{(\alpha_0+\alpha_1)}\left[\frac{1}{\alpha_1}\left(1-e^{-\alpha_1 i}(1+\alpha_1 i)\right)\right] \tag{IX.11}$$

and when λ = 0,

$$\left(\frac{WC}{y}\right)^{log-log} = \left(\frac{\alpha_1}{1-\alpha_1}\right)e^{\alpha_0} i^{1-\alpha_1}. \tag{IX.12}$$

As presented in Equations (IX.11) and (IX.12), we obtain the welfare cost of inflation to income ratio in the special cases of λ = 1 and λ = 0, although there is no closed form solution to the welfare cost of inflation estimated with the nonlinear form of money demand illustrated in Equation (IX.10). Despite the fact that the integral shown in Equation (IX.10) does not have an explicit solution, we can take the derivative of that expression with respect to the nominal rate of interest using Leibniz's rule[30] and find the marginal welfare cost of seigniorage for the most general form of money demand. It takes the following form,

$$MWC^{nonlinear} = \frac{\alpha_1 i^\lambda}{1-\alpha_1 i^\lambda}, \tag{IX.13}$$

since the marginal welfare cost of seigniorage (*dWC/dS*) equals $\frac{dWC/di}{dS/di}$ where

$dWC/di = \alpha_1 i^\lambda e^{\alpha_0-\alpha_1\left(\frac{i^\lambda-1}{\lambda}\right)}$, and $dS/di = (1-\alpha_1 i^\lambda)e^{\alpha_0-\alpha_1\left(\frac{i^\lambda-1}{\lambda}\right)}$.

By using the general form of the marginal welfare cost of revenue from money creation ($MWC^{nonlinear}$) presented in Equation (IX.13), we can derive the marginal welfare costs with the Cagan and the double-log form by substituting λ = 1 and λ = 0, respectively. When λ = 1, Equation (IX.13) reduces to

$$MWC^{Cagan} = \frac{\alpha_1 i}{1-\alpha_1 i}.$$

and when λ = 0,

$$MWC^{log\text{-}log} = \frac{\alpha_1}{1-\alpha_1}.$$

The marginal welfare cost per dollar revenue accruing to the creation of money derived from the nonlinear form of money demand can be evaluated at the sample mean of the seigniorage to real income ratio with the estimated values of α_1 = 0.034032 and λ = -0.820392 and the result

$MWC^{nonlinear}$ *(evaluated at* S_{mean} *= 0.39 %) = 50.38 %,*

is interpreted as 1 dollar increase in seigniorage revenue creates a welfare cost of 50 cents.

The least squares estimation of the Cagan function presented in Equation (VI.12) with λ = 1 yields the following values for the parameters

$\alpha_0^{Cagan} = -8.957849$

$\alpha_1^{Cagan} = 6.676507$.

Hence, the revenue maximizing nominal rate of interest obtained from the estimated Cagan money demand equation is i_{max} = 0.149779 = 14.98 % per annum. Using these estimated parameter values presented above and using Equation (IX.8b), we can calculate the maximum value of seigniorage as a share of income,

$(S/y)_{max} = 0.56\ \%$,

which is much greater than the estimated seigniorage revenue using the nonlinear form. However, the revenues from money creation evaluated at the sample mean of the nominal rate of interest using nonliear form and the Cagan form

$(S/y)^{nonlinear}$ *(evaluated at* $i_{mean} = 5.92\ \%) = 0.38\ \%$,

and $(S/y)^{Cagan}$ *(evaluated at* $i_{mean} = 5.92\ \%) = 0.41\ \%$ are not much different from each other.

The ordinary least squares estimation of the double-log function presented in Equation (VI.12) with $\lambda = 0$ yields the following parameter values

$\alpha_0^{\text{log-log}} = -3.937207$

$\alpha_1^{\text{log-log}} = 0.428218.$

As we discussed before, seigniorage revenue as a fraction of GDP continually increases as the nominal rate of interest rises. Therefore, there is no revenue maximizing interest rate with the double-log function. To compare the maximum revenue from money creation estimated by using the double-log form with the value of maximum seigniorage obtained from the Cagan and the nonlinear form, we compute the seigniorage to income ratio at $i_{max} =$ 0.01624 which is the revenue maximizing interest rate estimated with the nonlinear form. The result,

$(S/y)^{log\text{-}log}$ *(evaluated at* $i_{max} = 1.62\ \%) = 0.18\ \%$,

is very close to the value estimated with the nonlinear form,

$(S/y)^{nonlinear}$ *(evaluated at* $i_{max} = 1.62\ \%) = 0.23\ \%$.

Another way of testing the sensitivity of estimates of seigniorage revenue as a share of income to the specification of money is to compare the estimated values of the revenue evaluated at the sample mean of the interest rate. The revenue using the double-log form and evaluated at the mean value of the interest rate, is found to be

$$(S/y)_{mean}^{log\text{-}log} \text{ (evaluated at } i_{mean} = 5.92\ \%) = 0.39\ \%,$$

which is more or less the same as the value obtained from the nonlinear form,

$$(S/y)_{mean}^{nonlinear} \text{ (evaluated at } i_{mean} = 5.92\ \%) = 0.38\ \%.$$

By comparing the values of the revenues evaluated at i_{mean} and i_{max}, we can say that the estimates from the double-log form are much closer to the values obtained from the nonlinear form than the estimates from the Cagan form. The differences between the Cagan and the double-form become more clear when we look at the values of the welfare costs of inflation as a share of income derived from those two different money demand specifications. The welfare cost of inflation estimated with the double-log form evaluated at the sample mean of seigniorage as a fraction of GDP is

$$(WC/y)^{log\text{-}log} \text{ (evaluated at } S_{mean} = 0.39\ \%) = 0.29\ \%,$$

which is much bigger than the one estimated with the Cagan form,

$$(WC/y)^{Cagan} \text{ (evaluated at } S_{mean} = 0.39\ \%) = 0.08\ \%.$$

In addition to the differences between the estimates of the welfare costs of inflation, we also realize major discrepancies in estimates of the marginal and the average welfare costs of seigniorage. Even though the marginal welfare cost estimated with the double-log form and evaluated at the mean value of the seigniorage to real income ratio

$$MWC^{log\text{-}log} \text{ (evaluated at } S_{mean} = 0.39\ \%) = 74.89\ \%,$$

is not enormously different from the one estimated with the Cagan form,

$$MWC^{Cagan} \text{ (evaluated at } S_{mean} = 0.39\ \%) = 56.54\ \%,$$

there is a substantial difference in estimates of the average welfare costs of seigniorage. The estimated value of the average welfare cost with the double-log form is

$$AWC^{log\text{-}log} \textit{ (evaluated at } S_{mean} = 0.39\ \%) = 74.89\ \%,$$

whereas the one estimated with the Cagan form

$$AWC^{Cagan} \textit{ (evaluated at } S_{mean} = 0.39\ \%) = 20.45\ \%.$$

Therefore, the additional gain moving from zero inflation to the steady deflation of around 3 percent that would attain the Friedman (1969) optimum is much greater with the double-log function than with the constant semi-elasticity Cagan form. In other words, the double-log function assigns much greater benefit to the reduction of moderate inflations, and substantial gains from reducing the inflation rate to the steady deflation of 3 percent that would reduce the nominal rate of interest to zero. According to our estimates of the welfare costs of inflation, moving from the inflation rate of 2.92 % (from the nominal interest rate of 5.92 %) to the steady deflation of 3 % (to the zero nominal rate of interest), the double-log function assigns a substantial gain of 0.29 % of GDP whereas the Cagan form assigns 0.09 %. Thus, there is a huge difference between the double-log and the Cagan form in estimated values of the welfare gain of reducing inflation rate to the steady state deflation of around 3 % per annum. The reader may wish to consult Table 10 to see the sensitivity of estimates of the revenue maximizing nominal rate of interest, seigniorage revenue and the welfare cost of inflation as a share of real income, and the marginal as well as the average welfare cost of revenue from money creation.

X. SEIGNIORAGE REVENUE AND THE WELFARE COST OF INFLATION IN A CURRENCY-DEPOSIT MODEL

The role of reserve ratios in seigniorage has been discussed in some of the recent literature. It has been widely recognized that required reserve ratios have two conflicting effects on seigniorage collection: first, higher levels have a direct impact on the demand for the reserve component of the monetary base and therefore tend to increase seigniorage collection; second, higher required reserve ratios tend to widen the spread between bank lending rates and rates offered to depositors. Depending on the elasticities of the components of money demand, higher reserve ratios thus indirectly reduce the demand for reserves.

An empirical role for reserve ratios in determining seigniorage collection seems obvious for the U.S. economy. The statistical results presented below offer the t-statistics on the reserve ratio (μ_t) and the reserve ratio times the nominal rate of interest ($\mu_t i_t$) after each was added in a separate experiment in an estimated "base" equation in which the Phelps-Auernheimer type of seigniorage was the dependent variable:

$$i_t\,(MB/GDP)_t = b_0 + b_1\, i_t + b_2\, i_{t-1}\,(MB/GDP)_{t-1} + e_t, \qquad \text{(X.1)}$$

where i_t is the short-term interest rate (3-month t-bill rate), *MB* is the monetary base (reserves plus currency), *GDP* represents the nominal income level, and e_t is the error term. The OLS estimation of Equation (X.1) yields

$$i_t\,(MB/GDP)_t = \underset{(5.1702)}{0.0005} + \underset{(9.5600)}{0.0251}\, i_t + \underset{(9.1460)}{0.4807}\, i_{t-1}\,(MB/GDP)_{t-1}$$

$\bar{R}^2 = 0.9176$ F = 819.85 Log-likelihood = 946.47

The addition of the reserve ratio and the reserve ratio times the interest rate to the original equation leads to positive and large t-statistics on μ_t and $\mu_t i_t$, suggesting μ is a significant determinant of the U.S. revenue from money creation. As presented below, the statistical results also indicate that there are dramatically larger Log-likelihood and F-ratios as well as the adjusted-R^2 statistics upon inclusion of μ_t or $\mu_t i_t$. The inclusion of μ_t and $\mu_t i_t$ to the generic equation (X.1), respectively, yields the following results

$$i_t\,(MB/GDP)_t = \underset{(-2.1915)}{-0.0003} + \underset{(14.533)}{0.0376\, i_t} + \underset{(\mathbf{8.6041})}{0.0145\, \mu_t} + \underset{(7.1785)}{0.3318\; i_{t-1}\,(MB/GDP)_{t-1}}$$

$\bar{R}^2 = 0.9452$ F = 846.52 Log-likelihood = 977.17

$$i_t\,(MB/GDP)_t = \underset{(4.6881)}{0.0003} + \underset{(18.817)}{0.0302\, i_t} + \underset{(\mathbf{16.073})}{0.4671\, \mu_t\, i_t} + \underset{(5.3481)}{0.1939\; i_{t-1}\,(MB/GDP)_{t-1}}$$

$\bar{R}^2 = 0.9703$ F = 1602.71 Log-likelihood = 1022.51

As to the empirical value of including μ in our model of steady-state seigniorage, direct inference is of course impossible since steady-state seigniorage is unobservable. We have shown above that reserve ratios contribute substantially in improving the fit of an unconstrained forecast of short-run seigniorage revenue. Table 11 compares short-run forecasts of seigniorage from the Monetary Base (MB), M1, and our currency-deposit model. According to the statistical results presented in Table 11, in terms of the standard errors of regression from the short-run seigniorage estimates (or the standard deviation of the forecast error), the performance of our currency-deposit model is always superior. The reader may wish to consult Table 11 to see the performance of different money demand specifications for the short-run seigniorage estimates by comparing the means and the standard deviations of the forecast errors resulted from the Monetary Base, M1, and the currency-deposit model .

Thus, it seems fair to say that the reserve ratio is an important part of the story, at least in the short-run (which is all we have in order to estimate implied steady-states). In the following sections, we first develop and illustrate a currency-deposit model in which we separate money into its currency and deposit components and apply a required reserve ratio to the deposit component. After discussing the microfoundations of our currency-deposit model, we provide some insight into the implications of the analysis for the optimal rate of inflation and the required reserve ratio. Finally, we study the implications of different money demand specifications for steady state seigniorage and the social welfare costs of inflationary finance in our currency-deposit model.

The Currency-Deposit Model: We will assume that the economy is populated by infinitely lived consumers or dynasties who derive utility from the only consumption good, c, and from the real money stock which is composed of currency, m, and deposits, d, (total deposits, i.e., M2 minus currency). The households in the economy provide labor services in exchange for wages, receive interest income on assets, purchase goods for consumption, and save by accumulating additional monetary and non-monetary assets. We will further assume that the utility function is additively separable, that is

$$U = \int_0^\infty e^{-\rho t} N(t)(u[m(t),d(t)] + v[c(t)])dt\,, \tag{X.2}$$

where ρ is the subjective rate of time preference (or the personal discount rate), $N(t)$ is the total amount of people alive at time t and $N(t)$ will be assumed to grow at an exogenous rate n. By normalizing initial population to 1, we have $N(t) = e^{nt}$. It is also assumed that the rate of time preference is greater than the growth rate of population ($\rho > n$), which implies that the overall utility is bounded if m, d, and c are all constant over time. The instantaneous per capita utility, $u()$, is a function of the per capita real currency, $m(t)$, and deposits, $d(t)$, and the per capita utility, $v()$, is a function of the real consumption per capita, $c(t)$. Individuals will be assumed to maximize utility subject to the budget constraint

$$\frac{B}{N} + \frac{M}{PN} + \frac{D}{PN} = r\frac{B}{N} + i_d \frac{D}{PN} + \frac{H}{N} + \frac{W}{N} - c, \tag{X.3}$$

where B is real non-monetary assets such as capital and bonds, M is nominal currency, D is nominal deposits, P is the price level, r is the real interest rate on real assets B, i_d is the deposit rate on deposit holdings, W is the real wage rate and H is the net real lump-sum transfer from the government. Equation (X.3) says that per capita savings (nonconsumed resources) are equal to per capita investment plus money (currency and deposits) accumulation. If we define lower case variables as the real per capita versions of their capital letter counterparts ($b=B/N$, $m=M/PN$, $d=D/PN$, $h=H/N$, $w=W/N$) and we denote the per capita real assets by a ($a = b+m+d$), we can rewrite the household's budget constraint as

$$a = ra + h + w - c - na - im - (i - i_d)d \,, \tag{X.4}$$

where i is the nominal rate of interest on bonds, $i = r + \pi$, where π is the expected and the actual inflation rate. Equation (X.4) says that assets (monetary and nonmonetary) per person rise with per capita income, $ra+w$, and the net real lump-sum transfer (assuming $h > 0$) from the government, fall with per capita consumption, c, and the opportunity cost of holding per capita real currency, im, and deposits, $(i - i_d)d$, as well as the expansion of population, na.

We can set up the present-value Hamiltonian presented in Equation (X.5) and get the following set of first-order conditions:

$$J = e^{-(\rho-n)t}[u(m_t, d_t) + v(c_t)] + \Omega_t[(r_t - n)a_t + h_t + w_t - c_t - i_t m_t - (i_t - i_d(t))d_t], \quad \text{(X.5)}$$

$$e^{-(\rho-n)t} v_c = \Omega_t, \quad \text{(X.6a)}$$

$$e^{-(\rho-n)t} u_m = \Omega_t i_t, \quad \text{(X.6b)}$$

$$e^{-(\rho-n)t} u_d = \Omega_t (i_t - i_d(t)), \quad \text{(X.6c)}$$

$$\Omega_t (r_t - n) = -\dot{\Omega}_t, \quad \text{(X.6d)}$$

$$lim_{t\to\infty} \Omega_t a_t = 0 \quad \text{(X.6e)}$$

In order to obtain an explicit form of money demand, the next step is to define a specific form of utility function. The conventional, homothetic functions such as the Cobb-Douglas and CES imply double-log money demand functions. If we assume the following additively separable utility function in per capita consumption, currency and deposits,

$$u[m(t), d(t)] + v[c(t)] = \left[\left(\frac{\varphi}{1-\alpha}\right)(m_t^{1-\alpha} - 1) + \left(\frac{\phi}{1-\beta}\right)(d_t^{1-\beta} - 1) + \left(\frac{1}{1-\gamma}\right)(c_t^{1-\gamma} - 1)\right], \quad \text{(X.7)}$$

where α, β, γ, ϕ, and φ are all positive constants so that the utility function is a strictly concave with continuous first and second derivatives then the first order conditions (X.6a)-(X.6b)-(X.6c) presented above yield the following constant elasticity currency and deposit demand functions, respectively:

$$m_t = \varphi^{1\;\alpha}\, i_t^{-1\;\alpha}\, c_t^{\gamma\alpha}, \quad \text{(X.8a)}$$

$$d_t = \phi^{1/\beta} (i_t - i_d(t))^{-1/\beta} c_t^{\gamma\beta}, \tag{X.8b}$$

However, the Cagan form can be provided a microfoundation in this intertemporal optimizing framework with an unconventional utility function since its derivation in either "the shopping time" or "the money in the utility function" approach is somewhat ad hoc. Therefore, we have to impose a specific form on the utility function in order to arrive at the constant semi-elasticity Cagan-type demand for money. This utility function takes the following form:

$$u[m(t), d(t)] + v[c(t)] = \frac{m_t(1+\alpha_0 - \ln m_t)}{\alpha_1} + \frac{d_t(1+\beta_0 - \ln d_t)}{\beta_1} + v[c(t)], \tag{X.9}$$

where α_0, α_1, β_0, and β_1 are all positive constants. Using (X.9) and the first order conditions, we obtain the Cagan-form asset demand functions:

$$\ln m_t = \alpha_0 - \alpha_1 v_c i_t, \tag{X.10a}$$

$$\ln d_t = \beta_0 - \beta_1 v_c (i_t - i_d(t)), \tag{X.10b}$$

where we assume that the marginal utility of consumption, v_c, is a positive constant.

Steady-State Seigniorage and the Welfare Cost Analysis: In the currency-deposit model developed in the previous section where private agents hold cash balances and bank deposits, with the former asset bearing a zero interest rate and banks are subject to a fractional reserve requirement on total deposits, we have presented the microfoundations of the two special specifications of the money demand function; the double-log and the Cagan, respectively. In this section, to illustrate the implications of different money demand specifications for steady state seigniorage and the welfare cost of inflation in our currency-deposit model, we first introduce the most general (nonlinear) form of money demand then compare the results with the Cagan and the double-log specification. With the typical side restriction that the income elasticity equal unity, asset demand functions for currency m_t and bank deposits d_t can be written in the nonlinear form as

$$\left(\frac{m}{y}\right)_t^{nonlinear} = e^{\alpha_0 - \alpha_1\left(\frac{i_t^{\lambda_1} - 1}{\lambda_1}\right)}, \qquad \text{(X.11a)}$$

$$\left(\frac{d}{y}\right)_t^{nonlinear} = e^{\beta_0 - \beta_1\left(\frac{(i_t - i_d(t))^{\lambda_2} - 1}{\lambda_2}\right)}, \qquad \text{(X.11b)}$$

where i_t denotes the nominal lending rate and $i_d\,(t)$ is the deposit rate. If banks face no operating costs, the zero-profit condition under perfect competition yields

$$i_d\,(t) = i_t\,(1 - \mu), \qquad 0 < \mu < 1, \qquad \text{(X.12)}$$

where we can rewrite the opportunity cost of holding bank deposits as $i_t - i_d\,(t) = \mu_t\, i_t$ where μ_t denotes the required reserve ratio. The policymaker's objective is to maximize seigniorage revenue as a share of real income, which is given by

$$\left(\frac{S}{y}\right)_t^{nonlinear} = i_t\left[\left(\frac{m}{y}\right)_t^{nonlinear} + \mu_t\left(\frac{d}{y}\right)_t^{nonlinear}\right], \qquad \text{(X.13)}$$

with respect to the nominal interest rate, i_t, and the reserve ratio, μ_t. Solving this maximization problem yields

$$\frac{\partial (S/y)_t^{nonlinear}}{\partial \mu_t} = i_t e^{\beta_0 - \beta_1\left(\frac{(\mu_t i_t)^{\lambda_2} - 1}{\lambda_2}\right)}[1 - \beta_1(\mu_t i_t)^{\lambda_2}] = 0, \qquad \text{(X.14a)}$$

$$\frac{\partial (S/y)_t^{nonlinear}}{\partial i_t} = e^{\alpha_0 - \alpha_1\left(\frac{i_t^{\lambda_1} - 1}{\lambda_1}\right)}(1 - \alpha_1 i_t^{\lambda_1}) + \mu_t e^{\beta_0 - \beta_1\left(\frac{(\mu_t i_t)^{\lambda_2} - 1}{\lambda_2}\right)}[1 - \beta_1(\mu_t i_t)^{\lambda_2}] = 0, \qquad \text{(X.14b)}$$

Equation (X.14b) indicates that when the reserve ratio is zero, the revenue-maximizing interest rate is equal to $i^{*} = (1/\alpha_1)^{1/\lambda_1}$. If both instruments are used, Equation (X.13) has a maximum at

$$\left(\frac{S}{y}\right)_{\max}^{nonlinear} = \left(\frac{1}{\alpha_1}\right)^{1/\lambda_1} e^{\alpha_0 - \left(\frac{1-\alpha_1}{\lambda_1}\right)} + \left(\frac{1}{\beta_1}\right)^{1/\lambda_2} e^{\beta_0 - \left(\frac{1-\beta_1}{\lambda_2}\right)}, \qquad \text{(X.15)}$$

where the revenue-maximizing interest rate and the reserve ratio are

$$i_{\max} = (1/\alpha_1)^{1/\lambda_1}, \qquad \text{(X.16a)}$$

$$\mu_{\max} = \frac{(1/\beta_1)^{1/\lambda_2}}{(1/\alpha_1)^{1/\lambda_1}}, \qquad \text{(X.16b)}$$

provided that $\mu_{\max} \leq 1$. If $\mu_{\max} > 1$, the zero-profit condition implies $\mu_{\max} = 1$, $i_{\max}$ will be higher than $(1/\alpha_1)^{1/\lambda_1}$, and of course $\left(\frac{S}{y}\right)_{\max}^{nonlinear}$ is less than that given in Equation (X.15).

With the real interest rate r, the revenue-maximizing inflation rate can be obtained by using the fact that $i = \pi + r$. Hence,

$$\pi_{\max} = (1/\alpha_1)^{1/\lambda_1} - r. \qquad \text{(X.16c)}$$

If μ is exogenous then there is no analytical solution for $i_{\max}$ and $\left(\frac{S}{y}\right)_{\max}^{nonlinear}$, although these expressions can be evaluated numerically (as we do below). $\left(\frac{S}{y}\right)_{\max}^{nonlinear}$ will be less than that displayed in (X.15) and $i_{\max}$ will be higher or lower than $(1/\alpha_1)^{1/\lambda_1}$ (higher if $\mu < \mu_{\max}$, lower if $\mu > \mu_{\max}$).

When we consider the special case of $\lambda_1 = \lambda_2 = 1$, the nonlinear currency and deposit demand equations reduces to simple constant semi-elasticity (Cagan) form,

$$\left(\frac{m}{y}\right)_t^{Cagan} = e^{(\alpha_0+\alpha_1)-\alpha_1 i_t}, \tag{X.17a}$$

$$\left(\frac{d}{y}\right)_t^{Cagan} = e^{(\beta_0+\beta_1)-\beta_1 \mu_t i_t}. \tag{X.17b}$$

In this case, Equations (X.16a) and (X.16b) imply that the revenue-maximizing interest rate and the reserve ratio reduces to

$$i_{max} = 1/\alpha_1, \tag{X.18a}$$

$$\mu_{max} = \alpha_1/\beta_1. \tag{X.18b}$$

Bailey (1956), Calvo and Fernandez (1983), and Marty (1988,1994) each discuss the special case of the model ($\lambda_1 = \lambda_2 = 1$) presented above where currency and deposits are perfect substitutes so that when $i_d > 0$ no currency is held. In that case $(S/y)_t = (\mu i)\, e^{\beta_0 - \beta_1(\mu i)}$. The maximum seigniorage to real income ratio is $(S/y)_{max} = (1/\beta_1)\, e^{\beta_0 - 1}$ at nominal interest rate equal to $i_{max} = (1/\mu\beta_1)$. This $(S/y)_{max}$ can be achieved for any combination of i and μ along $i\mu = 1/\beta_1$. In these studies μ is treated as exogenous (but in the range $0 < \mu \leq 1$). Marty (1994) has two cases, one of which is deposit only model where currency and deposits are perfect substitutes and another is currency-deposit model with fixed currency-deposit ratio.

When $\lambda_1 = \lambda_2 = 0$, we get the double-log currency and deposit demand equations with a constant elasticities of $-\alpha_1$ and $-\beta_1$, respectively:

$$\left(\frac{m}{y}\right)_t^{\log-\log} = e^{\alpha_0} i_t^{-\alpha_1}, \tag{X.19a}$$

$$\left(\frac{d}{y}\right)_t^{\log-\log} = e^{\beta_0}(\mu_t i_t)^{-\beta_1}. \tag{X.19b}$$

As we discussed in Section IX, the double-log function does not generate a Laffer curve. In other words, the revenue from the creation of money continually increases as the nominal interest rate rises assuming that the interest rate elasticities of demand for currency and deposits are less than unity. To see this, note that seigniorage as a fraction of GDP is

$$\left(\frac{S}{y}\right)_t^{\log-\log} = e^{\alpha_0} i_t^{1-\alpha_1} + e^{\beta_0}(\mu_t i_t)^{1-\beta_1}, \tag{X.20}$$

where

$$\frac{\partial (S/y)_t^{\log-\log}}{\partial i_t} = (1-\alpha_1)e^{\alpha_0} i_t^{-\alpha_1} + (1-\beta_1)e^{\beta_0}\mu_t^{1-\beta_1} i_t^{-\beta_1} > 0, \tag{X.21a}$$

$$\frac{\partial (S/y)_t^{\log-\log}}{\partial \mu_t} = (1-\beta_1)e^{\beta_0} i_t^{1-\beta_1}\mu_t^{-\beta_1} > 0, \tag{X.21b}$$

assuming $0 < \alpha_1 < 1$ and $0 < \beta_1 < 0$. Hence, there is no revenue-maximizing interest rate or reserve ratio with the double-log function. To explain the implications of different money demand specifications for seigniorage revenue, we run through the empirical results for the U.S. economy in this section.

The nonlinear least squares estimation of real currency and deposit demand equations presented in (X.11a) and (X.11b) yield the following values for the parameters

$\alpha_0 = -3.311961$ $\beta_0 = -13.200982$

$\alpha_1 = 0.004715$ $\beta_1 = 10.288980$

$\lambda_1 = -1.351950$ $\lambda_2 = 0.807364.$

Hence, the maximum seigniorage revenue turns out to be 1.17 % of GDP. The $\{i_{max}, \mu_{max}\}$ set turns to be 0.243 and 1, respectively (these values are evaluated when both instruments are used). The means ($\bar{\pi}, \bar{\mu}$) for the period 1957:I - 1994:I are 0.0592 and 0.0431, respectively, while the sample mean of seigniorage as a fraction of GDP ($\bar{S}$) is 0.0039. The maximum seigniorage calculated at the sample mean of μ ($\bar{\mu}$) rather than μ_{max} and the interest rate required to attain it turn to be the same 1.17 % and somewhat lower 0.2006, respectively, than in the case where μ is freely chosen.

Table 12 presents these measures for the nonlinear form and shows alternative estimates of i_{max}, μ_{max}, and $(S/y)_{max}$ based on alternative specifications; the constant semi-elasticity (Cagan) and the double-log form. The least squares estimation of the Cagan form of currency and deposit demand equations presented in (X.17a) and (X.17b) offered the following long-run coefficients

$\alpha_0 + \alpha_1 = -2.860579$ $\beta_0 + \beta_1 = -0.464847$

$\alpha_1 = 3.762393$ $\beta_1 = 37.413776.$

According to these estimated values, the maximum revenue from the creation of money is 1.18 % (which is more or less the same as the one obtained from the nonlinear specification) where $i_{max} = 0.2658$ and $\mu_{max} = 0.1006$ when both instruments are used. The maximum seigniorage collection as a share of real income calculated at $\bar{\mu}$ rather than μ_{max} and the rate of interest corresponding to this maximum revenue turn out to be somewhat lower 1.08 % and somewhat higher 0.3749, respectively, than in the case where μ is flexible. As we mentioned before, since there is no revenue-maximizing interest rate and the reserve ratio with the double-log function we cannot display any statistical results with regard to the maximum seigniorage revenue, i_{max} or μ_{max} resulted from the double-log form although we will compare the welfare cost estimates obtained from the double-log specification with those calculated from the nonlinear and the Cagan form.

To test the sensitivity of estimates of the marginal and the average welfare cost of seigniorage to the specification of money demand, we will now introduce the quantitative analysis of the welfare cost of inflation evaluated with different forms of asset demand functions. Deadweight social costs of raising revenue via seigniorage collection can be interpreted as the lost areas under currency and deposit demand curves owing to inflation's effect on their respective opportunity costs. This is the traditional "shoe-leather" cost. In the nonlinear specification we are using, the loss (WC) is measured as the deadweight loss from positive i in reducing currency demand,

$$\left[\int_0^i e^{\alpha_0-\alpha_1\left(\frac{x^{\lambda_1}-1}{\lambda_1}\right)}dx - e^{\alpha_0-\alpha_1\left(\frac{i^{\lambda_1}-1}{\lambda_1}\right)}i\right],$$

plus the deadweight loss from positive μi in reducing deposit demand,

$$\left[\int_0^{\mu i} e^{\beta_0-\beta_1\left(\frac{x^{\lambda_2}-1}{\lambda_2}\right)}dx - e^{\beta_0-\beta_1\left(\frac{(\mu i)^{\lambda_2}-1}{\lambda_2}\right)}(\mu i)\right].$$

Hence, in the nonlinear specification of currency and deposit demand, the welfare cost of inflation as a fraction of GDP is defined as follows,

$$\left(\frac{WC}{y}\right)^{nonlinear} = \int_0^i e^{\alpha_0-\alpha_1\left(\frac{x^{\lambda_1}-1}{\lambda_1}\right)}dx - e^{\alpha_0-\alpha_1\left(\frac{i^{\lambda_1}-1}{\lambda_1}\right)}i + \int_0^{\mu i} e^{\beta_0-\beta_1\left(\frac{x^{\lambda_2}-1}{\lambda_2}\right)}dx - e^{\beta_0-\beta_1\left(\frac{(\mu i)^{\lambda_2}-1}{\lambda_2}\right)}(\mu i), \quad \text{(X.22)}$$

which does not have a closed form solution. However, when $\lambda_1 = \lambda_2 = 1$ where currency and deposit demand equations take the constant semi-elasticity (Cagan) form, Equation (X.22) reduces to

$$\left(\frac{WC}{y}\right)^{Cagan} = e^{(\alpha_0+\alpha_1)}\left[\frac{1}{\alpha_1}\left(1-e^{-\alpha_1 i}(1+\alpha_1 i)\right)\right]+e^{(\beta_0+\beta_1)}\left[\frac{1}{\beta_1}\left(1-e^{-\beta_1 \mu i}(1+\beta_1 \mu i)\right)\right], \quad \text{(X.23a)}$$

and when $\lambda_1 = \lambda_2 = 0$ where currency and deposit demand equations take the constant elasticity (double-log) form, Equation (X.22) reduces to

$$\left(\frac{WC}{y}\right)^{\log-\log} = \left(\frac{\alpha_1}{1-\alpha_1}\right)e^{\alpha_0} i^{1-\alpha_1} + \left(\frac{\beta_1}{1-\beta_1}\right)e^{\beta_0}(\mu i)^{1-\beta_1}. \quad \text{(X.23b)}$$

Although the integral shown in (X.22) does not have an explicit solution we can take the derivative of that expression with respect to the nominal rate of interest using Leibniz's rule and find the marginal welfare cost of seigniorage derived from the nonlinear from of asset demand functions. As presented below, expressions for marginal social burdens for the nonlinear form turn out to be very complicated:

$$MWC_{\bar{\mu}}^{nonlinear} = \frac{\alpha_1 i^{\lambda_1} e^{\alpha_0-\alpha_1\left(\frac{i^{\lambda_1}-1}{\lambda_1}\right)} + \mu\beta_1(\mu i)^{\lambda_2} e^{\beta_0-\beta_1\left(\frac{(\mu i)^{\lambda_2}-1}{\lambda_2}\right)}}{(1-\alpha_1 i^{\lambda_1})e^{\alpha_0-\alpha_1\left(\frac{i^{\lambda_1}-1}{\lambda_1}\right)} + \mu(1-\beta_1(\mu i)^{\lambda_2})e^{\beta_0-\beta_1\left(\frac{(\mu i)^{\lambda_2}-1}{\lambda_2}\right)}}, \quad \text{(X.24)}$$

since the marginal welfare cost of seigniorage $\left(\frac{\partial WC}{\partial S}\right)_{\bar{\mu}}^{nonlinear}$ equals $\frac{\partial WC/\partial i}{\partial S/\partial i}$ where

$$\left(\frac{\partial WC}{\partial i}\right)_{\bar{\mu}}^{nonlinear} = \alpha_1 i^{\lambda_1} e^{\alpha_0-\alpha_1\left(\frac{i^{\lambda_1}-1}{\lambda_1}\right)} + \mu\beta_1(\mu i)^{\lambda_2} e^{\beta_0-\beta_1\left(\frac{(\mu i)^{\lambda_2}-1}{\lambda_2}\right)}, \quad \text{(X.25a)}$$

$$\left(\frac{\partial S}{\partial i}\right)_{\bar{\mu}}^{nonlinear} = (1-\alpha_1 i^{\lambda_1})e^{\alpha_0-\alpha_1\left(\frac{i^{\lambda_1}-1}{\lambda_1}\right)} + \mu(1-\beta_1(\mu i)^{\lambda_2})e^{\beta_0-\beta_1\left(\frac{(\mu i)^{\lambda_2}-1}{\lambda_2}\right)}. \quad \text{(X.25b)}$$

and μ is assumed to be exogenous. By using the general form of marginal social burdens presented in (X.24), we can derive the marginal welfare costs with the Cagan and the double-log form by substituting $\lambda_1 = \lambda_2 = 1$ and $\lambda_1 = \lambda_2 = 0$, respectively. When $\lambda_1 = \lambda_2 = 1$, Equation (X.24) reduces to

$$MWC_{\bar{\mu}}^{Cagan} = \frac{\alpha_1 i e^{(\alpha_0+\alpha_1)-\alpha_1 i} + \beta_1 i \mu^2 e^{(\beta_0+\beta_1)-\beta_1 \mu i}}{(1-\alpha_1 i)e^{(\alpha_0+\alpha_1)-\alpha_1 i} + (1-\beta_1 i \mu)\mu e^{(\beta_0+\beta_1)-\beta_1 \mu i}}, \quad \text{(X.26a)}$$

and when $\lambda_1 = \lambda_2 = 0$,

$$MWC_{\bar{\mu}}^{\text{log-log}} = \frac{\alpha_1 e^{\alpha_0} i^{-\alpha_1} + \beta_1 e^{\beta_0} \mu^{1-\beta_1} i^{-\beta_1}}{(1-\alpha_1)e^{\alpha_0} i^{-\alpha_1} + (1-\beta_1)e^{\beta_0} \mu^{1-\beta_1} i^{-\beta_1}}, \quad \text{(X.26b)}$$

where μ is assumed to be exogenous.

In the case where currency and deposit demand equations take the Cagan form and μ is an instrument chosen only for revenue purposes, it can be shown that the optimal (i.e., welfare cost minimizing) value of μ for given seigniorage is a constant α_l/β_l - the same as the seigniorage-maximizing value of μ - provided that $\alpha_l/\beta_l < 1$ ($\mu_{optimal} = 1$ otherwise). In other words, a monetary authority that attempts to minimize the welfare cost of inflationary finance will choose the unique reserve ratio $\mu_{optimal} = \alpha_1/\beta_1$ and vary the inflation rate (or the nominal rate of interest) as revenue requirements change. Let us consider the problem of finding a cost-minimizing reserve ratio for a given level of seigniorage. We analyze this constrained minimization problem using the method of Lagrange multipliers. Begin by writing the Lagrangian

$$L(i,\ \mu,\ \lambda) = WC + \lambda(\bar{S} - e^{(\alpha_0+\alpha_1)-\alpha_1 i}\, i - e^{(\beta_0+\beta_1)-\beta_1 \mu i}(\mu i)) \quad \text{(X.27)}$$

and differentiate it with respect to each of the choice variables, i and μ, and the Lagrange multiplier, λ. The first-order conditions are

$$\left(\frac{\partial L}{\partial i}\right) = \frac{\partial WC}{\partial i} + \lambda(-e^{(\alpha_0+\alpha_1)-\alpha_1 i}(1-\alpha_1 i) - \mu(1-\beta_1 \mu i)e^{(\beta_0+\beta_1)-\beta_1 \mu i}) = 0 \quad \text{(X.28a)}$$

$$\left(\frac{\partial L}{\partial \mu}\right) = \frac{\partial WC}{\partial \mu} + \lambda(-i(1-\beta_1 \mu i)e^{(\beta_0+\beta_1)-\beta_1 \mu i}) = 0 \quad \text{(X.28b)}$$

$$\left(\frac{\partial L}{\partial \lambda}\right) = \bar{S} - e^{(\alpha_0+\alpha_1)-\alpha_1 i}\, i - e^{(\beta_0+\beta_1)-\beta_1 \mu i}(\mu i) = 0 \quad \text{(X.28c)}$$

where

$$\left(\frac{\partial WC}{\partial i}\right) = \alpha_1 i e^{(\alpha_0+\alpha_1)-\alpha_1 i} + \beta_1 i \mu^2 e^{(\beta_0+\beta_1)-\beta_1 \mu i}, \quad \text{(X.28d)}$$

$$\left(\frac{\partial WC}{\partial \mu}\right) = \beta_1 i^2 \mu e^{(\beta_0+\beta_1)-\beta_1 \mu i}. \quad \text{(X.28e)}$$

Equations (X.28b) and (X.28e) yields

$$\lambda = \frac{\beta_1 \mu i}{1-\beta_1 \mu i},$$

and substitution of λ into Equation (X.28a) gives the welfare cost minimizing reserve ratio, $\mu_{optimal} = \alpha_1/\beta_1$. From Equation (X.28a) λ turns to be the marginal welfare cost of seigniorage in the μ-exogenous case as presented in Equation (X.26a). Therefore, when μ is freely chosen and the monetary authority attempts to optimize (i.e., substitute $\mu_{optimal} = \alpha_1/\beta_1$ into λ) the expression for marginal social burden can be written simply in the form

$$MWC_{\min}^{Cagan} = \frac{\alpha_1 i}{1-\alpha_1 i}.$$

When $\alpha_I / \beta_I \leq 1$ and the government optimizes with respect to both i and μ it turns out that the average welfare cost of seigniorage (AWC^{Cagan}) is a "natural" constant $e-2 = 0.7183$...This is to say, whatever the values of the α's and β's, at the highest average social cost burden will be about 72 % of the revenue collected provided μ is an instrument; if μ is exogenous, AWC^{Cagan}_{max} will be higher. As presented in Table 12, with the Cagan form, the average deadweight loss at maximum seigniorage in the μ -exogenous case is 77.31 %. The average welfare cost evaluated at the sample mean of seigniorage (as a fraction of GDP) when μ is freely chosen (AWC^{Cagan}_{min}) is 0.0723, slightly higher when evaluated at the sample mean of (S/y) with μ constrained to equal its mean ($AWC^{Cagan}_{\bar{\mu}} = 0.0863$). The marginal welfare costs evaluated at the sample mean of (S/y) range from the lower hypothetical $MWC^{Cagan}_{min} = 0.1602$ (evaluated at $\mu_{optimal}$) to the somewhat higher value when evaluated at $\bar{\mu}$: $MWC^{Cagan}_{\bar{\mu}} = 0.1937$. Thus, according to these welfare cost measures estimated with the Cagan form of asset demand functions, the "optimal" average and marginal welfare costs of revenue from money creation are not far from the value obtained when μ is constrained to its sample mean ($\bar{\mu}$).

Empirically, we can say that there is almost no difference in marginal welfare costs estimated from the nonlinear and the Cagan form and the differences in marginal costs obtained from the nonlinear and the double-log form do not turn out to be major. The marginal welfare cost per dollar revenue accruing to the creation of money derived from the double-log form of currency and deposit demand equations can be evaluated at the sample mean of (S/y) and $\bar{\mu}$ with the estimated parameter values of

α_0 = -3.976671 β_0 = -0.814444

α_1 = 0.303992 β_1 = 0.043141,

and the result $MWC_{\bar{\mu}}^{log-log}$ = 0.2686 is slightly higher than the marginal social burden measured by the nonlinear form $MWC_{\bar{\mu}}^{nonlinear}$ = 0.2101.

Since the nonlinear forms of asset demand functions do not yield an explicit solution for the welfare cost of inflation we cannot find the average welfare cost of seigniorage collection with this form. However, the welfare costs of inflationary finance as well as the average social burdens of raising revenue via the inflation tax obtained from the Cagan and the double-log form turn out to be substantially different from each other. With the double-log form of currency and deposit demand equations, the welfare cost of inflation (as a fraction of GDP) evaluated at the sample mean of (S/y) when μ is constrained to its sample mean is $\left(\frac{WC}{y}\right)_{\bar{\mu}}^{log-log}$ = 0.12 % which is four times as high as the welfare costs measured by the Cagan form and evaluated also at the sample mean of (S/y) and $\bar{\mu}$ since $\left(\frac{WC}{y}\right)_{\bar{\mu}}^{Cagan}$ = 0.03 %. As we mentioned in the previous section, the additional gain moving from zero inflation to the steady state deflation of around 3 percent that would attain the Friedman (1969) optimum is therefore much greater with the double-log function than with the constant semi-elasticity Cagan form. This substantial difference in welfare cost measures also leads to a major difference in the average social costs of seigniorage. The average welfare cost calculated at the sample mean of (S/y) and also at $\bar{\mu}$ is $AWC_{\bar{\mu}}^{log-log} = 0.2982$ which is considerably higher than that estimated with the Cagan form $AWC_{\bar{\mu}}^{Cagan} = 0.0863$.

These measures of marginal and average social costs appear inside the plausible range for a country having an efficient explicit tax system. Ballard, et al. (1985) estimate marginal excess burdens of U.S. taxes with a large-scale numerical general equilibrium model and various assumptions about labor and savings elasticities. They find excess burdens of various taxes at existing tax levels ranging from 17 % to 56 % per dollar of extra revenue. These estimates do not include collection costs. According to our estimates, the marginal welfare costs evaluated at the sample mean of seigniorage revenue (as a fraction of GDP) range from 19.32 % (with the Cagan form) to 26.86 % (with the double-log form) which are slightly greater than the lower limit (17 %) and considerably less than the upper limit (56 %) of their range.

XI. A SQUARE ROOT FORMULA FOR THE WELFARE COST OF INFLATION

Bailey's (1956) original study says that the welfare cost of inflationary finance increase with the square of the interest rate, assigning large costs to very high rates of inflation but trivial costs to moderate inflation's. According to this approach, the objective of a zero inflation rate has negligible benefit, and an additional gain from moving zero inflation to the steady deflation of 3 percent that would attain the Friedman optimum is assigned an even smaller value.

In this section our objective is to argue that the traditional quadratic approximation to the welfare cost of inflation that Bailey (and many others) have used may not be suitable for this question, and to propose an alternative formula for evaluating the welfare cost of inflationary finance. The alternative is very close to a square root formula for welfare costs. It assigns much greater benefit to the reduction of moderate inflations, and substantial gains from reducing inflation to the steady deflation of around 3 percent that would reduce nominal interest rates to zero.

We will now compare two economies with the same preferences and technology, both in a deterministic steady state with constant rates of money growth and price inflation, and constant nominal interest rates equal to the common real rate plus the inflation rate. In one economy monetary policy induces a steady deflation at the Friedman optimal rate corresponding to a nominal interest rate of zero; in the other money grows so as to induce a positive nominal interest rate i. Then we let $w(i)$ be the fraction of income people would be willing to forego to move from the second economy to the first, and call this function $w(i)$ the welfare cost of the inflation rate corresponding to i. Following Lucas (1994), we will provide a general formula for evaluating the welfare cost of inflation as an alternative to Bailey's (1956) consumer's surplus argument that uses the area under a semi-log money demand function as an estimator of the welfare cost of inflation.

The Model: In order to derive a general formula for quantitative issues in the welfare cost analysis of inflation in the intertemporal optimizing framework, we will use the discrete time version of our macro model that was employed in Section III. The technique to provide both the quadratic approximation and the square root rule for assessing the welfare costs of inflation is adapted from Lucas's (1994) paper.

Consider a deterministic, representative agent model, in which a large number of identical households that live forever gain utility from the consumption c_t of a single, nonstorable good, and from their holdings m_t of real money balances. Household preferences are:

$$\sum_{t=0}^{\infty}(1+\rho)^{-t}u(c_t,m_t)\,. \qquad \text{(XI.1)}$$

In this model, each household is endowed with one unit of time, which is inelastically supplied to the market and which produces $y_t = y_0\,(1+\gamma)^t$ units of the consumption good in period t. Hence one equilibrium condition is

$$c_t = y_t = y_0\,(1+\gamma)^t\,. \qquad \text{(XI.2)}$$

Households begin period t with M_t units of money and pay a lump sum tax H_t (or, if $H_t < 0$, receive a lump sum transfer from the government). They can save by accumulating nominal money balances M_{t-1}, by investing in real capital k_{t-1}, and by buying governments bonds B_{t-1}. All three variables denote quantities held at the beginning of period $t+1$. The bonds that agents buy in period t, B_{t-1}, are sold at the nominal price q_t and yield one unit of money in period $t+1$ so that the nominal rate of return on bonds between t and $t+1$ is $i_t = (1-q_t)/q_t$. The real rate of return is therefore defined as $1 + r_t = (1 + i_t)/(1+\pi_t)$ where $\pi_t = (P_{t-1}-P_t)/P_t$ is the inflation rate between t and $t+1$. The number of bonds B_t divided by the price of the commodity P_t in that period is denoted by b_t.

In each period the households have the following budget constraint:

$$M_{t-1} + P_t\,k_{t-1} + q_t\,B_{t-1} = M_t + P_t\,k_t(1-\delta) + B_t + P_t\,y_t - P_t\,c_t - H_t \qquad \text{(XI.3a)}$$

in nominal terms and thus

$$(1+\pi_t)m_{t-1} + k_{t-1} + (1+r_t)^{-1}\,b_{t-1} = m_t + k_t(1-\delta) + b_t + y_t - c_t - h_t \qquad \text{(XI.3b)}$$

in real terms, where $h_t = H_t/P_t$, and δ is the depreciation rate.

Households maximize (XI.1) subject to (XI.3b). We can now begin by writing the time dependent Lagrangian and get the following set of first-order Euler conditions for the maximum problem:

$$J = (1+\rho)^{-(t-1)}\, u(c_{t-1}, m_{t-1}) + (1+\rho)^{-t}\, u(c_t, m_t)$$
$$+ (1+\rho)^{-(t-1)}\lambda_{t-1}\,[(1+\pi_{t-1})m_t + k_t + (1+r_{t-1})^{-1} b_t - m_{t-1} - k_{t-1}(1-\delta) - b_{t-1} - y_{t-1} + c_{t-1} + h_{t-1}]$$
$$+ (1+\rho)^{-t}\lambda_t\,[(1+\pi_t)m_{t-1} + k_{t-1} + (1+r_t)^{-1} b_{t-1} - m_t - k_t(1-\delta) - b_t - y_t + c_t + h_t]$$

$$\left(\frac{\partial J}{\partial c_t}\right) = (1+\rho)^{-t}\, u_c(c_t, m_t) + (1+\rho)^{-t}\lambda_t = 0 \qquad \text{(XI.4a)}$$

$$\left(\frac{\partial J}{\partial m_t}\right) = (1+\rho)^{-t}\, u_m(c_t, m_t) + (1+\rho)^{-(t-1)}(1+\pi_{t-1})\,\lambda_{t-1} - (1+\rho)^{-t}\lambda_t = 0 \qquad \text{(XI.4b)}$$

$$\left(\frac{\partial J}{\partial b_t}\right) = (1+\rho)^{-(t-1)}(1+r_{t-1})^{-1}\lambda_{t-1} - (1+\rho)^{-t}\lambda_t = 0. \qquad \text{(XI.4c)}$$

Dividing (XI.4b) by (XI.4a), we get

$$1 + \frac{u_m(c_t, m_t)}{u_c(c_t, m_t)} = (1+\rho)(1+\pi_{t-1})\frac{\lambda_{t-1}}{\lambda_t} = 1 + i_{t-1} \qquad \text{(XI.5)}$$

since $\dfrac{\lambda_{t-1}}{\lambda_t} = \dfrac{(1+r_{t-1})}{(1+\rho)}$ from (XI.4c) and $1 + r_{t-1} = (1 + i_{t-1})/(1 + \pi_{t-1})$.

In any equilibrium, $c_t = y_t$ for all t, from (XI.2). Now consider a balanced growth equilibrium in which the money growth rate is constant at μ, so that the inflation factor $(1 + \pi_{t-1})$ is constant at the value $(1+\mu)/(1+\gamma)$, and the real money balances-income ratio $m_t/y_t = m_t/c_t$ is constant at m. In this case, (XI.5) becomes:

$$u_m(1, m) = i\, u_c(1, m) \qquad \text{(XI.6)}$$

where m is the steady state level of real money balances. Let $m(i)$ denote the m value that satisfies (XI.6), expressed as a function of the nominal rate of interest. The flow utility enjoyed by the household in the steady state is $u(1, m(i))$. In this section, following Lucas (1994), we will define the welfare cost $w(i)$ of a nominal rate of interest i to be the percentage income compensation needed to leave the household indifferent between i and 0. That is, $w(i)$ is defined as the solution to:

$$u[1+w(i), m(i)] = u[1, m(0)]. \tag{XI.7}$$

Our objective is to use an estimated $m(i)$ to obtain a quantitative estimate of the function $w(i)$.

One way to do this is to take the second order Taylor series expansion of the function $w(i)$ about the zero nominal rate of interest. This is the method used by Bailey (1956) and Lucas (1994). The first two derivatives of $w(i)$ are found by using (XI.6) and (XI.7). Total differentiation of (XI.7) yields

$$u_c\, w'(i) + u_m\, m'(i) = 0$$

which can be written as

$$w'(i) = -\frac{u_m}{u_c}\, m'(i) = -\, i\, m'(i)$$

since $\frac{u_m}{u_c} = i$ from (XI.6). Then we find the second derivative of $w(i)$ with respect to i as

$$w''(i) = -\, m'(i) - i\, m''(i).$$

Thus, the Taylor series expansion of the function $w(i)$ about $i = 0$ gives

$$w(i) = w(i)|_{i=0} + w'(i)|_{i=0}\,(i - 0) + \frac{1}{2}\, w''(i)|_{i=0}\,(i - 0)^2$$

$$= -\frac{1}{2} m'(0)\, i^2 \qquad \text{(XI.8)}$$

since $w(i)|_{i=0} = w'(i)|_{i=0} = 0$ and $w''(i)|_{i=0} = -m'(0)$.

To quantify the right side of (XI.8), Bailey used estimates of the semi-elasticity $\frac{1}{m(i)} m'(i) = -\eta$. In terms of η, (XI.8) can be written as

$$w(i) = \frac{1}{2} m(0)\eta i^2 \qquad \text{(XI.9)}$$

where $m(0)$ is the inverse of the annual income velocity of money at $i = 0$ assuming that i is an annual nominal interest rate and income is normalized at unity.

To obtain an explicit form of $w(i)$, we need an estimate of the semi-elasticity (η). Here and below, we will use the estimate $\dot{\eta} = 6.68$, obtained from 1957:I-1994:I quarterly U.S. time series, using the monetary base as the monetary aggregate and the 3-month t-bill rate as the short-term interest rate and assuming the U.S. money demand takes the Cagan form. Then assuming that the function $m(i)$ passes through the point $(m_0, i_0) = (0.054, 0.0776)$ observed in 1990:I (these and other monetary aggregates and the interest rates used in this study are from the December 1994 IMF, International Financial Statistics database), and taking the semi-elasticity to be constant (Cagan from), we can obtain the estimate $m(0) = (0.054)\, e^{(6.68)(0.0776)} = 0.0906$ since $m(i) = m(0)\, e^{-\eta i}$, and hence from (130) the welfare cost estimate

$$w(i) = (0.302)\, i^2. \qquad \text{(XI.10)}$$

The second column of Table 13 presents some values of this quadratic function.

A second way to get the welfare cost estimates in the Sidrauski model is to parameterize the utility function $u(c_t, m_t)$ presented below in (III.2') in a particular way, estimate the parameters and calculate $w(i)$ employing the same three observations (m_0, i_0, η). We will use same constant intertemporal elasticity of substitution (CIES) utility function employed in Section III except here we assume the income elasticity of demand for money equal to one.

$$u(c_t, m_t) = \left(\frac{1}{1-\phi}\right)\left[\left(\frac{1}{1-\alpha}\right)(c_t^{1-\alpha} - 1) + \left(\frac{\beta}{1-\alpha}\right)(m_t^{1-\alpha} - 1)\right]^{1-\phi}, \qquad \text{(III.2')}$$

where ϕ, α, and β are positive constants so that $u(\ )$ is a strictly concave utility function with continuous first and second derivatives. With this CIES form, the equilibrium condition (XI.6) implies

$$m^{-\alpha}\beta = i \qquad \text{(XI.11)}$$

since

$$u_m = \left[\left(\frac{1}{1-\alpha}\right)(c_t^{1-\alpha} - 1) + \left(\frac{\beta}{1-\alpha}\right)(m_t^{1-\alpha} - 1)\right]^{-\phi} m_t^{-\alpha}\beta,$$

$$u_c = \left[\left(\frac{1}{1-\alpha}\right)(c_t^{1-\alpha} - 1) + \left(\frac{\beta}{1-\alpha}\right)(m_t^{1-\alpha} - 1)\right]^{-\phi} c_t^{-\alpha},$$

which yields $u_m/u_c = m^{-\alpha}\beta$ where $m = m_t/c_t$ as we mentioned above. The welfare cost is defined as the solution $w(i)$ to $u[1+w(i), m(i)] = u[1, \infty]$ (since in this case $m(0) = \infty$) which implies the following first-order differential equation from which we can derive the welfare cost estimate $w(i)$:

$$w'(i) = -\, i\, m'(i) \qquad \text{(XI.12)}$$

where $m'(i) = \dfrac{-\beta^{1/\alpha}}{\alpha} i^{(-1/\alpha)-1}$ from (XI.11). The solution to the first-order differential equation given in (XI.12) yields the following form for $w(i)$:

$$w(i) = \frac{\beta^{1/\alpha}}{\alpha - 1} i^{(\alpha-1)/\alpha}. \tag{XI.13}$$

We can calibrate this model, again using (m_0, i_0, η) = (0.054, 0.0776, 6.68). From (XI.11), the implied money demand semi-elasticity is

$$\frac{1}{m(i)} m'(i) = -\eta = -\frac{1}{i}\left(\frac{1}{\alpha}\right). \tag{XI.14}$$

Notice that with the utility function (III.2'), it is the elasticity of $m(i)$ that is constant, not the semi-elasticity. With $i = 0.0776$ and $\eta = 6.68$, (135) implies $\alpha = 1.93$ and $(\alpha-1)/\alpha = 0.482$. From (XI.11) with $m = 0.054$, we find $\beta^{1/\alpha} = m\, i^{1/\alpha} = (0.054)(0.0776)^{1/1.93} = 0.0144$. The implied welfare cost function is therefore:

$$w(i) = (0.0154)\, i^{0.482}. \tag{XI.15}$$

The third column of Table 13 gives the values of this function. Since the second and third columns of Table 13 provide estimates of the outcomes of the same experiment, based on the same model, their differences are especially striking. According to this table, the welfare cost of inflation which increases with the *square* of the interest rate presented in (XI.10) assigns large costs to very high rates of inflation but trivial costs to moderate inflation's whereas the *square root* formula for welfare costs given in (XI.15) assigns much greater benefit to the reduction of moderate inflation's and implies substantial gains from reducing inflation to the steady deflation of around 3 percent that would reduce nominal interest rates to zero.

According to our estimates with the monetary base and the 3-month t-bill rate, moving from the inflation rate of 2.92 % (from the nominal rate of interest of 5.92 %) to the steady deflation of 3 % (to the zero nominal rate of interest) the square root formula assigns a substantial gain of 0.39 % of GDP whereas the quadratic approximation assigns 0.11 %.

Lucas (1994) uses M1 as the relevant money and runs his welfare integrals from a zero nominal rate to a positive rate. This is a procedure that lumps together currency and interest bearing deposits and treats this composite as non-interest bearing. He also mentions that he has magnified the amount of welfare gains from reducing inflation to the steady deflation of around 3 percent by working with M1 rather than the monetary base since M1 is about 3 times the monetary base in the United States.

To compare the welfare cost estimates, we also use M1 with an estimate of the semi-elasticity $\dot{\eta} = 8.23$, obtained from 1957:I-1994:I quarterly U.S. time series, using M1 as the monetary aggregate and the 3-month t-bill rate as the short-term interest rate and assuming the U.S. money demand takes the Cagan form. Then assuming that the function $m(i)$ passes through the point $(m_0, i_0) = (0.1494, 0.0776)$ observed in 1990:I, we can obtain the welfare cost estimate

$$w(i) = (1.163)\, i^2. \qquad \text{(XI.16)}$$

The fourth column of Table 13 shows the values of this function to compare the results with the second column. We also used the second way to get the welfare cost of inflation by estimating the parameters and calculating $w(i)$ employing the three observations $(m_0, i_0, \eta) = (0.1494, 0.0776, 8.23)$. The resulting expression for $w(i)$ is

$$w(i) = (0.052)\, i^{0.362}. \qquad \text{(XI.17)}$$

The fifth column of Table 13 displays the values of this function to compare the results with the third column.

XII. THE QUANTITY ELASTICITY OF MONEY DEMAND:

In this section we first present the estimates of income and consumption elasticities of money demand then test the unitary quantity elasticity hypothesis. Seigniorage and welfare cost analysis are compatible with a steady state only if the velocity of money remains constant. In turn, constant velocity requires a unitary quantity elasticity of demand for money[31]. Therefore, it is crucial that the quantity elasticity is unity. If this were not the case, we could not use steady state analysis.

There has been a big debate about the size of the income elasticity of money demand among macroeconomists. In the theoretical literature, the predicted elasticities range between one-third and one: In the Baumol (1952) - Tobin (1956) model, when the transaction costs are assumed to be independent of income which is not completely realistic, the transactions demand for money predicts an income elasticity of one-half. In a similar kind of model, Karni (1973) shows that if the transaction costs are related to the time needed to go to the bank, then the cost is related to the wage rate, which, in turn, will be positively correlated with the aggregate level of income. In this case, the overall elasticity would be greater than one-half. The stochastic version of the model developed by Miller and Orr (1966) reduces the prediction of the income elasticity to about one-third. The elasticity predicted by the "cash-in-advance" model of Barro and Fischer (1976) is unity. At the empirical level, the elasticity estimates are even more erratic. Many empirical studies of money demand exist; their estimates vary widely. Most of these studies make use of the time-series data but Mulligan and Sala-i-Martin (1992) estimate money demand functions using cross-sections of U.S. states for the period 1929-1990 and find that the income elasticity of both demand deposits and a broader measure of money lies between 1.3 and 1.5.

According to our estimates. as presented in Table 17, the income elasticity of demand for the monetary base (0.5 with the semi-log from and 0.56 with the double-log form) is found to be slightly greater than the consumption elasticity (0.45 with the semi-log form and 0.53 with the double-log form) and these figures turn out to be larger when M1 is chosen as the monetary aggregate. With M1, the income elasticity (0.60 with the semi-log form and 0.66 with the double-log) is also greater than the consumption elasticity (0.51 with the semi-log form and 0.58 with the double-log form). In other words, the quantity (income or consumption) elasticity of demand for money (monetary base or M1) turn out to be well below one suggesting that the use of a unitary income (or consumption) elasticity assumption leads to overestimate the welfare cost of inflation. We utilize the Wald and the Likelihood Ratio tests to determine whether the unitary income (or consumption) elasticity assumption can be employed to estimate U.S. money demand. As shown in Table 18. for each functional form with monetary base and M1, the unitary income (or consumption) elasticity of demand for money is strongly rejected, implying that using steady state analysis in estimating money demand equations is improper.

The estimated welfare cost of inflation is proportional to the money stock, so the choice of an observed monetary aggregate to serve as "money" in the formula is crucial. Lucas (1994) chooses to use M1 as the relevant money and runs his welfare integral from a zero nominal rate to a positive rate -a procedure that lumps currency and interest bearing deposits together and treats this composite as non-interest bearing. In addition to Lucas (1994), Dotsey and Ireland (1996) use M1 to estimate the welfare cost of inflation in steady state assuming unitary income elasticity of demand for real M1. As is evident in this paper, identifying M1 instead of the monetary base as the relevant definition of money leads to overestimate the true welfare cost of inflation.

Lucas (1994) states "The fact that real balances are a minor "good" in the U.S. economy (and I have probably magnified their importance by working with M1 rather than with the monetary base), so the fiscal consequences of even sizable changes in inflation and interest rates are just not likely to be large". Within his framework, and with his preferred constant elasticity double-log function with income, he finds the welfare cost of a 10 % nominal interest rate is about 1.3 % of GDP which is very close to ours. In a general equilibrium monetary model, Dotsey and Ireland (1996) find the welfare cost of a sustained 4 % inflation rate is over 1% of GDP when M1 is defined as money. According to our welfare cost estimates, assuming unitary income elasticity of demand for M1, the double-log function yields the welfare cost of 1.27 % of GDP evaluated at the 10% nominal interest rate whereas when the monetary base is used as the relevant monetary aggregate, the same functional form with the same scale variable yields the welfare cost of 0.4 % of GDP which is about one-third of the welfare cost estimated with M1.

Our second observation concerns how the use of a unitary quantity (income or consumption) elasticity assumption affects the welfare cost estimates. As depicted in Charts 1 and 2, for each specification of money demand (with monetary base or M1) and each scale variable (income or consumption), the welfare cost of inflation is overestimated with the unitary quantity elasticity assumption since the actual quantity elasticity is estimated well below one. To determine the degree of overestimation, we calculate the mean of the welfare cost series and present them in Charts 1 and 2. For each functional form and each scale variable, when the monetary base (M1) is identified as the definition of money, the welfare cost of inflation estimated with the unitary quantity elasticity ($WC_{unitary}$) is about three (two) times the welfare cost estimated with the nonunitary quantity elasticity ($WC_{nonunitary}$).

XIII. THE EXACT WELFARE COST OF INFLATION IN A CURRENCY-DEPOSIT MODEL

The estimated welfare cost of inflation is proportional to the money stock, so the choice of an observed monetary aggregate to serve as "money" in the formula is crucial. Lucas (1993,1994) chooses to use M1 as the relevant money and runs his welfare integral from a zero nominal rate to a positive rate - a procedure that lumps currency and interest bearing deposits together and treats this composite as non-interest bearing. In this regard, it is relevant to return to Bailey's (1956) classic article and review his treatment of currency and deposits. Bailey used Cagan's (1956) hyperinflation data to estimate the welfare cost of inflation. He set the real rate of interest at zero ($r = 0$) so that the cost of holding currency was the anticipated inflation π (a proxy for the nominal interest rate i). He assumed competitive banks pay interest on deposits but subject to a sterile reserve requirement. A zero-profit condition was imposed: the revenue from a bank's interest bearing assets was completely dispersed in interest payments on deposits: $i_d = (1- \mu)i$ where μ is the required reserve ratio and $i=\pi$. The zero-profit condition under perfect competition implies that the opportunity cost of holding deposits is $(i - i_d) = \mu i$ (or $\mu\pi$). Since deposits are partially indexed, Bailey assumed that at very high rates of inflation, the public uses only deposits and all currency would be held as bank reserves. In this case, Bailey's welfare integrals run from zero to $\mu\pi$ to account of the partial indexing of the deposit rate to inflation. In contrast to Lucas, there is a substantial reduction in welfare cost of inflation when interest on deposits is partially indexed against inflation.

What is at issue is the polar choice of two very incomplete models. On the one hand, we can lump currency and deposits together and assume that both pay no interest. We can then run our integrals from zero to the nominal rate of interest. This is Lucas' choice. On the other hand, we could assume with Bailey that all high powered money is held as reserves and run the welfare integrals from a nominal rate of zero to μi, the opportunity cost of holding interest bearing deposits. Lucas' choice overstates the welfare loss of deviating from the Friedman (1969) rule[32]. The assumption in Bailey that all deposits pay interest underestimates the welfare loss. Without a theory of banking in which the distinctive roles of currency and deposits are modeled, it remains uncertain how large is the degree of overestimation or underestimation. As an alternative to these two polar cases, we try to model the distinctive roles of currency and deposits in Ramsey-Sidrauski infinite horizon model. We measure the welfare cost of inflation using the traditional approach, developed by Bailey (1956), by computing the appropriate areas under currency and deposit demand curves. In our currency-deposit model, it is possible to run separate integrals: that for currency running from a zero nominal rate to a positive rate, and the integral for deposits running from zero to μi. We also provide the quadratic approximation and the square root formula, originally developed by Lucas (1993) in a single-monetary-asset model, for evaluating the welfare cost of inflation in our currency-deposit model as an alternative to Bailey's (1956) consumer's surplus argument.

In addition to Lucas (1993,1994), Lucas (1981) and Dotsey and Ireland (1996) use M1 to estimate the welfare cost of inflation for the American economy. As will be discussed in this paper, identifying M1 instead of the monetary base as the relevant definition of money leads to overestimate the true welfare cost of inflation in the single-monetary-asset model since the estimated welfare cost of inflation is proportional to the money stock and M1 is about three times the monetary base in the United States. Lucas (1994) states "The fact that real balances are a minor "good" in the U.S. economy (and I have probably magnified their importance by working with M1 rather than with the monetary base), so the fiscal consequences of even sizable changes in inflation and interest rates are just not likely to be large". Within his framework, and with his preferred constant elasticity double-log function scaled with income, he finds the welfare cost of a 10 % nominal interest rate is about 1.3 % of GDP. In a general equilibrium monetary model, Dotsey and Ireland (1996) find the welfare cost of a sustained 4 % inflation rate is over 1% of GDP when M1 is defined as money. Fischer (1981) and Lucas (1981) find the cost of inflation to be surprisingly low. Fischer computes the deadweight loss from an increase in inflation from zero to 10 % as just 0.3 % of GNP identifying money with the monetary base. Lucas (1981) places the cost of a 10 % inflation at about 0.45 % of GNP using M1 as the measure of money. According to our welfare cost estimates with Bailey's (1956) consumer's surplus and Lucas's (1993) compensating variation approaches, the double-log function specified with M1 and scaled with income yields the welfare cost of 1.08 % of GDP when evaluated at the 10% nominal interest rate whereas when the monetary base is used as the appropriate monetary aggregate, the same functional form with the same scale variable yields the welfare cost of 0.37 % of GDP which is about one-third of the welfare cost estimated with M1.

This paper quantifies the welfare costs of deviating from a zero inflation policy and the costs of deviating from the Friedman optimal deflation rate for the U.S. economy in both a currency-deposit model and a single-monetary-asset model using recent data on money, income, consumption and the short-term interest rate. As will be evident in this paper, the magnitude of the estimated welfare cost of inflation is sensitive to the specification of the money demand function and to the definition of money. Thus, it is crucial to identify the proper specification of money demand as well as the appropriate monetary aggregate to serve as "money" in the formula. To determine what form the U.S. money demand takes, the Box-Cox transformation is applied to the nominal interest rate and the proper specification of the money demand function is identified with the aid of two different econometric tests. The empirical results indicate that compared to the constant semi-elasticity Cagan-type demand for money, the double-log function with constant elasticity is a more accurate characterization of the actual data over the range of observed interest rates. We measure the welfare cost of inflation first using the quadratic approximation (for the semi-log form) and the square root formula (for the double-log form) originally developed by Lucas (1993) in Sidrauski's (1967) general equilibrium single-monetary-asset model. We then estimate the welfare cost of inflation in Bailey's (1956) partial equilibrium framework by computing the appropriate area under the money demand curve (summing the areas under currency and deposit demand schedules for the currency-deposit model and computing the area under the monetary base and M1 demand curve separately for the single-monetary-asset model). The welfare cost measures obtained from Bailey's consumer's surplus and Lucas's compensating variation approaches imply that for each monetary aggregate (currency-deposit, monetary base, M1) and each scale variable (income, consumption), the double-log function with constant elasticity of less than one yields substantial welfare cost estimates for the U.S. economy compared to the semi-log function.

The Currency-Deposit Model: Several approaches can be found in the literature for introducing the role of money into the intertemporal optimizing framework[33]. To specify the distinctive roles of currency and deposits in an infinite horizon model, we modify Sidrauski's (1967) general equilibrium single-monetary-asset model by separating M1 into its currency and demand deposit components and apply a required reserve ratio to the deposit component. In this currency-deposit model, we think about the money stock as "making life easier" since it allows people to get consumption goods without having to go to the bank and transform bonds into consumption goods all the time.

We will assume that the economy is populated by infinitely lived consumers or dynasties who derive utility from the only consumption good c_t and from real money stock which is composed of non-interest bearing real currency m_t and interest bearing real demand deposits d_t. Each household is assumed to have access to a production function that is homogenous of degree one in its two inputs, capital and labor. We will also assume that labor is supplied inelastically so that the production function with diminishing marginal product of capital can be written as

$$y_t = f(k_t) \qquad \text{(XIII.1)}$$

where y_t is real output, k_t is real capital and $dy_t/dk_t = f'(k_t)>0$, $d^2y_t/dk_t^2 = f''(k_t)<0$. Here capital is considered as output that is not consumed, so its price is the same as that of consumption good. The households in the economy begin period t with M_t units of money, pay a lump sum tax H_t (or, if $H_t<0$, receive a lump sum transfer from the government).

They can save by accumulating nominal currency M_{t-1} and nominal demand deposits D_{t-1}, by investing in real capital k_{t-1}, and by buying governments bonds B_{t-1}. All four variables denote quantities held at the beginning of period $t+1$. The bonds that agents buy in period t, B_{t-1}, are sold at the nominal price z_t and yield one unit of money in period $t+1$ so that the nominal rate of return on bonds between t and $t+1$ is $i_{t-1} = (1 - z_t)/z_t$. The real rate of return on bonds is therefore defined as $1 + r_{t-1} = (1 + i_{t-1})/(1 + \pi_{t-1})$ where $\pi_{t-1} = (P_{t-1} - P_t)/P_t$ is the inflation rate between t and $t+1$. The number of bonds B_t divided by the price of the commodity P_t in that period is denoted by b_t. We will assume that competitive banks pay interest on deposits but subject to fractional reserve requirement on demand deposits. Assuming banks face no operating costs, the zero-profit condition under perfect competition yields $i_d(t) = (1-\mu_t)i_t$ implying that the opportunity cost of holding demand deposit is $\mu_t i_t$: the difference between i_t and the nominal interest on demand deposits $i_d(t)$. In each period the households have the following budget constraint:

$$M_{t-1} + D_{t-1} + P_t\, k_{t-1} + z_t B_{t-1} = M_t + D_t(1+i_d(t)) + P_t\, k_t(1-\delta) + B_t + P_t\, y_t - P_t\, c_t - H_t \quad \text{(XIII.2)}$$

in nominal terms and thus

$$(1 + \pi_{t-1})m_{t-1} + (1 + \pi_{t-1})d_{t-1} + k_{t-1} + (1 + r_{t-1})^{-1} b_{t-1} = m_t + d_t(1+i_d(t)) + k_t(1-\delta) + b_t + f(k_t) - c_t - h_t \quad \text{(III.3)}$$

in real terms, where $h_t = H_t/P_t$, and δ is the depreciation rate.

We will further assume that the household preferences are time separable, that is

$$\sum_{t=0}^{\infty}(1+\rho)^{-t}\,u(c_t, m_t, d_t) \quad \text{(XIII.4)}$$

where ρ is the subjective rate of time preference. Households maximize (XIII.4) subject to (XIII.3) and the first-order Euler conditions for the maximum problem can be written as equalities holding for each period:

$$u_c(c_t, m_t, d_t) - \Omega_t = 0, \qquad \text{(XIII.5a)}$$

$$u_m(c_t, m_t, d_t) + \Omega_t - (1+\rho)(1+\pi_t)\,\Omega_{t-1} = 0, \qquad \text{(XIII.5b)}$$

$$u_d(c_t, m_t, d_t) + (1 + i_d(t))\,\Omega_t - (1+\rho)(1+\pi_t)\,\Omega_{t-1} = 0, \qquad \text{(XIII.5c)}$$

$$\Omega_t - (1+\rho)(1+r_t)^{-1}\,\Omega_{t-1} = 0, \qquad \text{(XIII.5d)}$$

$$\Omega_t\,(1 + f'(k_t) - \delta) - (1+\rho)\Omega_{t-1} = 0, \qquad \text{(XIII.5e)}$$

where Ω_t and Ω_{t-1} are the Lagrangian multipliers attached to the budget constraints at time t and $t-1$, respectively. From (XIII.5d) and (XIII.5e) we have $\frac{\Omega_{t-1}}{\Omega_t} = \frac{(1+r_t)}{(1+\rho)} = \frac{(1+f'(k_t)-\delta)}{(1+\rho)}$ which implies that the real rate of return on capital equals the real rate of interest on bonds: $(f'(k_t) - \delta) = r_t$. Dividing (XIII.5b) and (XIII.5c) by (XIII.5a) and using the fact that $\frac{\Omega_{t-1}}{\Omega_t} = \frac{(1+r_t)}{(1+\rho)}$ where $(1+r_t) = \frac{1+i_t}{1+\pi_t}$, we get the following conditions

$$\frac{u_m(c_t, m_t, d_t)}{u_c(c_t, m_t, d_t)} = -1 + (1+\rho)(1+\pi_t)\frac{\Omega_{t-1}}{\Omega_t} = i_t, \qquad \text{(XIII.6a)}$$

$$\frac{u_d(c_t, m_t, d_t)}{u_c(c_t, m_t, d_t)} = -(1+i_d) + (1+\rho)(1+\pi_t)\frac{\Omega_{t-1}}{\Omega_t} = i_t - i_d(t). \qquad \text{(XIII.6b)}$$

which imply

$$u_m(c_t, m_t, d_t) = i_t\, u_c(c_t, m_t, d_t), \qquad \text{(XIII.7a)}$$

$$u_d(c_t, m_t, d_t) = (i_t - i_d(t))\, u_c(c_t, m_t, d_t). \qquad \text{(XIII.7b)}$$

In order to find a closed from solution for real currency and real deposit demand functions, suppose that the period utility function takes the following CES-isoelastic form:

$$U(c_t, m_t, d_t) = \frac{\left\{\left[\gamma_1^{1/\theta} c_t^{(\theta-1)/\theta} + \gamma_2^{1/\theta} m_t^{(\theta-1)/\theta} + \gamma_3^{1/\theta} d_t^{(\theta-1)/\theta}\right]^{\theta/(\theta-1)}\right\}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}, \qquad \text{(XIII.8)}$$

where θ>0 is the intertemporal substitution elasticity. Imposing CES-isoelastic preferences on first-order conditions, (XIII.5a), (XIII.5b) and (XIII.5c), yields a generalization of asset demand functions which take the following double-log form with constant elasticity of θ:

$$m_t = \left(\frac{\gamma_2}{\gamma_1}\right) i_t^{-\theta} c_t \qquad \text{(XIII.9a)}$$

$$d_t = \left(\frac{\gamma_3}{\gamma_1}\right) (i_t - i_d(t))^{-\theta} c_t \qquad \text{(XIII.9b)}$$

where the consumption elasticities of demand for real currency and real deposits equal one. As presented above, standard utility functions yield double-log demands for money although the semi-log form has been widely used in money demand studies, more a matter of convenience than theoretical or empirical attractiveness. The convenience comes from its simplicity, the fact that steady state seigniorage is easily defined at a constant nominal interest rate and the fact that the seigniorage-maximizing interest rate as well as the maximum seigniorage are easily derived. However, the semi-logarithmic form has its drawbacks. The micro-foundation of the semi-log specification in either "the shopping time" or "the money-in-the-utility function" approach is somewhat ad hoc. We have to impose a specific form on the utility function [see Calvo and Leiderman (1992)] or on the shopping time technology [see Brock (1989)] in order to arrive at the constant semi-elasticity Cagan-type demand for money. The utility function and the similarly shaped shopping time technology take the following unusual forms, respectively:

$$u(c_t, m_t, d_t) = \frac{m_t(1+\varphi-\ln m_t)}{\alpha} + \frac{d_t(1+\omega-\ln d_t)}{\beta} + v(c_t), \qquad u_m>0,\ u_d>0,\ u_c>0$$

$$\psi(c_t, m_t, d_t) = \frac{m_t(\ln m_t-\varphi-1)}{\alpha} + \frac{d_t(\ln d_t-\omega-1)}{\beta} + v(c_t), \qquad \psi_m<0,\ \psi_d<0,\ \psi_c>0$$

which imply the following Cagan-type demand for currency and deposits:

$$\ln m_t = \varphi - \alpha v_c i_t,$$

$$\ln d_t = \omega - \beta v_c (i_t - i_d(t)),$$

where the marginal utility of consumption is assumed to be a positive constant, $v_c > 0$.

The Welfare Cost of Inflation: In the theoretical literature, it is a widely accepted idea that seigniorage is a certain amount of revenue a government collects, with the aid of its central bank, from the issue of monetary base[34]. Seigniorage can be viewed as a tax on private agents' domestic currency holdings since money creation causes inflation, thereby lowering the real value of nominal assets. In other words, inflation imposes a tax on money holdings because it is the rate at which individuals lose the purchasing power of a dollar. Therefore, individuals change their holdings and their use of money to lower the total cost of holding money when inflation rises. Their efforts to do so, however, reduce total services from real money balances, thereby lowering individuals' welfare. This loss is the welfare cost of inflation. In this section, we attempt to quantify the welfare loss of deviating from a zero inflation policy and answer the question: How much welfare does the U.S. gain in moving from zero inflation to the Friedman optimal deflation rate needed to bring nominal interest rates to zero? We also measure the sensitivity of the estimated welfare cost of inflation to the specification of money demand and to the definition of money in both the currency-deposit model and the single-monetary-asset model.

We first use the quadratic approximation (for the semi-log function) and the square root formula (for the double-log function), originally introduced by Lucas (1993) in a single-asset monetary model, then the traditional approach (consumer's surplus argument) developed by Bailey (1956). Bailey's original study uses the constant semi-elasticity Cagan-type demand for money and as will be demonstrated in this section the semi-log function yields a quadratic formula for the welfare cost of inflation. In this case, the welfare cost of inflation increases with the square of the interest rate, assigning large costs to very high rates of inflation but trivial costs to moderate inflation's. That is, with the semi-logarithmic form, the objective of a zero inflation rate has negligible benefit, and an additional gain in moving from a zero inflation rate to a zero nominal interest rate that would attain the Friedman (1969) optimum is assigned an even smaller value. We will also provide a square root formula for the welfare cost of inflation as an alternative to Bailey's consumer's surplus argument that uses the appropriate area under a semi-log money demand function as an estimator of the welfare cost of inflation.

Lucas compares two economies with the same preferences and technology, both in a deterministic steady state with constant rates of money growth and price inflation, and constant nominal interest rates equal to the common real rate plus the inflation rate. He assumes that in one economy monetary policy induces a steady deflation at the Friedman optimal rate corresponding to a nominal interest rate of zero; in the other money grows so as to induce a positive nominal interest rate. Then the welfare cost of inflation is defined as the fraction of income people would be willing to forego to move from the second economy to the first.

We will now following Lucas assume that each household in the model described in section II is endowed with one unit of time, which is inelastically supplied to the market and which produces $y_t = y_0 (1 + \Phi)^t$ units of the consumption good in period t. Hence one equilibrium condition is $c_t = y_t = y_0 (1 + \Phi)^t$. Now consider a balanced growth equilibrium in which the money growth rate is constant at $\xi = (M_{t-1} - M_t)/M_t$, so that the inflation factor $(1 + \pi)$ is constant at the value $(1+\xi)/(1+\Phi)$, the real currency-income ratio $m_t/y_t = m_t/c_t$ is constant at $\bar{m}$, and the real deposit-income ratio $d_t/y_t = d_t/c_t$ is constant at $\bar{d}$. In this case, Equations (XIII.7a) and (XIII.7b) become:

$$u_m(1, \bar{m}, \bar{d}) = i\, u_c(1, \bar{m}, \bar{d}) \qquad \text{(XIII.10a)}$$

$$u_d(1, \bar{m}, \bar{d}) = (i - i_d)\, u_c(1, \bar{m}, \bar{d}) \qquad \text{(XIII.10b)}$$

where $\bar{m}$ and $\bar{d}$ are the steady state levels of currency and deposits (both as a share of income).

respectively. Let $\bar{m}(i)$ denote the $\bar{m}$ value that satisfies (XIII.10a), expressed as a function of the nominal rate of interest and $\bar{d}(\mu i)$ denote the $\bar{d}$ value that satisfies (XIII.10b), expressed as a function of the difference between the nominal rate of interest and the nominal deposit rate: $i - i_d = \mu i$. The flow utility enjoyed by the household in the steady state is $u(1, \bar{m}(i), \bar{d}(\mu i))$. Provided $\bar{m}'(i)$ [$= \partial\bar{m}(i)/\partial(i)$] < 0 and $\bar{d}'(\mu i)$ [$= \partial\bar{d}(\mu i)/\partial(i)$] < 0, this utility is maximized over nonnegative nominal interest rates at $i = 0$:

the Friedman rule of a deflation equal to the real rate of interest. We define the welfare cost $w(i)$ of a nominal rate of interest i to be the percentage income compensation needed to leave the household indifferent between i and 0. That is, $w(i)$ is defined as the solution to:

$$u[1+w(i), \bar{m}(i), \bar{d}(\mu i)] = u[1, \bar{m}(0), \bar{d}(0)]. \qquad \text{(XIII.11)}$$

Our objective is to use an estimated $\bar{m}(i)$ and $\bar{d}(\mu i)$ to obtain a quantitative estimate of the function $w(i)$. One way to do this is to take the second order Taylor series expansion of the function $w(i)$ about the zero nominal interest rate. The first two derivatives of $w(i)$ are found by using (XIII.10a), (XIII.10b) and (XIII.11). Total differentiation of (XIII.11) yields: $u_c w'(i) + u_m \bar{m}'(i) + u_d \mu \bar{d}'(\mu i) = 0$, which can be written as: $w'(i) = -\frac{u_m}{u_c}\bar{m}'(i) - \frac{u_d}{u_c}\mu\bar{d}'(\mu i) = -i\bar{m}'(i) - i\mu^2\bar{d}'(\mu i)$ since $\frac{u_m}{u_c} = i$ from (XIII.10a) and $\frac{u_d}{u_c} = (i - i_d) = \mu i$ from (XIII.10b). Then we find the second derivative of $w(i)$ with respect to i: $w''(i) = -\bar{m}'(i) - i\bar{m}''(i) - \mu^2[\bar{d}'(\mu i) + i\mu\bar{d}''(\mu i)]$. Thus, the Taylor series expansion of the function $w(i)$ about $i = 0$ gives

$$w(i) = w(i)|_{i=0} + w'(i)|_{i=0}(i - 0) + \frac{1}{2}w''(i)|_{i=0}(i - 0)^2 = \frac{1}{2}i^2[-\bar{m}'(0) - \mu^2\bar{d}'(0)] \qquad \text{(XIII.12)}$$

since $w(i)|_{i=0} = w'(i)|_{i=0} = 0$ and $w''(i)|_{i=0} = -\bar{m}'(0) - \mu^2\bar{d}'(0)$. To quantify the right side of (XIII.12) for the semi-log function used by Cagan (1956) and Bailey (1956), we will use estimates of the semi-elasticity $\frac{1}{\bar{m}(i)}\bar{m}'(i) = \frac{1}{\bar{d}(\mu i)}\bar{d}'(\mu i) = -\eta$ assuming that the demand for currency and deposits have the same semi-elasticity. In terms of η, (XIII.12) can be written as[35]

$$w(i) = \frac{1}{2}\eta i^2 \, [\bar{m}(0) + \mu^2 \bar{d}(0)] \qquad \text{(XIII.13)}$$

where $\bar{m}(0)$ and $\bar{d}(0)$ are the inverses of the annual income velocities of currency and deposits at i=0, respectively, assuming that i is an annual nominal interest rate and income is normalized at unity.

To obtain an explicit form of $w(i)$, we need an estimate of the semi-elasticity (η). Here and below, we will use the parameter estimates obtained from 1957:I-1997:II quarterly U.S. time series. We will now illustrate the procedure for finding a general formula for the semi-log function in the currency-deposit model. Assuming that the functions $\bar{m}(i)$ and $\bar{d}(\mu i)$ pass through the points (m_0, i_0) = (0.0509, 0.0520) and $(d_0, \mu_0 i_0)$ = (0.1034, 0.0042) observed in 1997:II and using the constant semi-elasticity[36] η=7.8365 (α_1 in footnote 36), we obtain the estimate $\bar{m}(0) = (0.0509)e^{(7.8365)(0.0520)}$ = 0.0765 since $\bar{m}(i) = \bar{m}(0)\,e^{-\eta i}$, and $\bar{d}(0) = (0.1034)e^{(7.8365)(0.0042)}$ = 0.1074 since $\bar{d}(\mu i) = \bar{d}(0)\,e^{-\eta(\mu i)}$. Hence from (XIII.13) the estimated welfare cost of inflation evaluated at the mean of the reserve ratio μ_{mean} = 0.1373 is[37]

$$w(i) = (0.301)\, i^2. \qquad \text{(XIII.14)}$$

The second column of Table 20 presents some values of this quadratic function which yields a welfare cost of a 10 % nominal interest rate at about 0.30 % of GDP. Assuming the real rate of interest is around 3 %, reduction to a 3 % nominal rate (about the rate associated with zero inflation) reduces the cost to about 0.03 % of GDP, for a gain of about 0.27 % of GDP. The additional gain in moving from a zero inflation to the Friedman rate is negligible (0.03% of GDP) with the quadratic welfare cost function given in (XIII.14).

Lucas uses M1 as the relevant money and mentions that he has magnified the amount of welfare gains in moving from zero inflation to zero nominal rate by working with M1 rather than the monetary base since M1 is about three times the monetary base in the United States. To set out the degree of overestimation in the single-monetary-asset model, we follow the same procedure to find $w(i)$ for M1, using $\eta = 7.2649$ (α_1 in Table 19) and the point $(m_0, i_0) = (0.1550, 0.0520)$ observed in 1997:II. The welfare cost estimate for M1 turns out to be[38].

$$w(i) = (0.821)\, i^2, \tag{XIII.15}$$

which is about three times $w(i)$ for the currency-deposit specification presented in (XIII.14). The fourth column of Table 20 shows some values of this quadratic function which implies a welfare cost of a 10% nominal interest rate at about 0.82 % of GDP. Reduction to a zero inflation rate (or a 3 % nominal rate) reduces the cost to about 0.07 %, for a gain of about 0.75 % of GDP. The welfare gain in moving from zero inflation to the steady deflation of around 3 % is again trivial, 0.07 % of GDP, with the quadratic approximation (using M1) of the welfare cost of inflation.

To compare the welfare cost measures obtained from the currency-deposit specification with those estimated with the most relevant definition of money (the monetary base) in the single-asset monetary model, we follow the same procedure for finding $w(i)$ for the base, using $\eta = 7.0645$ (α_1 in Table 19) and the point $(m_0, i_0) = (0.0596, 0.0520)$ observed in 1997:II. The welfare cost estimate for the monetary base is

$$w(i) = (0.304)\, i^2. \tag{XIII.16}$$

The sixth column of Table 20 gives some values of this quadratic function which yields a welfare cost of a 10 % nominal interest rate at about 0.30 % of GDP. Reduction to 3 % nominal rate reduces the cost to about 0.03 %, for a gain of about 0.27 % of GDP. The additional gain in moving from zero inflation to the Friedman rate is negligible (0.03% of GDP). The welfare cost of inflation estimated with the monetary base is found to be almost the same as the welfare cost

computed in the currency-deposit model and both of these estimates are about one-third of the welfare cost estimated with M1.

A second way to obtain the welfare cost of inflation in the currency-deposit model is to parameterize the flow utility function $u(c_t, m_t, d_t)$ in a particular way, estimate the parameters and calculate $w(i)$ exactly. We use the CES utility function presented in (XIII.8) which yields the double-log asset demand functions with unitary income elasticities shown in (XIII.9a) and (XIII.9b). With (XIII.9a) and (XIII.9b), the equilibrium conditions (XIII.10a) and (XIII.10b) imply

$$\bar{m}(i) = \left(\frac{m_t}{c_t}\right) = \left(\frac{\gamma_2}{\gamma_1}\right) i_t^{-\theta}. \tag{XIII.17a}$$

$$\bar{d}(\mu i) = \left(\frac{d_t}{c_t}\right) = \left(\frac{\gamma_3}{\gamma_1}\right) (\mu_t i_t)^{-\theta}. \tag{XIII.17b}$$

The welfare cost is defined as the solution $w(i)$ to $u[1+w(i), \bar{m}(i), \bar{d}(\mu i)] = u(1, \infty, \infty)$. With the CES utility assumed in (XIII.8), the right side is finite provided $\theta<1$, which is to say, provided the elasticity of substitution between goods consumption and real currency and the elasticity of substitution between goods consumption and real deposits are less than one. Using (XIII.17a), (XIII.17b) and the utility function in (XIII.8), we find that[39]:

$$w(i) = \left[1 - \left(\frac{\gamma_2}{\gamma_1}\right) i_t^{1-\theta} - \left(\frac{\gamma_3}{\gamma_1}\right) \mu_t^{1-\theta} i_t^{1-\theta}\right]^{\theta/(\theta-1)} - 1. \qquad \text{(XIII.18)}$$

We will calibrate the exact welfare cost of inflation $w(i)$ given in (XIII.18), employing the estimated parameter values. For the double-log form with the currency-deposit specification[40], $\theta = 0.4240$ (or $\beta_1 = 0.4240$ in footnote 40), $\gamma_2/\gamma_1 = 0.0134$ (or $e^{\beta_0} = e^{-4.3109}$ in footnote 40) and $\gamma_3/\gamma_1 = 0.0149$ (or $e^{\beta_2} = e^{-4.2063}$ in footnote 40). The implied welfare cost function evaluated at the mean of the reserve ratio $\mu_{mean} = 0.1373$ is thus[41]

$$w(i) = (1 - 0.0182\, i_t^{0.5760})^{-0.7362} - 1. \qquad \text{(XIII.19)}$$

The third column of Table 20 presents some values of this function which yields a welfare cost of a 10% nominal interest rate at about 0.36 % of GDP. Reduction to a zero inflation rate (or a 3 % nominal rate) reduces the cost to about 0.18 %, for a gain of about 0.18 % of GDP. Therefore, compared to the semi-log function, the double-log function yields substantial welfare gain (0.18 % of GDP) of moving from zero inflation to the steady deflation of around 3 %.

To determine the degree of overestimation due to the identification of money with M1, the same procedure is followed to find $w(i)$ for M1, using $\theta = 0.4454$ (or $\beta_1 = 0.4454$ in Table 19) and $\gamma_2/\gamma_1 = 0.0478$ (or $e^{\beta_0} = e^{-3.0405}$ in Table 19). The exact welfare cost $w(i)$ for M1 is therefore[42]

$$w(i) = (1 - 0.0478\, i_t^{0.5546})^{-0.8031} - 1. \qquad \text{(XIII.20)}$$

The fifth column of Table 20 displays some values of this function which yields a substantial welfare cost of a 10 % nominal interest rate at about 1.08 % of GDP. Reduction to a zero inflation rate reduces the cost to about 0.55 %, for a considerable gain of about 0.53 % of GDP.

We now present the exact welfare cost of inflation $w(i)$ for the monetary base using the estimated parameter values θ=0.4388 (or β_1 = 0.4388 in Table 19) and γ_2/γ_1 = 0.0172 (or $e^{\beta_0} = e^{-4.0635}$ in Table 19). The implied welfare cost function is thus

$$w(i) = (1 - 0.0172\, i_t^{0.5612})^{-0.7819} - 1. \tag{XIII.21}$$

The last column of Table 20 demonstrates some values of this function which yields a welfare cost of 10 % nominal interest rate at about 0.37 % of GDP. Reduction to a zero inflation rate reduces the cost to about 0.19 %, for a gain of about 0.18 % of GDP. Therefore, compared to the semi-log function, the double-log function with the monetary base yields substantial welfare gain (0.19 % of GDP) of moving from zero inflation to the steady deflation of around 3 %.

The welfare cost estimates in this paper may seem small for people who are unfamiliar with quantitative analysis of the welfare cost of inflation. But the United States is a \$8 trillion economy, so to convert the stocks and flows to dollars, we need to multiply all of them by \$8 $\times 10^{12}$. For example, while comparing the additional welfare gains (estimated with the double-log function) in moving from π=0 to i=0: 0.55 % of GDP (with M1) as opposed to 0.19 % or 0.18 % of GDP (with the base or with the currency-deposit), we imply \$44 billion as opposed to \$15.2 billion or \$14.4 billion. The estimates with the semi-log function: 0.07 % of GDP (with M1) as opposed to 0.03 % of GDP (with the base or with the currency-deposit) imply \$5.6 billion as opposed to \$2.4 billion.

We will now introduce Bailey's consumer surplus argument to estimate the welfare cost of inflation in the currency-deposit model using the semi-log and the double-log forms of asset demand functions. Deadweight social costs of raising revenue via seigniorage collection can be interpreted as the lost areas under currency and deposit demand curves owing to inflation's effect on their respective opportunity costs. This is the traditional "shoe-leather" cost developed by Bailey (1956). With the general form of currency and deposit demand functions, the loss (WC) is measured as the deadweight loss from positive i in reducing currency demand,

$$\int_0^i f(x)dx - if(i),$$

plus the deadweight loss from positive μi in reducing deposit demand,

$$\int_0^{\mu i} g(x)dx - \mu i g(\mu i).$$

Hence, with the semi-logarithmic form of currency and deposit demand, the welfare cost of inflation (as a fraction of GDP) is defined as follows[43],

$$WC_{Bailey}^{semi-\log} = \int_0^i e^{\alpha_0-\alpha_1 x}dx - e^{\alpha_0-\alpha_1 i}i + \int_0^{\mu i} e^{\beta_0-\beta_1 x}dx - e^{\beta_0-\beta_1 \mu i}(\mu i)$$

$$WC_{Bailey}^{semi-\log} = \frac{e^{\alpha_0}}{\alpha_1}(1-e^{-\alpha_1 i}(1+\alpha_1 i)) + \frac{e^{\beta_0}}{\beta_1}(1-e^{-\beta_1 \mu i}(1+\beta_1 \mu i)),$$

and with the double-log form,

$$WC_{Bailey}^{double-\log} = \int_0^i e^{\alpha_0}x^{-\alpha_1}dx - e^{\alpha_0}i^{1-\alpha_1} + \int_0^{\mu i} e^{\beta_0}x^{-\beta_1}dx - e^{\beta_0}(\mu i)^{1-\beta_1}$$

$$WC_{Bailey}^{double-log} = \left(\frac{\alpha_1}{1-\alpha_1}\right) e^{\alpha_0} i^{1-\alpha_1} + \left(\frac{\beta_1}{1-\beta_1}\right) e^{\beta_0} (\mu i)^{1-\beta_1} .$$

Our first important observation is that there is almost no difference between the magnitudes of the welfare cost of inflation estimated in Bailey's partial equilibrium and Lucas's general equilibrium frameworks. In other words, we obtain the same result that for each monetary aggregate (currency-deposit, monetary base, M1), the double-log function implies sizable benefits in moving from zero inflation to the steady deflation of around 3 % that would attain the Friedman optimum, while under semi-log demand these benefits are trivial. The reader may wish to consult Tables 21 and 22 to see the sensitivity of the estimated welfare cost of inflation to the specification of money demand and to the definition of money with Bailey's approach. Second, the welfare cost measures in the currency-deposit model turn out to be almost the same as those in the single-monetary-asset model when the monetary base is identified as the relevant definition of money. Defining M1 without modeling the distinctive roles of currency and deposits as the appropriate money overestimates the true welfare cost of inflation in both Bailey's consumer's surplus and Lucas's compensating variation approaches. We also estimate the welfare cost of inflation using consumption as the scale variable. The welfare cost estimates as a percentage of consumption turn out to be proportionately higher (without affecting our conclusions) since the ratio of consumption to GDP is around two-thirds in the United States.

XIV. CONCLUSIONS

There are several objectives of this paper. First, we focus on the microfoundations of the double-log form of the money demand function for the U.S. economy in an intertemporal model of money and consumption. In order to find a steady state solution to the model in which the real rate of interest and the growth rate of various quantities are constant and derive its implications for seigniorage and the welfare cost of inflation, we use constant intertemporal elasticity of substitution (CIES) utility function which yields the constant elasticity double-log money demand function.

Then since the specification of the money demand function has important implications for seigniorage and the welfare cost of inflation, we reexamine two aspects of this specification for the U.S. economy. The first aspect is the choice of a scale variable: Consumption or income. The second is the choice of the functional form of money demand: Cagan (constant semi-elasticity) or the double-log (constant elasticity). While GDP is the standard scale variable for money demand in macroeconomic models, we argue that both theoretical and empirical considerations suggests that it may not be the right choice. We find that consumer expenditure is a more empirically successful scale variable in estimated money demand equations since the empirical results presented in this paper indicate that consumption spending produces more stable measures of monetary velocity and outperforms GDP in money demand estimates for the United States.

To determine what form the U.S. money demand takes, the Box-Cox transformation is applied to the nominal interest rate. As a result of the one-dimensional grid-search method, the Likelihood ratio test, the Lagrange multiplier test for the Box-Cox model, and the Wald test, we end up with the same conclusion that the demand for real money balances per capita is logarithmic in the nominal interest rate, approving the double-log form of the money demand function as the proper specification for the U.S. economy.

We also utilized the augmented Dickey-Fuller test to determine whether or not the variables employed in our steady state money demand estimates contain a unit root after selecting the most appropriate lag length using Akaike information criterion (AIC) and Schwartz Bayesian criterion (SBC). The statistical results suggest that the U.S. time series used in this study are all characterized as nonstationary I(1) variables since we find that there is a unit root in the level of the variables but there is none in their first differences. Moreover, in order to determine if the variables employed in estimated money demand equations are cointegrated, we first obtain the residual sequence from each money demand regression equation and then perform a Dickey-Fuller test on these residuals to determine whether the residual series contain a unit root. According to the Engle-Granger cointegration test, the U.S. time series used in our analysis are cointegrated of order (1.1).

In addition to the standard single-asset (monetary base or M1) specification, this paper develops a currency-deposit model that provides some insight into the implications of the analysis for the optimal rate of inflation and required reserve ratio. Our purpose in this section is to test the sensitivity of estimates of the revenue maximizing inflation rate, maximum seigniorage revenue, and the marginal welfare cost of seigniorage to the specification of the money demand function (Box-Cox, Cagan, Double-log) in the more conventional single-asset monetary model and in the currency-deposit model which accounts for variation in reserve ratios. The empirical results implies that even though the marginal welfare cost of seigniorage estimated with the double-log form is not considerably different from the one estimated with the Cagan form, there is a substantial difference in estimates of the welfare cost of inflation and in turn the average welfare cost of revenue from money creation. We also find that these measures of marginal and average social costs appear inside the plausible range for a country having an efficient explicit tax system. Taking into consideration of our estimates, the marginal welfare costs evaluated at the sample mean of seigniorage revenue (as a fraction of GDP) range from 19.32 % (with the Cagan form) to 26.86 % (with the double-log form) in the currency-deposit model.

To find the exact welfare cost of inflation in the standard single-monetary-asset model, we derive the quadratic approximation and the square root formula for welfare costs in the context of Sidrauski's intertemporal optimizing framework. We argue that the traditional quadratic approximation to the welfare cost of inflation that Bailey (and many others) have used may not be suitable for this question, and to propose an alternative formula for evaluating the welfare cost of inflationary finance. The alternative is very close to a square root formula for welfare costs. It assigns much greater benefit to the reduction of moderate inflations, and substantial gains from reducing inflation to the steady deflation of around 3 percent that would reduce nominal interest rates to zero. According to our estimates with the monetary base and the 3-month t-bill rate, moving from the inflation rate of 2.92 % (from the nominal rate of interest of 5.92 %) to the steady deflation of 3 % (to the zero nominal rate of interest) the square root formula assigns a substantial gain of 0.39 % of GDP whereas the quadratic approximation assigns 0.11 % of GDP.

This paper emphasizes that it is crucial to identify the proper specification of money demand as well as the appropriate monetary aggregate to find the exact welfare cost of inflation. The empirical results obtained from the nonlinear form of money demand with Box-Cox restriction indicate that the double-log function with constant elasticity of less than one is a more accurate characterization of the actual data since the semi-logarithmic form is rejected but the double-log form is not at the 5% level of significance. This paper also quantifies the welfare loss of deviating from a zero inflation policy and determines how much welfare the U.S. gains in moving from zero inflation to the Friedman optimal deflation rate. The welfare cost measures from Bailey's consumer's surplus and Lucas's compensating variation approaches turn out to be almost the same and compared to the constant semi-elasticity Cagan-type demand for money, the constant elasticity double-log function yields substantial costs of deviating from both a zero inflation policy and an optimal deflation policy in both the currency-deposit model and the single-asset monetary model. Using the double-log form of currency and deposit demand functions, we find the welfare cost of a sustained 4 % inflation is around 0.29 % of GDP and the welfare gain in moving from a 4 % to a zero inflation rate is computed as around 0.11 % of GDP. In the single-monetary-asset model, the welfare cost measures are not affected when the monetary base is used as the relevant definition of money but the measures obtained from the currency-deposit specification or from the monetary base are inflated to 0.89 % of GDP (welfare cost of a 4 % inflation) and 0.34 % of GDP (welfare gain in moving from $\pi = 4$ % to $\pi = 0$) when M1 is identified as the appropriate monetary aggregate. With the currency-deposit specification or with the monetary base, the additional gain in moving from zero inflation to zero nominal rate that would attain the Friedman optimum is found to be around 0.18 % of GDP whereas with M1 it is overestimated as 0.55 % of GDP.

The welfare cost estimates in this paper may seem small for people who are unfamiliar with quantitative analysis of the welfare cost of inflation. But the United States is a $8 trillion economy, so to convert the stocks and flows to dollars, we need to multiply all of them by \$8 $\times 10^{12}$. For example, while comparing the additional welfare gains (estimated with the double-log function) in moving from π=0 to i=0: 0.55 % of GDP (with M1) as opposed to 0.19 % or 0.18 % of GDP (with the base or with the currency-deposit), we imply \$44 billion as opposed to \$15.2 billion or \$14.4 billion. The estimates with the semi-log function: 0.07 % of GDP (with M1) as opposed to 0.03 % of GDP (with the base or with the currency-deposit) imply \$5.6 billion as opposed to \$2.4 billion.

ENDNOTES :

---

[1] For a detailed discussion about the procedure of introducing money into the intertemporal optimizing framework, see Obstfeld and Rogoff (1996), Turnovsky (1995), Calvo and Leiderman (1992), Eckstein and Leiderman (1992), Den Haan (1990), Blanchard and Fischer (1989), Poterba and Rotemberg (1987), McCallum (1987), Drazen (1979), Levhari and Patinkin (1968), Sidrauski (1967).

[2] Suppose that the basic utility function is $\bar{u}(c_t, l_t)$ where l_t is leisure in period t. Also suppose that shopping time in a period t is $\psi(c_t, m_t)$ with $\psi_c>0$ and $\psi_m<0$. Then if the total time available per period is normalized at 1 and x_t is used to denote labor time $l_t = 1 - x_t - \psi(c_t, m_t)$ and substitution yields $\bar{u}(c_t, 1 - x_t - \psi(c_t, m_t)) \equiv u(c_t, m_t, x_t)$.

[3] This condition implies that

$$U_{cc} = -s\, c_t^{-2a}\left[\left(\frac{1}{1-a}\right)(c_t^{1-a}-1)+\left(\frac{d}{1-b}\right)(m_t^{1-b}-1)\right]^{-s-1}$$

$$-a\, c_t^{-a-1}\left[\left(\frac{1}{1-a}\right)(c_t^{1-a}-1)+\left(\frac{d}{1-b}\right)(m_t^{1-b}-1)\right]^{-s}<0, \qquad \text{(A1)}$$

$$U_{mm} = -s\, d^2 m_t^{-2b}\left[\left(\frac{1}{1-a}\right)(c_t^{1-a}-1)+\left(\frac{d}{1-b}\right)(m_t^{1-b}-1)\right]^{-s-1}$$

$$-d\, b\, m_t^{-b-1}\left[\left(\frac{1}{1-a}\right)(c_t^{1-a}-1)+\left(\frac{d}{1-b}\right)(m_t^{1-b}-1)\right]^{-s}<0, \qquad \text{(A2)}$$

$$U_{mc} = -s\, d\, c_t^{-a} m_t^{-b}\left[\left(\frac{1}{1-a}\right)(c_t^{1-a}-1)+\left(\frac{d}{1-b}\right)(m_t^{1-b}-1)\right]^{-s-1}<0, \qquad \text{(A3)}$$

which yields

$$U_{cc}U_{mm} - U_{mc}^2 = dc_t^{-a}m_t^{-b}\left[\left(\frac{1}{1-a}\right)(c_t^{1-a} - 1) + \left(\frac{d}{1-b}\right)(m_t^{1-b} - 1)\right]^{-2s-1} \cdot$$

$$\left(s(bc_t^{-a}m_t^{-1} + adc_t^{-1}m_t^{-b}) + abc_t^{-1}m_t^{-1}\left[\left(\frac{1}{1-a}\right)(c_t^{1-a} - 1) + \left(\frac{d}{1-b}\right)(m_t^{1-b} - 1)\right]\right) > 0. \qquad \text{(A4)}$$

[4] This restricted utility function ($s=0$) is the one considered by Altonji (1981) and Blinder (1974).

[5] When $s=0$, the original utility function reduces to

$$u[(c_t), (m_t)] = \left[\left(\frac{1}{1-a}\right)(c_t^{1-a} - 1) + \left(\frac{d}{1-b}\right)(m_t^{1-b} - 1)\right], \qquad \text{(A5)}$$

where, by L'Hopital's rule,

$$lim_{a\to 1}\left[\left(\frac{1}{1-a}\right)(c_t^{1-a} - 1)\right] = lim_{a\to 1}\frac{d(c_t^{1-a} - 1)/da}{-1}$$

$$= lim_{a\to 1}\frac{-c_t^{1-a}\ln c_t}{-1}$$

$$= ln\ c_t.$$

$$lim_{b\to 1}\left[\left(\frac{d}{1-b}\right)(m_t^{1-b} - 1)\right] = d\left[\lim_{b\to 1}\frac{d(m_t^{1-b} - 1)/db}{-1}\right]$$

$$= d\left[\lim_{b\to 1}\frac{-m_t^{1-b}\ln m_t}{-1}\right]$$

$$= d\ ln\ m_t.$$

Therefore, when $a=1$, $b=1$, $s=0$, a logarithmic utility function can be written as

$$u[(c_t), (m_t)] = ln\ c_t + d\ ln\ m_t. \qquad \text{(A6)}$$

[6] In the special case of $s=0$, the marginal utility of consumption and money, respectively, are defined as $U_c = c_t^{-a} > 0$, $U_m = dm_t^{-b} > 0$ which results in $U_{cc} = -ac_t^{-a-1} < 0$, $U_{mm} = -$

$dbm_t^{-b-i} < 0$. Then the instantaneous elasticity of intertemporal substitution in consumption and money are, respectively, written as $-U_{cc}c_t/U_c = 1/a$, $-U_{mm}m_t/U_m = 1/b$.

7 We can change the capital letter version of the household's budget constraint (III.3) to the lower case (or the real per capita) version of it (III.4) as follows:

$$m = \frac{M}{PN} \Rightarrow \dot{m} = \frac{\dot{M}}{PN} - (n+\pi)\frac{M}{PN} \Rightarrow \frac{\dot{M}}{PN} = \dot{m} + (n+\pi)m.$$

$$f = \frac{F}{N} \Rightarrow \dot{f} = \frac{\dot{F}}{N} - nf \Rightarrow \frac{\dot{F}}{N} = \dot{f} + nf.$$

Substitution of $\frac{\dot{M}}{PN}$ and $\frac{\dot{F}}{N}$ (rewritten in terms of the real per capita variables above) to the capital letter version of the private budget constraint yields

$$\dot{f} + nf + \dot{m} + (n+\pi)m = r(1-\tau)f + h + w(1-\tau) - c.$$

By adding and subtracting $r(1-\tau)m$ to the right hand-side of this equation, we can get

$$\dot{f} + \dot{m} = r(1-\tau)(f+m) + h + w(1-\tau) - c - n(f+m) - m(r(1-\tau)+\pi),$$

and using the fact that $\dot{v} = \dot{f} + \dot{m}$ when $v = f + m$, we obtain the real per capita version of the private budget constraint

$$\dot{v} = r(1-\tau)v + h + w(1-\tau) - c - nv - im.$$

8 Following Barro (1989), when the marginal tax rate, τ, applies to all forms of income-including interest income and returns to capital, as well as labor earnings. Then, if the tax rate is constant over time, the after-tax real rate is $(1-\tau)r_t$, and the after-tax nominal rate is

---

$i = r(1-\tau)+\pi$. A steady state requires equality between $(1-\tau)r_t$ and the rate of time preference, ρ, thus implying that

$$r = \rho / (1-\tau). \tag{A7}$$

It follows that an increase in τ leads to a rise in r.

9 For a detailed discussion about dynamic optimization using optimal control theory, see Kamien and Schwartz (1981) pp. 111, and Chiang (1992) pp. 161.

10 See Kamien and Schwartz (1981) pp. 53, and Chiang (1992) pp. 59.

11 Taking the natural logarithm of Eq. (III.5a') we get the following expression.

$$-(\rho - n)t - a \ln c_t = \ln \Omega_t \tag{A8}$$

then the time derivative of this expression yields

$$-(\rho - n) - a\frac{c_t}{c_t} = \frac{\Omega_t}{\Omega_t} \tag{A9}$$

and using the fact that $\frac{\Omega_t}{\Omega_t} = n - r(1-\tau)$ from Equation (III.5d), we obtain the basic condition for choosing consumption over time:

$$\frac{c_t}{c_t} = \frac{r(1-\tau)-\rho}{a} \quad . \tag{A10}$$

12 For a detailed discussion about the production technology and the returns to scale, see Varian (1992) pp. 2.

13 The Inada conditions imply

$$lim_{k\to 0}[f'(k)] = \infty \quad \text{and} \quad lim_{k\to \infty}[f'(k)] = 0.$$

[14] One simple production function that is often thought to provide a reasonable description of actual economies is the Cobb-Douglas function.

$$Y = AK^{\alpha}N^{1-\alpha} \qquad \text{(A11)}$$

where $A > 0$ is the level of technology which is assumed to be constant over time, and α is a constant with $0 < \alpha < 1$. The Cobb-Douglas function can be written in intensive form as

$$y = Ak^{\alpha}. \qquad \text{(A12)}$$

Note that $f'(k) = A\alpha k^{\alpha-1} > 0$, $f''(k) = -A\alpha(1-\alpha)k^{\alpha-2} < 0$.

$$lim_{k\to 0}[f'(k)] = \infty \quad \text{and} \quad lim_{k\to\infty}[f'(k)] = 0.$$

Thus, the Cobb-Douglas form satisfies the properties of a neoclassical production function.

[15] This simplifying assumption was also used by Roubini and Sala-i-Martin (1992).

[16] Since, in this study, we do not allow the government to issue debt, we do not consider the taxation of interest earnings on bonds.

[17] Substituting the government budget constraint (III.24),

$$m = g + h - r\tau k - w\tau - m(\pi + n),$$

into the household's budget constraint (III.4), where $v = m + k$ since we assume $v = m + k$,

$$k + m = r(1-\tau)k + h + w(1-\tau) - c - nk - m(n+\pi),$$

we can get the social budget constraint as follows,

$$g + h - r\tau k - w\tau - m(\pi + n) + k = r(1-\tau)k + h + w(1-\tau) - c - nk - m(n+\pi),$$

thus implying the equation,

$$k = rk + w - c - g - nk, \quad \text{(A13)}$$

where

$$r = f'(k) - \delta,$$

$$w = f(k) - kf'(k)$$

$$g = \varepsilon c.$$

Plugging the firm's first order condition and the relation between the government spending and the private consumption presented above into (A13) yields the social budget constraint.

$$k = f(k) - (n + \delta)k - (1 + \varepsilon)c.$$

[18] For a detailed discussion about the Box-Cox transformation, see Greene (1997) pp. 479, Hamilton (1994) pp.126, Amemiya (1985) pp. 249, Spitzer (1982a), Spitzer (1982b), Poirier and Melino (1978).

[19] For a detailed discussion about the grid search method, see Greene (1997) p. 480, Hamilton (1994) pp. 133.

[20] For a detailed discussion about the nonlinear least squares methods, see Greene (1997) pp. 450, Amemiya (1985) pp. 127.

[21] Amemiya (1985) discusses several well-known methods that are especially suitable for obtaining the maximum likelihood and the nonlinear least squares estimator. The methods of iteration discussed in Amemiya (1985) are Newton-Raphson Method pp.137, Gauss-Newton Method pp. 137, Marquardt p.140, Hartley p.140, and Gallant p.141.

[22] See Greene (1997) pp. 596, Johnston and DiNardo (1997) pp.182, Gujarati (1995) pp. 605, Maddala (1992) pp. 249, Amemiya (1985) pp. 195.

[23] For a detailed discussion about the Likelihood ratio test, see Greene (1997) pp. 486, Johnston and DiNardo (1997) pp. 147, Hamilton (1994), pp. 296, Maddala (1992) pp. 119, Amemiya (1985) pp. 141.

[24] See Greene (1997) pp. 489, Johnston and DiNardo (1997) pp. 149, Hamilton (1994) pp. 430, Maddala (1992) pp. 120, Amemiya (1985) pp.142.

[25] See Greene (1997) pp. 490.

[26] For a detailed discussion about the Wald test, see Greene (1997) pp. 162, Johhnston and DiNardo (1997) pp. 148, Hamilton (1994) pp. 429, Maddala (1992) pp. 121, Amemiya (1985) pp. 144.

[27] The average number of times a unit of money is used annually is called the *income velocity (V) of money*. Velocity is computed by dividing nominal GDP by the money supply : $V = PY/M$. Multiplying both sides of this equation by M yields $MV = PY$, a result called *the equation of exchange*. This equation is interpreted as the quantity of money times its velocity is equal to the price level times real output, which equals nominal GDP. If we redefine the equation of exchange by replacing real output with the real income per capita, $MV = PyN$, we can get the per capita income elasticity of demand for money as

$$\beta = \frac{d\ln(y/V)}{d\ln y} = \frac{d\ln y - d\ln V}{d\ln y} \qquad \text{(A14)}$$

since $M/PN = y/V$ from the equation of exchange. Equation (A14) implies that the growth rate of the velocity of money equals the growth rate of output per man times one minus the income elasticity,

$dlnV/dt = (1-\beta)\ [dlny/dt]$. (A15)

Equation (A15) says that the per capita income elasticity of demand for real money balances must be unity so that the equilibrium level of velocity is independent of the growth of output.

28 This is no more than to say that a monopolist facing a demand curve with a constant elasticity less than unity can always increase her total revenue by raising her relative price.

29 For a detailed discussion on the welfare cost of inflation, see Lucas (1994).

30 Leibniz's rule says that if f is continuous on [a, b], and $u(x)$ and $v(x)$ are differentiable functions of x whose values lie in [a, b], then

$$\frac{d}{dx}\int_{u(x)}^{v(x)} f(t)dt = f(v(x))\frac{dv}{dx} - f(u(x))\frac{du}{dx}.$$

31 The income velocity of money defined as the average number of times a unit of money is used annually is computed by dividing real income (denoted by "y") by the real money balances (denoted by "m"): $V = y/m$. The income elasticity of demand can thus be written as

$$\eta = \frac{d\ln m}{d\ln y} = \frac{d\ln(y/V)}{d\ln y}$$ which implies

$$\frac{d\ln V}{dt} = (1-\eta)\frac{d\ln y}{dt}.$$

Therefore, the income elasticity must be unity so that the equilibrium level of velocity is independent of the growth of output.

[32] The rule can be implemented by having prices fall at the fixed real rate of interest so that the nominal rate of interest is zero. This fall in prices requires the authorities to implement the appropriate growth of the money supply.

[33] Within the infinite horizon model three approaches have been adopted to incorporate the role of money. The first is to incorporate its role as a medium of exchange through the so-called cash-in-advance constraint, originally proposed by Clower (1967) and then generalized by Lucas and Stokey (1983) to include "cash" and "credit" goods. The second is the "transactions-time" approach developed by McCallum (1983) and McCallum and Goodfriend (1987). The third, originally due to Sidrauski (1967), is to introduce money directly into the utility function.

[34] With the base money stock M_t and the price level P_t, the revenue from money creation is traditionally defined as the change in the nominal stock of money deflated by the price level $(M_{t-1} - M_t)/P_t$, or equivalently, the stock of real money balances held M_t/P_t times the growth rate of the monetary base $(M_{t-1} - M_t)/M_t$. This implies that when the nominal money stock is not growing, no revenue accrues to the government. Alternatively, the Phelps (1973) - Auernheimer (1974) definition uses the nominal rate of interest (rather than the growth rate of the monetary base) times stock of real balances as the revenue accruing from the authorities. The rationale for this definition is as follows. Less revenue is spent by the authorities because governments have monopoly power to hold part of their debt to the private sector in the form of non-interest bearing money rather than as interest bearing bonds. As long as the nominal rate of interest is greater than zero, this monopoly power implies that on the government debt in the form of money, no interest is

paid. This saving of interest payment is measured by the stock of non-interest bearing money times the nominal rate of interest.

[35] Assuming different semi-elasticities for currency and deposit demand schedules, $\frac{1}{\bar{m}(i)}\bar{m}'(i) = -\eta$ for currency and $\frac{1}{\bar{d}(\mu i)}\bar{d}'(\mu i) = -\varepsilon$ for deposits, the welfare cost of inflation for the semi-log function in the currency-deposit model can be written as: $w(i) = \frac{1}{2}i^2[\eta\bar{m}(0) + \varepsilon\mu^2\bar{d}(0)]$.

[36] In the restricted case where we assume the demand for currency and deposits have the same semi-elasticity, the long-run regression coefficients, obtained from the estimated asset demand equations with the semi-log form: $log\,(m/y)_t=\alpha_0-\alpha_1 i_t$ and $log\,(d/y)_t=\alpha_2-\alpha_1(\mu_t i_t)$, are found to be α_0 = -2.5990 (-9.5959), α_1 = 7.8365 (1.8886) and α_2 = -2.0713 (-15.578) with t-ratios in parentheses.

[37] Following the same procedure for finding $w(i)$ with different semi-elasticities of demand for currency (η= 3.693) and for deposits (ε = 72.317), the estimated welfare cost evaluated at the mean of the reserve ratio μ_{mean} = 0.1373 turn out to be $w(i) = (0.209)\,i^2$.

[38] In the single-monetary-asset model, the welfare cost of inflation $w(i)$ is defined as the solution to: $u[1+w(i),\ \bar{m}(i)] = u[1,\ \bar{m}(0)]$ where $\bar{m}(i)$ denote the $\bar{m}$ value that satisfies $u_m(1,\ \bar{m}) = i\,u_c(1,\ \bar{m})$. Following the same procedure illustrated in the currency-deposit model, the taylor series expansion of the function $w(i)$ about $i = 0$ gives $w(i) = \frac{1}{2}\bar{m}(0)\eta i^2$.

[39] The CES utility function presented in (XIII.8) implies

$$u(1, \infty, \infty) = \frac{\left[\left(\gamma_1^{1/\theta}\right)^{\theta/(\theta-1)}\right]^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}$$

$$u[1+w(i), \bar{m}(i), \bar{d}(\mu i)] = \frac{\left\{\left[\gamma_1^{1/\theta}(1+w(i))^{(\theta-1)/\theta} + \gamma_2^{1/\theta}\bar{m}(i)^{(\theta-1)/\theta} + \gamma_3^{1/\theta}\bar{d}(\mu i)^{(\theta-1)/\theta}\right]^{\theta/(\theta-1)}\right\}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}$$

where $\bar{m}(i) = \left(\frac{\gamma_2}{\gamma_1}\right) i_t^{-\theta}$ and $\bar{d}(\mu i) = \left(\frac{\gamma_3}{\gamma_1}\right)(\mu_t i_t)^{-\theta}$. Solving $u[1+w(i), \bar{m}(i), \bar{d}(\mu i)] = u(1, \infty, \infty)$ for $w(i)$ yields the exact welfare cost of inflation presented in (XIII.18).

[40] In the restricted case where we assume the demand for currency and deposits have the same elasticity, the long-run regression coefficients, obtained from the estimated asset demand equations with the double-log form: $\log (m/y)_t = \beta_0 - \beta_1 \log i_t$ and $\log (d/y)_t = \beta_2 - \beta_1 \log (\mu_t i_t)$, are found to be β_0 = -4.3109 (-9.1084), β_1 = 0.4240 (2.6511) and β_2 = -4.2063 (-5.0218) with t-ratios in parentheses.

[41] Lucas (1993) uses an estimate of the semi-elasticity (η=7) of demand for M1 in the single-monetary-asset model obtained from 1900-1985 annual U.S. time series to obtain the interest elasticity ($\theta = \eta i$) which turn out be about 0.5. Then he finds the welfare cost estimate as: $w(i) = (1 - 0.045 i_t^{0.5})^{-1} - 1 \cong 0.045 i_t^{0.5}$ which he calls the square-root formula.

[42] In the single-monetary-asset model, the exact welfare cost of inflation $w(i)$ is defined as the solution to: $u[1+w(i), \bar{m}(i)] = u(1, \infty)$. With the following CES utility function,

$$U(c_t, m_t) = \frac{\left\{\left[\gamma_1^{1/\theta} c_t^{(\theta-1)/\theta} + \gamma_2^{1/\theta} m_t^{(\theta-1)/\theta}\right]^{\theta/(\theta-1)}\right\}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}},$$

we find that $w(i) = \left[1 - \left(\frac{\gamma_2}{\gamma_1}\right) i_t^{1-\theta}\right]^{\theta/(\theta-1)} - 1$.

[43] Bailey (1956) assumes that real currency and real deposit demand schedules are the same, i.e., they have the same functional form and the same elasticity. As presented in Table 21, we estimate the welfare cost of inflation in the currency-deposit model with and without Bailey's assumption. We also present in Table 22 the welfare cost measures estimated in the single-monetary-asset model with Bailey's consumer's surplus approach.

TABLE 1 : THE STANDARD DEVIATION OF VELOCITY, 1957:1 - 1994:I

MONETARY AGGREGATES	CONSUMPTION (LEVEL)	CONSUMPTION (FIRST-DIFFERENCE)	INCOME (LEVEL)	INCOME (FIRST-DIFFERENCE)
M1	2. 4344	6. 7211	2. 6247	6. 9514
M2	1. 9558	3. 8947	2. 1670	4. 2456
MB	6. 4115	11. 7438	6. 5878	11. 9974
CURRENCY	22. 8210	38. 4551	22. 9484	38. 6967
DEMAND DEPOSITS	5. 7834	7. 2807	5. 8473	7. 3878
TIME DEPOSITS	1. 7249	11. 5307	1. 8512	11. 6153
RESERVES	7.7207	6. 2242	7. 8610	6. 3992

Note: Level is detrended by regressing the log of velocity on time and the square of time; entries are standard errors of the regression, multiplied by 100, and can thus be interpreted as percentages. Similarly, first-differenced columns are the standard deviation of the change in the log of velocity multiplied by 100. The data are quarterly.

TABLE 2 : ONE DIMENSIONAL GRID-SEARCH METHOD: (SSR for λ from 0 to 1)

λ	SSR(CONSUMPTION)	SSR(INCOME)
0.00	**0.531735**	**0.545763**
0.01	0.532009	0.546087
0.02	0.532284	0.546411
0.03	0.532559	0.546737
0.04	0.532835	0.547063
0.05	0.533111	0.547391
0.06	0.533388	0.547719
0.07	0.533665	0.548048
0.08	0.533942	0.548378
0.09	0.534220	0.578709
0.10	0.534498	0.549040
0.20	0.537286	0.552374
0.30	0.540063	0.555719
0.40	0.542800	0.559034
0.50	0.545472	0.562284
0.60	0.548061	0.565445
0.70	0.550552	0.568495
0.80	0.552935	0.571420
0.90	0.555204	0.574209
1.00	**0.557355**	**0.576856**

Note : To determine the proper specification of the money demand function for steady state seigniorage and the welfare cost analysis, one-dimensional grid-search method is applied to our fundamental money demand equation (44) using the nonlinear least squares estimation technique. The nonlinear estimation of (44) is carried out by using monetary base (MB) and 3-month t-bill rate. The results indicate that the double-log specification with consumption gives the best fit for the U.S. economy.

TABLE 3 : STEADY STATE MB DEMAND ESTIMATES, 1957:I - 1994:I

TBILL RATE	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	INCOME (SEMI-LOG)	INCOME (LOG-LOG)
Interest Elasticity	- 0.0064 (-2.973)	- 0.0517 (- 4.031)	- 0.7511 (- 3.298)	- 0.0607 (- 4.445)
Lagged MB	0.9144 (33.895)	0.8927 (32.605)	0.8875 (28.984)	0.8583 (27.474)
SSR	0.5574	0.5317	0.5769	0.5457
Log Likelihood	203.0477	206.5300	200.5028	204.6030
D.W	2.044	2.110	1.976	2.043
h statistic	- 0.283	- 0.710	0.157	- 0.283
10-YEAR GOVBOND YIELD	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	INCOME (SEMI-LOG)	INCOME (LOG-LOG)
Interest Elasticity	- 1.0506 (- 4.037)	- 0.0998 (- 5.208)	- 1.2118 (-4.422)	- 0.1141 (- 5.678)
Lagged MB	0.8663 (27.308)	0.8316 (25.690)	0.8300 (23.188)	0.7880 (21.672)
SSR	0.5316	0.4981	0.5464	0.5073
Log Likelihood	206.5492	211.2577	204.5105	210.0085
D.W	2.035	2.074	1.963	2.002
h statistic	- 0.231	- 0.490	0.250	- 0.014

Note : Since our main purpose is to set out a more accurate procedure to use in estimating seigniorage revenue and the welfare cost of inflation as a fraction of GDP in steady state, the per capita income elasticity of demand for real money balances is assumed to be unity. Therefore, the monetary aggregate MB is simply deflated by the nominal income. This table presents the steady state estimates of the demand for monetary base (reserves plus currency) using a short-term (3-month t-bill rate) and a long-term interest rate (10-year government bond yield) between 1957:I and 1994:I for the U.S. economy.

TABLE 4 : STEADY STATE M1 DEMAND ESTIMATES, 1957:I - 1994:I

TBILL RATE	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	INCOME (SEMI-LOG)	INCOME (LOG-LOG)
Interest Elasticity	- 0.3437 (- 3.755)	- 0.0280 (-4.984)	- 0.3879 (- 3.673)	- 0.0324 (-4.892)
Lagged M1	0.9625 (87.614)	0.9487 (82.824)	0.9528 (70.864)	0.9338 (65.080)
SSR	0.0787	0.0737	0.0918	0.0861
Log Likelihood	347.9070	352.7394	336.5289	341.2507
D.W	2.106	2.241	1.924	2.038
h statistic	- 0.651	- 1.48	0.469	- 0.235
10-YEAR GOVBOND YIELD	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	INCOME (SEMI-LOG)	INCOME (LOG-LOG)
Interest Elasticity	-0.4351 (-3.366)	-0.0553 (-5.093)	-0.4801 (-3.175)	- 0.0668 (-5.067)
Lagged M1	0.9478 (62.894)	0.9104 (51.881)	0.9373 (50.131)	0.8835 (39.107)
SSR	0.0801	0.0732	0.0938	0.0852
Log Likelihood	346.6090	353.2197	334.9213	342.0117
D.W	2.063	2.136	1.887	1.923
h statistic	- 0.390	- 0.847	0.706	0.487

Note : Since seigniorage and the welfare cost analysis are compatible with a steady state only if the velocity of money remains constant which requires a unitary per capita income elasticity of demand for real balances. Thus, the monetary aggregate M1 is simply deflated by the nominal income. This table presents the steady state estimates of the demand for M1 (currency plus demand deposits and checkable deposits) using a short-term (3-month t-bill rate) and a long-term interest rate (10-year government bond yield) between 1957:I and 1994:I for the U.S. economy.

TABLE 5 : LONG-RUN STEADY STATE INTEREST RATE SEMI-ELASTICITIES AND ELASTICITIES

TBILL RATE	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	INCOME (SEMI-LOG)	INCOME (LOG-LOG)
MB	- 0.0747	- 0.4820	- 6.6765	- 0.4282
M1	- 9.1572	- 0.5448	- 8.2261	- 0.4898
10-YEAR GOVBOND YIELD	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	INCOME (SEMI-LOG)	INCOME (LOG-LOG)
MB	- 7.8571	- 0.5927	- 7.1297	- 0.5384
M1	- 8.3289	- 0.6173	- 7.6535	- 0.5736

Note : This table presents the long-run interest rate semi-elasticities (for the Cagan form) and elasticities (for the double-log form) of demand for the monetary aggregates (MB and M1).

TABLE 6 : MB DEMAND ESTIMATES, 1957:I - 1994:I

TBILL RATE	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	INCOME (SEMI-LOG)	INCOME (LOG-LOG)
Interest Elasticity	- 0.6685 (- 3.474)	- 0.0447 (-3.772)	- 0.6908 (- 3.448)	- 0.0480 (- 3.802)
Quantity Elasticity	0.0359 (1.382)	0.0518 (1.875)	0.0418 (1.411)	0.0641 (1.994)
Lagged MB	0.7743 (20.254)	0.7800 (20.899)	0.7711 (19.824)	0.7744 (20.406)
RSS	0.4449	0.4388	0.4446	0.4374
Log Likelihood	219.7303	220.7488	219.7703	220.9781
D.W.	2.077	2.123	2.074	2.122
h statistic	- 0.529	- 0.840	- 0.511	- 0.837
10-YEAR GOVBOND YIELD	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	INCOME (SEMI-LOG)	INCOME (LOG-LOG)
Interest Elasticity	- 0.9099 (- 3.817)	- 0.0759 (- 3.954)	- 0.9421 (- 3.754)	- 0.0814 (- 3.933)
Quantity Elasticity	0.0734 (2.350)	0.0988 (2.788)	0.0828 (2.306)	0.1170 (2.805)
Lagged MB	0.7498 (18.766)	0.7626 (19.887)	0.7446 (18.115)	0.7545 (19.109)
RSS	0.4378	0.4349	0.4384	0.4346
Log Likelihood	220.9095	221.4030	220.8078	221.4488
D.W.	2.057	2.080	2.050	2.072
h statistic	- 0.397	- 0.550	- 0.347	- 0.496

Note : The empirical results presented in this paper suggests that consumption is a better scale variable in the money demand function than GDP since consumer spending produces more stable measures of monetary velocity and outperforms GDP in estimated money demand equations. Therefore, in addition to estimating U.S. money demand with the assumption of a unitary income elasticity for seigniorage and the welfare cost analysis in steady state, we estimate the most general regression equation given by Equation (43) to compare the elasticity of demand for money with respect to income and consumption. This table presents the quantity (income and consumption) elasticity of demand for MB.

TABLE 7 : M1 DEMAND ESTIMATES, 1957:I - 1994:I

TBILL RATE	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	GDP (SEMI-LOG)	GDP (LOG-LOG)
Interest Elasticity	- 0.0066 (-7.0037)	- 0.0395 (- 7.239)	- 0.6826 (- 6.924)	- 0.0424 (-7.276)
Quantity Elasticity	0.0399 (4.920)	0.0507 (4.838)	0.0442 (3.972)	0.0595 (4.880)
Lagged MB	0.8576 (32.458)	0.8769 (36.280)	0.8544 (31.821)	0.8709 (35.469)
RSS	0.0622	0.0611	0.0625	0.0609
Log Likelihood	365.3694	366.6420	364.9225	366.8226
D.W.	2.071	2.213	2.053	2.208
h statistic	- 0.456	- 1.355	- 0.341	- 1.326
10-YEAR GOVBOND YIELD	CONSUMPTION (SEMI-LOG)	CONSUMPTION (LOG-LOG)	GDP (SEMI-LOG)	GDP (LOG-LOG)
Interest Elasticity	- 0.8700 (- 5.754)	- 0.0748 (- 6.694)	- 0.8442 (- 5.2300)	- 0.0779 (- 6.256)
Quantity Elasticity	0.0700 (4.829)	0.0989 (5.960)	0.0712 (4.223)	0.1105 (5.485)
Lagged MB	0.8153 (22.862)	0.8194 (26.065)	0.8220 (22.163)	0.8158 (24.463)
RSS	0.0678	0.0635	0.0700	0.0655
Log Likelihood	358.9914	363.7245	356.5254	361.4494
D.W.	1.876	1.982	1.835	1.924
h statistic	0.837	0.118	1.125	0.512

Note : The empirical results presented in this paper suggests that consumption is a better scale variable in the money demand function than GDP since consumer spending produces more stable measures of monetary velocity and outperforms GDP in estimated money demand equations. Therefore, in addition to estimating U.S. money demand with the assumption of a unitary income elasticity for seigniorage and the welfare cost analysis in steady state, we estimate the most general regression equation given by Equation (43) to compare the elasticity of demand for money with respect to income and consumption. This table presents the quantity (income and consumption) elasticity of demand for M1.

TABLE 8a: UNIT ROOT TEST FOR THE LEVEL OF THE VARIABLES

VARIABLES (LEVEL)	LAGGED DIFFERENCES (1)			LAGGED DIFFERENCES (4)			LAGGED DIFFERENCES (8)		
	t-ratio	*AIK*	*SBC*	*t-ratio*	*AIK*	*SBC*	*t-ratio*	*AIK*	*SBC*
log(MB/GDP)*	-5.7726	-5.5995	-5.5181	-1.1711	-7.2116	-7.0672	0.5918	-7.9256	-7.6944
log(MB/CON)*	-6.1910	-5.6844	-5.6031	-1.3649	-7.3105	-7.1662	0.6926	-8.1115	-7.8804
log(M1/GDP)	0.6415	-7.2538	-7.2131	1.7040	-8.0499	-7.9468	0.7930	-8.1072	-7.9181
log(M1/CON)	0.9701	-7.4009	-7.3602	1.7816	-8.2560	-8.1528	0.6568	-8.2762	-8.0871
t-bill rate	-0.9397	-9.5535	-9.5128	-0.7522	-9.6683	-9.5652	-0.6853	-9.7384	-9.5493
log(t-bill rate)*	-2.4907	-3.9739	-3.9128	-2.6350	-4.3183	-4.1946	-1.6303	-4.5651	-4.3550
10govbond	-0.3161	-10.582	-10.542	-0.3200	-10.549	-10.446	-0.1203	-10.525	-10.335
log(10govbond)	-0.7679	-5.6238	-5.5831	-1.1137	-5.6572	-5.5540	-0.8105	-5.6773	-5.4882
MacKinnon	5 %		1 %	5 %		1 %	5 %		1 %
Critical values	-1.9420		-2.5796	-1.9421		-2.5799	-1.9421		-2.5803

* Since we used trend and intercept in log(MB/GDP) and log(MB/CON) and only intercept in log(t-bill rate) the critical values for log(MB/GDP) and log(MB/CON) are -4.0228 (1%), -3.4409 (5%) for the first lagged difference and -4.0241 (1%), -3.4415 (5%) for the 4th lagged differences and -4.0259 (1%), -3.4424 (5%) for 8th lagged differences and for log(t-bill rate) -3.4758 (1%), -2.8811 (5%) for the first lagged difference and -3.4767 (1%), -2.8815 (5%) for the 4th lagged differences and -3.4779 (1%), -2.8821 (5%).

TABLE 8b: UNIT ROOT TEST FOR THE FIRST DIFFERENCE OF THE VARIABLES

VARIABLES (LEVEL)	LAGGED DIFFERENCES (1)			LAGGED DIFFERENCES (4)			LAGGED DIFFERENCES (8)		
	t-ratio	*AIK*	*SBC*	*t-ratio*	*AIK*	*SBC*	*t-ratio*	*AIK*	*SBC*
log(MB/GDP)	-8.8868	-5.4081	-5.3672	-13.779	-7.3398	-7.2362	-3.1051	-7.9180	-7.7280
log(MB/CON)	-8.8100	-5.4621	-5.4213	-14.278	-7.4576	-7.3540	-2.9251	-8.0695	-7.8795
log(M1/GDP)	-7.6870	-7.2502	-7.2093	-5.5702	-8.0420	-7.9384	-2.6324	-8.1518	-7.9618
log(M1/CON)	-7.7017	-7.3981	-7.3572	-5.2847	-8.2419	-8.1383	-2.1930	-8.3251	-8.1351
t-bill rate	-10.345	-9.6321	-9.5912	-4.2215	-9.7076	-9.6040	-4.3771	-9.7334	-9.5434
log(t-bill rate)	-10.915	-4.1130	-4.0722	-5.9913	-4.3391	-4.2355	-4.4714	-4.5559	-4.3659
10govbond	-7.7397	-10.579	-10.538	-5.5373	-10.571	-10.468	-4.2214	-10.519	-10.329
log(10govbond)	-8.3091	-5.6356	-5.5947	-5.4820	-5.6529	-5.5493	-4.1781	-5.6695	-5.4795
MacKinnon Critical values	5 % -1.9420		1 % -2.5797	5 % -1.9421		1 % -2.5800	5 % -1.9422		1 % -2.5804

TABLE 9: COINTEGRATION TEST FOR THE VARIABLES EMPLOYED IN STEADY STATE ANALYSIS

VARIABLES (LEVEL)	LAGGED DIFFERENCES (0)			LAGGED DIFFERENCES (1)			LAGGED DIFFERENCES (4)		
	t-ratio	*AIK*	*SBC*	*t-ratio*	*AIK*	*SBC*	*t-ratio*	*AIK*	*SBC*
log(MB/GDP) t-bill rate	-12.018	-5.5337	-5.5133	-8.8589	-5.5086	-5.4473	-8.9602	-7.0909	-6.9873
log(MB/GDP) log(t-bill rate)	-12.489	-5.5949	-5.5746	-9.2563	-5.5906	-5.5497	-6.7934	-6.9829	-6.8793
log(MB/GDP) 10govbond	-12.079	-5.5994	-5.5790	-9.1431	-5.6006	-5.5597	-7.9579	-7.1347	-7.0311
log(MB/GDP) log(10govbond)	-12.551	-5.6933	-5.6729	-9.7818	-5.7185	-5.6776	-6.5268	-7.1442	-7.0406
log(M1/GDP) t-bill rate	-11.632	-7.3666	-7.3463	-7.6635	-7.3536	-7.3127	-6.3426	-8.2443	-8.1407
log(M1/GDP) log(t-bill rate)	-12.314	-7.4294	-7.4090	-8.0505	-7.4139	-7.3731	-5.8111	-8.2361	-8.1326
log(M1/GDP) 10govbond	-11.417	-7.3467	-7.3263	-7.4413	-7.3375	-7.2966	-5.8625	-8.2159	-8.1123
log(M1/GDP) log(10govbond)	-11.650	-7.4426	-7.4222	-7.3611	-7.4423	-7.4015	-4.9142	-8.2743	-8.1707
log(MB/CON) t-bill rate	-12.418	-5.5672	-5.5469	-8.8727	-5.5526	-5.5118	-9.8110	-7.1958	-7.0922
log(MB/CON) log(t-bill rate)	-12.882	-5.6209	-5.6005	-9.2201	-5.6113	-5.5705	-7.5290	-7.0653	-6.9617
log(MB/CON) 10govbond	-12.463	-5.6230	-5.6026	-9.0459	-5.6176	-5.5767	-8.4940	-7.1985	-7.0949
log(MB/CON) log(10govbond)	-12.890	-5.7041	-5.6838	-9.5423	-5.7156	-5.6747	-6.8234	-7.1746	-7.0710
log(M1/CON) t-bill rate	-12.741	-7.5219	-7.5016	-8.0131	-7.5090	-7.4682	-6.4855	-8.4583	-8.3547
log(M1/ CON) log(t-bill rate)	-13.648	-7.5993	-7.5789	-8.5782	-7.5826	-7.5417	-5.8998	-8.4615	-8.3579
log(M1/ CON) 10govbond	-12.481	-7.5029	-7.4825	-7.7543	-7.4945	-7.4536	-6.0126	-8.4394	-8.3358
log(M1/ CON) log(10govbond)	-12.970	-7.5982	-7.5779	-7.9182	-7.5939	-7.5530	-5.1483	-8.4940	-8.3904
MacKinnon Critical values	5 % -1.9420		1 % -2.5797	5 % -1.9420		1 % -2.5797	5 % -1.9421		1 % -2.5800

TABLE 10 : STEADY STATE SEIGNIORAGE AND WELFARE COST ESTIMATES

FUNCTIONAL FORM	i_{mean}	i_{max}	$(S/y)_{mean}$	$(S/y)_{max}$	$(WC/y)_{Smean}$	AWC_{Smean}	MWC_{Smean}	(S/y) at i_{mean}	(WC/y) at i_{mean}	AWC at i_{mean}	MWC at i_{mean}
Nonlinear	0.059189	0.016236	0.003850	0.002283	------------	------------	0.503819	0.003751	------------	------------	0.529153
Double-log	0.059189	------------	0.003850	------------	0.002884	0.748917	0.748917	0.003873	0.002901	0.748917	0.748917
Cagan	0.059189	0.149779	0.003850	0.005628	0.000787	0.204458	0.565433	0.004072	0.000922	0.226404	0.653373

i_{mean} : sample mean of the nominal interest rate

i_{max} : revenue maximizing nominal rate of interest

$(S/y)_{mean}$: sample mean of seigniorage revenue as a share of real income

$(S/y)_{max}$: seigniorage revenue as a share of real income evaluated at i_{max}

$(WC/y)_{Smean}$: welfare cost of inflation as a share of real income evaluated at $(S/y)_{mean}$

AWC_{Smean} : average welfare cost of seigniorage evaluated at $(S/y)_{mean}$

MWC_{Smean} : marginal welfare cost of seigniorage evaluated at $(S/y)_{mean}$

(S/y) at i_{mean} : seigniorage revenue as a share of real income evaluated at i_{mean}

(WC/y) at i_{mean} : welfare cost of inflation as a share of real income evaluated at i_{mean}

AWC at i_{mean} : average welfare cost of seigniorage evaluated at i_{mean}

MWC at i_{mean} : marginal welfare cost of seigniorage evaluated at i_{mean}

TABLE 11: EFFECT OF RESERVE REQUIREMENT ON ESTIMATED SEIGNIORAGE

FUNCTIONAL FORM	*Mean of ε* *Monetary Base*	*Std. Dev. of ε* *Monetary Base*	*Mean of ε* *M1*	*Std. Dev. of ε* *M1*	*Mean of ε* Currency & Deposits	*Std. Dev. of ε* Currency & Deposits
Nonlinear	0.0000081	0.000146	0.0000464	0.000269	0.0000222	0.000138
Double-log	0.0000102	0.000148	0.0000470	0.000273	0.0000231	0.000142
Cagan	0.0000039	0.000147	0.0000447	0.000273	0.0000138	0.000140

ε : forecast error of estimated seigniorage; $\varepsilon = S_{actual} - S_{estimated}$.

The smallest variance (or standard deviation) of the forecast error is obtained from the estimation which utilizes currency and deposit demand instead of monetary base or M1.

TABLE 12 : STEADY STATE SEIGNIORAGE AND WELFARE COST ESTIMATES

IN A MODEL OF CURRENCY & DEPOSITS

Functional Form	i_{mean}	i_{max}	μ_{mean}	μ_{max}	S_{mean}	S_{max}	i_{max} at μ_{mean}	S_{max} at μ_{mean}	AWC_{max} at μ_{mean}	AWC at S_{mean}	MWC at S_{mean}	AWC_{min} at S_{mean}	MWC_{min} at S_{mean}
Nonlinear	0.0592	0.2430	0.0431	1	0.0039	0.0117	0.2006	0.0117	———	———	0.2101	———	6.3823
Double-log	0.0592	———	0.0431	———	0.0039	———	———	———	———	0.2982	0.2686	———	———
Cagan	0.0592	0.2658	0.0431	0.1006	0.0039	0.0118	0.3749	0.0108	0.7731	0.0863	0.1937	0.0723	0.1602

i_{mean} : sample mean of the nominal interest rate

i_{max} : revenue maximizing nominal rate of interest when μ is flexible

μ_{mean} : sample mean of the required reserve ratio

μ_{max} : revenue maximizing required reserve ratio

S_{mean} : sample mean of seigniorage revenue as a share of real income

S_{max} : seigniorage revenue as a share of real income evaluated at i_{max}

i_{max} **at** μ_{mean} : revenue maximizing nominal rate of interest evaluated at μ_{mean}

S_{max} **at** μ_{mean} : seigniorage revenue as a share of real income evaluated at i_{max} and μ_{mean}

AWC_{max} **at** μ_{mean} : average welfare cost of seigniorage evaluated at i_{max} and μ_{mean}

AWC **at** S_{mean} : average welfare cost of seigniorage evaluated at S_{mean} and μ_{mean}

MWC **at** S_{mean} : marginal welfare cost of seigniorage evaluated at S_{mean} and μ_{mean}

AWC_{min} **at** S_{mean} : minimum average welfare cost of seigniorage evaluated at S_{mean} and μ_{mean}

MWC_{min} **at** S_{mean} : minimum marginal welfare cost of seigniorage evaluated at S_{mean} and μ_{mean}

TABLE 13: WELFARE COST OF INFLATION AS A PERCENTAGE OF GDP

Interest Rate	Quadratic Formula (MB)	Square Root (MB)	Quadratic Formula (M1)	Square Root (M1)
i	$100\,(0.302)\,i^2$	$100\,(0.0154)\,i^{0.482}$	$100\,(1.163)\,i^2$	$100\,(0.052)\,i^{0.362}$
0.00	0.0000	0.0000	0.0000	0.0000
0.0102*	0.0031	0.1689	0.0121	0.9888
0.02	0.0121	0.2337	0.0465	1.2618
0.03	0.0272	0.2841	0.1047	1.4612
0.04	0.0483	0.3264	0.1861	1.6216
0.05	0.0755	0.3634	0.2908	1.7581
0.0592*	0.1058	0.3942	0.4074	1.8688
0.07	0.1480	0.4274	0.5698	1.9858
0.08	0.1933	0.4558	0.7443	2.0841
0.09	0.2446	0.4825	0.9420	2.1749
0.10	0.3020	0.5076	1.1630	2.2595
0.1509*	0.6877	0.6189	2.6482	2.6223
0.25	1.8875	0.7895	7.2688	3.1482
0.50	7.5500	1.1026	29.075	4.0460
1.00	30.200	1.5400	116.30	5.2000
2.00	120.80	2.1509	465.20	6.6831

* $i_{min} = 0.0102$, $i_{mean} = 0.0592$, $i_{max} = 0.1509$

Note: This table presents the welfare cost estimates obtained from the quadratic and the square root formula using Monetary Base (MB) and M1.

TABLE 14A: STEADY STATE MB AND M1 DEMAND ESTIMATES (η=1)

Money/Scale	Specification	α_0 β_0	α_1 β_1	R^2	Log-L	h-stat.	ARCH(1)
MB/GDP	Semi-log	-2.2773 (-20.022)	6.7554 (3.8861)	0.9166	206.0628	0.4518	0.0039
MB/GDP	Double-log	-3.9583 (-16.702)	0.4351 (5.4397)	0.9215	210.5427	0.0167	0.0004
MB/CONS	Semi-log	-1.7986 (-11.937)	7.5873 (3.2680)	0.9304	208.7483	0.0074	0.0032
MB/CONS	Double-log	-3.6945 (-11.442)	0.4917 (4.5121)	0.9339	212.5693	-0.4244	0.0001
M1/GDP	Semi-log	-1.2184 (-11.363)	8.0686 (4.7291)	0.9847	321.8426	-1.4212	0.0354
M1/GDP	Double-log	-3.1056 (-14.785)	0.4830 (6.9064)	0.9856	326.0551	-2.0757	0.0049
M1/CONS	Semi-log	-0.7579 (-5.8907)	8.9142 (4.2439)	0.9878	329.4777	-2.7851	0.0017
M1/CONS	Double-log	-2.8392 (-10.556)	0.5336 (5.9982)	0.9884	333.4981	-3.5311	0.0785

Note: This table presents the steady state estimates of the demand for the monetary base (MB) and M1 using quarterly data for the period 1957:I and 1994:II. The quantity (income or consumption) elasticity of money demand (η) is assumed unity. α's are the coefficients in the semi-log function and β's are the coefficients in the double-log function. t-ratios are shown in parentheses. Since the h-statistics are less than 1.96 or greater than -1.96, there is no serial correlation in estimated money demand equations with the monetary base but there is with M1 at the 5% level of significance. The critical value for testing ARCH(1) is $\chi_{(0.01)}$ = 6.63 which implies we do not have ARCH(1) errors in our estimations at the 1% level of significance.

TABLE 14B: NONLINEAR STEADY STATE MB AND M1 DEMAND ESTIMATES (η=1)

Money/Scale	α_0	α_1	λ	R^2	Log-L	h-stat.	ARCH(1)
MB/GDP	-3.1244 (-16.037)	0.0292 (0.7910)	-0.8749 (-2.2898)	0.9243	213.2518	-1.1954	0.1347
MB/CONS	-2.7222 (-11.855)	0.0265 (0.7000)	-0.9440 (-2.2025)	0.9362	215.1750	-1.6341	0.1414
M1/GDP	-2.4682 (-6.8241)	0.1254 (0.8925)	-0.4430 (-1.2436)	0.9858	326.9109	-2.6306	0.2889
M1/CONS	-2.0838 (-5.1686)	0.1182 (0.8154)	-0.4946 (-1.2753)	0.9886	334.4365	-4.1342	0.6659

Note: This table presents the nonlinear steady state estimates of the demand for the monetary base (MB) and M1 using quarterly data for the period 1957:I and 1994:II. The quantity (income or consumption) elasticity of money demand (η) is assumed unity. t-ratios are shown in parentheses. Since the h-statistics are less than 1.96 or greater than -1.96, there is no serial correlation in estimated money demand equations with the monetary base but there is with M1 at the 5% level of significance. The critical value for testing ARCH(1) is $\chi_{(0.01)}$ = 6.63 which implies we do not have ARCH(1) errors in our estimations at the 1% level of significance.

TABLE 15A: MB AND M1 DEMAND ESTIMATES ($\eta \neq 1$)

Money/S	Specificat.	α_0 / β_0	α_1 / β_1	α_2 / β_2	R^2	Log-L	h-stat.	ARCH
MB/GDP	Semi-log	9.6911 (5.9697)	2.9509 (3.8743)	0.4995 (7.3944)	0.9025	224.1146	-0.1004	0.9965
MB/GDP	Double-Log	7.4395 (4.2164)	0.2122 (4.4653)	0.5592 (8.0884)	0.9062	226.9691	-0.5350	1.2211
MB/CON	Semi-log	11.125 (6.9536)	2.9274 (3.6237)	0.4492 (6.6361)	0.9040	225.5792	-0.4211	1.0802
MB/CON	Double-Log	8.2471 (4.0623)	0.2179 (3.9117)	0.5266 (6.6207)	0.9055	227.2676	-0.8073	1.1497
M1/GDP	Semi-log	8.3566 (5.2310)	4.9269 (5.7710)	0.6016 (8.9438)	0.9709	342.5964	-2.0678	0.0810
M1/GDP	Double-Log	5.7295 (3.4765)	0.3146 (6.7444)	0.6594 (10.364)	0.9729	347.9468	-3.1593	1.2113
M1/CON	Semi-log	10.659 (9.7439)	4.3839 (6.9501)	0.5149 (11.013)	0.9702	346.4225	-2.4887	1.0262
M1/CON	Double-Log	7.8895 (5.3592)	0.2955 (6.4402)	0.5837 (10.143)	0.9709	348.2458	-3.4382	3.9469

Note: This table presents the long-run estimates of the demand for the monetary base (MB) and M1 using quarterly data for the period 1957:I and 1994:II. The quantity (income or consumption) elasticity of money demand (η) is not unity. α's are the coefficients in the semi-log function and β's are the coefficients in the double-log function. t-ratios are shown in parentheses. Since the h-statistics are less than 1.96 or greater than -1.96, there is no serial correlation in estimated money demand equations with the monetary base but there is with M1 at the 5% level of significance. The critical value for testing ARCH(1) is $\chi_{(0.01)}$ = 6.63 which implies we do not have ARCH(1) errors in our estimations at the 1% level of significance.

TABLE 15B: NONLINEAR MB AND M1 DEMAND ESTIMATES ($\eta \neq 1$)

Money/S	α_0	α_1	α_2	λ	R^2	Log-L	h-stat.	ARCH
MB/GDP	7.3431 (4.0913)	0.0249 (0.7374)	0.5779 (8.0341)	-0.7124 (-1.7244)	0.9077	228.2060	-1.3207	1.6690
MB/CON	9.0591 (4.8455)	0.0330 (0.6299)	0.5163 (6.7747)	-0.6127 (-1.2406)	0.9068	227.8463	-1.2970	1.3932
M1/GDP	5.7077 (3.4030)	0.2829 (1.2824)	0.6621 (9.7945)	-0.0373 (-0.1390)	0.9729	347.9574	-3.2379	1.3525
M1/CON	7.9819 (5.8505)	0.6511 (1.1762)	0.5623 (10.014)	0.2850 (0.9319)	0.9711	348.7152	-3.1590	2.5537

Note: This table presents the nonlinear long-run estimates of the demand for the monetary base (MB) and M1 using quarterly data for the period 1957:I and 1994:II. The quantity (income or consumption) elasticity of money demand (η) is not unity. t-ratios are shown in parentheses. Since the h-statistics are less than 1.96 or greater than -1.96, there is no serial correlation in estimated money demand equations with the monetary base but there is with M1 at the 5% level of significance. The critical value for testing ARCH(1) is $\chi_{(0.01)}$ = 6.63 which implies we do not have ARCH(1) errors in our estimations at the 1% level of significance.

TABLE 16: TESTING THE PROPER SPECIFICATION OF MONEY DEMAND

	WALD	WALD	LR	LR	LM	LM
Money/Scale	Double-log	Semi-log	Double-log	Semi-log	Double-log	Semi-log
MB/GDP	2.9736	17.183	2.4738	8.1828	2.1078	4.1824
MB/CONS	1.5390	10.663	0.5787	4.5342	0.8728	4.2315
M1/GDP	0.0193	14.939	0.0212	10.722	0.0226	9.0597
M1/CONS	0.8685	5.4643	0.9388	4.5854	0.9605	4.1658

Note: This table presents the Wald, the Likelihood Ratio (LR) and the Lagrange Multiplier (LM) test statistics for determining whether the proper specification of demand for real money balances for the U.S. economy is the double-log or the semi-log form. The semi-logarithmic form is rejected but the double-log form is not at the 5 % level of significance since the critical value is $\chi_{(0.05)} = 3.84$.

TABLE 17: THE QUANTITY ELASTICITY OF DEMAND FOR MONEY

	Semi-log	Double-log	Nonlinear	Semi-log	Double-log	Nonlinear
Scale variable	MB	MB	MB	M1	M1	M1
GDP	0.4995 (7.3944)	0.5592 (8.0884)	0.5779 (8.0341)	0.6016 (8.9438)	0.6594 (10.634)	0.6621 (9.7945)
CONS	0.4492 (6.6361)	0.5266 (6.6207)	0.5163 (6.7747)	0.5149 (11.013)	0.5837 (10.143)	0.5623 (10.014)

Note: This table presents the estimates of the quantity (income or consumption) elasticity of demand for the monetary base (MB) and M1 using quarterly data for the period 1957:I and 1994:II for the U.S. economy. For each specification and each monetary aggregate, the quantity elasticity of money demand is estimated well below one. t-ratios are shown in parentheses.

TABLE 18: TESTING THE UNITARY QUANTITY ELASTICITY OF MONEY DEMAND

MONEY/SCALE	SPECIFICATION	WALD TEST	LR TEST
MB/GDP	Semi-log	54.896	36.104
MB/GDP	Double-log	40.654	32.853
MB/GDP	Nonlinear	34.438	29.908
MB/CONS	Semi-log	66.239	33.662
MB/CONS	Double-log	35.437	28.398
MB/CONS	Nonlinear	40.276	25.343
M1/GDP	Semi-log	35.083	41.508
M1/GDP	Double-log	28.661	43.783
M1/GDP	Nonlinear	24.979	42.093
M1CONS	Semi-log	107.62	33.890
M1CONS	Double-log	60.784	28.557
M1/CONS	Nonlinear	52.335	29.495

Note: This table presents the Wald and the Likelihood Ratio (LR) test statistics for testing the unitary quantity (income or consumption) elasticity of demand for the monetary base (MB) or M1. For each specification and each monetary aggregate, the unitary quantity elasticity of demand for MB or M1 is rejected at the 1% level of significance since the critical value is $\chi_{(0.01)} = 6.63$.

TABLE 19: STEADY STATE MONEY DEMAND ESTIMATES

Money/Scale	Specification	α_0 β_0	α_1 β_1	R^2	Log-L	h-stat.	ARCH(1)
CURRENCY/ GDP	Semi-log	-2.8738 (-63.750)	3.6930 (5.3220)	0.8963	301.8255	-1.3580	0.0418
CURRENCY/ GDP	Double-log	-3.7786 (-44.488)	0.2335 (8.1587)	0.9034	307.5597	-1.5507	0.1479
CURRENCY/ CONS	Semi-log	-2.4167 (-44.056)	3.8785 (4.5741)	0.9100	307.8868	-2.2179	0.1426
CURRENCY/ CONS	Double-log	-3.3707 (-31.810)	0.2466 (6.9084)	0.9157	313.1096	-2.5618	0.4442
DEPOSITS/ GDP	Semi-log	-1.6788 (-6.1138)	72.317 (1.8196)	0.9914	365.1336	-2.3997	1.0901
DEPOSITS / GDP	Double-log	-5.1160 (-2.8566)	0.5743 (1.7077)	0.9914	365.0295	-2.4957	1.2193
DEPOSITS / CONS	Semi-log	-1.1866 (-3.1346)	99.852 (1.5569)	0.9931	370.2916	-3.8533	6.2435
DEPOSITS / CONS	Double-log	-6.0607 (-1.9604)	0.8161 (1.4377)	0.9931	370.2090	-3.9499	6.5574
MB/GDP	Semi-log	-2.3799 (-17.014)	7.0645 (3.0977)	0.9877	376.4112	0.1634	0.0240
MB/GDP	Double-log	-4.0635 (-13.353)	0.4388 (4.3868)	0.9880	378.4363	-0.0202	0.0446
MB/CONS	Semi-log	-1.8921 (-10.596)	8.6147 (2.7901)	0.9915	391.9322	-1.0842	0.0249
MB/CONS	Double-log	-3.9166 (-9.2631)	0.5277 (3.8580)	0.9917	394.0756	-1.3411	0.0343
M1/GDP	Semi-log	-1.3326 (-10.383)	7.2649 (3.5370)	0.9903	383.2190	-1.9808	0.0864
M1/GDP	Double-log	-3.0405 (-10.642)	0.4454 (4.7679)	0.9905	384.8668	-2.1844	0.0985
M1/CONS	Semi-log	-0.8704 (-5.2453)	8.6020 (3.0690)	0.9926	391.7311	-3.7070	0.4429
M1/CONS	Double-log	-2.8662 (-7.1013)	0.5205 (4.0315)	0.9927	393.2633	-3.9471	0.6125

Note: This table presents the steady state estimates of the demand for currency, deposits, MB, and M1 using quarterly data for the period 1957:I and 1997:II. The income (or consumption) elasticity of money demand is assumed unity. α's are the coefficients in the semi-log function and β's are the coefficients in the double-log function. t-ratios are shown in parentheses. Since the h-statistics are less than 1.96 or greater than -1.96, there is no serial correlation in estimated money demand equations with currency (with income) and MB but there is with demand deposits, currency (with consumption) and M1 at the 5% level of significance. The critical value for testing ARCH(1) is $\chi_{(0.01)}$ =6.63 which implies we do not have ARCH(1) errors at the 1% level of significance.

TABLE 20: THE ESTIMATED WELFARE COST OF INFLATION (LUCAS)

Interest	Quadratic Cur.-Dep.	Square Root Cur.-Dep.	Quadratic M1	Square Root M1	Quadratic MB	Square Root MB
i	$(0.301)\,i^2$	$[(1-0.0182i^{0.5760})^{-0.7362}-1]$	$(0.821)\,i^2$	$[(1-0.0478i^{0.5546})^{-0.8031}-1]$	$(0.304)\,i^2$	$[(1-0.0172i^{0.5612})^{-0.7819}-1]$
0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.01	0.0030	0.0944	0.0082	0.2996	0.0031	0.1015
0.02	0.0120	0.1408	0.0328	0.4408	0.0122	0.1499
0.03	0.0271	0.1779	0.0739	0.5526	0.0273	0.1882
0.04	0.0481	0.2101	0.1314	0.6489	0.0486	0.2213
0.05	0.0752	0.2389	0.2052	0.7351	0.0760	0.2509
0.06	0.1082	0.2655	0.2956	0.8140	0.1094	0.2780
0.07	0.1473	0.2902	0.4023	0.8874	0.1489	0.3032
0.08	0.1924	0.3135	0.5254	0.9563	0.1945	0.3269
0.09	0.2435	0.3356	0.6650	1.0216	0.2461	0.3493
0.10	0.3006	0.3567	0.8210	1.0838	0.3039	0.3707
0.20	1.2026	0.5328	3.2839	1.6010	1.2154	0.5481
0.30	2.7058	0.6740	7.3888	2.0138	2.7347	0.6892
0.40	4.8102	0.7966	13.136	2.3715	4.8616	0.8111
0.50	7.5160	0.9071	20.524	2.6934	7.5963	0.9204
0.60	10.823	1.0087	29.555	2.9896	10.939	1.0207
0.70	14.731	1.1035	40.228	3.2662	14.889	1.1142
0.80	19.241	1.1930	52.542	3.5273	19.447	1.2020
0.90	24.352	1.2780	66.500	3.7755	24.612	1.2854
1.00	30.064	1.3592	82.097	4.0130	30.385	1.3649

Note: This table presents the estimated welfare cost of inflation in both a currency-deposit model and a single-monetary-asset model using Lucas' compensating variation approach with two different functional forms (semi-log, double-log) and three different monetary aggregates (currency-deposit, monetary base, M1). Entries are multiplied by 100 and can thus be interpreted as percentages.

TABLE 21: THE ESTIMATED WELFARE COST OF INFLATION IN A CURRENCY-DEPOSIT MODEL (BAILEY)

Interest Rate	Semi-log (Cur.-Dep.)[a]	Double-log (Cur.-Dep.)[b]	Semi-log (Cur.-Dep.)[c]	Double-log (Cur.-Dep.)[d]
0.00	0.0000	0.0000	0.0000	0.0000
0.01	0.0029	0.0943	0.0352	0.0694
0.02	0.0109	0.1406	0.1322	0.1005
0.03	0.0233	0.1775	0.2791	0.1255
0.04	0.0394	0.2095	0.4658	0.1474
0.05	0.0586	0.2383	0.6839	0.1672
0.06	0.0804	0.2646	0.9258	0.1855
0.07	0.1042	0.2892	1.1854	0.2028
0.08	0.1298	0.3123	1.4573	0.2191
0.09	0.1566	0.3343	1.7372	0.2347
0.10	0.1845	0.3552	2.0213	0.2496
0.20	0.4730	0.5294	4.6332	0.3780
0.30	0.7136	0.6687	6.3766	0.4849
0.40	0.8905	0.7892	7.3815	0.5803
0.50	1.0200	0.8975	7.9449	0.6681
0.60	1.1203	0.9968	8.2704	0.7504
0.70	1.2040	1.0894	8.4703	0.8284
0.80	1.2787	1.1764	8.6014	0.9029
0.90	1.3485	1.2590	8.6922	0.9746
1.00	1.4154	1.3378	8.7574	1.0439

Note: This table presents the estimated welfare cost of inflation in a currency-deposit model using Bailey's consumer's surplus argument with two different functional forms (semi-log, double-log). Entries are multiplied by 100 and can thus be interpreted as percentages.

[a] Restricted case where we following Bailey assume that the demand for currency and deposits have the same semi-elasticity.
[b] Restricted case where we following Bailey assume that the demand for currency and deposits have the same elasticity.
[c] Unrestricted case where we assume that the demand for currency and deposits have different semi-elasticities.
[d] Unrestricted case where we assume that the demand for currency and deposits have different elasticities.

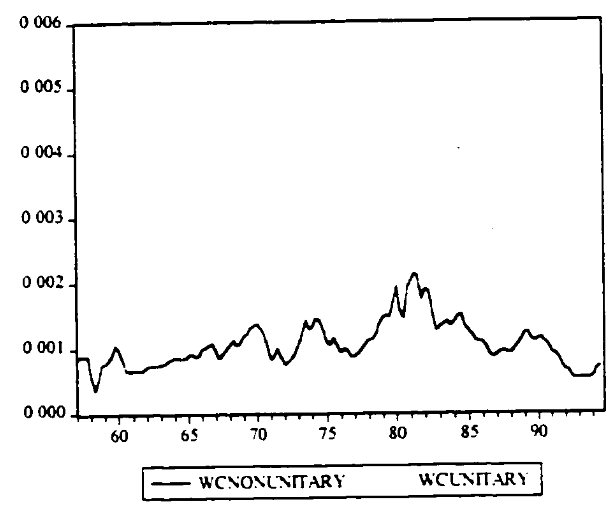
TABLE 22: THE ESTIMATED WELFARE COST OF INFLATION IN A SINGLE-MONETARY ASSET MODEL (BAILEY)

Interest Rate (i)	Semi-log M1	Double-log M1	Semi-log MB	Double-log MB
0.00	0.0000	0.0000	0.0000	0.0000
0.01	0.0091	0.2986	0.0031	0.1014
0.02	0.0348	0.4386	0.0119	0.1496
0.03	0.0747	0.5492	0.0256	0.1878
0.04	0.1266	0.6442	0.0434	0.2207
0.05	0.1887	0.7290	0.0648	0.2502
0.06	0.2594	0.8066	0.0892	0.2771
0.07	0.3370	0.8786	0.1160	0.3022
0.08	0.4203	0.9461	0.1449	0.3257
0.09	0.5081	1.0100	0.1753	0.3479
0.10	0.5994	1.0708	0.2071	0.3691
0.20	1.5480	1.5727	0.5406	0.5447
0.30	2.3253	1.9693	0.8193	0.6838
0.40	2.8553	2.3099	1.0132	0.8037
0.50	3.1861	2.6142	1.1366	0.9109
0.60	3.3821	2.8924	1.2112	1.0090
0.70	3.4943	3.1506	1.2548	1.1002
0.80	3.5570	3.3927	1.2796	1.1858
0.90	3.5914	3.6218	1.2935	1.2668
1.00	3.6100	3.8397	1.3012	1.3440

Note: This table presents the estimated welfare cost of inflation in a single-monetary-asset model using Bailey's consumer's surplus argument with two different functional forms (semi-log, double-log) and two different monetary aggregates (monetary base, M1). Entries are multiplied by 100 and can thus be interpreted as percentages

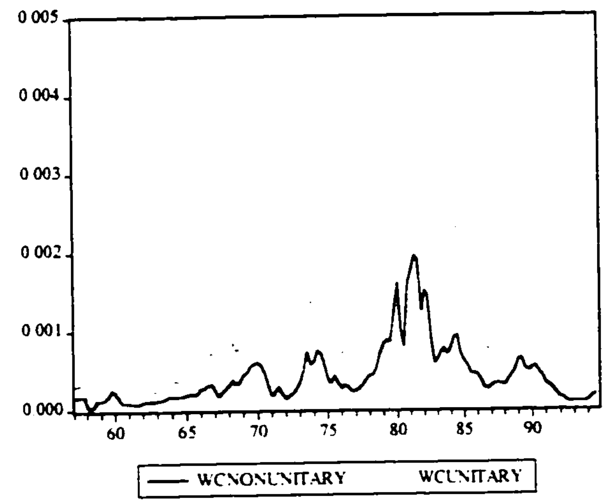
WELFARE COST ESTIMATES: MONETARY BASE & INCOME

Double-log



$WC_{unitary}$ = 0.2894 %
$WC_{nonunitary}$ = 0.1044 %

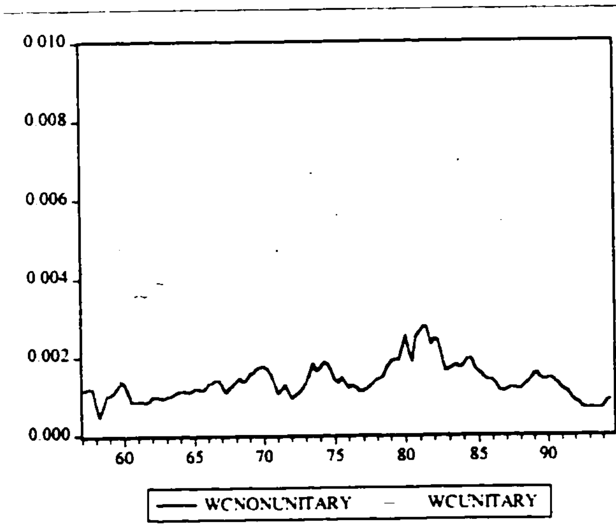
Semi-log



$WC_{unitary}$ = 0.1026 %
$WC_{nonunitary}$ = 0.0422 %

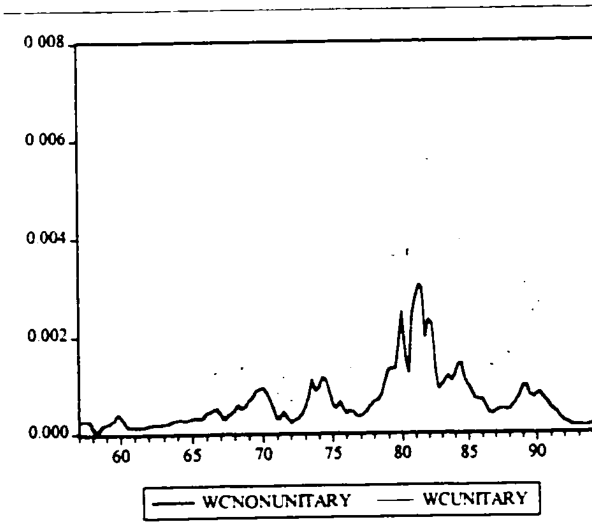
WELFARE COST ESTIMATES: MONETARY BASE & CONSUMPTION

Double-log



$WC_{unitary}$ = 0.5550 %
$WC_{nonunitary}$ = 0.1348 %

Semi-log

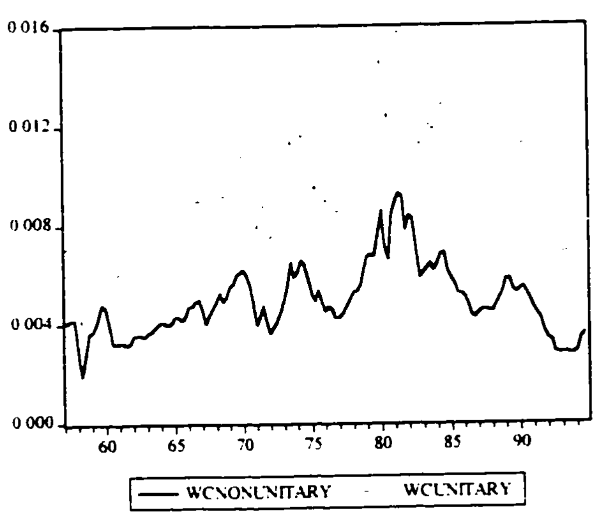


$WC_{unitary}$ = 0.1782 %
$WC_{nonunitary}$ = 0.0645 %

Note: The figures below the charts represent the mean of the estimated welfare cost of inflation assuming unitary ($WC_{unitary}$) and nonunitary ($WC_{unitary}$) quantity elasticity of money demand.

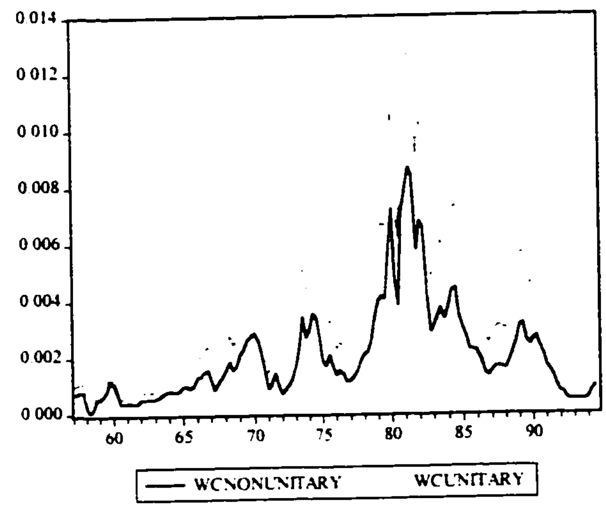
WELFARE COST ESTIMATES: M1 & INCOME

Double-log



$WC_{unitary}$ = 0.9419 %
$WC_{nonunitary}$ = 0.4923 %

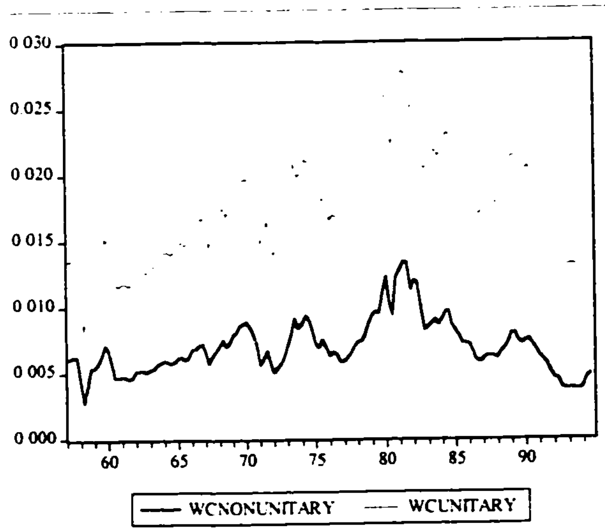
Semi-log



$WC_{unitary}$ = 0.3304 %
$WC_{nonunitary}$ = 0.2022 %

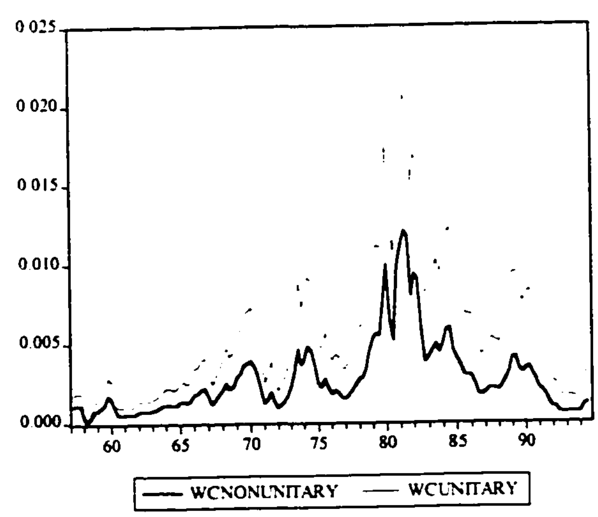
WELFARE COST ESTIMATES: M1 & CONSUMPTION

Double-log



$WC_{unitary}$ = 1.7386 %
$WC_{nonunitary}$ = 0.6933 %

Semi-log



$WC_{unitary}$ = 0.5546 %
$WC_{nonunitary}$ = 0.2722 %

Note: The figures below the charts represent the mean of the estimated welfare cost of inflation assuming unitary ($WC_{unitary}$) and nonunitary ($WC_{unitary}$) quantity elasticity

BIBLIOGRAPHY:

Altonji, Joseph G. "The Intertemporal Substitution Model of Labor Market Fluctuations: An Empirical Analysis." *Review of Economic Studies*, XLIX (1982), 783-824.

Amemiya, T. *Advanced Econometrics*. Cambridge: Harvard University Press. 1985.

Auernheimer, Leonardo. "The Honest Government's Guide to the Revenue from the Creation of Money." *Journal of Political Economy* 82 (May/June 1974), 598-606.

Ballard, Charles L., John B. Shoven, and John Whalley. "General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States." *The American Economic Review* (March 1985), 128-38.

Bailey, Martin J. "The Welfare Cost of Inflationary Finance." *Journal of Political Economy* 64 (April 1956), 93-110.

Barro, R. J. "The Neoclassical Approach to Fiscal Policy." In *Modern Business Cycle Theory*, R. J. Barro, ed. Cambridge, MA: Harvard University Press, 1989.

Baumol, William J. "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics*, Vol. 66, No. 4, 1952, pp. 545-56.

Blanchard, O. J., and Stanley Fischer. *Lectures on Macroeconomics*. Cambridge: MIT Press, 1990.

Blinder, Alan S. *Toward an Economic Theory of Income Distribution*. Cambridge, MA: MIT Press, 1974.

Brock, Philip L. "Reserve Requirements and the Inflation Tax," *Journal of Money, Credit and Banking* Vol. 21, No. 1, (February 1989), pp. 106-21.

Cagan, Phillip. "The Monetary Dynamics of Hyperinflation." In *Studies in the Quantity Theory of Money*, edited by M. Friedman. Chicago: University of Chicago Press, 1956.

Calvo, Guillermo A., and Leonardo Leiderman. "Optimal Inflation Tax Under Precommitment: Theory and Evidence." *American Economic Review* (March 1992), 179-94.

Calvo, Guillermo A., and Roque B. Fernandez. "Competitive Banks and the Inflation Tax." *Economic Letters* 12 (1983), 313-17.

Chiang, Alpha C. *Elements of Dynamic Optimization*. McGraw-Hill, Inc., 1992.

Clower, Robert W. "A Reconsideration of the Microfoundations of Monetary Theory." *Western Economic Journal* 6 (December 1967), 1-8.

Den Haan, W. J. "The Optimal Inflation Path in a Sidrauski-Type Model with Uncertainty." *Journal of Monetary Economics* 25 (1990), 389-409.

Dickey, David, and Wayne A. Fuller. "Distribution of the Estimates for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74 (June 1979), pp. 427-31.

__________. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica* 49 (July 1981), pp. 1057-72.

Dotsey, Michael, and Peter Ireland. "The Welfare Cost of Inflation in General Equilibrium," *Journal of Monetary Economics*, Vol. 37 (1996), pp. 29-47.

Drazen, Allan. "The Optimal Rate of Inflation Revisited." *Journal of Monetary Economics* 5 (1979), 231-48.

Dwyer. Gerald P. Jr., and Thomas R. Saving. "Government Revenue from Money Creation with Government and Private Money," *Journal of Monetary Economics* (March 1986). pp. 239-49.

Eckstein. Zvi. and Leonardo Leiderman. "Seigniorage and the Welfare Cost of Inflation: Evidence from an Intertemporal Model of Money and Consumption." *Journal of Monetary Economics* 29 (1992). 389-410.

Engle. Robert E., and Clive W. J. Granger. "Cointegration and Error-Correction: Representation. Estimation. and Testing." *Econometrica* 55 (March 1987). pp. 251-76.

__________, and B. Yoo. "Forecasting and Testing in Cointegrated Systems." *Journal of Econometrics* 35 (1987). pp. 143-59.

Faig. Miquel. "Seasonal Fluctuations and the Demand for Money." *Quarterly Journal of Economics* 104 (4) (1989). 847-61.

Feenstra. Robert C. "Functional Equivalence Between Liquidity Costs and the Utility of Money." *Journal of Monetary Economics* 17 (March 1986). 271-91.

Fischer. S., 1981, Towards an understanding of the costs of inflation: II, Carnegie-Rochester Conference Series on Public Policy 15. 5-41.

Friedman. Milton. *The Optimum Quantity of Money and Other Essays*. Chicago: Aldine, 1969.

Friedman, Milton, and Anna Jacobson Schwartz. "A Monetary History of the United States, 1867-1960," Princeton, N.J.: *Princeton University Press* 1963.

__________. "Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867-1975," Chicago: *University of Chicago Press* 1982.

Greene, William H. *Econometric Analysis*. New Jersey: Prentice Hall, 1997.

Gujarati, Damodar N. *Basic Econometrics*. New York: McGraw-Hill Inc., 1995.

Hamilton, James D. *Time Series Analysis*. Princeton: Princeton University Press, 1994.

Hwang, H. "Test of the adjustment process and linear homogeneity in a stock adjustment model of money demand." *The Review of Economics and Statistics*, LXVII (1985), 689-92.

Johnston, Jack, and John DiNardo. *Econometric Methods*. McGraw-Hill Inc., 1997.

Kamien, Morton I., and Nancy L. Schwartz. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*. Elsevier North Holland, Inc., 1981.

Karni, Edi. "The Transactions Demand for Cash: Incorporation of the Value of Time into the Inventory Approach." *Journal of Political Economy* Vol. 18, No. 5, 1973, 1216-25.

Levhari, David, and Don Patinkin. "The Role of Money in a Simple Growth Model." *American Economic Review* 58 (September 1968), 713-753.

Lucas, R.E., Jr., 1981, Discussion of: Stanley Fischer, 'Towards an understanding of the costs of inflation: II', Carnegie-Rochester Conference Series on Public Policy 15, 43-52.

Lucas, R.E., Jr., 1993, On the welfare cost of inflation, Manuscript (Department of Economics, University of Chicago, Chicago, IL).

Lucas. R.E.. Jr.. 1994. On the welfare cost of inflation. Working Paper 394. (Stanford University. Center for Economic Policy Research).

Lucas. R.E.. Jr. and N.L. Stokey. 1983. Optimal fiscal and monetary policy in an economy without capital. Journal of Monetary Economics 12. 55-93.

Maddala, G. S. *Introduction to Econometrics*. New York: Macmillan Publishing Company. 1992.

Mankiw. N. Gregory. and Lawrence H. Summers. "Money Demand and the Effects of Fiscal Policies." *Journal of Money. Credit and Banking* 18 (4). (1986). 415-29.

Marty. Alvin L. "Growth and the Welfare Cost of Inflationary Finance." *Journal of Political Economy* 75 (February 1967). 71-76.

Marty. Alvin L. "A Note on the Welfare Cost of Money Creation." *Journal of Monetary Economics* (January 1976). pp. 122-4.

Marty. Alvin L. and Frank J. Chaloupka. "Optimal Inflation Rates: A Generalization." *Journal of Money. Credit and Banking* (February 1988). pp. 141-4.

Marty. Alvin L. "The Inflation Tax and the Marginal Welfare Cost in a World of Currency and Deposits." *Federal Reserve Bank of St. Louis* (July/August 1994). pp. 25-9.

McCallum. B.T.. 1983, The role of overlapping-generations models in monetary economics, Carnegie-Rochester Conference Series on Public Policy 18, 9-44.

McCallum, B.T. and M.S. Goodfriend, 1988, Demand for money: Theoretical studies, in: The New Palgrave: A Dictionary of Economics (London, Macmillan Press, 1987), reprinted in Federal Reserve Bank of Richmond Economic Review 74. 16-24.

McCallum, B. T. "Inflation: Theory and Evidence." in *Handbook of Monetary Economics*, Volume II, edited by B.M. Friedman and F.H. Hahn. Amsterdam: North Holland, 1987.

Miller, Merton, and Daniel Orr. "A Model of the Demand for Money by Firms." *Quarterly Journal of Economics* Vol. 80, No. 3, 1966, pp. 413-35.

Mulligan, Casey B., and Xavier Sala-i-Martin. "U.S. Money Demand: Surprising Cross-sectional Estimates." *Brookings Paper on Economic Activity*, 2: 1992, pp. 285-343.

Obstfeld, Maurice, and Kenneth Rogoff. *Foundations of International Macroeconomics*. Cambridge: MIT Press, 1996.

Phelps, Edmund S. "Inflation in the Theory of Public Finance." *Swedish Journal of Economics* 75 (March 1973), 67-82.

Poirier, D., and A. Melino. "A Note on the Interpretation of Regression Coefficients Within a Class of Truncated Distributions." *Econometrica* 46 (1978), 1207-1209.

Poterba, J. M., and J. Rotemberg. "Money in the Utility Function: An Empirical Implementation." In *New Approaches to Monetary Economics*, edited by W. A. Barnett and K. J. Singleton. Cambridge: Cambridge University Press, 1987.

Roubini, Nouriel, and Xavier Sala-i-Martin. "Financial Repression and Economic Growth." *Journal of Development Economics* 39 (1992), 5-30.

Sidrauski, Miguel. "Rational Choice and Patterns of Growth in a Monetary Economy." *American Economic Review* 57 (May 1967a), 534-44.

_____________. "Inflation and Economic Growth." *Journal of Political Economy* 75 (1967b), 796-810.

Spitzer, J. "A Fast and Efficient Method for Estimation of Parameters in Models with the Box-Cox Transformation." *Journal of the American Statistical Association* 77 (1982a), 760-66.

________. "A Primer on the Box-Cox Estimation." *Review of Economics and Statistics* 64 (1982b), 307-313.

Stock, James. "Asymptotic Properties of Least-Squares Estimators of Cointegrating Vectors." *Econometrica* 55 (1987), pp. 1035-56.

Tobin, James. "The Interest-Elasticity of Transactions Demand for Cash." *Review of Economics and Statistics* Vol. 38, No 3, 1956, pp. 241-47.

Turnovsky, Stephen J. *Methods of Macroeconomic Dynamics*. Cambridge: MIT Press, 1995.

Varian, Hal R. *Microeconomic Analysis*. W. W. Norton & Company, 1992.

Implementation of the Heath-Jarrow-Morton and the Black-Derman-Toy Interest Rate Models for Pricing Options on Eurodollar Futures

by

Turan Gokcen Bali

INTRODUCTION

This sections explains arbitrage-free term structure models used for pricing fixed income securities and interest rate options with particular emphasis on the Heath-Jarrow-Morton model and its applications. This new approach utilizes option pricing theory and was developed in a sequence of papers by Heath-Jarrow-Morton (1990, 1991, 1992). It was motivated by the earlier work of Ho and Lee (1986) on this same topic. The approach taken in this paper to study the pricing of fixed income securities and interest rate options follows the standard binomial approach so often used to analyze the pricing equity options. This basic approach is independent of the Heath-Jarrow-Morton (HJM) model. It can, in fact, be used to approximate any arbitrage-free term-structure model, e.g., that of Black, Derman, and Toy (1990), Hull and White (1990), Cox, Ingersoll, and Ross (1985), or Vasicek (1977).

In this paper, we concentrate on valuing options on Eurodollar Futures using two different models of interest rates that values interest-rate sensitive derivative securities. The first model is the binomial approximation to the continuous trading term structure model of Heath-Jarrow-Morton (HJM) which uses the martingale measure approach and provides arbitrage-free prices that do not explicitly depend on the "market price of risk," but rather depend on an exogenously specified initial forward rate curve. The second model is Black-Derman-Toy (BDT) which uses today's term structure (the estimated yields for a set of zero-coupon Treasury securities and the corresponding estimated yield volatilities for these same securities) to determine the expected means and standard deviations of future short rates.

The short-term rate of interest is fundamental to much of theoretical and empirical finance, yet no consensus has emerged on the dynamics of its volatility. In this paper, we show that models which define volatility only as a function of interest rate levels tend to overemphasize the sensitivity of volatility to levels and fail to model adequately the serial correlation in conditional variances. On the other hand, serial correlation based models like GARCH models fail to capture adequately the relationship between interest rate levels and volatility. We introduce and test a model for the dynamics of short-term interest rate volatility, which allows volatility to depend on both the interest rate levels and information shocks.

THE HEATH-JARROW-MORTON MODEL

I. TERM STRUCTURE OF INTEREST RATES

This section introduces the notation, terminology, and assumptions used in the remainder of this paper.

The Economy: Following the original HJM model, we consider a *frictionless, competitive,* and *discrete trading* economy. By frictionless, they mean that there are no transaction costs in buying and selling financial securities, there are no bid/ask spreads, there are no restrictions on trade such as margin requirements or short sale restrictions, and there are no taxes. The markets are assumed to be *competitive*; i.e., each trader believes that she can buy/sell as many shares of a traded security as she desires without influencing its price. Last, we consider a *discrete trading* economy with trading dates {0, 1, 2,..., τ}. This assumption is not very restrictive because it is a reasonable approximation to actual security markets, especially if τ is large and the time interval between trading periods is small. The alternative, continuous trading, provides similar results but with significantly more complicated mathematics.

The Traded Securities: The traded securities in this economy are zero-coupon bonds of all maturities {0,..... τ} and a *money market account*. The price of a zero-coupon bond at time t that pays a sure dollar at time $T \geq t$ is denoted by $P(t,T)$. All zero-coupon bonds are assumed to be default free and have strictly positive prices. Table 1 provides three different hypothetical zero-coupon bond price curves. Panel A gives these prices for a flat term structure, panel B is for a downward sloping term structure, and panel C is for an upward sloping term structure. The numbers are constructed so that each time period corresponds to half a year.

The money market account is initialized at time 0 with a dollar investment, and its time t value is denoted $B(t)$. Thus, by convention, $B(0) = 1$.

Table 1: Hypothetical zero-coupon bond prices, forward rates, forward rates, and yields for various term structures.

Panel A: Flat Term Structure

Time to Maturity (T)	*Zero-coupon bond prices P(0,T)*	*Forward rates f(0,T)*	*Yields y(0,T)*
0	1.000000	1.02	
1	0.980392	1.02	1.02
2	0.961168	1.02	1.02
3	0.942322	1.02	1.02
4	0.923845	1.02	1.02
5	0.905730	1.02	1.02
6	0.887971	1.02	1.02
7	0.870560	1.02	1.02
8	0.853490	1.02	1.02
9	0.836755		1.02

Panel B: Downward Sloping Term Structure

Time to Maturity (T)	*Zero-coupon bond prices P(0,T)*	*Forward rates f(0,T)*	*Yields y(0,T)*
0	1.000000	1.024431	
1	0.976151	1.023342	1.024431
2	0.953885	1.022701	1.023886
3	0.932711	1.022319	1.023491
4	0.912347	1.022025	1.023198
5	0.892686	1.021794	1.022963
6	0.873645	1.021627	1.022768
7	0.855150	1.021544	1.022605
8	0.837115	1.020748	1.022472
9	0.820099		1.022281

Panel C: Upward Sloping Term Structure

Time to Maturity (*T*)	*Zero-coupon bond prices* *P(0,T)*	*Forward rates* *f(0,T)*	*Yields* *y(0,T)*
0	1.000000	1.016027	
1	0.984225	1.016939	1.016027
2	0.967831	1.017498	1.016483
3	0.951187	1.017836	1.016821
4	0.934518	1.018102	1.017075
5	0.917901	1.018312	1.017280
6	0.901395	1.018465	1.017452
7	0.885052	1.018542	1.017597
8	0.868939	1.019267	1.017715
9	0.852514		1.017887

Interest Rates: Various different interest rates play significant roles in the term structure models. We will define the most important of these which will be used in our analysis: *yields, forward rates,* and *spot rates*. For simplicity of notation and exposition, all rates will be denoted as one plus a percentage (these are sometimes called dollar returns). This convention greatly simplifies all the subsequent formulas.

The *yield* at time t on a T-maturity zero-coupon bond, denoted by $y(t,T)$, is defined by the expression (1):

$$y(t,\ T) = \left[\frac{1}{P(t,T)}\right]^{1/(T-t)}. \tag{I.1}$$

The yield is one plus the percentage return earned per period by holding the T-maturity zero-coupon bond from time t until its maturity. It is often called the ***holding period return***. Alternatively written, expression (I.1) implies

$$P(t,\ T) = \frac{1}{[y(t,T)]^{(T-t)}}. \tag{I.2}$$

The time *t* *forward rate* for the period $[T,\ T+1]$, denoted $f(t,T)$, is defined by

$$f(t,T) = \frac{P(t,T)}{P(t,T+1)}. \tag{I.3}$$

This corresponds to the rate contracted at time *t* for a riskless loan over the time period $[T,\ T+1]$.

To see this interpretation, consider forming the following portfolio at time *t*: (*i*) buy one zero-coupon bond maturing at time *T*, and (*ii*) sell $P(t,T)/P(t,T+1)$ zero-coupon bonds maturing at time $T+1$. Hold each zero-coupon bond until maturity. The cash flows to this portfolio are given in Table 2. The initial cash flow from forming this portfolio at time *t* is zero. Indeed, the cash outflow is $P(t,T)$ dollars, but the cash inflow is $[P(t,T)/P(t,T+1)]P(t,T+1) = P(t,T)$ dollars. These net to zero. The first cash flow of one dollar occurs at time *T*. It is an inflow. In addition, there is a cash outflow of $[P(t,T)/P(t,T+1)]$ dollars at time $T+1$. This pattern of cash flows is the same as that obtained from a dollar loan over $[T,\ T+1]$, contracted at time *t*. The implicit rate on this loan is $[P(t,T)/P(t,T+1)]$, the forward rate. This completes the argument.

Table 2: Cash flows to a portfolio generating a cash flow equal to the time *t* forward rate for date T, $f(t,T)$.

Time	*t*	*T*	*T+1*
Buy bond with maturity	$-P(t,T)$	+1	
Sell $P(t,T)/P(t,T+1)$ bonds with maturity $T+1$.	$+\frac{P(t,T)}{P(t,T+1)}P(t,T+1)$		$-\frac{P(t,T)}{P(t,T+1)}$

From expression (I.3) we can derive an expression for the bond's price in terms of the various maturity forward rates:

$$P(t,T) = \frac{1}{\prod_{j=t}^{T-1} f(t,j)}. \tag{I.4}$$

DERIVATION OF EXPRESSION (I.4):

Step 1: $f(t,t) = \dfrac{P(t,t)}{P(t,t+1)} = \dfrac{1}{P(t,t+1)}$ since $P(t,t) = 1$. So

$$P(t,t+1) = \frac{1}{f(t,t)}.$$

Step 2: Next, $f(t,t+1) = \dfrac{P(t,t+1)}{P(t,t+2)}$. So, $P(t,t+2) = \dfrac{P(t,t+1)}{f(t,t+1)}$. Substitution yields

$$P(t,t+2) = \frac{1}{f(t,t)f(t,t+1)}.$$ Continuing, we get

$$P(t,t+j) = \frac{1}{f(t,t)f(t,t+1)f(t,t+2)\cdots\cdots f(t,t+j-1)}.$$

Expression (4) shows that the bond's price is equal to a dollar received at time T and discounted by the different maturity forward rates.

Last, the *spot rate*, denoted $r(t)$, is defined as the rate contracted at time t on a one-period riskless loan starting immediately. By definition, therefore,

$$r(t) = f(t,t). \tag{I.5}$$

From Table 1 we see that the spot rate $r(0) = f(0,0)$ is 1.02 for the flat term structure, 1.024431 for the downward sloping term structure, and 1.016027 for the upward sloping term structure. Alternatively, using expressions (I.1) and (I.3), the spot rate is seen to be the holding period return on the shortest maturity bond, i.e.,

$$r(t) = y(t,t+1). \tag{I.6}$$

We can now return to clarify the return on the money market account. By construction,

$$B(t) = B(t-1)\, r(t-1) = \prod_{j=0}^{t-1} r(j) \tag{I.7}$$

Since, in this paper, we concentrate on pricing options on Eurodollar Futures, the following section explains the Eurodollar Futures contract and compares it with T-bill futures.

Eurodollar Futures Contract: The Eurodollar futures contract traded on the Chicago Mercantile Exchange (CME) is a very popular contract. A Eurodollar is a dollar deposited in a U.S. or foreign bank outside the United States. The Eurodollar interest rate is the rate of interest earned on Eurodollars deposited by one bank with another bank. It is also known as the London Interbank Offer Rate (LIBOR). Eurodollar interest rates are generally higher than the corresponding Treasury bill interest rates because a bank has to pay a higher rate of interest than the Federal government on borrowed funds. The interest rate underlying the Eurodollar futures contract is a 90-day rate.

On the surface, a Eurodollar futures contract appears to be structurally the same as the Treasury bill futures contract. Suppose the quoted futures price is F. The formula for calculating the value of one contract from the quoted futures is the same as that used for Treasury bill futures. The quote of 93.96 for example corresponds to a Eurodollar interest rate quote of 6.04 and a contract price of

$$10{,}000[100 - 0.25(100 - 93.96)] = \$984{,}900.$$

However, there are some important differences between the Treasury bill and Eurodollar futures contracts. For a Treasury bill, the contract price converges at maturity to the price of a 90-day $1 million face value Treasury bill, and if a contract is held until maturity, this is the instrument delivered. A Eurodollar futures contract is settled in cash on the second London business day before the third Wednesday of the month. The final marking to market sets the contract price equal to

$$10{,}000\,(100 - 0.25R)$$

where R is the quoted Eurodollar rate at that time. This quoted Eurodollar rate is the actual 90-day rate on Eurodollar deposits with quarterly compounding. It is not a discount rate. The Eurodollar futures contract is therefore a futures contract on an interest rate, whereas the Treasury bill futures contract is a futures contract on the price of a Treasury bill or a discount rate.

Futures prices and forward prices may not be the same for long-dated contracts. This point is particularly relevant to Eurodollar futures contracts since they have maturities up to 10 years. Eurodollar futures are regularly used to calculate zero-coupon LIBOR rates. For contracts lasting only a year or two it is reasonable to assume that the futures price is the forward price, or equivalently, that the rate calculated from the futures price is a forward interest rate. For longer-dated contracts this is far less reasonable.

II. THE EVOLUTION OF THE TERM STRUCTURE OF INTEREST RATES

Arbitrage pricing theory is often called a *relative pricing theory*. It is called a relative pricing theory because it takes the prices of a primary set of traded assets as given, as well as their stochastic evolution, and then it prices a secondary set of traded assets. It prices a secondary traded asset by constructing a portfolio of the primary assets such that the portfolio's cash flows and value replicate the cash flows and value of the traded secondary asset. To prevent riskless profit opportunities, or *arbitrage*, the cost of this replicating portfolio is known, since the prices of the primary traded assets are given. In this study, the primary set of traded assets comprises zero-coupon bonds and money market account. The secondary set of traded assets comprises either other zero-coupon bonds and interest rate options. The basic inputs to this relative pricing theory are the prices of the zero-coupon bonds and their stochastic evolution through time.

The State Space Process: At time 0, one of two possible outcomes can occur over the next time interval: "up," denoted by *u*, and "down," denoted by *d*. The up state occurs with probability $q_0>0$, and the down state occurs with probability $1-q_0>0$. At this point, the states up and down have no real economic interpretation and they are introduced to characterize the only uncertainties influencing the term structure of interest rates. At time 1, one of two possible states exists: $\{u,d\}$. We denote the generic state at time 1, s_1. Thus, $s_1 \in \{u,d\}$. Over the next interval, one of two possible outcomes can occur again: up, denoted by *u*, or down, denoted by *d*. The resulting state at time 2 is $s_1 u$ or $s_1 d$. The up state occurs with probability $q_1(s_1)>0$, and the down state occurs with probability $1-q_1(s_1)>0$. At time 2, therefore, there are four possible states ($s_1 u$, $s_1 d$ for each $s_1 \in \{u,d\}$), namely, *uu, ud, du, dd*. The ordering in which the ups and downs occur is important; *ud* is different from *du*. The process continues in this up-and-down fashion until τ. For an arbitrary time *t*, there are 2^t possible states. The possible states at time *t* corresponds to all possible *t*-sequences of *u* and *d*. We let s_t denote a generic state at time *t*, so $s_t \in$ {all possible *t*-sequences of *u*'s and *d*'s}. Over the next time period $[t,t+1]$, one of two possible outcomes can occur, up and down. The resulting state at time $t+1$ (s_{t+1}) is $s_t u$ or

$s_t d$. The up outcome occurs with probability $q_t(s_t)>0$, and the down outcome occurs with probability $1-q_t(s_t)>0$. At the last date τ all uncertainty is resolved. As the state space process is constructed, the entire history at any node may be important in determining the probabilities of the next outcome in the tree. This is indicated by making the probabilities at each date t dependent on the state s_t, i.e., $q_t(s_t)$. Thus, the state process is said to *path dependent* and the tree is often called *bushy*. Such a tree ensures the most flexibility for modeling term structure evolutions.

The Bond Price Process: The state process describes the uncertainty underlying and generating the evolution of all the zero-coupon bond prices. The evolution of the zero-coupon bond prices, in turn, determines the evolution of the forward rates and the spot rates. We now describe these stochastic processes. Formally, we indicate this state process's influence on the zero-coupon bond prices by expanding the notation, letting $P(t,T;s_t)$ be the T-maturity zero-coupon bond's price at time t under state s_t. Similarly, we expand the notation for yields, forward rates, spot rates, and so on. We assume that $P(T,T;s_T) = 1$ for all T and s_T. This assumption formalizes the statement that zero-coupon bonds are default-free, i.e., that they are worth a dollar at maturity under all possible states. In addition, we assume that $P(t,T;s_t)>0$ for all $t \leq T$ and s_t. This ensures that one cannot get something for nothing, i.e., that a sure dollar costs something. To understand the zero-coupon bond price process, we now introduce a new notation for the returns on the zero-coupon bond in the up state and in the down state:

$$u(t,T;s_t) = \frac{P(t+1,T;s_t u)}{P(t,T;s_t)} \quad \text{for all } t+1 \leq T \qquad \text{(II.1a)}$$

$$\text{and} \quad d(t,T;s_t) = \frac{P(t+1,T;s_t d)}{P(t,T;s_t)} \quad \text{for all } t+1 \leq T \tag{II.1b}$$

where $u(t,T;s_t) > d(t,T;s_t)$ for all $t < T-1$ and s_t. Expression (II.1a) defines $u(t,T;s_t)$ as the return at time t on the T-maturity zero-coupon bond in the up state. Similarly, expression (II.1b) defines $d(t,T;s_t)$ as the return at time t on the T-maturity zero-coupon bond in the down state. Since the zero-coupon bond pays a sure dollar, i.e.; $P(T,T;s_T) = 1$ for all T and s_T, expressions (II.1a) and (II.1b) implies the following:

$$u(t,t+1;s_t) = d(t,t+1;s_t) = \frac{1}{P(t,t+1;s_t)} = r(t;s_t) \quad \text{for all } t \text{ and } s_t. \tag{II.2}$$

Over the last interval in the bond's life, the returns in the up and down states are identical, nonrandom, and equal to the spot rate.

We can summarize the evolution of the one-factor zero-coupon bond price curve analytically as in the expression (II.3):

$$P(t+1,T;s_{t-1}) = \begin{cases} u(t,T;s_t)\, P(t,T;s_t) & \text{if } s_{t-1} = s_t u \text{ (with probability } q_t(s_t)>0) \\ d(t,T;s_t)\, P(t,T;s_t) & \text{if } s_{t-1} = s_t d \text{ (with probability } 1-q_t(s_t)>0) \end{cases} \tag{II.3}$$

where $u(t,T;s_t)>d(t,T;s_t)$ for $t < T-1$

$$u(t,t+1;s_t)\, P(t,t+1;s_t) = d(t,t+1;s_t)\, P(t,t+1;s_t) = 1.$$

The Forward Rate Process: To understand the forward curve evolution, we introduce a new notation for the rate of change in the forward rate over any interval of time $[t,t+1]$ conditional upon the history at time t; i.e.,

$$\alpha(t,T;s_t) = \frac{f(t+1,T;s_t u)}{f(t,T;s_t)} \quad \text{for all } t+1 \leq T \leq \tau-1 \tag{II.4a}$$

$$\beta(t,T;s_t) = \frac{f(t+1,T;s_t d)}{f(t,T;s_t)} \quad \text{for all } t+1 \le T \le \tau-1 \qquad \text{(II.4b)}$$

Expression (II.4) defines the rate of change in the T-maturity forward rate over $[t,t+1]$ to be $\alpha(t,T;s_t)$ in the up state and to be $\beta(t,T;s_t)$ in the down state. These rates of change depend on the history s_t. We can summarize the evolution of the one-factor zero-coupon bond price curve analytically as in the expression (II.5):

$$f(t+1,T;s_{t-1}) = \begin{cases} \alpha(t,T;s_t)\, f(t,T;s_t) & \text{if } s_{t-1} = s_t u \text{ (with probability } q_t(s_t)>0) \\ \beta(t,T;s_t)\, f(t,T;s_t) & \text{if } s_{t-1} = s_t d \text{ (with probability } 1-q_t(s_t)>0) \end{cases} \qquad \text{(II.5)}$$

where $t+1 \le T \le \tau-1$. Finally, at time $\tau-1$ only one zero-coupon bond remains in the market and only one forward rate exists, the spot rate. At time τ, when the last zero-coupon bond matures, no additional forward rates can be defined, and the model is terminated. Using the definition of a forward rate in (I.3), the relation between the zero-coupon bond price process and the forward rate process can be easily deduced:

$$f(t+1,T;s_{t-1}) = \frac{P(t+1,T;s_{t+1})}{P(t+1,T+1;s_{t+1})} \qquad \text{(II.6)}$$

Letting $s_{t-1} = s_t u$, we can rewrite expression (II.6) in return form:

$$f(t+1,T;s_t u) = \frac{P(t,T;s_t)u(t,T;s_t)}{P(t,T+1;s_t)u(t,T+1;s_t)} \qquad \text{(II.7)}$$

Using the definition of a forward rate, expression (I.3),again yields

$$f(t+1,T;s_t u) = f(t,T;s_t)\frac{u(t,T;s_t)}{u(t,T+1;s_t)} \qquad \text{(II.8a)}$$

A similar analysis for $s_{t-1} = s_t d$ yields

$$f(t+1,T;s_t d) = f(t,T;s_t)\frac{d(t,T;s_t)}{d(t,T+1;s_t)} \qquad \text{(II.8b)}$$

Comparison with expression (II.5) gives the final result:

$$\alpha(t,T;s_t) = \frac{u(t,T;s_t)}{u(t,T+1;s_t)} \qquad \text{for } t+1 \le T \le \tau-1 \qquad \text{(II.9a)}$$

and $$\beta(t,T;s_t) = \frac{d(t,T;s_t)}{d(t,T+1;s_t)} \qquad \text{for } t+1 \le T \le \tau-1 \qquad \text{(II.9b)}$$

Expression (II.9) relates the forward rate's rate of change parameters to the zero-coupon bond price process's rate of return parameters in the up (II.9a) and down (II.9b) states, respectively. *Expression* (II.9) *is useful when we parameterizes the bond price process first and then want to deduce the forward rate process from it*. The parameterization can also work in the reverse direction. We can alternatively deduce the zero-coupon bond price process's rate of return parameters in the up and down states. These relations are given in expression (II.10):

$$u(t,T;s_t) = \frac{r(t;s_t)}{\prod_{j=t+1}^{T-1}\alpha(t,j;s_t)} \qquad \text{for } t+2 \le T \le \tau-1 \qquad \text{(II.10a)}$$

and $$d(t,T;s_t) = \frac{r(t;s_t)}{\prod_{j=t+1}^{T-1}\beta(t,j;s_t)} \qquad \text{for } t+2 \le T \le \tau-1 \qquad \text{(II.10b)}$$

Expression (II.10) *is useful when we parameterizes the forward rate process first and then want to deduce the bond price process from it.*

The Spot Rate Process: The spot rate curve starts at $r(0)$, and it moves "up" to $r(1;u)$ probability $q_0>0$ and "down" to $r(1;d)$ probability $1-q_0>0$. Since the spot rate moves inversely to the zero-coupon bond price curve movement, u indicates that spot rates move *down* and *d* indicates that spot rates move *up*. At time t, under state s_t, the spot rate $r(t;s_t)$ moves down to $r(t+1;s_t u) = u(t+1,t+2;s_t u)$ with probability $q_t(s_t)>0$ and up to $r(t+1;s_t d) = d(t+1,t+2;s_t d)$ with probability $1-q_t(s_t)>0$. We can summarize the spot rate's stochastic process as

$$r(t+1;s_{t-1}) = \begin{cases} u(t+1,t+2;s_t u) & \text{with probability } q_t(s_t)>0 \\ d(t+1,t+2;s_t d) & \text{with probability } 1-q_t(s_t)>0 \end{cases} \tag{II.11}$$

for all s_t and $t+1 \leq T-1$.

III. CONTINUOUS-TIME LIMITS

This section discusses the empirical implementation of the interest rate option models originally developed by Heath, Jarrow and Morton (1991). As a discrete-time model, its approximation to reality is expected to be good when the number of periods (τ) is quite large. In that case the discrete-time model is an approximation to the continuous trading limit. In fact, for purposes of empirical estimation, it is convenient to reparameterize the discrete-time model in terms of its continuous-time limit. The primary purpose of this section is to study this parameterization and the resulting continuous-time limit. A secondary purpose is to demonstrate how to construct arbitrage-free zero-curve evolutions.

To parameterize the forward rate process in terms of its continuous limit, we change the time scale in the discrete-time model. As it is currently constructed, there are τ time periods $t = 0,1,2,.....,\tau$. These time periods are arbitrarily specified. In order to take limits, let us fix a future date $\bar{\tau}$ (say, January 1, 2030), and divide the time horizon 0 to $\bar{\tau}$ into subperiods of equal length Δ. Thus, in terms of calendar time, the discrete periods

$0,1,2,\ldots,\tau$ correspond to the dates $0, \Delta, 2\Delta, 3\Delta, \ldots, \tau\Delta = \bar{\tau}$. We are interested in studying the various discrete-time economies when the number of trading dates becomes large (i.e., $\tau \to \infty$) or equivalently, when the time between trades becomes small (i.e., $\Delta \to 0$).

From the continuous-time perspective, the evolution of *observed* zero-coupon bond prices and forward rates are generated by a ***continuous empirical economy*** with parameters (*i*) $\mu^*(t,T)$, the expected change in (the logarithm of) the forward rates per unit time, and (*ii*) $\sigma(t,T)$, the standard deviation of the change in (the logarithm of) the forward rates per unit time.

We now construct an approximating discrete-time empirical economy such that as the step size shrinks ($\Delta \to 0$), the discrete-time economy approaches the continuous-time economy. A discrete-time empirical economy is characterized by (*i*) the probability of movements of forward rates $q_t^{\Delta}(s_t)$ and (*ii*) the (one plus) percentage changes in forward rates across the various states $[\alpha_\Delta(t,T;s_t), \beta_\Delta(t,T;s_t)]$. It is a fact from probability theory [see He (1990)] that under mild technical conditions, this approximation can be obtained by choosing

$$\frac{E_t\{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)}{\Delta} \to \mu^*(t,T) \qquad \text{as } \Delta \to 0$$

and

$$\frac{Var_t\{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)}{\Delta} \to \sigma(t,T)^2 \qquad \text{as } \Delta \to 0$$

where the expectations and variance are obtained using $q_t^{\Delta}(s_t)$. Under these conditions, for small Δ, the two economies will be similar, and the discrete-time economy will be a good approximation to the limit economy (and conversely).

Given the discrete-time empirical economy constructed above, the assumption of no arbitrage gives the existence of *unique* pseudo probabilities $\pi_\Delta(t;s_t)$, which are used for valuation of contingent claims. The discrete-time pseudo economy is characterized by (*i*) the probability of movements of forward rates $\pi_\Delta(t;s_t)$ and (*ii*) the (one plus) percentage changes in forward rates across the various states $[\alpha_\Delta(t,T;s_t), \beta_\Delta(t,T;s_t)]$. The percentage changes in forward rates are *identical* across the two discrete-time economies; only the likelihoods of the movements differ. However, we need $\pi_\Delta(t;s_t)>0$ if and only if $q_t^\Delta(s_t)>0$. Those states with positive probability in the pseudo economy must have positive probability in the empirical economy, and conversely.

Analogous to the discrete-time case, the assumption of no arbitrage in the continuous-time model gives the existence of unique pseudo probabilities, which are used for the valuation of contingent claims (see Heath, Jarrow, Morton (1992) for technical details). These no-arbitrage restrictions imply that the limit pseudo economy has the parameters (*i*) $\mu(t,T)$, the expected change in (the logarithm of) the forward rates per unit time, and (*ii*) $\sigma(t,T)$, the standard deviation of the change in (the logarithm of) the forward rates per unit time. The standard deviations of change in (the logarithm of) the forward rates are identical across the two limit economies; only the likelihoods (and, therefore, the expected changes of (the logarithm of) the forward rates) can differ. The construction is complete if the discrete-time pseudo economy also converges to the limit pseudo economy, because then for small Δ, contingent claim values as computed in the discrete-time model will be good approximations to contingent claim values as computed in the continuous-time model (and conversely). This construction can be obtained if we choose

$$\frac{\tilde{E}_t\{\log f_\Delta(t+\Delta,T)-\log f_\Delta(t,T)\}}{\Delta}\rightarrow\mu(t,T) \qquad \text{as } \Delta\rightarrow 0$$

and $$\frac{\tilde{Var}_t\{\log f_\Delta(t+\Delta,T)-\log f_\Delta(t,T)\}}{\Delta}\rightarrow\sigma(t,T)^2 \qquad \text{as } \Delta\rightarrow 0$$

where the expectations and variance are obtained using $\pi_\Delta(t;s_t)$.

There are numerous ways of constructing the discrete-time economies. This follows because we require only that the limiting systems match as $\Delta \rightarrow 0$. From among these constructions, we would like to select one that makes the computation of contingent claims values as simple as possible. This simplicity of computation occurs, for example, if the pseudo probability satisfies $\pi_\Delta(t;s_t) = 1/2$ for all Δ, t, and s_t.

Consider the forward rate process described in the previous section,

$$f_\Delta(t+\Delta,T;s_{t-\Delta}) = \begin{cases} \alpha_\Delta(t,T;s_t)f_\Delta(t,T;s_t) & \text{if } s_{t-\Delta}=s_t u \text{ (with probability } q_t^\Delta(s_t)>0) \\ \beta_\Delta(t,T;s_t)f_\Delta(t,T;s_t) & \text{if } s_{t-\Delta}=s_t d \text{ (with probability } 1-q_t^\Delta(s_t)>0) \end{cases} \qquad \text{(III.1)}$$

where $t+\Delta \leq T \leq \tau\Delta-\Delta$, and both t and T are integer multiples of Δ. Let us now parameterize (III.1) in terms of three new stochastic process, $\mu(t,T;s_t)$, $\sigma(t,T;s_t)$, and $\phi(t;s_t)$, implicitly defined as follows:

$$\alpha_\Delta(t,T;s_t) = e^{[\mu(t,T;s_t)\Delta-\sigma(t,T;s_t)\sqrt{\Delta}]} \qquad \text{(III.2a)}$$

$$\beta_\Delta(t,T;s_t) = e^{[\mu(t,T;s_t)\Delta+\sigma(t,T;s_t)\sqrt{\Delta}]} \qquad \text{(III.2b)}$$

$$q_t^\Delta(s_t) = \frac{1}{2}+\frac{1}{2}\phi(t;s_t)\sqrt{\Delta} \qquad \text{(III.2c)}$$

Substitution (III.2) into (III.1) gives

$$f_\Delta(t+\Delta,T;s_{t-\Delta}) = \begin{cases} f_\Delta(t,T;s_t)\, e^{[\mu(t,T,s_t)\Delta - \sigma(t,T,s_t)\sqrt{\Delta}]} & \text{if } s_{t-\Delta} = s_t u \;\; (q_t^u = \frac{1}{2} + \frac{1}{2}\phi(t;s_t)\sqrt{\Delta} > 0) \\ f_\Delta(t,T;s_t)\, e^{[\mu(t,T,s_t)\Delta + \sigma(t,T,s_t)\sqrt{\Delta}]} & \text{if } s_{t-\Delta} = s_t d \;\; (q_t^d = \frac{1}{2} - \frac{1}{2}\phi(t;s_t)\sqrt{\Delta} > 0) \end{cases} \quad \text{(III.3)}$$

Next, take the natural logarithm of both sided of expression (III.3) to obtain

$$\log f_\Delta(t+\Delta,T;s_{t-\Delta}) - \log f_\Delta(t,T;s_t) = \begin{cases} \mu(t,T;s_t)\Delta - \sigma(t,T;s_t)\sqrt{\Delta} & \text{with } q_t^u = \frac{1}{2} + \frac{1}{2}\phi(t;s_t)\sqrt{\Delta} > 0 \\ \mu(t,T;s_t)\Delta + \sigma(t,T;s_t)\sqrt{\Delta} & \text{with } q_t^d = \frac{1}{2} - \frac{1}{2}\phi(t;s_t)\sqrt{\Delta} > 0 \end{cases} \quad \text{(III.4)}$$

The mean and variance of the changes in the logarithms of forward rates can be computed to be

$$E_t\{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)\} = [\mu(t,T) - \phi(t)\,\sigma(t,T)]\,\Delta \quad \text{(III.5a)}$$

and $$Var_t\{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)\} = \sigma(t,T)^2\,\Delta - \phi(t)^2\,\sigma(t,T)^2\,\Delta^2 \quad \text{(III.5b)}$$

DERIVATION OF EXPRESSION (III.5): For simplicity of notation, define

$$\Delta \log f_\Delta(t,T) = \log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T).$$

Using expression (III.4),

$$E_t\{\Delta \log f_\Delta(t,T)\} = \left(\frac{1}{2} + \frac{1}{2}\phi(t)\sqrt{\Delta}\right)[\mu(t,T)\Delta - \sigma(t,T)\sqrt{\Delta}]$$

$$+\left(\frac{1}{2} - \frac{1}{2}\phi(t)\sqrt{\Delta}\right)[\mu(t,T)\Delta + \sigma(t,T)\sqrt{\Delta}]$$

$$= [\mu(t,T) - \phi(t)\,\sigma(t,T)]\,\Delta.$$

$$Var_t\{\Delta \log f_\Delta(t,T)\} = E_t\{[\Delta \log f_\Delta(t,T)]^2\} - [E_t\{\Delta \log f_\Delta(t,T)\}]^2$$

$$=\left(\frac{1}{2}+\frac{1}{2}\phi(t)\sqrt{\Delta}\right)[\mu(t,T)\Delta - \sigma(t,T)\sqrt{\Delta}\,]^2$$

$$+\left(\frac{1}{2}-\frac{1}{2}\phi(t)\sqrt{\Delta}\right)[\mu(t,T)\Delta+\sigma(t,T)\sqrt{\Delta}\,] - [\mu(t,T)\Delta - \phi(t)\,\sigma(t,T)\,\Delta]^2$$

$$= \sigma(t,T)^2\,\Delta - \phi(t)^2\,\sigma(t,T)^2\,\Delta^2.$$

Dividing by Δ and taking limits of these quantities as $\Delta \to 0$ give

$$lim_{\Delta\to 0}\ \frac{E_t\{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)\}}{\Delta} = [\mu(t,T) - \phi(t)\,\sigma(t,T)] \qquad \text{(III.6a)}$$

and $$lim_{\Delta\to 0}\ \frac{Var_t\{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)\}}{\Delta} = \sigma(t,T)^2. \qquad \text{(III.6b)}$$

Expression (III.6a) represents the drift $\mu^*(t,T) = \mu(t,T) - \sigma(t,T)\,\phi(t)$ of the empirical process and $\sigma(t,T)$ represents the volatility of the empirical process. Both $\mu^*(t,T)$ and $\sigma(t,T)$ can in principle be estimated from past observations of forward rates. The stochastic process $\phi(t)$ is interpreted as a *risk premium*, i.e., a measure of the excess expected return (above the spot rate) per unit of standard deviation for the zero-coupon bonds. We next investigate the discrete-time pseudo economy implied by expression (III.2) and the assumption of no arbitrage.

Arbitrage-Free Restrictions: This section studies the restrictions implied by no arbitrage on the above system. This system is arbitrage-free if and only if there exist unique pseudo probabilities $\pi_\Delta(t;s_t)$ such that

$$\pi_\Delta(t;s_t) = \frac{r_\Delta(t;s_t) - d_\Delta(t,T;s_t)}{u_\Delta(t,T;s_t) - d_\Delta(t,T;s_t)} \qquad \text{(III.7)}$$

for all s_t, $0 \le t \le T-\Delta$, and $T \le \tau\Delta$, where both t and T are integer multiples of Δ. Using expression (III.2) in (II.10) we get:

$$u_\Delta(t,T;s_t) = r_\Delta(t;s_t)\, e^{-\sum_{j=t+\Delta}^{T-\Delta} \mu(t,j;s_t)\Delta + \sum_{j=t+\Delta}^{T-\Delta} \sigma(t,j;s_t)\sqrt{\Delta}} \qquad \text{(III.8a)}$$

and $$d_\Delta(t,T;s_t) = r_\Delta(t;s_t)\, e^{-\sum_{j=t+\Delta}^{T-\Delta} \mu(t,j;s_t)\Delta - \sum_{j=t+\Delta}^{T-\Delta} \sigma(t,j;s_t)\sqrt{\Delta}} \qquad \text{(III.8b)}$$

Substitution of these expressions into (III.7) and simplification yields

$$\pi_\Delta(t;s_t) = \frac{1 - \left\{ e^{-\sum_{j=t+\Delta}^{T-\Delta} \mu(t,T;s_t)\Delta - \sum_{j=t+\Delta}^{T-\Delta} \sigma(t,T;s_t)\sqrt{\Delta}} \right\}}{\left\{ e^{-\sum_{j=t+\Delta}^{T-\Delta} \mu(t,T;s_t)\Delta + \sum_{j=t+\Delta}^{T-\Delta} \sigma(t,T;s_t)\sqrt{\Delta}} \right\} - \left\{ e^{-\sum_{j=t+\Delta}^{T-\Delta} \mu(t,T;s_t)\Delta - \sum_{j=t+\Delta}^{T-\Delta} \sigma(t,T;s_t)\sqrt{\Delta}} \right\}} \qquad \text{(III.9)}$$

for all s_t, $0 \le t \le T-\Delta$, and $T \le \tau\Delta$, where both t and T are integer multiples of Δ. Expression (III.9) gives the cross-restrictions on the drifts and volatilities of the forward rate process that are both necessary and sufficient for the existence of the pseudo probabilities. Using these pseudo probabilities the change in the logarithm of forward rates is represented as

$$\log f_\Delta(t+\Delta, T; s_{t-\Delta}) - \log f_\Delta(t,T;s_t) = \begin{cases} \mu(t,T;s_t)\Delta - \sigma(t,T;s_t)\sqrt{\Delta} & \text{with } \pi_\Delta(t;s_t)>0 \\ \mu(t,T;s_t)\Delta + \sigma(t,T;s_t)\sqrt{\Delta} & \text{with } 1-\pi_\Delta(t;s_t)>0 \end{cases} \qquad \text{(III.10)}$$

The mean and the variance of the changes in forward rates can be computed:

$$\tilde{E}_t\, \{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)\} = \mu(t,T)\,\Delta - (1 - 2\,\pi_\Delta(t))\,\sigma(t,T)\sqrt{\Delta} \qquad \text{(III.11a)}$$

and $$\tilde{Var}_t\, \{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)\} = 4\,\sigma(t,T)^2\,\Delta\,\pi_\Delta(t)\,(1 - \pi_\Delta(t)) \qquad \text{(III.11b)}$$

DERIVATION OF EXPRESSION (III.11): For simplicity of notation, define

$$\Delta \log f_\Delta(t,T) = \log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T).$$

Using expression (III.10),

$$\tilde{E}_t\ \{\Delta \log f_\Delta(t,T)\} = \pi_\Delta(t)\ [\mu(t,T)\Delta - \sigma(t,T)\sqrt{\Delta}\ ] + (1 - \pi_\Delta(t))\ [\mu(t,T)\Delta + \sigma(t,T)\sqrt{\Delta}\ ]$$

$$= \mu(t,T)\ \Delta + (1 - 2\ \pi_\Delta(t))\ \sigma(t,T)\sqrt{\Delta}$$

$$\tilde{Var}_t\ \{\Delta \log f_\Delta(t,T)\} = \tilde{E}_t\ \{[\Delta \log f_\Delta(t,T)]^2\} - [\tilde{E}_t\ \{\Delta \log f_\Delta(t,T)\}]^2$$

$$= \pi_\Delta(t)\ [\mu(t,T)\Delta - \sigma(t,T)\sqrt{\Delta}\ ]^2 + (1 - \pi_\Delta(t))\ [\mu(t,T)\Delta + \sigma(t,T)\sqrt{\Delta}\ ]^2$$

$$- [\mu(t,T)\ \Delta + (1 - 2\ \pi_\Delta(t))\ \sigma(t,T)\sqrt{\Delta}\ ]^2$$

$$= 4\ \sigma(t,T)^2\ \Delta\ \pi_\Delta(t)\ (1 - \pi_\Delta(t))\ .$$

We also require that as $\Delta \rightarrow 0$

$$lim_{\Delta\rightarrow 0}\ \frac{\tilde{E}_t\{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)\}}{\Delta} = \mu(t,T) \qquad \text{(III.12a)}$$

and $$lim_{\Delta\rightarrow 0}\ \frac{\tilde{Var}_t\{\log f_\Delta(t+\Delta,T) - \log f_\Delta(t,T)\}}{\Delta} = \sigma(t,T)^2. \qquad \text{(III.12b)}$$

This implies the added restriction that

$$\pi_\Delta(t;s_t) = \frac{1}{2} + 0(\sqrt{\Delta}) \qquad \text{(III.13)}$$

where $lim_{\Delta\rightarrow 0}\ \frac{0(\sqrt{\Delta})}{\sqrt{\Delta}} = 0$.

DERIVATION OF EXPRESSION (III.13): From expressions (III.11) and (III.12), we obtain

$$lim_{\Delta\rightarrow 0}\frac{(1-2\pi_\Delta(t))\sigma(t,T)}{\sqrt{\Delta}} = 0$$

and $lim_{\Delta\rightarrow 0}\ 4\ \pi_\Delta(t)\ (1 - \pi_\Delta(t)) = 1$

The first of these gives (III.13), since it implies that

$$lim_{\Delta\to 0} \frac{(\frac{1}{2} - \pi_\Delta(t))}{\sqrt{\Delta}} = 0\,.$$

For computational efficiency, it is convenient to set the pseudo probabilities to 1/2:

$\pi_\Delta(t;s_t) = \frac{1}{2}$ for all t and s_t. This is a special case of expression (III.13). With this

restriction, we get from expression (III.9) that

$$\frac{1}{2}\left[\left\{e^{-\sum_{j=t+\Delta}^{T-\Delta}\mu(t,j;s_t)\Delta + \sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}}\right\} - \left\{e^{-\sum_{j=t+\Delta}^{T-\Delta}\mu(t,j;s_t)\Delta - \sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}}\right\}\right] =$$

$$1 - \left\{e^{-\sum_{j=t+\Delta}^{T-\Delta}\mu(t,j;s_t)\Delta - \sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}}\right\}.$$

This is true if and only if

$$e^{\sum_{j=t+\Delta}^{T-\Delta}\mu(t,j;s_t)\Delta} = \frac{1}{2}\left[e^{\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}} + e^{-\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}}\right] = \cosh\left(\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}\right) \qquad \text{(III.15)}$$

where $cosh(x) = \frac{1}{2}(e^x + e^{-x})$.

Computation of the Arbitrage-Free Term Structure Evolutions: This section shows how to use the previous expressions to compute an arbitrage-free term structure evolution. First, we compute the evolution of the zero-coupon bond prices. Substitution of expression (III.15) into (III.8) gives

$$u_\Delta(t,T;s_t) = r_\Delta(t;s_t)\left[\cosh\left(\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}\right)\right]^{-1} e^{\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}} \qquad \text{(III.16a)}$$

and $$d_\Delta(t,T;s_t) = r_\Delta(t;s_t)\left[\cosh\left(\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}\right)\right]^{-1} e^{-\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}} . \qquad \text{(III.16b)}$$

Using expression (II.3) presented in the previous section, substitution of expression (III.16) yields

$$P_\Delta(t+\Delta,T;s_{t-\Delta}) = \begin{cases} P_\Delta(t,T;s_t)\, r_\Delta(t;s_t)\left[\cosh\left(\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}\right)\right]^{-1} e^{\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}} & \text{if } s_{t-\Delta} = s_t\, u \\ P_\Delta(t,T;s_t)\, r_\Delta(t;s_t)\left[\cosh\left(\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}\right)\right]^{-1} e^{-\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta}} & \text{if } s_{t-\Delta} = s_t\, d \end{cases} \qquad \text{(III.17)}$$

for all s_t, $0 \le t \le T-\Delta$, and $T \le \tau\Delta$, where both t and T are integer multiples of Δ. *Under the empirical probabilities* $\frac{1}{2}+\frac{1}{2}\phi(t;s_t)$, this bond price process converges to the limiting empirical process for the bond's price. *Under pseudo probabilities* $\frac{1}{2}$, this converges to the limiting pseudo process for the bond's price. Because the pseudo economies are all that are relevant to application of the contingent claim valuation theory, we never have to estimate the stochastic process $\phi(t;s_t)$ which can be interpreted as a risk premium, i.e., a measure of the excess expected return (above the spot rate) per unit of standard deviation for the zero-coupon bonds. Therefore, to apply this technology to price contingent claims, we never have to estimate a zero-coupon bond's risk premium. This is an important characteristic of the model.

To construct the forward rate process evolution, from (III.15) we get

$e^{\mu(t,t+\Delta;s_t)\Delta} = \cosh(\sigma(t,t+\Delta;s_t)\sqrt{\Delta})$ and

$$e^{\mu(t,T,s_t)\Delta} = \frac{\cosh(\sum_{j=t+\Delta}^{T}\sigma(t,j;s_t)\sqrt{\Delta})}{\cosh(\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta})}$$

for $T \geq t+2\Delta$ where both t and T are integer multiples of Δ. Substitution into expression (III.3) yields

$$f_\Delta(t+\Delta,T;s_{t-\Delta}) = \begin{cases} f_\Delta(t,T;s_t)\left[\dfrac{\cosh(\sum_{j=t+\Delta}^{T}\sigma(t,j;s_t)\sqrt{\Delta})}{\cosh(\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta})}\right]e^{-\sigma(t,T,s_t)\sqrt{\Delta}} & \text{if } s_{t-\Delta} = s_t\, u \\ f_\Delta(t,T;s_t)\left[\dfrac{\cosh(\sum_{j=t+\Delta}^{T}\sigma(t,j;s_t)\sqrt{\Delta})}{\cosh(\sum_{j=t+\Delta}^{T-\Delta}\sigma(t,j;s_t)\sqrt{\Delta})}\right]e^{+\sigma(t,T,s_t)\sqrt{\Delta}} & \text{if } s_{t-\Delta} = s_t\, d \end{cases} \qquad \text{(III.19)}$$

for all s_t, $0 \leq t \leq T-\Delta$, and $T+\Delta \leq \tau\Delta$, where both t and T are integer multiples of Δ, and where as a notational convenience we define

$$\cosh(\sum_{j=t+\Delta}^{t}\sigma(t,j;s_t)\sqrt{\Delta}) = 1 .$$

Thus, expression (III.19) gives the evolution of the discrete-time forward rate curve.

Under the empirical probabilities $\frac{1}{2}+\frac{1}{2}\phi(t;s_t)$, this bond price process converges to the limiting empirical process for the bond's price. *Under pseudo probabilities* $\frac{1}{2}$, this converges to the limiting pseudo process for the bond's price. The computation of contingent claims values is done using the pseudo probabilities. Note that under the pseudo probabilities, a specification of the volatility structure of forward rates,

$$\begin{bmatrix} \sigma(t,t+\Delta;s_t) \\ \sigma(t,t+2\Delta;s_t) \\ \sigma(t,t+3\Delta;s_t) \\ \cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots \\ \sigma(t,\tau\Delta-\Delta;s_t) \end{bmatrix}$$

for all $0 \le t \le \tau - \Delta$ and s_t is sufficient to determine the evolution of the forward rate curve. Two functional forms of the volatility function $\sigma(t,T;s_t)$ have received special attention in the literature.

1. **Deterministic Volatility Function**: The first case is that in which the volatility $\sigma(t,T;s_t)$ is a deterministic function, independent of the state s_t. This restriction on the volatility function implies that forward rates can go negative. This case includes as special cases Ho and Lee's (1986) model [$\sigma(t,T;s_t)$ is a constant] and a discrete-time approximation in the HJM framework to Vasicek's (1977) and Hull and White's (1990) models [$\sigma(t,T;s_t) = \xi\, e^{-\eta(T-t)}$ for ξ, $\eta > 0$ constants].
2. **Nearly Proportional Volatility Function**: The volatility is $\sigma(t,T;s_t) = \eta(t,T)$ $\min[f(t,T)-1,M]$, where $\eta(t,T)$ is a deterministic function and $M > 0$ is a large positive constant. In other words, in the second case $\sigma(t,T;s_t)$ is approximately proportional to the current value of the forward rate $f(t,T)$ less one. The proportionality factor is $\eta(t,T)$. This proportionality implies that forward rates are always nonnegative. Nonnegativity of forward rates is a condition usually required in models, because negative interest rates for zero-coupon bonds are inconsistent with the existence of cash currency, which can be stored costlessly at zero interest rates. In case 2, the larger the forward rate the larger the volatility. If the forward rate becomes too large,

however, the volatility is bounded by $\eta(t,T)M$. This upper bound guarantees that forward rates do not explode with positive probability as $\Delta \rightarrow 0$.

The Econometric Approach to Estimation of $\eta(t,T)$: In this section, we describe the econometric approach used in estimating the volatility parameter $\eta(t,T)$. We use the following discrete-time approximation to the continuous time forward rate specification:

$$f_t - f_{t-1} = \alpha + \beta f_{t-1} + \varepsilon_t$$

$$E(\varepsilon_t) = 0, \qquad \sigma(t,T,f_{t-1}) = E(\varepsilon^2_t | I_{t-1}) = \eta(t,T)^2 f^2_{t-1}.$$

This discrete-time model has the advantage of allowing the variance of forward rate changes to depend directly on the level of the forward rate in a way consistent with the continuous-time model. In these models, $\alpha + \beta f_{t-1}$ is the drift and $\eta(t,T)^2 f^2_{t-1}$ is the variance of unexpected forward rate changes. Also, rewriting $\alpha + \beta f_{t-1}$ as $\beta(f - \alpha^*)$ reveals that β can be viewed as a measure of the speed of mean reversion in rate levels. The more negative β is, the faster f respond to deviations from α^*. The volatility parameter $\eta(t,T)^2$ is simply a scale factor for the variance of unexpected forward rate changes. If $\eta(t,T)^2$ doubles, then the variance doubles. The estimated values of $\eta(t,T)$ for illustration of the model are: $\eta(0) = 0.1176$, $\eta(1)=0.0883$, $\eta(2)=0.0686$ and $f(0,1) = 1.02$.

IV. MODEL SOLUTIONS

Assume that M = 1,000,000 then

$\sigma(0,1) = \eta(1)[f(0,1)-1]$ $\quad \sigma(1,2)^u = \eta(1)[f(1,2)^u-1]$ $\quad \sigma(2,3)^{uu} = \eta(1)[f(2,3)^{uu}-1]$

$\sigma(0,2) = \eta(2)[f(0,2)-1]$ $\quad \sigma(1,2)^d = \eta(1)[f(1,2)^d-1]$ $\quad \sigma(2,3)^{ud} = \eta(1)[f(2,3)^{ud}-1]$

$\sigma(0,3) = \eta(3)[f(0,3)-1]$ $\quad \sigma(1,3)^u = \eta(2)[f(1,3)^u-1]$ $\quad \sigma(2,3)^{du} = \eta(1)[f(2,3)^{du}-1]$

$\sigma(1,3)^d = \eta(2)[f(1,3)^d-1]$ $\quad \sigma(2,3)^{dd} = \eta(1)[f(2,3)^{dd}-1]$

$$f(1,1)^{u} = f(0,1)\left[\frac{\cosh[\sum_{j=1}^{T=1}\sigma(0,j)]}{\cosh[\sum_{j=1}^{0}\sigma(0,j)]}\right]e^{-\sigma(0,1)} = f(0,1)\cosh[\sigma(0,1)]\, e^{-\sigma(0,1)}$$

$$f(1,1)^{d} = f(0,1)\left[\frac{\cosh[\sum_{j=1}^{T=1}\sigma(0,j)]}{\cosh[\sum_{j=1}^{0}\sigma(0,j)]}\right]e^{+\sigma(0,1)} = f(0,1)\cosh[\sigma(0,1)]\, e^{-\sigma(0,1)}$$

where $\cosh[\sum_{j=t+\Delta}^{t}\sigma(t,j;s_t)\sqrt{\Delta}] = 1 \Rightarrow \cosh[\sum_{j=1}^{0}\sigma(0,j)] = 1$.

$$f(1,2)^{u} = f(0,2)\left[\frac{\cosh[\sum_{j=1}^{T=2}\sigma(0,j)]}{\cosh[\sum_{j=1}^{T=1}\sigma(0,j)]}\right]e^{-\sigma(0,2)} = f(0,2)\left[\frac{\cosh[\sigma(0,1)+\sigma(0,2)]}{\cosh[\sigma(0,1)]}\right]e^{-\sigma(0,2)}$$

$$f(1,2)^{d} = f(0,2)\left[\frac{\cosh[\sum_{j=1}^{T=2}\sigma(0,j)]}{\cosh[\sum_{j=1}^{T=1}\sigma(0,j)]}\right]e^{+\sigma(0,2)} = f(0,2)\left[\frac{\cosh[\sigma(0,1)+\sigma(0,2)]}{\cosh[\sigma(0,1)]}\right]e^{-\sigma(0,2)}$$

$$f(1,3)^{u} = f(0,3)\left[\frac{\cosh[\sum_{j=1}^{T=3}\sigma(0,j)]}{\cosh[\sum_{j=1}^{T=2}\sigma(0,j)]}\right]e^{-\sigma(0,3)} = f(0,3)\left[\frac{\cosh[\sigma(0,1)+\sigma(0,2)+\sigma(0,3)]}{\cosh[\sigma(0,1)+\sigma(0,2)]}\right]e^{-\sigma(0,3)}$$

$$f(1,3)^{d} = f(0,3)\left[\frac{\cosh[\sum_{j=1}^{T=3}\sigma(0,j)]}{\cosh[\sum_{j=1}^{T=2}\sigma(0,j)]}\right]e^{+\sigma(0,3)} = f(0,3)\left[\frac{\cosh[\sigma(0,1)+\sigma(0,2)+\sigma(0,3)]}{\cosh[\sigma(0,1)+\sigma(0,2)]}\right]e^{-\sigma(0,3)}$$

$$f(2,2)^{uu} = f(1,2)^{u}\left[\frac{\cosh[\sum_{j=2}^{T=2}\sigma(1,j,u)]}{\cosh[\sum_{j=2}^{T=1}\sigma(1,j,u)]}\right]e^{-\sigma(1,2,u)} = f(1,2)^{u}\cosh[\sigma(1,2,u)]\, e^{-\sigma(1,2,u)}$$

$$f(2,2)^{ud} = f(1,2)^{u}\left[\frac{\cosh[\sum_{j=2}^{T=2}\sigma(1,j,u)]}{\cosh[\sum_{j=2}^{T=1}\sigma(1,j,u)]}\right]e^{+\sigma(1,2,u)} = f(1,2)^{u}\cosh[\sigma(1,2,u)]\, e^{-\sigma(1,2,u)}$$

where $\cosh[\sum_{j=2}^{T=1}\sigma(1,j,u)] = 1$.

$$f(2,2)^{du} = f(1,2)^{d}\left[\frac{\cosh[\sum_{j=2}^{T=2}\sigma(1,j,d)]}{\cosh[\sum_{j=2}^{T=1}\sigma(1,j,d)]}\right]e^{-\sigma(1,2,d)} = f(1,2)^{d}\cosh[\sigma(1,2,d)]\, e^{-\sigma(1,2,d)}$$

$$f(2,2)^{dd} = f(1,2)^{d}\left[\frac{\cosh[\sum_{j=2}^{T=2}\sigma(1,j,d)]}{\cosh[\sum_{j=2}^{T=1}\sigma(1,j,d)]}\right]e^{+\sigma(1,2,d)} = f(1,2)^{d}\cosh[\sigma(1,2,d)]\, e^{-\sigma(1,2,d)}$$

where $\cosh[\sum_{j=2}^{T=1}\sigma(1,j,d)] = 1$.

$$f(2,3)^{uu} = f(1,3)^{u}\left[\frac{\cosh[\sum_{j=2}^{T=3}\sigma(1,j,u)]}{\cosh[\sum_{j=2}^{T=2}\sigma(1,j,u)]}\right]e^{-\sigma(1,3,u)} = f(1,3)^{u}\left[\frac{\cosh[\sigma(1,2,u)+\sigma(1,3,u)]}{\cosh[\sigma(1,2,u)]}\right]e^{-\sigma(1,3,u)}$$

$$f(2,3)^{ud} = f(1,3)^{u}\left[\frac{\cosh[\sum_{j=2}^{T=3}\sigma(1,j,u)]}{\cosh[\sum_{j=2}^{T=2}\sigma(1,j,u)]}\right]e^{+\sigma(1,3,u)} = f(1,3)^{u}\left[\frac{\cosh[\sigma(1,2,u)+\sigma(1,3,u)]}{\cosh[\sigma(1,2,u)]}\right]e^{-\sigma(1,3,u)}$$

$$f(2,3)^{du} = f(1,3)^{d}\left[\frac{\cosh[\sum_{j=2}^{T=3}\sigma(1,j,d)]}{\cosh[\sum_{j=2}^{T=2}\sigma(1,j,d)]}\right]e^{-\sigma(1,3,d)} = f(1,3)^{d}\left[\frac{\cosh[\sigma(1,2,d)+\sigma(1,3,d)]}{\cosh[\sigma(1,2,d)]}\right]e^{-\sigma(1,3,d)}$$

$$f(2,3)^{dd} = f(1,3)^{d}\left[\frac{\cosh[\sum_{j=2}^{T=3}\sigma(1,j,d)]}{\cosh[\sum_{j=2}^{T=2}\sigma(1,j,d)]}\right]e^{+\sigma(1,3,d)} = f(1,3)^{d}\left[\frac{\cosh[\sigma(1,2,d)+\sigma(1,3,d)]}{\cosh[\sigma(1,2,d)]}\right]e^{-\sigma(1,3,d)}$$

$$f(3,3)^{uuu} = f(2,3)^{uu}\left[\frac{\cosh[\sum_{j=3}^{T=3}\sigma(2,j,uu)]}{\cosh[\sum_{j=3}^{T=2}\sigma(2,j,uu)]}\right]e^{-\sigma(2,3,uu)} = f(2,3)^{uu}\cosh\,[\sigma(2,3,uu)]\; e^{-\sigma(2,3,uu)}$$

$$f(3,3)^{uud} = f(2,3)^{uu}\left[\frac{\cosh[\sum_{j=3}^{T=3}\sigma(2,j,uu)]}{\cosh[\sum_{j=3}^{T=2}\sigma(2,j,uu)]}\right]e^{+\sigma(2,3,uu)} = f(2,3)^{uu}\cosh\,[\sigma(2,3,uu)]\; e^{-\sigma(2,3,uu)}$$

where $\cosh[\sum_{j=3}^{T=2}\sigma(2,j,uu)] = 1.$

$$f(3,3)^{udu} = f(2,3)^{ud}\left[\frac{\cosh[\sum_{j=3}^{T=3}\sigma(2,j,ud)]}{\cosh[\sum_{j=3}^{T=2}\sigma(2,j,ud)]}\right]e^{-\sigma(2,3,ud)} = f(2,3)^{ud}\cosh\,[\sigma(2,3,ud)]\; e^{-\sigma(2,3,ud)}$$

$$f(3,3)^{udd} = f(2,3)^{ud}\left[\frac{\cosh[\sum_{j=3}^{T=3}\sigma(2,j,ud)]}{\cosh[\sum_{j=3}^{T=2}\sigma(2,j,ud)]}\right]e^{+\sigma(2,3,ud)} = f(2,3)^{ud}\cosh\,[\sigma(2,3,ud)]\; e^{-\sigma(2,3,ud)}$$

where $\cosh[\sum_{j=3}^{T=2}\sigma(2,j,ud)] = 1.$

$$f(3,3)^{duu} = f(2,3)^{du}\left[\frac{\cosh[\sum_{j=3}^{T=3}\sigma(2,j,du)]}{\cosh[\sum_{j=3}^{T=2}\sigma(2,j,du)]}\right]e^{-\sigma(2,3,du)} = f(2,3)^{du}\cosh[\sigma(2,3,du)]\ e^{-\sigma(2,3,du)}$$

$$f(3,3)^{dud} = f(2,3)^{du}\left[\frac{\cosh[\sum_{j=3}^{T=3}\sigma(2,j,du)]}{\cosh[\sum_{j=3}^{T=2}\sigma(2,j,du)]}\right]e^{+\sigma(2,3,du)} = f(2,3)^{du}\cosh[\sigma(2,3,du)]\ e^{-\sigma(2,3,du)}$$

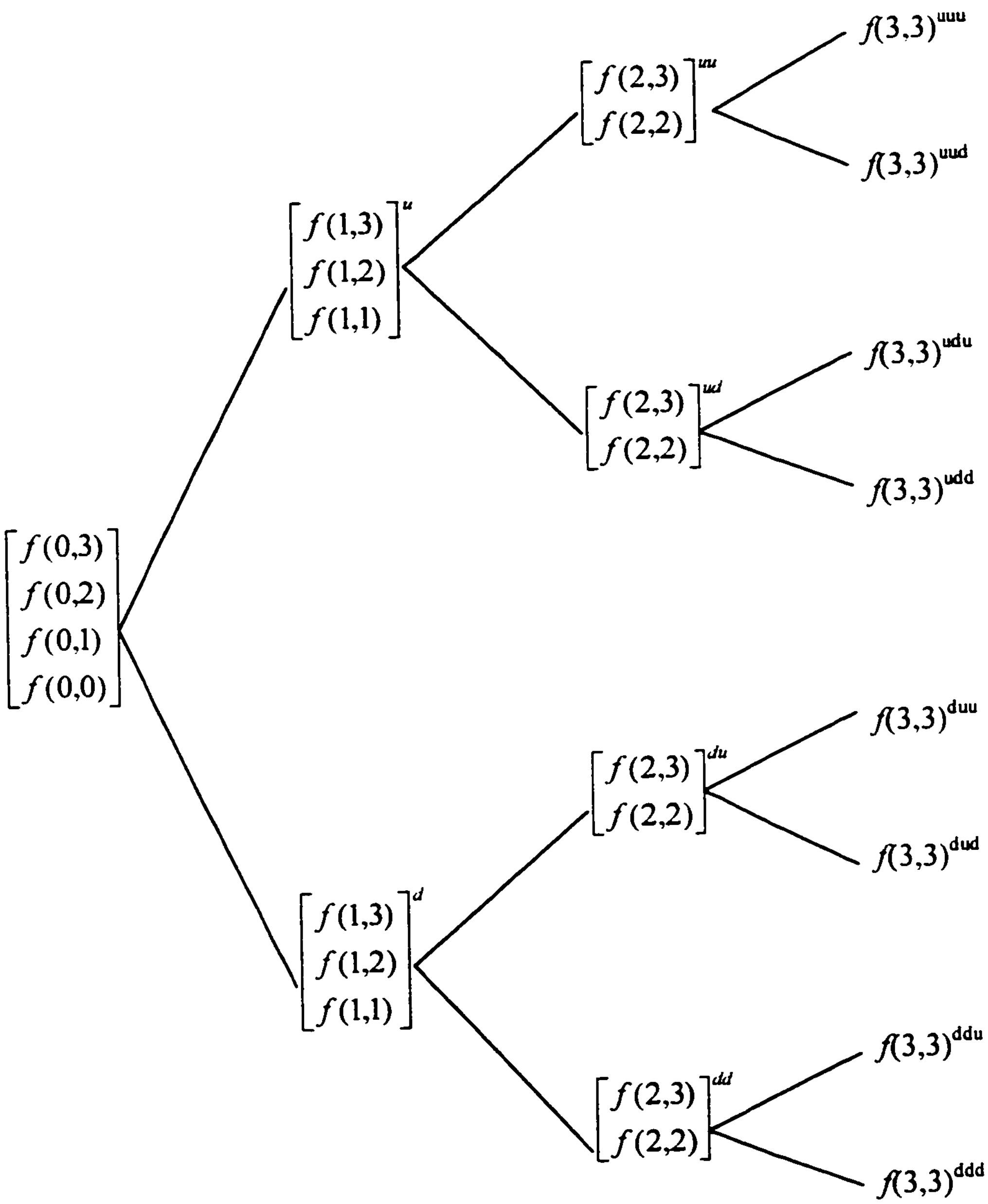
where $\cosh[\sum_{j=3}^{T=2}\sigma(2,j,du)] = 1$.

$$f(3,3)^{ddu} = f(2,3)^{dd}\left[\frac{\cosh[\sum_{j=3}^{T=3}\sigma(2,j,dd)]}{\cosh[\sum_{j=3}^{T=2}\sigma(2,j,dd)]}\right]e^{-\sigma(2,3,dd)} = f(2,3)^{dd}\cosh[\sigma(2,3,dd)]\ e^{-\sigma(2,3,dd)}$$

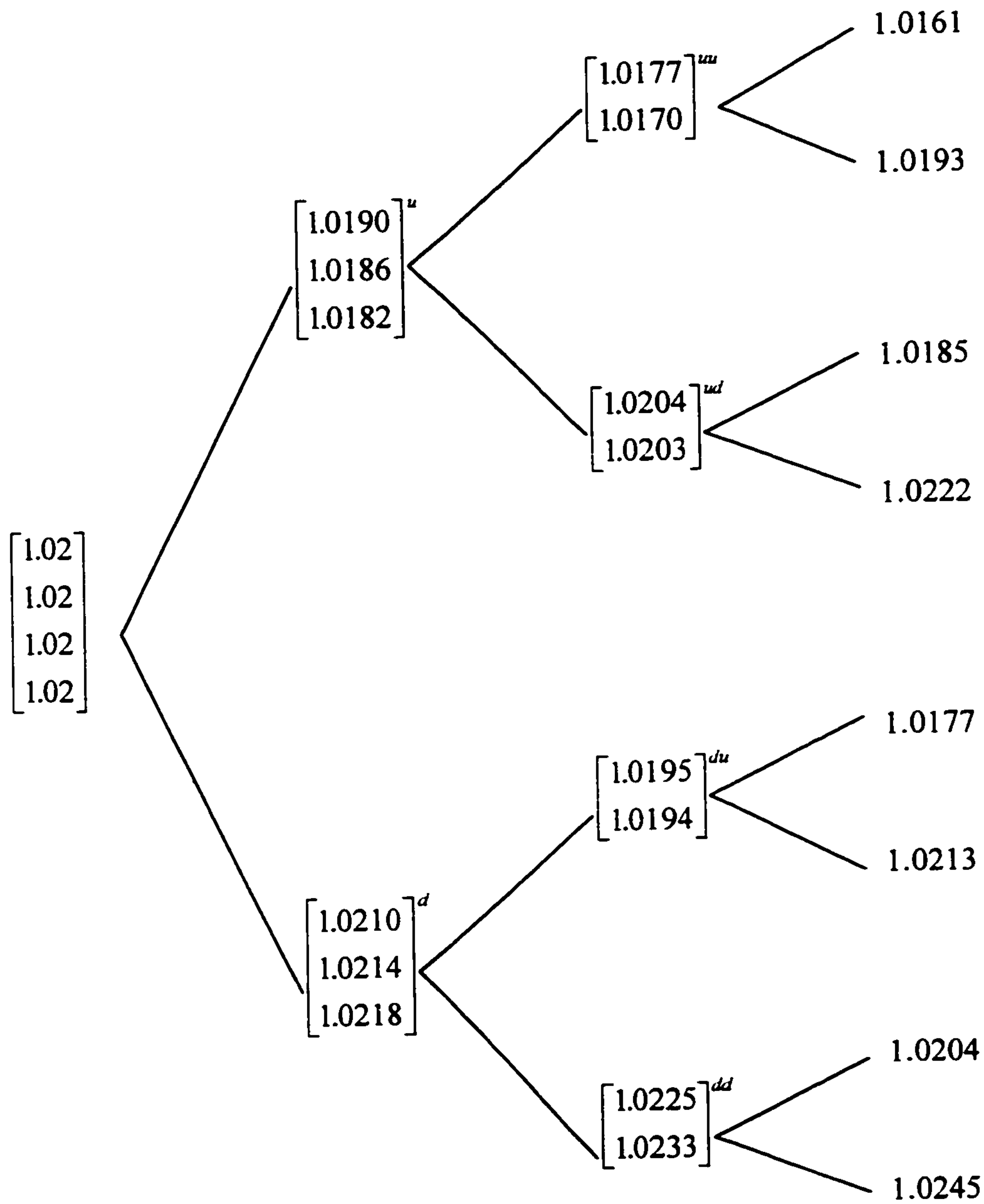
$$f(3,3)^{ddd} = f(2,3)^{dd}\left[\frac{\cosh[\sum_{j=3}^{T=3}\sigma(2,j,dd)]}{\cosh[\sum_{j=3}^{T=2}\sigma(2,j,dd)]}\right]e^{+\sigma(2,3,dd)} = f(2,3)^{dd}\cosh[\sigma(2,3,dd)]\ e^{-\sigma(2,3,dd)}$$

where $\cosh[\sum_{j=3}^{T=2}\sigma(2,j,dd)] = 1$.

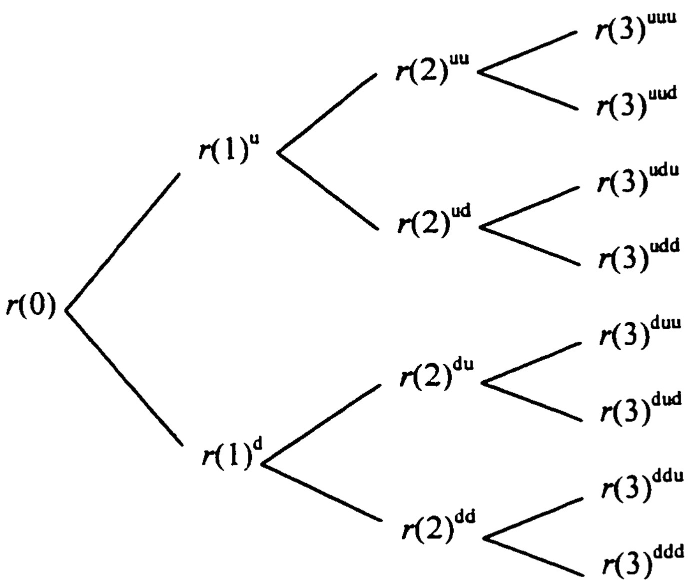
FORWARD RATE TREE

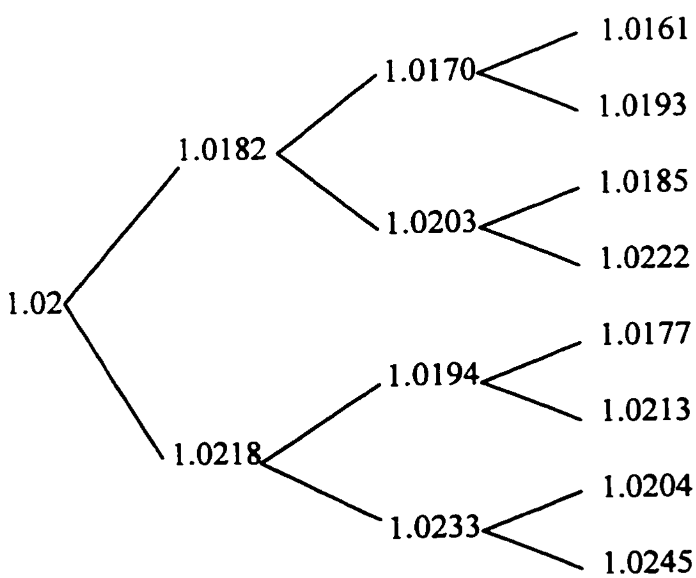


FORWARD RATE TREE

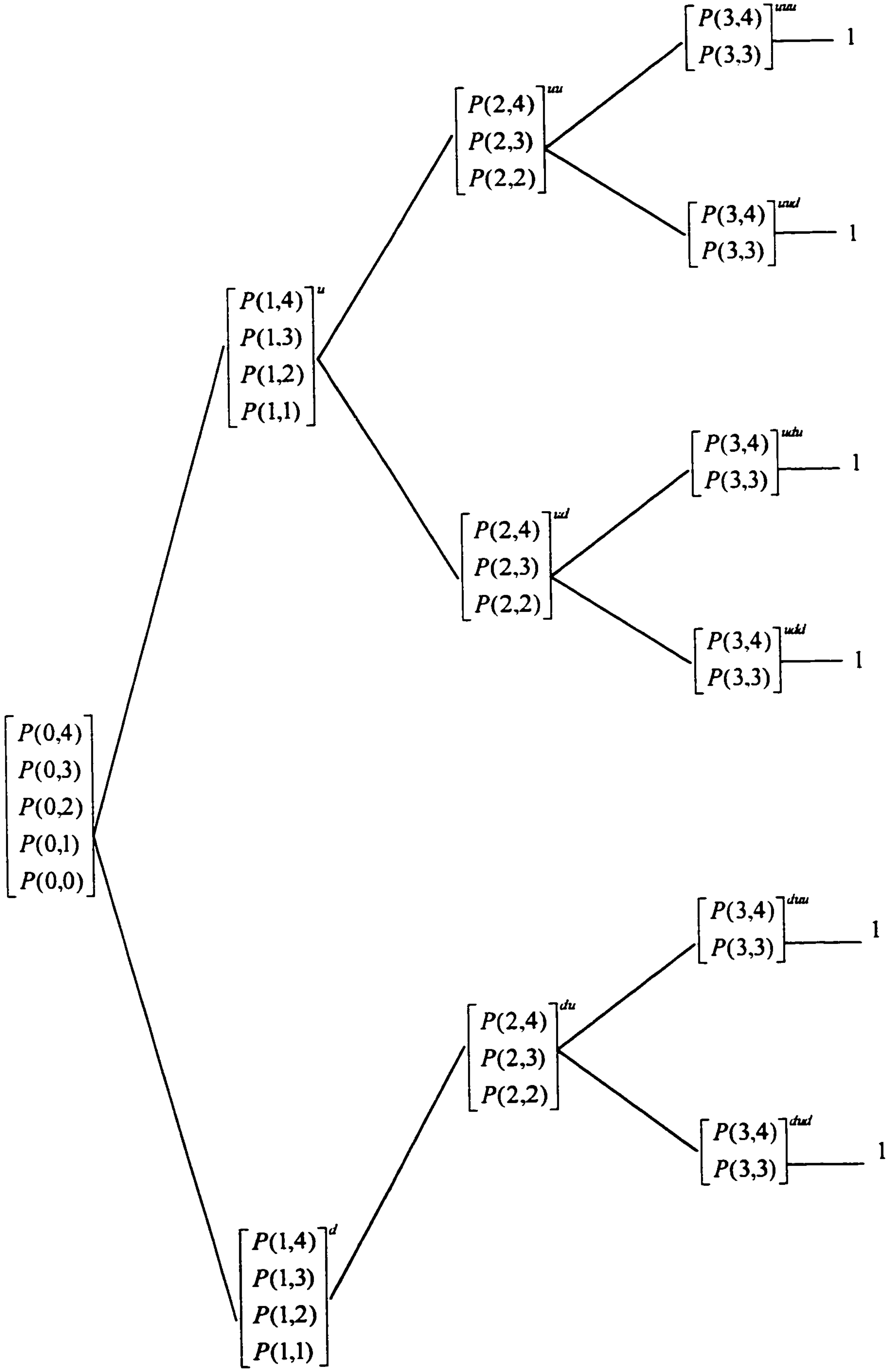


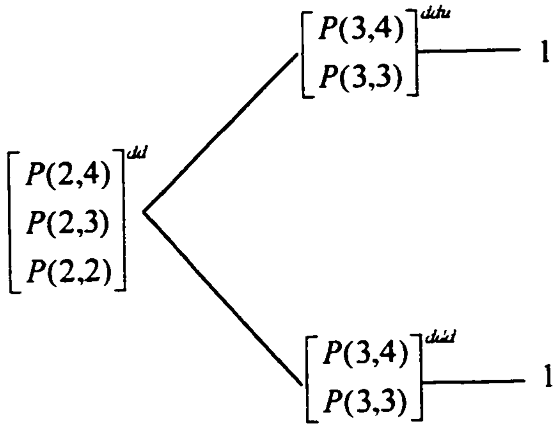
SPOT RATE TREE



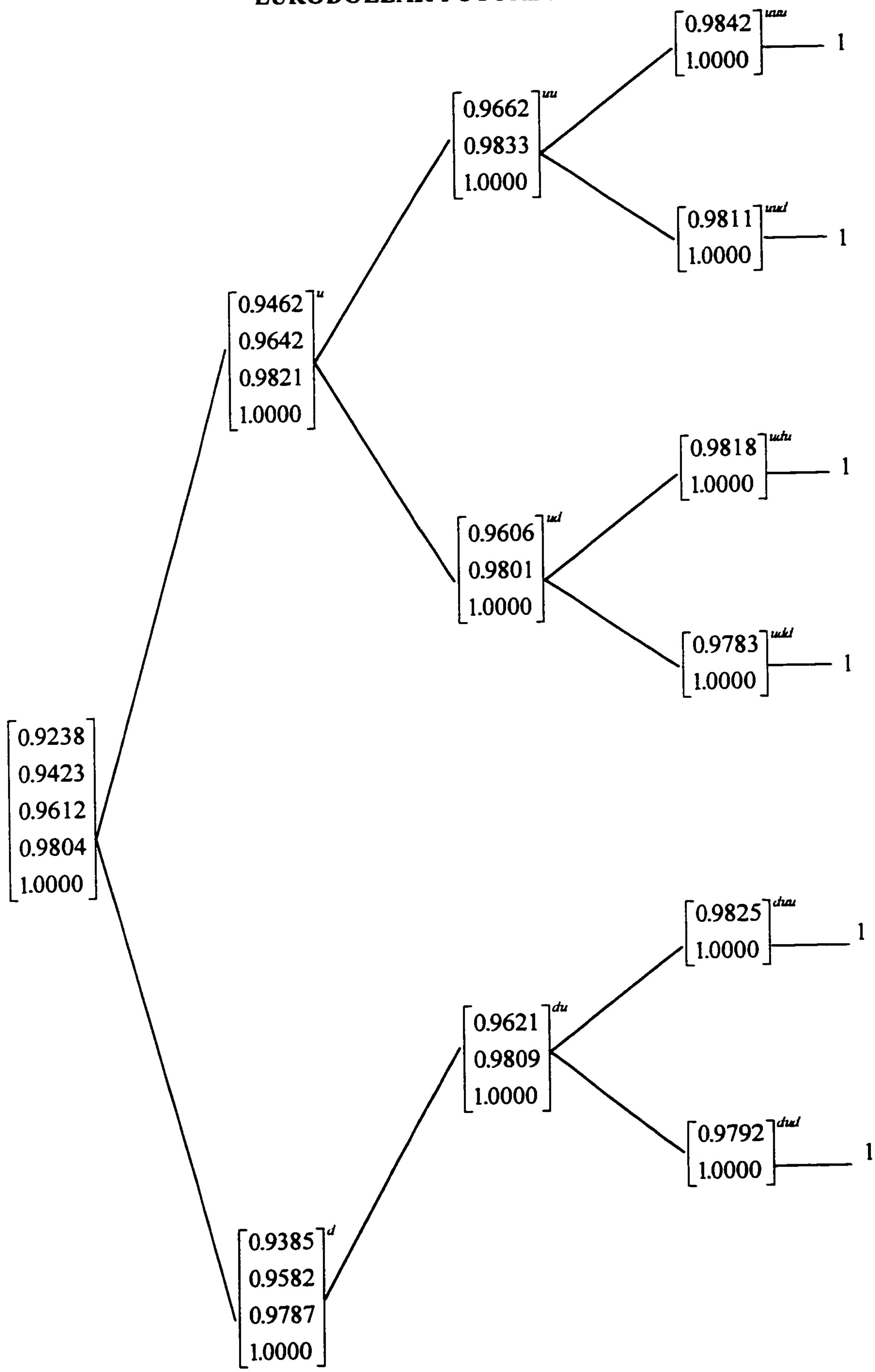


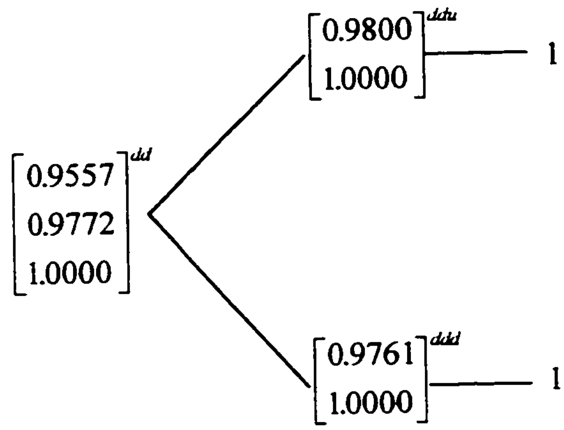
EURODOLLAR FUTURES PRICE TREE




P(3,4)
P(3,3)
ddu
1
P(2,4)
P(2,3)
P(2,2)
dd
P(3,4)
P(3,3)
ddd
1

EURODOLLAR FUTURES PRICE TREE





0.9800
1.0000
ddu
1
0.9557
0.9772
1.0000
dd
0.9761
1.0000
ddd
1

THE BLACK-DERMAN-TOY MODEL

I. INTRODUCTION

This section concentrates on pricing options on Eurodollar futures with a different class of volatility estimators using the single-factor Black, Derman, and Toy (1990) (henceforth BDT) interest rate model. The BDT model is based on common sense about the relation between today's long rates and future short rates. That is, it assumes that yields for long-term interest rate sensitive security reflect the market's expectation for future short rates. In this model, interest rates are assumed to be pulled back to some long-run average level over time. This phenomenon is known as *mean reversion*. When the interest rate is high, mean reversion tends to cause it to have a negative drift; when it is low, mean reversion tends to cause it to have a positive drift[1]. The main limitation of this model is that it assumes all changes in interest rates are caused by changes in short-term rates. As a result, a portfolio of bonds is no more diversified than a single bond, since both are influenced only by the short-term rate. Bond portfolios in this model are therefore probably too sensitive to interest rate changes.

We show that pricing Eurodollar futures options is sensitive to the volatility model used. The models utilized in our analysis for estimating the volatility of short term interest rates are *moving average models* such as historical and exponentially smoothed volatility, *time-series models* like GARCH and GARCH-X, and *implied volatility from explicit option pricing models*. The short-term interest rate driving the changes in the entire term structure is the fundamental variable in the BDT model. This makes the choice of a class of volatility estimation model for short rates crucial to pricing interest rate sensitive derivative securities.

This paper is organized as follows. Section II describes the BDT model and explains how it is used for pricing options on Eurodollar futures. Section III presents the models employed for forecasting the volatility of short-term interest rates. Section IV describes yield curve smoothing, the technique utilized in this paper for extracting one-month forward rates from the observable market data for Eurodollar futures. Section V provides the option prices estimated with different volatility structures. Section VI concludes the paper.

II. THE BLACK-DERMAN-TOY MODEL

In this section we explain how the current interest rates and their volatilities are used in the BDT model as input to determine its output, the value of an option on Eurodollar futures. We use two sets of input: a *yield curve* (the estimated yields from the market data for Eurodollar futures) and a *volatility curve* (the corresponding estimated yield volatilities for these same derivatives). These two curves together are called the *term structure of interest rates* in the BDT model which uses today's term structure to determine the expected means and standard deviations of future short rates.

In this model yields for long-term Eurodollar futures are assumed to reflect the market's expectations for future short-term rates. That is, when people expect future short-term rates to be high (low), they demand higher (lower) long-term yields today. The model quantifies this notion by deriving a *valuation formula* that lets us precisely compute expected long-term yields (or prices) for Eurodollar futures today using expected future short-term rates.

If people feel a great deal of uncertainty about future short-term rates, they will be correspondingly uncertain about how much change to expect in today's long-term yields. This means that current long-term yields will be highly volatile. The model quantifies this idea by deriving a *sensitivity formula* that predicts the expected change in any long-term yield today when the short-term rate changes. The current volatility of any Eurodollar futures is obtained from this formula.

The BDT model assumes that today's long rate reflects expected future short rates, and today's long rate volatility reflects expected future short-rate volatilities. Empirically, the term structure of volatilities slopes downward. In other words, long-term Eurodollar futures have lower yield (or forward rate) volatilities than short-term ones. One explanation for this phenomenon is as follows. Short-term rates are determined by current economic conditions while forward rates are determined by expectations of future short rates or, equivalently, future economic conditions. Furthermore, expectations or forward rates are much less volatile than current economic conditions or short-term rates: While current business conditions can change rapidly, one's expectation about the short-term rate in 20 years does not change much from day to day. Long-term rates are complex averages of short-term rates and forward rates. But since forward rates are less volatile than short-term rates, long-term rates will also be less volatile than short-term rates.

Following the traditional modeling of uncertainty of interest rates, the fundamental variable of this model - the short-term interest rate - is described as a random variable: given its value today, it will be uncertain tomorrow and more uncertain the next day. The model assumes that changes in short-term interest rates follow a lognormal probability distribution which implies that the evolution of the short rate R in discrete-time can be described by:

$$\Delta r(t) = [\delta_0(t) - \delta_1(t)\, r(t)]\ \lambda + \sigma(t)\ \Delta W(t), \qquad (1)$$

where $r(t) = \ln[R(t)]$ is the logarithm of the short-term interest rate , $\Delta r(t) \equiv r(t+\lambda) - r(t)$, λ is the length of the time interval, $\delta_0(t)$ and $\delta_1(t)$ are parameters that must be estimated, $\sigma(t)$ is the volatility at time t, and $\Delta W(t)$ is a normally distributed random variable with zero mean and variance of λ. We can approximate the process presented in equation (1) via the binomial representation:

$$r(t+\lambda) - r(t) = \begin{Bmatrix} [\delta_0(t) - \delta_1(t) r(t)]\lambda + \sigma(t)\sqrt{\lambda}; probability 1/2 \\ [\delta_0(t) - \delta_1(t) r(t)]\lambda - \sigma(t)\sqrt{\lambda}; probability 1/2 \end{Bmatrix} \qquad (2)$$

Since $\mu(r, t) = \delta_0(t) - \delta_1(t)\, r(t)$ is added to the log-interest rate in both the "up" and "down" states, $\mu(r, t)$ is called the drift of the logarithm of the short-rate. In this model, $\delta_1(t)$ can be treated as a measure of the speed of mean reversion in log-rate levels. At time t, if the log-interest rate is $r(t)$, then one period later, using equation (2), the interest rate will be either:

$$[r(t+\lambda)^{up} - r(t)] = [\delta_0(t) - \delta_1(t)\, r(t)]\ \lambda + \sigma(t)\sqrt{\lambda}$$

or

$$[r(t+\lambda)^{down} - r(t)] = [\delta_0(t) - \delta_1(t)\, r(t)]\ \lambda - \sigma(t)\sqrt{\lambda}\,.$$

Subtracting the above two equations yields the *sensitivity formula* which shows how the term structure of volatilities determines the spread between the two possible log-interest rates at time $t+\lambda$:

$$[r(t+\lambda)^{up} - r(t+\lambda)^{down}]/2 = \sigma(t)\sqrt{\lambda}\,. \qquad (3)$$

The BDT model uses one of the most popular hypothesis of term structure: *local expectations hypothesis*, the only hypothesis consistent with no-arbitrage[2]. If we apply the local expectations hypothesis (LEH) to the valuation of Eurodollar futures the LEH says that all Eurodollar futures provide the same expected rate of return over very small holding periods. Formally, this hypothesis is states as

$$\frac{E_t[P(t+1,T)]}{P(t,T)} = 1 + R_t .$$

The expected one-period return during the interval $[t, t+1]$ on a Eurodollar futures maturing at date T is equal to the one-period return R_t on a Eurodollar futures at date t maturing at date $t+1$.

This form of expectations yields the *valuation formula* used in the BDT model. To illustrate this idea, suppose that the price of a T- maturity Eurodollar futures starts out as $P(t,T)$ at time t and after one year moves up to $P(t+1,T)^{up}$ or down to $P(t+1,T)^{down}$ at date $t+1$ with equal probability of 1/2. The price at time t is determined by discounting the possible $t+1$ prices by the short rate prevailing at time t:

$$P(t,T) = \frac{E_t[P(t+1,T)]}{1+R_t} = \frac{(1/2)P(t+1,T)^{up} + (1/2)P(t+1,T)^{down}}{1+R_t} .$$

Having estimated the volatility of short rates, using the valuation and the sensitivity formula we construct a binomial tree of expected future short rates. This tree is then used to find the value and yield of Eurodollar futures of any maturity out to 6 time intervals where the length of each interval is 3 months. Given a market term structure and the binomial tree of short rates, the model is employed to value options on Eurodollar futures. First the expected future prices of Eurodollar futures at various points in time are calculated. These prices are then utilized to determine the option's value at expiration. Finally, the model's valuation formula is used to compute the discounted present value of the expected future cash flows of the option at expiration.

When we run the BDT model on a computer, we divide the 3-month time steps into three (i.e., each step becomes 1-month) after extracting the one-month forward rates from the observable market data for the prices of Eurodollar futures using the *maximum smoothness forward rates approach* which will be explained in Section IV. In this way we model the fluctuations in interest rates more realistically, and in so doing get a more accurate estimate for the value of Eurodollar futures today.

III. VOLATILITY ESTIMATION MODELS

This section describes the models - *moving average models, time-series models, implied volatility from explicit asset pricing models* - used in our analysis for estimating the volatility of short term interest rates. Which of the three classes of models best captures the dynamics of short rates is important for pricing Eurodollar futures options since the valuation of interest rate options is sensitive to the volatility model used.

A key point is that we are interested in valuing options with three different classes of volatility estimators, all of them (except implied volatility) based on a weighted sum of squared deviations from the mean for historical interest rates. We use a unified framework for estimating volatility that includes, as special cases, *historical volatility, exponential smoothed volatility, GARCH,* and *GARCH-X.* As discussed by Boudoukh, Richardson and Whitelaw (1997), these methods differ only in weights given to changes in the level of short term interest rates[3]. We also use *implied volatility*, obtained from the Black's (1976) formula, which sometimes provides a better estimate because the approach is forwardlooking in nature. We now describe the above-mentioned techniques which are employed for forecasting the volatility of short term interest rates.

MOVING AVERAGE MODELS

A moving average is an average taken over a *rolling window* of a fixed number of data points. Each time the window is rolled one point is knocked off behind and another is added at the end, so that the sample size remains fixed. Moving averages have been a useful tool in financial forecasting for many years. However, recent advances in time series analysis allow a more critical view of the efficacy of such methods, which will be discussed in the following section.

Historical Volatility: As mentioned earlier, the BDT model assumes that the short term interest rate (denoted R) is lognormally distributed. The simplest way to forecast the volatility of short rates in the BDT model is to find the usual estimate of the standard deviation of log-interest rate (denoted $r = \log R$):

$$\sigma_T^2 = \frac{\sum_{t=1}^{n}(r_t - \bar{r})^2}{n-1} \tag{4}$$

where *n-period historic volatility* denoted by σ_T depends on time to maturity T and is the unbiased estimate of standard deviation over a sample size n. That is, the traditional "historic" volatility in this study has a term structure because it has been used to value Eurodollar futures options of different maturities. This method places no structure on how volatility might evolve through time and puts constant weights $1/(n-1)$ on past observations. Hence, the historic volatility is taking no account of dynamic properties of interest rates, such as autocorrelation - it is essentially a static model which has been forced into a time-varying framework.

Exponentially Smoothed Volatility: There are two other criticisms of historical volatility. First, if interest rate volatility clusters, it follows that more recent rates should be given more weight. Interest rates in the last period provide more information about current volatility than rates some time ago. Second, the choice of the length of the observation period is somewhat arbitrary. Weight is given to observations that occur within the most recent *n* periods, while no weight is given to observations beyond this window. This procedure is ad hoc at best.

An *exponentially smoothed volatility* places more weight on more recent observations and this weighting is done by using a *smoothing constant* λ: the larger the value of λ the more weight is placed on past observations and so the smoother the series becomes. An *n*-period exponentially smoothed volatility estimator (ESVE) is defined as

$$\frac{(r_{t-1}-\bar{r})^2+\lambda(r_{t-2}-\bar{r})^2+\lambda^2(r_{t-3}-\bar{r})^2+\ldots\ldots+\lambda^{n-1}(r_{t-n}-\bar{r})^2}{1+\lambda+\lambda^2+\ldots\ldots+\lambda^{n-1}}.$$

Since the denominator converges to $1/(1-\lambda)$ as $n \to \infty$, an infinite ESVE can be written as

$$\sigma_t^2 = (1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}(r_{t-i}-\bar{r})^2 \qquad (5)$$

where $0 < \lambda < 1$. Of course, the researcher must truncate the interest rate series since only a finite number of observations are available for estimation[4]. It is this type of ESVE that is used for volatility forecasts in J.P. Morgan's RiskMetrics™. RiskMetrics forecasts of volatility over the next day are calculated by taking $\lambda = 0.94$. The infinite ESVE model is equivalent to an IGARCH model, which will be explained briefly in the next section, although the parameter λ is estimated in a sub-optimal way in RiskMetrics. So, the next day forecasts from RiskMetrics can look very similar to GARCH forecasts.

TIME-SERIES MODELS

GARCH Models: To analyze the variety of models for the short-term interest rate and its volatility, Chan, Karolyi, Longstaff and Sanders[5] (1992) consider the following discrete-time approximation to the continuous time short rate specification[6]:

$$R_t - R_{t-1} = \delta_0 + \delta_1 R_{t-1} + \varepsilon_t, \tag{6}$$

$$E(\varepsilon_t \mid I_{t-1}) = 0,\ E(\varepsilon_t^2 \mid I_{t-1}) \equiv \sigma_t^2 = \phi^2 R_{t-1}^{2\omega},$$

where R_t is the interest rate level at time t, $\delta_0 + \delta_1 R_{t-1}$ is the conditional mean and $\phi^2 R_{t-1}^{2\omega}$ is the conditional variance of unexpected interest rate changes. The drift parameter δ_1 can be viewed as a measure of the speed of mean reversion in rate levels. The more negative δ_1 is, the faster R corresponds to the deviations from the mean. The volatility parameter, ϕ^2, is simply a scale factor for the variance of unexpected interest rate changes. The parameter ω allows the volatility of interest rates to depend on the level of the interest rate. At higher ω's, the volatility is more sensitive to interest rate levels.

This model restricts the volatility to be a function of interest rate levels only, and not of the news arrival process. An alternative model that addresses this is the GARCH model, in which this period's volatility is a function of last period's unexpected news. For example, in a GARCH(1,1) model[7],

$$\sigma_t^2 \equiv E(\varepsilon_t^2 \mid I_{t-1}) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \tag{7}$$

where ε_{t-1} is the unexpected shock to R_t.

When $\alpha_1 + \beta = 1$ we can rewrite the GARCH(1,1) model as

$$\sigma_t^2 \equiv E(\varepsilon_t^2 \mid I_{t-1}) = \alpha_0 + (1-\lambda)\varepsilon_{t-1}^2 + \lambda \sigma_{t-1}^2, \qquad 0 < \lambda < 1. \tag{8}$$

Note that the unconditional variance $\sigma^2 = 1/(1 - \alpha_1 - \beta)$ is undefined - indeed we have a nonstationary GARCH model called the *Integrated* GARCH (IGARCH) model. Our main interest in the IGARCH model is that it is equivalent to an infinite ESVE, such as those used by RiskMetrics. This may be seen by repeated substitution in (8):

$$\sigma_t^2 = \alpha_0 + (1-\lambda)\,\varepsilon_{t-1}^2 + \lambda\,(\alpha_0 + (1-\lambda)\,\varepsilon_{t-2}^2 + \lambda\,(\alpha_0 + (1-\lambda)\,\varepsilon_{t-3}^2 + \lambda(.......$$

$$= \alpha_0/(1-\lambda) + (1-\lambda)(\,\varepsilon_{t-1}^2 + \lambda\,\varepsilon_{t-2}^2 + \lambda^2\,\varepsilon_{t-3}^2 + ...) \qquad (9)$$

where $\varepsilon_{t-1} = r_{t-1} - E(r_{t-1})$ and in the ESVE model $E(r_{t-1})$ is the sample mean of r. Comparing the ESVE model (5), the representation (9) is equivalent when $\alpha_0 = 0$.

Brenner, Harjes and Kroner (1996) show that models which parameterize volatility only as a function of interest rate levels tend to overemphasize the sensitivity of volatility to levels and fail to model adequately the serial correlation in conditional variances. On the other hand, serial correlation based models like GARCH models fail to capture adequately the relationship between interest rate levels and volatility. They introduce a new class of models for the dynamics of short-term interest rate volatility, which allows volatility to depend on both interest rate levels and information shocks. In the original BDT model, the volatility of log-interest rates is assumed to depend only on time and not on the short rate itself. Therefore, we first use GARCH(1,1) model which does not relate the volatility of short rates to the interest rate levels. In addition to GARCH(1,1), following Brenner et. al, we also use a modified GARCH model (GARCH-X) that captures both the serial correlation in conditional variances and the dependence of variances on interest rate levels.

The BDT model incorporates mean reversion of short term interest rates which are assumed to be lognormally distributed. We therefore choose to mean-adjust the logarithm of the series to allow for the effect of mean reversion in interest rates. Mean-adjusted log-interest rates (denoted $\tilde{r}_t = \log R_t - \overline{\log R_t}$) can be decomposed into a forecastable part $E(\tilde{r}_t \mid I_{t-1})$ (i.e., the conditional mean, where $E(\ldots|I_{t-1})$ represents conditioning on the information available at time t-1), and an unpredictable portion, ε_t:

$$\tilde{r}_t = E(\tilde{r}_t \mid I_{t-1}) + \varepsilon_t = \delta_1 \tilde{r}_{t-1} + \varepsilon_t, \tag{10}$$

where $E(\varepsilon_t \mid I_{t-1}) = 0$ and the conditional variance of the mean-adjusted log-interest rates in GARCH-X model , σ_t^2 is

$$\sigma_t^2 \equiv E(\varepsilon_t^2 \mid I_{t-1}) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma R_{t-1}^2, \tag{11}$$

where the autoregressive conditional heteroscedasticity enters through both the squared residuals and the squared interest rate levels.

IMPLIED VOLATILITY

In this paper, selecting the volatility of the short-term interest rate by an implied method means that the volatility is selected so that the model price for Eurodollar futures matches its market price. The hope here is that forcing the model to match a particular market price makes the model more realistic.

The one parameter in the Black's (1976) pricing formulas that cannot be observed directly is the volatility of the interest rate. In the previous section, we discussed how this can be estimated from a history of the interest rate. It is now appropriate to mention an alternative approach that uses what is termed an *implied volatility*. When an explicit analytic pricing formula is available (such as the Black's formula) the quoted prices of Eurodollar futures options, along with known variables such as the interest rates, time to maturity, exercise prices, can be used in an implicit formula for volatility. The result is called the implied volatility. It is a volatility forecast which is implicit in the quoted price of the option. The forecast horizon is therefore given by the maturity of the option.

Typically there are several options of different maturities and strikes on Eurodollar futures, and different options give different volatility forecasts for the same implied volatility. In general at-the-money options yield lower implied volatilities than in-the-money or out-of-the-money options, and this is called the volatility smile. Since a single implied volatility is required for pricing Eurodllar futures options, rather than the whole smile, we take a vega weighted average of all volatilities across the smile.

We calculate volatility estimates of daily changes in the annualized spot and forward rates obtained from Eurodollar futures. Each estimate is calculated using 90 daily observations on a rolling basis. In other words, the first 90 observations in the sample are used to compute an estimate for the following day (day 91), using each of the four methods (*historical volatility, exponentially smoothed volatility, GARCH and GARCH-X*) described above. Changes in log-interest rates are mean-adjusted using the sample mean from the 90-day estimation period. The analysis is performed over the period 1987-1998.

IV. YIELD CURVE SMOOTHING

This section is devoted to yield curve smoothing, the technique used for extracting forward rates and zero coupon bond yields from observable market data with different maturities than the maturities needed for analysis. Adams and van Deventer (1994) introduced a mathematical measure of smoothness and show that the yield curve with the smoothest possible forward rate function, consistent with the observable data, is related to, but significantly different from, the popular *cubic spline approach* to the smoothing of yields and discount bond prices. The yield curve that produces the smoothest possible forward rates consistent with given zero coupon bond prices has a quadratic forward rate function. This contrasts with the cubic polynomial used to fit either yields or discount bond prices in the cubic spline approach. There are two principal problems associated with the use of cubic yield (or price) splines:

(a) Cubic spline smoothing produces forward rate curves that are not "smooth" in that the forward rate curve is not twice differentiable at the knot point.

(b) In addition, the forward rate curves associated with a cubic spline-based yield curve tend to be volatile, particularly on the right-hand side of the yield curve, to such a degree that their use can lead to implausible forward rate curves.

For these reasons, the *maximum smoothness forward rate approach* offers a number of advantages.

***Maximum Smoothness Forward Rates*:** Adams and van Deventer attempt to remedy the two problems of cubic spline smoothing with a new approach that addresses both problems directly:

(a) They seek to derive a forward rate curve that is continuous and twice differentiable.

(b) They derive the curve in such a way that the forward rate curve is the smoothest curve of any of the family of curves that are continuous, twice differentiable, and consistent with the observable data.

Adams and van Deventer show that the smoothest possible forward rate curve consists of a quadratic forward rate function that is fitted between each knot point. This technique is utilized in our analysis for extracting one-month forward rates from observable data for Eurodollar futures as follows.

The term structure $f(T)$, $0 \le T \le \tau$, of forward rates that satisfies the maximum smoothness criterion

$$\min \int_0^{\tau} [f''(s)]^2 \, ds$$

while fitting the observed prices $P(T_1)$, $P(T_2)$, ..., $P(T_m)$ of Eurodollar futures with maturities T_1, T_2, ...T_m is a fourth-order spline given by

$$f(T) = a_i + b_i T + c_i T^2 + d_i T^3 + e_i T^4 \quad \text{for } T_{i-1} < T < T_i, \ i = 1, 2, ..., m+1$$

where $0 = T_0 < T_1 < T_2 < ..... < T_m < T_{m-1} = \tau$.

The use of the maximum smoothness forward rate function is somewhat more complex than the use of the cubic spline approach because there is a larger number of parameters to be determined. In the case where we have n observable data points, we have $5n$ unknowns since we need to find a, b, c, d, and e for each of the n segments of the forward rate curve. We have the following constraints that are essential to ensure the reasonableness of the resulting forward rate $f(T)$, yield $y(T)$, and price $P(T)$ curves:

(1) n-1 equations requiring that the forward rates be equal at each knot:

$$a_i + b_i T_i + c_i T_i^2 + d_i T_i^3 + e_i T_i^4 = a_{i-1} + b_{i-1} T_i + c_{i-1} T_i^2 + d_{i-1} T_i^3 + e_{i-1} T^4 \quad \text{for } i=1, 2, ..., n-1$$

(2) n-1 equations requiring that the first derivative of the forward rate curve be equal at each knot

$$b_i + 2c_i T_i + 3d_i T_i^2 + 4e_i T_i^3 = b_{i-1} + 2c_{i-1} T_i + 3d_{i-1} T_i^2 + 4e_{i-1} T_i^3 \quad \text{for } i = 1, 2, ..., n-1$$

(3) n-1 equations requiring that the second derivative of the forward rate curve be equal at each knot

$$2c_i + 6d_i T_i + 12e_i T_i^2 = 2c_{i-1} + 6d_{i-1} T_i + 12e_{i-1} T_i^2 \quad \text{for } i = 1, 2, ..., n-1$$

(4) n-1 equations requiring that the third derivative of the forward rate curve be equal at each knot

$$6d_i + 24e_iT_i = 6d_{i-1} + 24e_{i-1}T_i \qquad \text{for } i = 1, 2, \ldots, n\text{-}1$$

(5) n constraints that the forward rate curves be consistent with observable data[8]

$$a_i(T_i - T_{i-1}) + \frac{1}{2}b_i(T_i^2 - T_{i-1}^2) + \frac{1}{3}c_i(T_i^3 - T_{i-1}^3) + \frac{1}{4}d_i(T_i^4 - T_{i-1}^4) + \frac{1}{5}e_i(T_i^5 - T_{i-1}^5) = -\ln\left(\frac{P(T_i)}{P(T_{i-1})}\right)$$

for the n observable data points from $i = 1$ to n, noting that P for $T = T_0$ is 1.

So far, these constraints give us 4(n-1) + n or 5n-4 equations. We require two other constraints of economic significance

That the forward rate curve be consistent with an observable short rate $y(0)$, or

$$a_1 = y(0) = f(0)$$

That the slope of the forward rate curve at the right-hand side of the yield curve to be zero

(i.e., $f' = 0$) $\qquad b_n + 2c_nT_n + 3d_nT_n^2 + 4e_nT_n^3 = 0$

We can complete the system of 5n equations in 5n unknowns by imposing the additional constraints that the forward rate curve be instantaneously straight at both the left-hand and right-hand side of the curve, so that

$$f''(T_0) = 2c_1 + 6d_1T_0 + 12e_1T_0^2 = 0$$

$$f''(T_n) = 2c_n + 6d_nT_n + 12e_nT_n^2 = 0$$

We can then solve for each of the n sets of a, b, c, d, and e using matrix inversion.

V. RESULTS

This paper explains arbitrage-free term structure models used for pricing options on Eurodollar Futures with particular emphasis on the Black-Derman-Toy (BDT) and the

Heath-Jarrow-Morton (HJM) interest rate models and their applications. In this essay, I concentrate on valuing options on Eurodollar futures using the BDT and the HJM models *with different volatility structures*. I compare the estimated option prices of Eurodollar futures with the actual values to determine which of these two models [specified with different dynamics of interest rate volatility: historical volatility, exponentially smoothed volatility, implied volatility, GARCH, GARCH-X] perform better in pricing interest rate sensitive derivative securities.

First, I have attempted to explain the binomial approximation to the continuous trading term structure models of Heath-Jarrow-Morton and Black-Derman-Toy. Second, I have defined the dynamics of interest rate volatility which is fundamental to these two interest rate models. Finally, I have presented the models solutions in 4-time-steps to save space. However, the actual study in this paper is based on comparing the estimated prices of Eurodollar futures options computed by using HJM and BDT in 3-time-steps, 6-time-steps, 9-time-steps, 12-time-steps, 15-time-steps, and 18-time-steps with the actual prices of options. My purpose here is to determine which of these two models (HJM and BDT) specified with different dynamics of interest rate volatility (historical volatility, exponentially smoothed volatility, implied volatility, GARCH, GARCH-X) perform better in pricing interest rate sensitive derivative securities.

To evaluate the accuracy of the forecasts of Eurodollar futures options, I calculate mean squared errors (MSE) for each model (HJM or BDT) specified with different volatility structures and compare them. MSE provides a quadratic loss function:

$$\text{MSE} = \frac{1}{n}\sum_{t=1}^{n}(P_t - P_t^f)^2$$

where n is the total number of forecasts computed (in this paper, n is around 1100 depending on the type of model used), P_t and P_t^f represent the actual and estimated prices of Eurodollar futures options at time t, respectively.

I first use the original data set which includes the 3-month Eurodollar futures prices and compare the estimated option prices with the actual values without increasing the number of steps. Having extracted the 1-month Eurodollar futures prices from the 3-month Eurodollar futures prices employing the "Maximum Smoothness Forward Rates Approach" discussed in the previous section, I use the 1-month instead of the 3-month Eurodollar futures prices and increase the accuracy of forecasts for each model which is reflected in lower Mean-Squared-Errors.

Forecasting Results:

Grouping	Moving Average		Time Series		Implied Volatility
	Historic	ESVE	GARCH	GARCH-X	Black's formula
All	0.2581	0.3466	0.4584	0.4325	0.3875
Calls	0.1153	0.1445	0.4442	0.4281	0.1518
Puts	0.3967	0.5196	0.4718	0.4508	0.5109
Out-of-the-money	0.1279	0.1751	0.4615	0.4279	0.1933
At-the-money	0.1163	0.2568	0.2186	0.1856	0.2688
In-the-money	0.4575	0.5526	0.5471	0.5163	0.5866
3-month	0.0745	0.1239	0.2378	0.1795	0.1514
6-month	0.1933	0.2549	0.4065	0.3859	0.2699
9-month	0.3826	0.5339	0.5576	0.5632	0.5882
12-month	0.4724	0.6543	0.6195	0.5741	0.6599
15-month	0.6135	0.8058	0.7157	0.6932	0.7782
18-month	0.3730	0.6207	0.5700	0.5412	0.5687

The table shown above presents my forecasting results with categorizing the estimated option prices. First the whole estimates are compared without any grouping.

The results suggest that Moving Average Models outperform Time-series and Implied Volatility Models. Then there are three categories used for evaluating the forecasts: Calls-Puts is the first category, Moneyness (out-of-the-money, at-the-money, in-the-money) is the second category, and 3-month-steps is the third category.

For call options, Moving Average Models again outperform Time-series and Implied Volatility Models but for put options, there is no obvious volatility model which outperforms the others. However, when the mean squared errors are compared Time-series models yields slightly better price estimates of Eurodollar futures options than Moving Average and Implied Volatility Models.

The second grouping is "moneyness". Options are categorized as out-of-the-money, at-the-money, and in-the-money options. According to the results for out-of-the-money options, Moving Average Models have slightly higher performance than Implied Volatility Models but both yield much better estimates than Time-series models. For at-the-money options, Moving Average and Time-series models have better forecasting performance than Implied Volatility Models. For in-the-money options, all models have almost the same forecasting capability.

For the time-steps, in general Time-series and Implied Volatility Models yield almost the same quadratic loss values and these are greater than those from Moving Average Models. The least mean squared errors are obtained from the 3-month step since it is more liquid than the longer-term options like 15-month-step or 18-month-step options. Therefore, it is easier to price the shorter-term options with the BDT interest rate model using Moving Average Volatility Models.

ENDNOTES:

---

[1] Economic explanation of this is as follows: When rates are higher, the economy tends to slow down and there is less requirement for funds on the part of borrowers. As a result, rates decline. When rates are low, there tends to a high demand for funds on the part of borrowers. As a result, rates tend to rise.

[2] A number of hypotheses - such as unbiased expectations hypothesis, liquidity premium hypothesis, segmented market hypothesis, and local expectations hypothesis - have been advanced concerning the term structure of interest rates. They attempt to relate interest rates on pure discount bonds of differing maturities with forward rates, expected spot rates, and so on. The LEH used by the BDT model is the only hypothesis that is free from arbitrage.

[3] Boudoukh et al., find density estimation and RiskMetrics™ forecasts to be more accurate than GARCH and historical volatility for forecasting short-term interest rate volatility.

[4] Using one year of past data (250 daily observations) and a smoothing parameter of 0.96, for example, the truncation is done at a point where the weight on all the omitted observations would have been 0.000037.

[5] In addition to Chan, Karolyi, Longstaff and Sanders (1992), Marsh and Rosenfeld (1983), Dietrich-Campbell and Schwartz (1986) define the volatility of interest rates as a function of the level of the interest rate.

[6] Chan et. al, use a generalized continuous time short rate specification,

$$dR = (\delta_0 + \delta_1 R)\, dt + \phi R^{\omega}\, dZ.$$

They show that with appropriate restrictions on δ_0, δ_1, ϕ and ω, many popular interest rate models can be obtained. For example, setting $\omega=0$ gives the Vasicek (1977) model. The Merton (1973) model can be nested within the Vasicek model by the parameter restriction $\delta_1=0$. Setting $\omega=1/2$ gives the Cox, Ingersoll, and Ross (1985) square root model, and setting $\delta_0=0$ gives the Cox (1975) constant elasticity of variance model. Setting $\omega=1$ yields the Brennan-Schwartz (1980) model and the familiar geometric Brownian motion (GBM) process of Black and Scholes (1973) is nested within the Brennan-Schwartz model by the parameter restriction $\delta_0=0$. In turn, the Dothan (1978) model is nested within the GBM model by the parameter restriction $\delta_1=0$.

[7] See Engle, Lilien, and Robins (1987) and Evans (1989) for applications of this model to interest rate data, or Bollerslev, Chou, and Kroner (1992) for a comprehensive survey.

[8] The next set of constraints comes from the fact that

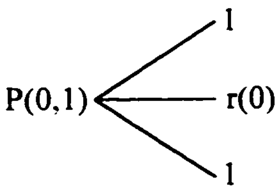
$$P(T) = e^{-\int_0^T f(s)ds} .$$

Since we are using a forward rate function broken into quadratic segments and we have observable data, we can write

$$P(T_i) = P(T_{i-1})\, e^{-\int_{T_{i-1}}^{T_i} f(s)ds} .$$

Rearranging this equation and expressing it as a linear function of the parameters a, b, c, d, and e gives the next set of constraints presented in the text.

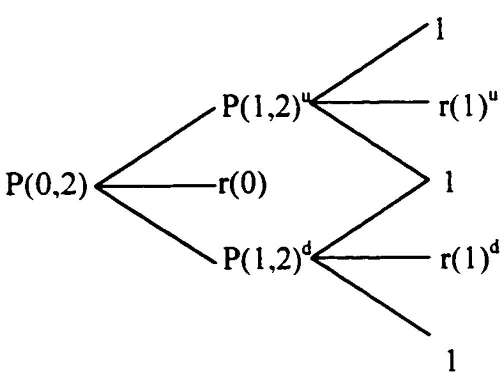
Bond Price Process for P(0,1):



$$\pi^{up} = \pi^{down} = 0.5$$

$$P(0,1) = \left(\frac{1}{1+r(0)}\right)(0.5 \times 1 + 0.5 \times 1) = \frac{1}{1+r(0)}$$

Bond Price Process for P(0,2):



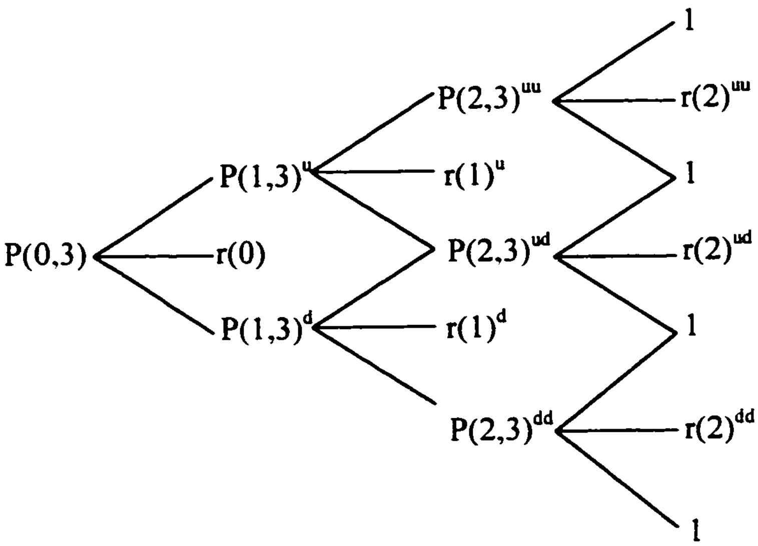
$$P(1,2)^u = \left(\frac{1}{1+r(1)^u}\right)(0.5 \times 1 + 0.5 \times 1) = \frac{1}{1+r(1)^u}$$

$$\sigma_2 = \frac{\ln(r(1)^u) - \ln(r(1)^d)}{2}$$

$$P(1,2)^d = \left(\frac{1}{1+r(1)^d}\right)(0.5 \times 1 + 0.5 \times 1) = \frac{1}{1+r(1)^d}$$

$$P(0,2) = \left(\frac{1}{1+r(0)}\right)\left(0.5 \times P(1,2)^u + 0.5 \times P(1,2)^d\right)$$

Bond Price Process for P(0,3):



$$P(2,3)^{uu} = \left(\frac{1}{1+r(2)^{uu}}\right)(0.5\times 1 + 0.5\times 1) = \frac{1}{1+r(2)^{uu}}$$

$$\sigma_3 = \frac{\ln(r(2)^{uu}) - \ln(r(2)^{ud})}{2}$$

$$P(2,3)^{ud} = \left(\frac{1}{1+r(2)^{ud}}\right)(0.5\times 1 + 0.5\times 1) = \frac{1}{1+r(2)^{ud}}$$

$$\sigma_3 = \frac{\ln(r(2)^{ud}) - \ln(r(2)^{dd})}{2}$$

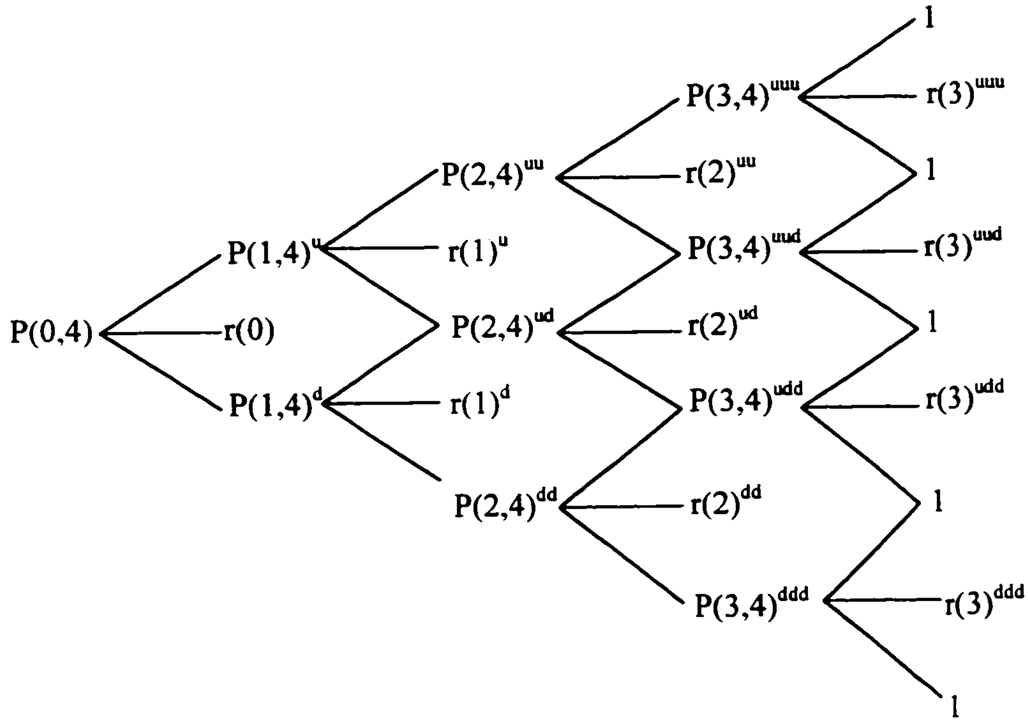
$$P(2,3)^{dd} = \left(\frac{1}{1+r(2)^{dd}}\right)(0.5\times 1 + 0.5\times 1) = \frac{1}{1+r(2)^{dd}}$$

$$P(1,3)^{u} = \left(\frac{1}{1+r(1)^{u}}\right)\left(0.5\times P(2,3)^{uu} + 0.5\times P(2,3)^{ud}\right)$$

$$P(1,3)^{d} = \left(\frac{1}{1+r(1)^{d}}\right)\left(0.5\times P(2,3)^{ud} + 0.5\times P(2,3)^{dd}\right)$$

$$P(0,3) = \left(\frac{1}{1+r(0)}\right)\left(0.5\times P(1,3)^{u} + 0.5\times P(1,3)^{d}\right)$$

Bond Price Process for P(0,4):



$$P(3,4)^{uuu} = \left(\frac{1}{1+r(3)^{uuu}}\right)(0.5\times 1 + 0.5\times 1) = \frac{1}{1+r(3)^{uuu}} \qquad \sigma_4 = \frac{\ln(r(3)^{uuu}) - \ln(r(3)^{uud})}{2}$$

$$P(3,4)^{uud} = \left(\frac{1}{1+r(3)^{uud}}\right)(0.5\times 1 + 0.5\times 1) = \frac{1}{1+r(3)^{uud}} \qquad \sigma_4 = \frac{\ln(r(3)^{uud}) - \ln(r(3)^{udd})}{2}$$

$$P(3,4)^{udd} = \left(\frac{1}{1+r(3)^{udd}}\right)(0.5\times 1 + 0.5\times 1) = \frac{1}{1+r(3)^{udd}} \qquad \sigma_4 = \frac{\ln(r(3)^{udd}) - \ln(r(3)^{ddd})}{2}$$

$$P(3,4)^{ddd} = \left(\frac{1}{1+r(3)^{ddd}}\right)(0.5\times 1 + 0.5\times 1) = \frac{1}{1+r(3)^{ddd}}$$

$$P(2,4)^{uu} = \left(\frac{1}{1+r(2)^{uu}}\right)\left(0.5\times P(3,4)^{uuu} + 0.5\times P(3,4)^{uud}\right)$$

$$P(2,4)^{ud} = \left(\frac{1}{1+r(2)^{ud}}\right)\left(0.5\times P(3,4)^{uud} + 0.5\times P(3,4)^{udd}\right)$$

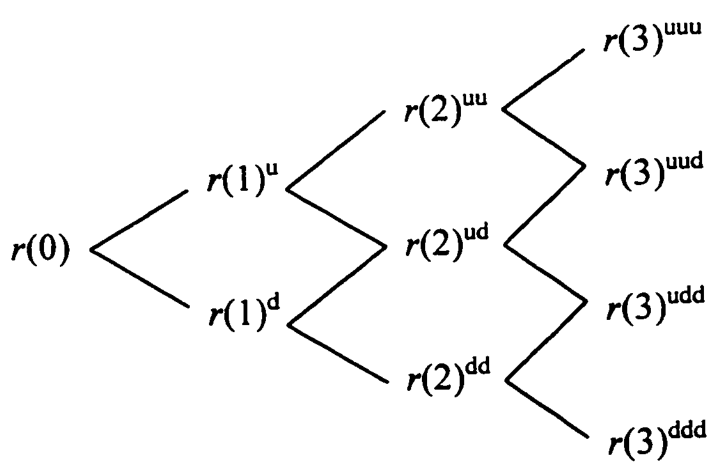
$$P(2,4)^{dd} = \left(\frac{1}{1+r(2)^{dd}}\right)\left(0.5 \times P(3,4)^{udd} + 0.5 \times P(3,4)^{ddd}\right)$$

$$P(1,4)^{u} = \left(\frac{1}{1+r(1)^{u}}\right)\left(0.5 \times P(2,4)^{uu} + 0.5 \times P(2,4)^{ud}\right)$$

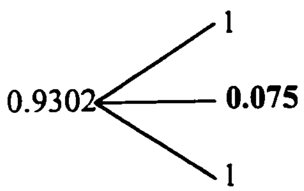
$$P(1,4)^{d} = \left(\frac{1}{1+r(1)^{d}}\right)\left(0.5 \times P(2,4)^{ud} + 0.5 \times P(2,4)^{dd}\right)$$

$$P(0,4) = \left(\frac{1}{1+r(0)}\right)\left(0.5 \times P(1,4)^{u} + 0.5 \times P(1,4)^{d}\right)$$

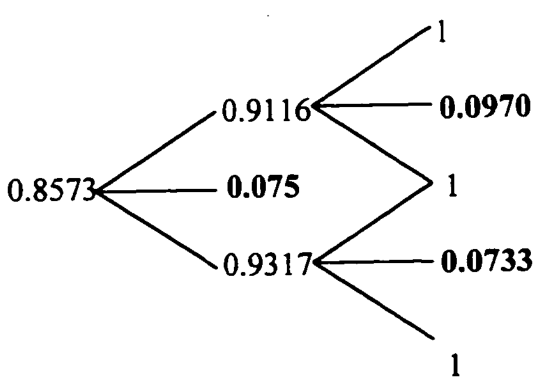
Spot Rate Process:



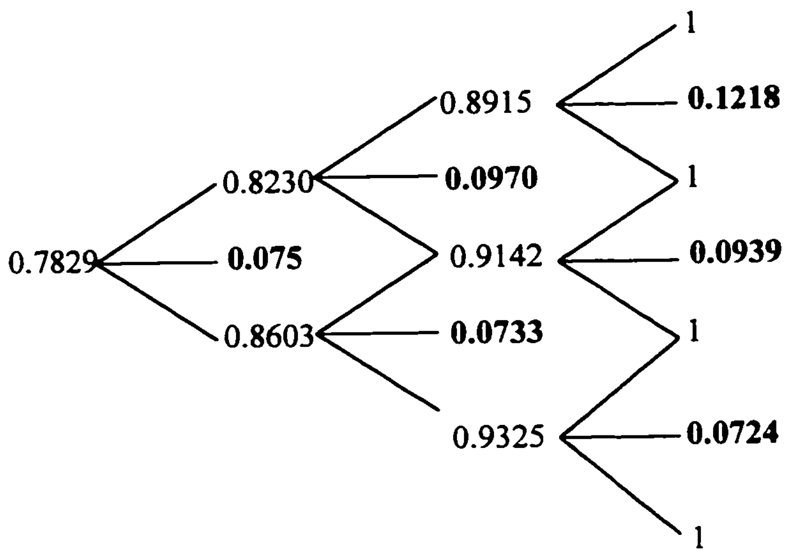
Bond Price Process for P(0,1):



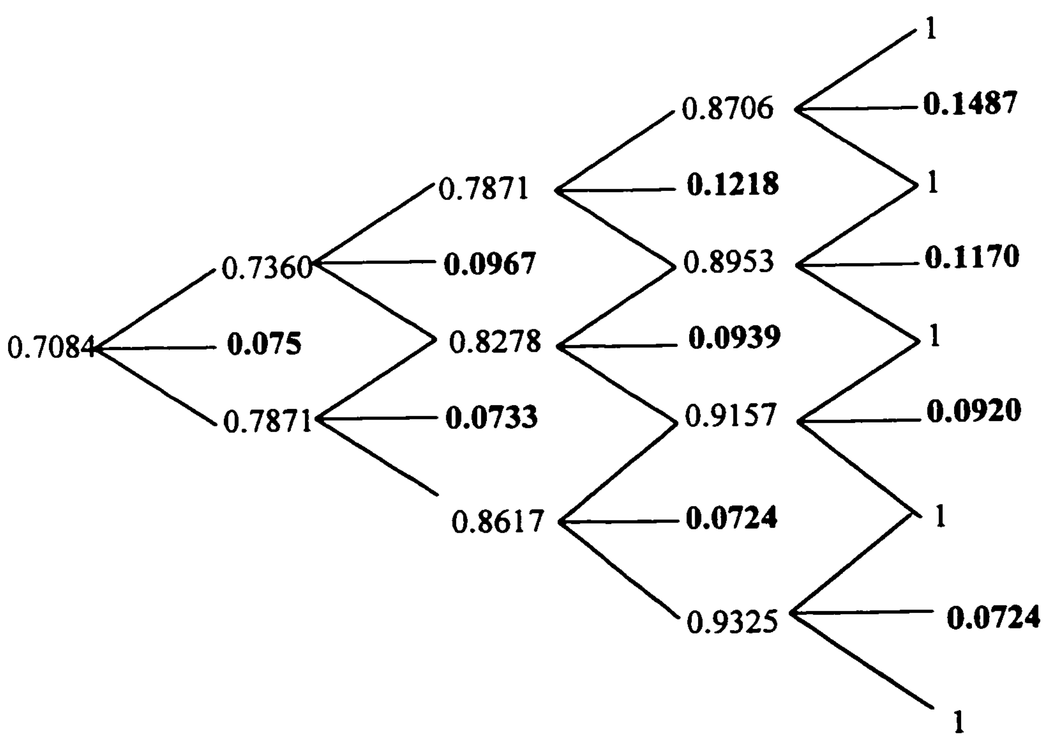
Bond Price Process for P(0,2):



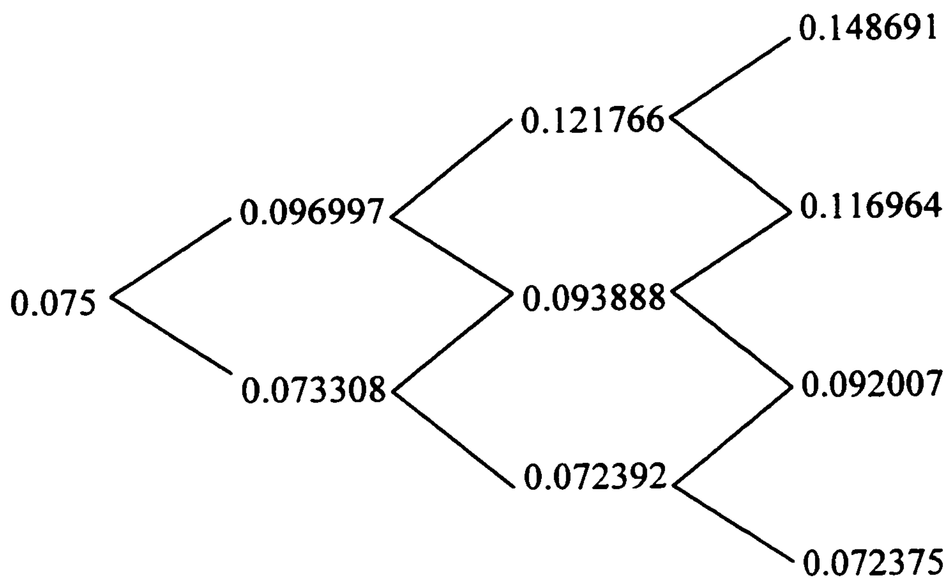
Bond Price Process for P(0,3):



Bond Price Process for P(0,4):



Spot Rate Process:



REFERENCES:

Black, F., E. Derman, and W. Toy, 1990. "A One Factor Model of Interest Rates and Its Applications to Treasury Bond Options." *Financial Analyst Journal* 46, 33-39.

Black, F., and Myron Scholes, 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, 81, 637-654.

Brennan , Michael J., and E. S. Schwartz, 1977. "Savings Bonds, Retractable Bonds, and Callable Bonds." *Journal of Financial Economics*, 3, 133-155.

________, 1980. "Analyzing Convertible Bonds." *Journal of Financial and Quantitative Analysis*, 15, 907-929.

________, 1982. "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency." *Journal of Financial and Quantitative Analysis*, 17, 75-100.

Courtadon, G., 1982. "The Pricing of Options on Default-Free Bonds." *Journal of Financial and Quantitative Analysis*, 17, 75-100.

Cox, John C., 1975. "Notes on Option Pricing I: Constant Elasticity of Variance Diffusions." Working paper, Stanford University.

Cox, J., J. Ingersoll, and S. Ross, 1980. "An Analysis of Variable Rate Loan Contracts." *Journal of Finance* 35, 389-403.

________, 1985. "A Theory of the Term Structure of Interest Rates." *Econometrica* 53, 385-407.

________, and Stephen A. Ross, 1976. "The Valuation of Options for Alternative Stochastic Processes." *Journal of Financial Economics*, 3, 145-166.

Dietrich-Campbell, Bruce, and E. S. Schwartz, 1986. "Valuing Debt Options: Empirical Evidence." *Journal of Financial Economics*, 16, 321-343.

Dothan, Uri L., 1978. "On the Term Structure of Interest Rates." *Journal of Financial Economics*, 6, 59-69.

Dunn, K. B., and J. J. McConnell, 1981. "Valuation of GNMA Mortgage-Backed Securities." *Journal of Finance*, 36, 599-616.

Gibson, R. and E. S. Schwartz, 1990. "Stochastic Convenience Yield and the Pricing of Oil Contingent Claims." *Journal of Finance*, 45, 959-976.

Fama, Eugene F., 1984. "Term Premiums in Bond Returns." *Journal of Financial Economics*, 13, 529-546.

He, Hua, 1990. "Convergence from Discrete-to-Continuous-Time Contingent Claims Prices." *Review of Financial Studies* 3 (4), 523-546.

Heath, D., R. Jarrow, and A. Morton, 1990. "Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation." *Journal of Financial and Quantitative Analysis* 25, 419-440.

Heath, D., R. Jarrow, and A. Morton, 1991. "Contingent Claim Valuation with a Random Evolution of Interest Rates." *Review of Futures Markets* 9 (1), 54-76.

Heath, D., R. Jarrow, and A. Morton, 1992. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation." *Econometrica* 60 (1), 77-105.

Ho, T.S., and S. Lee, 1986. "Term Structure Movements and Pricing Interest Rate Contingent Claims." *Journal of Finance* 41, 1011-1028.

Hull, J., and A. White, 1990. "Pricing Interest Rate Derivative Securities." *Review of Financial Studies* 3 (4), 573-592.

Jamshidian, Farshid, 1989. "An Exact Bond Pricing Formula." *Journal of Finance*, 44, 205-209.

Longstaff, F. A., 1990. "The Valuation of Options on Yields." *Journal of Financial Economics*, 26, 97-122.

Marsh, Terry A., and Eric R. Rosenfeld, 1983. "Stochastic Processes for Interest Rates and Equilibrium Bond Prices." *Journal of Finance*, 38, 635-646.

Merton, Robert C., 1973. "Theory of Rational Option Pricing." *Bell Journal of Economics and Management Science*, 4, 141-183.

Ramaswamy, K., and S. M. Sundaresan, 1986. "The Valuation of Floating-Rate Instruments." *Journal of Financial Economics*, 17, 251-272.

Sanders, A. B., and Haluk Unal, 1988. "On the Intertemporal Stability of the Short-term Rate of Interest." *Journal of Financial and Quantitative Analysis*, 23, 417-423.

Sundaresan, S. M., 1989. "Valuation of Swaps." Working paper, Columbia University.

Vasicek, O., 1977. "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics* 5, 177-188.

Wooldridge, J. M., 1990. "A Unified Approach to Robust, Regression-Based Specification Tests." *Econometric Theory*, 6, 17-43.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989