NBER WORKING PAPER SERIES

ESTIMATING THE CONTINUOUS TIME CONSUMPTION BASED ASSET PRICING MODEL

Sanford J. Grossman

Angelo Melino

Robert J. Shiller

Working Paper No. 1643

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1985

The research reported here is part of the NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER WORKING PAPER #1643
June 1985

Estimating the Continuous Time Consumption Based Asset Pricing Model

ABSTRACT

The consumption based asset pricing model predicts that excess yields are determined in a fairly simple way by the market's degree of relative risk aversion and by the pattern of covariances between per capita consumption growth and asset returns. Estimation and testing is complicated by the fact that the model's predictions relate to the instantaneous flow of consumption and point-in-time asset values, but only data on the integral or unit average of the consumption flow is available. In our paper, we show how to estimate the parameters of interest consistently from the available data by maximum likelihood. We estimate the market's degree of relative risk aversion and the instantaneous covariances of asset yields and consumption using six different data sets. We also test the model's overidentifying restrictions.

Sanford J. Grossman
Department of Economics
Dickinson Hall
Princeton University
Princeton, N.J. 08544
(609) 452-4035

Angelo Melino
Department of Economics
University of Toronto
150 St. George Street
Toronto, Canada M5S 1A1
(416) 978-6541

Robert J. Shiller
Cowles Foundation for Research in Economics
Yale University
Box 2125 Yale Station
New Haven CT 06520
(203) 432-4134

I. Introduction

In this paper, we provide an empirical test of the continuous time intertemporal capital asset pricing model, first proposed by Merton[1971]. The model as clarified by Breeden[1979] implies that an asset will be priced so that the expected return required will increase with its covariability with per capita consumption growth. Previous tests of this theory (e.g. Grossman-Shiller[1980], Hansen-Singleton[1983]) have examined discrete time versions of the model under the assumption that the timing interval of the model matches exactly the sampling interval for available data on per capita consumption. That is, if we have data on quarterly consumption, then the time period is assumed to be 1-quarter of a year. We show that if the true model is a continuous time model, and time averaged data (such as quarterly consumption) is used to test it, then substantial biases may be introduced unless the estimation procedure is corrected to take account of the effects of time averaging. We provide a procedure for obtaining consistent estimates with time averaged data. We then estimate and test the model using data on per capita consumption and the cumulated real returns to holding portfolios of stocks, bonds, and short-term paper.

II The Model

It is useful to review the Merton model. Our discussion follows closely the exposition of its generalization in Grossman-Shiller[1982]. In a discrete time model, each consumer is assumed to maximize a time-additive utility function over a single consumption good

$$(2.1) \qquad U = \sum_{j=0}^{T/h} \delta^{hj} u(c(j))h$$

where T is his time horizon, c(t) is consumption at time t and δ^h is the discount factor between utility at time t and t+h. For the purposes of this paper, we assume that the period utility function is of the constant relative risk aversion (or isoelastic) form

$$(2.2) \qquad u(c) = c^{1-A}/(1-A) .$$

Let $v_i(t)$ denote the value of asset i at time t including any accrued cash disbursements (such as dividends or coupons) earned between t-h and t. Assume that asset i is freely tradeable. A standard argument shows that

$$(2.3) \qquad E_t \delta^h u'(c(t+h))v_i(t+h) = u'(c(t))v_i(t)$$

where the expectation is conditioned on all the information possessed by the trader at time t. Using (2.2) and iterating (2.3), we can write

$$(2.4) \qquad E_t \left(\frac{c(\tau)}{c(t)}\right)^{-A} \frac{v_i(\tau)}{v_i(t)} = \left(\frac{1}{\delta}\right)^{\tau-t} \qquad \text{for } \tau = t+h, t+2h, \ldots$$

If we take the limit to continuous time and apply Ito's Lemma, we obtain

$$(2.5) \qquad E_t \frac{dv_i}{v_i} + \frac{1}{2} * A(A+1) * Var\left(\frac{dc}{c}\right) - A * E_t \frac{dc}{c} + \ln \delta dt = A * Cov\left(\frac{dc}{c}, \frac{dv_i}{v_i}\right)$$

where Var and Cov denote the variance and covariance operators.

Note that (2.5) holds for an individual. Under various assumptions about heterogeneity of information and wealth, (2.5) can be aggregated over individuals so that c can be interpreted as per capita consumption and A is replaced by a particular weighted average of the individual consumer's A (see Grossman-Shiller[1982]). Clearly (2.5) holds for all tradeable assets. If R_i^e is defined as the excess rate of return of asset i over say short-term paper, then (2.5) can be used for these two assets to yield

$$(2.6) \qquad ER_i^e = A*Cov(R_i^e, dc/c) \ .$$

The aggregate parameter of relative risk aversion can be computed by (2.6) given data on mean excess returns and the covariances between excess returns and per capita consumption growth. Table 1 provides some estimates of A based on the descriptive statistics from Table 3. The various data sets and variable definitions are described more fully in Section III. At this point, we simply wish to draw attention to one of the important empirical anomalies associated with the model and the potential role for time averaging as an explanation. The table shows that the mean excess return on stocks is associated with a relatively small covariance with consumption changes. Therefore this can be justified only by an implausibly high estimate of the risk aversion parameter. Similiar conclusions are reached by examining the excess returns on bonds.

One explanation for this is based on the idea that a time averaged variable is smoother than the same point sampled variable. In particular, if the true model holds in continuous time then the instantaneous rates of change in consumption can be more variable (and also covariable with returns) than is the average consumption change across years or quarters.

An example

To understand this effect, consider the following very simple process for v_i^e, the value of asset i in excess of asset 1, and consumption:

$$(2.7a) \qquad dc = \rho dt + d\epsilon$$

$$(2.7b) \qquad dv_i^e = \mu_i dt + d\eta_i$$

where ϵ, η_i are correlated Brownian motions with $\mathrm{Cov}(d\epsilon, d\eta_i) = \sigma_i dt$. Let $\bar{c}(t)$ and $\bar{v}_i^e(t)$ be the time averaged values of $c(t)$ and $v_i^e(t)$, i.e.,

$$\bar{c}(t) = T^{-1}\int_0^T c(t+s)ds \qquad \bar{v}_i^e = T^{-1}\int_0^T v_i^e(t+s)ds \ .$$

We will show that

$$(2.8) \quad \mathrm{Cov}(\bar{c}, \bar{v}_i^e) \equiv E[(\bar{c}(t)-\bar{c}(t-T))(\bar{v}_i^e(t)-\bar{v}_i^e(t-T))] - \rho T \mu_i T$$
$$= 2/3T\sigma_i \ .$$

If we normalize T=1, then the covariance of time averaged consumptions changes and price changes is 2/3 of the instantaneous value σ_i. Roughly speaking, this would lead us to overestimate A by 50%.

To understand (2.8) just note that

$$(2.9) \quad \bar{c}(t) - \bar{c}(t-T) = \int_0^T \rho ds + T^{-1}\int_0^T\int_{t+s-T}^{t+s} d\epsilon(\tau)ds$$
$$= \rho T + T^{-1}\int_0^T\int_{t+s-T}^{t} d\epsilon(\tau)ds$$
$$+ T^{-1}\int_0^T\int_t^{t+s} d\epsilon(\tau)ds$$
$$= \rho T + T^{-1}\int_{t-T}^{t} (T-t+s)d\epsilon(s)$$
$$+ T^{-1}\int_t^{t+T} (t+T-s)d\epsilon(s) \ .$$

A similiar expression may be derived for $\bar{v}_i^e(t)-\bar{v}_i^e(t-T)$. Hence,

$$(2.10)\quad E[(\bar{c}(t)-\bar{c}(t-T))(\bar{v}_i^e(t)-\bar{v}_i^e(t-T))]$$

$$=\rho T\mu_i T + T^{-2}E[\int_{t-T}^{t}(T-t+s)^2 d\epsilon d\eta_i + \int_{t}^{t+T}(t+T-s)^2 d\epsilon d\eta_i].$$

Equation (2.8) is easily derived from the last expression.

The purpose of this example is to give the reader a relatively simple view of the effect of time averaging in generating a stochastic process which is "smoother" than the instantaneous process. This suggests the possibility that assets appear to have a low risk (i.e. low covariance with consumption changes) because measured consumption changes are less variable than instantaneous consumption changes. Since it is the covariance with instantaneous consumption changes that is the relevant measure of an asset's risk, this leads us to overestimate A. In our simple example, A is overestimated by 50%. As we shall see below, for certain processes, the bias can be arbitrarily large.

Multivariate Model

In our empirical work, we postulate a slightly more complicated stochastic process for consumption and asset values. Define Y(t) according to

$$(2.11)\qquad Y(t) = \begin{bmatrix} \ln c(t) - k_c - g_c t \\ \ln v_1(t) - k_1 - g_1 t \\ \ln v_2(t) - k_2 - g_2 t \\ \ln v_3(t) - k_3 - g_3 t \end{bmatrix} \equiv \begin{bmatrix} C(t) \\ V_1(t) \\ V_2(t) \\ V_3(t) \end{bmatrix}.$$

We assume that Y(t) satisfies the stochastic differential equation

$$(2.12)\qquad dY = BYdt + \Sigma^{1/2}dZ$$

where B and Σ are (4x4) matrices and Z(t) is a vector of standard independent Wiener processes. Σ is assumed to be symmetric and positive definite. Without any loss of generality, $\Sigma^{1/2}$ can be taken to be lower triangular with positive diagonal elements. Let σ denote the vector of nontrivial parameters in $\Sigma^{1/2}$.

Switching to logarithms and applying Ito's Lemma we can rewrite (2.5) in terms of the Y(t) process as

$$(2.13) \qquad E_t[dV_i(t) - A*dC(t)] + (g_i - A*g_c + \ln\delta)dt + 1/2*A^2*Var(dC) - A*Cov(dC,dV_i) + 1/2*Var(dV_i) = 0.$$

If this is to hold at all points in time in m.s., then

$$(2.14) \qquad E_t[dV_i(t) - A*dC(t)] = 0.$$

The reason is that, according to our assumptions, the remaining terms in the expression are not functions of information. Since the model is homogeneous, the only way this sum can be constant is if it is zero.

Therefore (2.5) imposes the following restrictions on our model:

$$(2.15a) \qquad J_iB = 0$$

$$(2.15b) \qquad g_i - A*g_c + 1/2*J_i\Sigma J_i' + \ln\delta = 0 \quad i=1,2,3$$

where $J_i=(-A\ e_i')$ and $e_i \epsilon R^3$ is the vector with unity in component i and zero elsewhere.

Suppose that the process Y(t) is sampled at regular intervals. It is straightforward (see Bergstrom[1984]) to show that the point sampled process has the representation

$$(2.16) \qquad Y(t) = \phi Y(t-1) + u(t)$$

where $\phi = e^{B}$, the matrix exponential of B, and u(t) is the random variable $\int_{t-1}^{t} e^{B(t-s)}\Sigma^{1/2}dZ(s)$. Let $\bar{Y}(t)$ denote the time average of the Y(t) process, i.e. $\bar{Y}(t)=\int_{t-1}^{t} Y(s)ds$. Upon integrating both sides of (2.16) we obtain

(2.17) $$\bar{Y}(t) = \phi\bar{Y}(t-1) + \bar{u}(t)$$

where $\bar{u}(t)$ is the random variable $\int_{t-1}^{t}\int_{\tau-1}^{\tau} e^{B(\tau-s)}dZ(s)d\tau$.

Let f and g denote two "smooth" real-valued functions and z(s) a univariate Wiener process. Using the definition of the Ito integral, the following two results can be established:

(2.18) $$\int_0^T f(t)[\int_0^t g(s)dz(s)]dt = \int_0^T[\int_s^T f(t)dt]g(s)dz(s)$$

(2.19) $$E[\int_{t1}^{t2} f(s)dz(s)][\int_{t3}^{t4} g(\tau)dz(\tau)] = \int_M f(s)g(s)ds$$

$$\text{where } M \equiv [t_1,t_2]\cap[t_3,t_4]$$

and where the equality is understood in the mean square sense. Applying (2.18) element by element and other standard properties of the Ito integral allow us to write

(2.20) $$\bar{u}(t) = \int_{t-2}^{t-1}\int_s^{t-1} e^{B(\tau-s)}d\tau dZ(s) + \int_{t-1}^{t}\int_s^t e^{B(\tau-s)}d\tau dZ(s).$$

Define $\Omega_\tau \equiv E\, \bar{u}(t)\bar{u}(t-\tau)$ and $F(r,w) \equiv e^{Br}\Sigma e^{B'w}$. Applying (2.19) and standard change of variable rules, we obtain

(2.21a) $$\Omega_0 = \int_0^1\int_s^1\int_s^1 F(r,w)drdwds + \int_0^1\int_0^s\int_0^s F(r,w)drdwds$$

(2.21b) $$\Omega_1 = \int_0^1\int_s^1\int_0^s F(r,w)drdwds$$

(2.21c) $$\Omega_\tau = 0 \quad \tau\geq 2.$$

We conclude that $\bar{Y}(t)$ is a vector ARMA(1,1) process. Phillips[1978] and Bergstrom[1984] develop similiar results although the latter only considers the case where B is invertible. We can therefore write

(2.22) $$\bar{Y}(t) = \phi\bar{Y}(t-1) + \epsilon(t) + \theta\epsilon(t-1)$$

where the innovations $\epsilon(t)$ have mean zero and covariance matrix S, and Θ is a matrix with spectral radius not exceeding unity.

Define $y(t) = (\ln c(t) \;\; \ln v_1(t) \;\; \ln v_2(t) \; \ln v_3(t))'$, and let $\bar{y}(t)$ be its unit average. Eq (2.22) can be rewritten as

$$(2.23) \qquad \bar{y}(t) = \gamma_0 + \gamma_1 \bar{t} + \phi \bar{y}(t-1) + \epsilon(t) + \Theta\epsilon(t-1)$$

where $\gamma_0 = (I-\phi)k+\phi g$, $\gamma_1 = (I-\phi)g$, and $\bar{t} = \int_0^1 (t+s)ds$. The restrictions (2.15) are easily shown to imply $J_i \gamma_1 = J_i(I-\phi) = 0$. In particular, it also follows that the vector k cannot be identified uniquely. We therefore impose the identification restriction $k_i = A*k_c$ in our estimation. A tedious argument also shows that

$$(2.24) \quad (J_i-J_j)(\bar{y}(t)-\bar{y}(t-1)) = (J_i-J_j)\gamma_0 + (J_i-J_j)\epsilon(t) + .268*(J_i-J_j)\epsilon(t-1)$$

so that the time averaged excess returns on asset i over j follows an MA(1) process with coefficient .268.

To gain further intuition about the possible consequences of time averaging suppose $B = \mathrm{diag}(\lambda_1,\lambda_2,\lambda_3,\lambda_4)$. Then it can be shown that $\Omega_0(i,j) = h(\lambda_i,\lambda_j)\cdot\Sigma(i,j)$, where

$$h(\lambda_i,\lambda_j) = (\lambda_i\lambda_j)^{-1}[1 + e^{\lambda i+\lambda j} - 2*(\lambda_i+\lambda_j)^{-1}(1-e^{\lambda i+\lambda j}) + (1 + e^{\lambda i})(1 - e^{\lambda j})/\lambda_j + (1 + e^{\lambda j})(1 - e^{\lambda i})/\lambda_i]$$

and equal to the obvious limits as λ_i or λ_j goes to 0. Our simple example corresponds to the case $h(0,0) = 2/3$. If the process were stationary around trend, the eigenvalues of B would have negative real parts. Sampling a few values, we see that $h(-.1,-.1) = .60$, $h(-.5,-.5) = .45$, $h(-1,-1) = .28$, and h goes to zero as λ_i and λ_j both go to minus infinity. The bias in the estimate of A using time averaged data and (2.5) can therefore be arbitrarily large.[1]

III DATA ANALYSIS

Data Description

The data are fully described in an appendix to this paper which is available from the authors. Here we shall give only a broad description of the data to indicate how they were assembled and to show that they correspond as much as possible to the concepts represented in the model above.

Six separate data sets were prepared, each intended to represent a series of observations on the four-element vector $\bar{y}$. The data sets differ in sample period, sources and assumptions about taxation. Table 2 summarizes the important differences.

Data sets one and two are long historical annual time series beginning in the year 1890. These data sets are based on those used in Grossman and Shiller[1981] and described also in Shiller[1982]. Data sets three through six are quarterly time series. Data sets three and four begin in the second quarter of 1953. Data sets five and six begin in the second quarter of 1947. The use of annual and quarterly time series was dictated

---

[1]A more relevant comparison might be the ratio of A that would be obtained using time averaged data to that using point sampled data. Although details differ, it is easily shown that this ratio also can be arbitrarily large.

by the existing consumption data. Long time series data on consumption are available only on an annual basis. Quarterly consumption data are available only for the post-war period. Monthly consumption data are available starting in 1959. We did not use those data here because of some concern as to the accuracy of the monthly data and because of the somewhat shorter sample period that such data would impose.

In all data sets, the first element of $\bar{y}$ is the log of real per capita seasonally-adjusted consumption on nondurables and services. For years beginning with 1929 these data are from the National Income and Product Accounts of the United States. Earlier data are the Kuznets-Kendrick series. Since the published consumption series are total consumption over the period, the first element of $\bar{y}$ departs somewhat from that hypothesized in the paper: it is the log of the integral rather than the integral of the log.[2] Note that we use a physical measure of consumption directly and do not deflate nominal consumption by a price index that is averaged over the year, which would have introduced another departure from the assumptions of our model.

In all data sets the second element of $\bar{y}$ is a measure of the interval averaged log cumulated real return on corporate stocks, the third element is a measure of the interval averaged log cumulated real return on short debt and the fourth element

---

[2]Some Monte Carlo simulations indicate that the biases introduced by using the log of the average instead of the average of the log are extremely small, at least for our data.

is a measure of the interval averaged log cumulated real return on long-term bonds. The even-numbered data sets are based on after-tax returns. In constructing these series, the (after-tax in even-numbered cases) nominal returns were first computed on a monthly basis. At that point, a choice had to be made whether to use the consumption deflator to convert nominal returns to real returns or to use one of the monthly price indices for this purpose. The consumption deflator has the advantage that it corresponds to the measure of consumption that is supposed to enter the utility function. The monthly price indices have the advantage that we can use them to produce a monthly real series, so that our interval average will correspond more closely to the integral of the log of the real portfolio value as represented in our model. It was decided to use the consumption deflator for data sets one through four and the monthly consumer price index for data sets five and six. Thus, for example, the second through fourth elements of the $\bar{y}$ vector in data set two were constructed by first producing monthly series representing the cumulated after-tax nominal returns of the assets. Each series represented the nominal value of the portfolio of an individual who reinvests all after-tax income from the asset in the same asset.[3] The average for the year of the log of the monthly portfolio values was used to construct an annual series. Finally, the log of the consumption deflator was subtracted

[3]Let $(1+r_{im})$ denote the monthly after-tax nominal return on asset i, and let v^n_{iL} denote the cumulated after-tax return in month L. We set $v^n_{iL} = (1+r_{i1})(1+r_{i2})\cdots(1+r_{iL})$.

from each series to convert to a real series. With data set five, the first step in the construction of the second through fourth elements of $\bar{y}$ was essentially the same. We first produced a monthly series of cumulated returns of the assets. However, in data set five, this monthly series was subsequently deflated by dividing by the consumer price index, and a quarterly series was produced as the average for the three months of the quarter of the log of this monthly real series.

With data sets five and six another adjustment was also made before the average log cumulated real portfolio value was entered into the vector $\bar{y}$. In constructing the series, there was great concern that the data be aligned properly. The Ibbotson-Sinquefield returns data for each month are measured from the end of the preceding month to the end of the current month. This provides four point sampled observations on the log cumulated real portfolio for each quarter. These were connected by straight lines and the integral under the straight line interpolation was used to estimate the corresponding component of $\bar{y}$.

For data sets one and two, the return on corporate stocks is computed from the Standard and Poor's Composite Stock Price Index and associated dividend series. The return on short-term debt is computed from the prime commercial paper rate and the return on long-term debt is computed using the Macaulay railroad bond yield data for the first part of the sample and the Moody Aaa bond yield average for the years after 1936.

For data sets three and four, all return data come from series on the CITIBASE data library. Stock returns are again computed using the monthly Standard and Poor's Composite Stock Price Index, while the return on short debt is taken from the return on three-month treasury bills and the return on long debt is based on yields of twenty-year treasury notes.

For data sets five and six, return data come from Ibbotson and Sinquefield[1982]. The stock return series is their series common stocks, total returns; the short debt series is their series U.S. Treasury bills, total returns; the long debt return series is their series long-term corporate bonds, total returns.

For after-tax series, the assumed marginal income tax rate for 1918 to 1980 was that implicit in the spread between municipal and corporate bond yields. Before 1918, the marginal income tax rate was set to zero. Since the Ibbotson and Sinquefield data do not allow a decomposition of returns into capital gains and income components, it was assumed for data set six that all returns were taxed each month as income. For data sets two and four, however, capital gains were assumed taxed each month at a long-term capital gains rate. For the years 1946 to 1978, the effective rate on long-term capital gains was one-half the marginal income tax rate. For earlier years, the effective rate on long-term capital gains was computed from the marginal income tax rate using tax rate data in Seltzer[1951].

Preliminaries

Before considering formal estimation and testing, it is useful to review some of the broad features of the six data sets which our model must explain. Some descriptive statistics are provided in Table 3.

For all six data sets, we observe that stock portfolios gave the highest average real return, approximately 6% p.a. on a pre-tax basis or 4% after-tax. Short-term paper yields averaged about 2% p.a. on a pre-tax basis over our longest historical sample, but the average yield fell to about zero in the post-war period. After-tax real returns to holding short-term paper have been slightly negative. Long-term bonds, by contrast, have averaged essentially a zero real return over the last century, on both a pre- and after-tax basis. During the post-war period, however, pre-tax returns have been slightly negative. On an after-tax basis, bondholders have seen the real value of their portfolios shrink by over 2% p.a.

According to the consumption based asset pricing model, these persistent differences in average yields must be accounted for by the insurance provided by the different portfolios against events which impinge adversely on consumption. Useful evidence about this hypothesis is obtained by looking at the covariance structure of measured portfolio yields and changes in consumption. Some caution is necessary since the model's predictions pertain to the covariance structure of the instantaneous returns and our data are constructed from differences of unit averaged values. However, if $B \approx 0$, the latter can provide a reliable guide to the sign and order of magnitude of the instantaneous covariance matrix.

Several empirical regularities emerge. As measured by the variance, the change in consumption is the smoothest series, followed closely by the yield on short-term paper. Long-term bond yields have been fairly stable over our longest sample, whereas the variance of returns to holding a portfolio of stocks has been several orders of magnitude larger. In the post-war period, real returns to holding long-term bonds have been much more volatile with a variance almost as large as the return to holding common stocks.

Of more interest are the covariance properties. According to our model, it is not the variance but the covariance with consumption that is the relevant measure of a portfolio's risk. We find, uniformly across the six data sets, that stock yields have the largest covariance with changes in consumption, followed by short-term paper yields and then yields on long-term bonds. Qualitatively, this is exactly what the model requires given the ordering of the average yields. It indicates that the basic idea that insurance against adverse movements in consumption can account for observed yield differentials has some empirical promise.

Evidence of potential difficulties is provided by the autocovariance structure of excess returns on bonds and stocks over short-term paper. Given our assumptions about the probabilistic structure of consumption and portfolio values and the form of preferences, we expect the point sampled difference in yields between any two portfolios to be serially uncorrelated. As equation (2.24) shows, the time averaged difference in yields should have an MA(1) component with coefficient about .268. This particular prediction is independent of the mean or covariance of returns or the degree of relative risk aversion.

Table 3 shows that it is important to take into account the consequences of time averaging. The Box-Ljung statistics clearly indicate that the excess yields that are constructed from our data are not white noise. The adjusted excess returns referred to in Table 3 are filtered to remove the time dependence that is induced by unit averaging. Judging from the Box-Ljung statistics, the adjusted excess returns are indeed less serially correlated . Nonetheless, the autocorrelations of the adjusted excess returns to stocks remain statistically significant from zero in four of the six data sets.

Some Econometric Issues

It is demonstrated above that the vector of time averaged observations has a representation of the form

(3.1) $$\bar{y}(t) = \gamma_0(\alpha) + \gamma_1(\alpha)\bar{t} + \phi(\alpha)\bar{y}(t-1) + \epsilon(t) + \theta(\alpha)\epsilon(t-1)$$

where the disturbances $\epsilon(t)$ are distributed independently and identically as $MVN(0,S(\alpha))$. In our application, we can set $\alpha' \equiv (k_c, g_c, A, \delta, B_{1.}, \sigma)$, where $B_{1.}$ denotes the first row of the B matrix.

Linear Gaussian processes have been studied extensively by econometricians and statisticians. Nonetheless, there are several features of our model which put it outside of the standard assumptions in the literature used to prove laws of large numbers or central limit theorems. First, the model contains a time trend so that sample autocovariances of the exogenous variables, i.e. $T^{-1}\sum x_t x'_{t-j}$ where $x'_t = (1\ t)$, do not converge to well defined limits. Secondly, the model imposes restrictions not only across the autoregressive and moving average matrices, but across these and the contemporaneous covariance matrix as well. Finally, our model imposes the restriction that B be of rank one, so that $\phi(\alpha)$ will have three eigenvalues on the unit circle. To our knowledge, there are no laws of large numbers or central limit theorems that cover all three of these features. Application of the standard large sample procedures to estimate and test our model must be considered tentative.

Although all the features of our model have not been treated together in the literature, we can use available results to form a reasonable guess about the sampling properties of the approximate (conditional) maximum likelihood estimator described below. For example, it appears that a law of large numbers which would allow for all three of the features noted above would be a modest extension of the literature. Hannan et al.[1980] provide a law of large numbers for vector ARMAX models allowing for very general restrictions and, in particular, dependence across the covariance matrix of innovations and the other parameters of the model. Their assumptions about the error process are clearly satisfied by our model, but they rule out time trends as regressors and require all roots of the autoregressive polynomial to be outside the unit circle. In the absence of complicated restrictions or unit roots, the assumption that sample covariances converge to well defined limits can be replaced by the weaker Grenander conditions (see Hannan[1971]) which do allow for time trends as regressors. Similarly, in the absence of time trends and other restrictions, strong laws of large numbers can be established even if the autoregressive process is explosive. Individually, therefore, each of the three features of the model highlighted above is not an impediment to establishing a law of large numbers.

It is well known that unrestricted estimates of ϕ will not be asymptotically normal if there are unit roots in the autoregressive polynomial. A case for a central limit theorem can be made only if the estimation procedure exploits the prior knowledge of the structure of ϕ. Our restrictions imply that $\bar{y}_t$ is a co-integrated process (see Granger-Engle[1982]). These processes have had a long history in applied empirical research under the name of "error-correction" models. However, only recently has there been any serious investigation of the sampling properties of the MLE or its approximants. Available theorems do not allow for a time trend or moving average terms but these complications do not appear to present any conceptual difficulties. The main result is that the integrating factor[4] is estimated consistently by ML with a sampling error that is $o_p(T^{-1/2})$. The ML estimators for the remaining parameters are consistent and asymptotically normal with a covariance matrix that is estimated consistently by the usual formula. In our model, the integrating factor is just $B_1.$, appropriately scaled. Since we are never concerned with testing restrictions on the components of $B_1.$, the rapid convergence of the estimated integrating factor does not appear to present a problem.

We will proceed formally as if the standard large sample procedures for inference are valid under the maintained hypothesis that B is of rank one. As the preceding discussion makes clear, however, some scepticism is in order.

Estimation Strategy

Several strategies for the estimation of models with MA errors have been proposed.[5] In the time domain, it is natural to consider the maximum likelihood estimator, or one of its various approximants.

Put $e(1) \equiv 0$ and for any admissable α define $e(t)$ recursively according to

---

[4] A nonstochastic vector c such that $c'\bar{y}_t$ is stationary is called an integrating factor. In our application, it is any normalized basis vector for the row space of $I-\phi$.

[5] See Osborne (1977) for a survey of the unconstrained case.

$$(3.2) \qquad e(t) = \bar{y}(t) - \gamma_0(\alpha) - \gamma_1 \bar{t} - \phi(\alpha)\bar{y}(t-1) - \Theta(\alpha)e(t-1).$$

Following Wilson (1973), we choose as our estimator $\hat{\alpha}_a$, the admissable vector α which maximizes the approximate (conditional) log likelihood function

$$L_a(\alpha) \equiv -\frac{(T-1)}{2} \ln|S(\alpha)| - \frac{1}{2} \operatorname{tr} S^{-1}(\alpha)M$$

$$\text{where} \quad M \equiv \sum_{t=2}^{T} e(t)e(t)'.$$

Since ϕ has unit roots, we have little choice but to condition on the first observation $\bar{y}(1)$. The assumption that $e(1)=0$, by contrast, is made solely out of convenience. If $B=0$, the spectral radius of Θ is about .268, so the sampling distribution of $\hat{\alpha}_a$ will not be very sensitive to this assumption about the initial innovation. Putting $e(1)=0$ does simplify the computations somewhat. In particular, analytic derivatives can be easily and quickly computed using the method of adjoints and a straightforward application of the chain rule.

Several features of $L_a(\alpha)$ make the evaluation of $\hat{\alpha}_a$ challenging. As with any model with MA errors, it is not possible to reduce the data through sufficient statistics and we have to deal with a likelihood function that is not guaranteed to be globally concave. Our model poses several difficulties in addition to these standard ones. For example, it is not possible to concentrate out the covariance matrix, since S is functionally related to the regression parameters of (3.1). Also, some effort is required to evaluate $(\phi(\alpha), S(\alpha), \Theta(\alpha))$. Details are provided in Melino[1985], so we will give only a brief overview here.

Define the matrices

$$(3.4)\quad C \equiv \begin{bmatrix} -B & I & 0 & 0 \\ 0 & -B & \Sigma & 0 \\ 0 & 0 & 0 & I \\ 0 & 0 & 0 & B' \end{bmatrix} \qquad e^{C} \equiv \begin{bmatrix} F_1 & G_1 & H_1 & K_1 \\ 0 & F_2 & G_2 & H_2 \\ 0 & 0 & F_3 & G_3 \\ 0 & 0 & 0 & F_4 \end{bmatrix}.$$

Put $\tilde{C} = -C$ and denote the blocks of $e^{\tilde{C}}$ by $\tilde{F}_1$, $\tilde{G}_1$ etc. It can be shown that

$$(3.5)\qquad \Omega_0 = F_4'K_1 + K_1'F_4 - \tilde{K}_1\tilde{F}_1' - \tilde{F}_1\tilde{K}_1'$$

$$(3.6)\qquad \Omega_1 = \tilde{H}_1G_3 + \tilde{K}_1\tilde{F}_1' + \tilde{F}_1\tilde{K}_1'.$$

It is also useful to note that $\phi = F_4'$. Although the expressions appear to be unappetizing, they are straightforward to implement given an algorithm for computing the matrix exponential. We used a routine based on a diagonal Padé approximation that has very nice numerical properties.[6]

Solving for (S,Θ) given (Ω_0,Ω_1) turned out to be much easier than conjectured by Bergstrom[1984]. Wilson[1972] provides a general algorithm for factoring the autocovariance function of a multivariate MA process. We adapted his suggestion to our special case and applied Newton's method to find the matrix Θ with spectral radius no greater than unity which is a root of the polynomial

$$(3.7)\qquad \Omega_1 - \Theta\Omega_0 + \Theta\Omega_1\Theta' = 0.$$

Given an initial guess, $\Theta(0)$, this leads to the iterative scheme

---

[6]We would like to thank Dr. R.C. Ward of the Union Carbide Laboratory in Oak Ridge for kindly providing us with this code.

(3.8) $\Theta(n+1)[\Omega_0-\Omega_1\Theta'(n)] - \Theta(n)\Omega_1\Theta'(n+1) = \Omega_1 - \Theta(n)\Omega_1\Theta'(n).$

This scheme exhibits quadratic convergence and turns out to be quite fast. On average, less than three iterations were required to find Θ given (Ω_0,Ω_1). In fact, we found that this scheme rarely required more than 5 iterations. Given Θ, it is straightforward to solve for S using $S = \Omega_0 - \Omega_1\Theta'$.

Evaluation of $L_a(\alpha)$ and its analytic derivatives is fairly quick and easy. The main difficulty in computing $\hat{\alpha}_a$ turned out to be the extraordinary large number of iterations required to refine its location.

Parameter Estimates

Table 4 presents the estimated parameters of the constrained model for each of the the six data sets.[7] The estimates obtained using before- and our constructed after-tax yields are remarkably similiar, but there are considerable differences in the estimates across the three different sample periods.

Consider first the estimates of Σ, the covariance matrix of the instantaneous innovations. Once again, correlations are displayed above the diagonal, and the lower triangular elements are covariances. The estimates of Σ from the quarterly data sets are all similiar. However, there are some sharp contrasts

---

[7]Estimates were obtained using the GQOPT3 package provided by Professor Quandt of Princeton University. Various algorithms were required to refine the location of $\hat{\alpha}_a$. The reported standard errors, however, are always calculated by inverting the matrix of second derivatives evaluated at the optimum. The Hessian was computed using symmetrical numerical differences of analytic first derivatives.

with the estimates from the annual samples which cast doubt on our assumption that Σ has been constant over time. Consumption innovations appear to have had a much smaller variance in the post-war period, as have had the innovations to the value of short-term paper. By contrast, the innovations to stock market values have been slightly smoother, and those for long-term bonds are roughly comparable. The covariances of the innovations to portfolio values with consumption have the same ranking in all six data sets, but they are much smaller in the post-war period.

All six data sets yield small estimates of $B_{1\cdot}$, the first row of the B matrix. This indicates that the change in consumption has only a very small predictable component, aside from trend. The trend in consumption is estimated to be about 3% p.a. using the two long historical samples, about 2.5% using data sets three and four, and about 1.6% p.a. using data sets five and six. The corresponding point estimates for δ indicate, respectively, a substantial preference for present consumption, a substantial preference for future consumption, and indifference. These apparent differences can't be taken too seriously since the estimated standard errors indicate substantial uncertainty.

The differences in the estimated parameters of relative risk aversion are extremely interesting. Using our two longest historical samples, we obtain estimates of A of just over 20. This is too large to be plausible. Nonetheless, as anticipated, accounting for unit averaging of consumption results in a substantial reduction.[8]

Data sets five and six produce a very plausible estimate of A of just over 2. By contrast, data sets three and four produce an estimate of A over 150! The difference of the parameter estimates obtained using these very similiar post-war quarterly data sets is very large, and some clarification is in order.

The estimates of A presented in Table 1 are derived from restrictions which relate the unconditional means to the covariances of consumption changes and portfolio returns. However, the model provides us with further sources of information about A. Equation (2.14) tells us that the predictable change in the value of any portfolio is equal to a multiple of the predictable change in consumption, up to a constant. Since the multiple is just the parameter of relative risk aversion, this gives us another estimate of A based on the conditional information in the sample. The maximum likelihood estimator is usefully viewed as suitably pooling the disparate estimates based on conditional and unconditional information.

It turned out that the predictable change in consumption around its mean using the lagged information in data sets 1-4 was essentially zero. As a result, the maximum likelihood estimate of A closely reflects the estimates in Table 1 adjusted for unit averaging. In data sets 5 and 6, however, the predictable

---

[8]For data set 1, we also estimated the model as if the data was actually point sampled. We obtained an estimate $\hat{A} = 27.24$, with a standard error of about 11.2.

change in consumption about its mean, while still small, was large enough to provide a fairly accurate estimate of A. The maximum likelihood estimate reported for these data sets reflects the conditional information in the sample.[9]

Our model imposes restrictions on the time averaged representation of $\bar{y}$. In turn, the time averaged representation imposes additional structure on the parameters of the ARMA(1,1) representation. Table 5 contains the log likelihoods, denoted L_1, L_2, and L_3 respectively, for the fully restricted time averaged estimates (Model 1), the unconstrained time averaged estimates (Model 2), and the unconstrained ARMA(1,1) estimates (Model 3). For the reasons discussed above, Models 2 and 3 were estimated under the maintained hypothesis that B is of rank one and ø is the sum of the identity and a rank one matrix. For completeness, the log likelihoods for the totally unconstrained time averaged and ARMA(1,1) models, L_2' and L_3' respectively, are also reported.

The tests of the overidentifying restrictions imposed by the model are rejected with very high confidence when compared against either Model 2 or Model 3. Curiously, data sets three and four which produced the least plausible parameter estimates,

---

[9]Imposing only the restrictions implied by (2.14), we obtained for the six data sets:

$\hat{A}$ =	663.54	949.86	439.67	983.73	2.07	2.60
(s.e.)	(370.56)	(370.56)	(*)	(*)	(0.92)	(1.21)

Because the Hessian was singular, we are unable to provide standard errors for the estimates from data sets three and four.

Hansen-Singleton[1983] also report a very sharp difference in the estimate of A depending upon whether or not conditioning information is used. (See their Table 5) provide the weakest evidence against the overidentifying restrictions. Finally, a comparison of L_2 and L_3 indicates that there is some difficulty in accounting for the autocovariances of $\bar{y}$ by time averaging a first order process.

Why is the model rejected?

There are strong a priori reasons for linking consumption and portfolio choices. Moreover, the sample means and covariances of portfolio yields and changes in consumption lend qualitative support to the notion of assets being priced in accordance with the insurance they afford against adverse movements in consumption. Yet the various goodness of fit tests reported above as well as the implausibly high estimates of relative risk aversion from data sets one through four appear to constitute an overwhelming rejection of the model. What should we conclude?

A response that cannot be dismissed is that the assumed distribution of the goodness of fit tests is simply misleading. As we noted above, we cannot rely on the standard central limit theorems to establish the asymptotic distribution. Moreover, even if the large sample results obtain, as we conjecture, there is no guarantee that the asymptotic distribution provides a close approximation for samples of the size we have examined. Unfortunately, establishing the small sample distribution either analytically or by Monte Carlo methods is infeasible. We choose to take the evidence against the model seriously and to focus attention on the specific sources of predictive failure.

One is naturally led to examine more closely the various auxiliary assumptions that are being tested jointly alongside the hypothesis that agents behave as described by (2.3). The two most obvious are the stochastic process assumed to describe the evolution of consumption and portfolio values and the specific form of preferences. We will concentrate on the former.

The stochastic differential equation (2.12) imposes many overidentifying restrictions. One of them is that the time averaged vector has an ARMA(p,q) representation with p=q=1. To test this, the autocorrelations of the prediction errors from Model 3 were calculated. Box-Ljung tests did not indicate any need for considering a higher order process.

Although the evidence suggests that an ARMA(1,1) representation for $\bar{y}$ is a reasonable approximation, there are problems in accepting the restrictions that time averaging a first-order process imposes on this representation. Phillips(1978) shows that if $B \approx 0$ then $\phi \approx I+B$ and $\theta \approx .268(I+(B-\Sigma B'\Sigma^{-1})/4)$. Our unconstrained ARMA(1,1) estimates of ϕ suggest that B is indeed small. There is little difficulty in accepting the restrictions which a small B matrix and time averaging impose on ϕ. However, this combination imposes a great deal of structure on θ which is at odds with the data. For all six data sets, we found that both the constrained and unconstrained time averaged models produced estimates of $\theta \approx .268I$. The unconstrained ARMA(1,1) estimates of θ differed from .268I in several respects. The most noticeable discrepancy was that the unrestricted estimate of the row of the moving average matrix pertaining to the consumpton equation was essentially zero, in all six data sets. In fact, in data sets 1-2 and 5-6, the MA coefficient for the innovation in consumption was more than two standard deviations below .268. Failure to explain the MA component of consumption in and of itself would lead to rejection of the model at the 5% level for these data sets.

One possible explanation for the apparent absence of a moving average component in the consumption equation is measurement error. Suppose the unit average of consumption is measured with an error that is serially uncorrelated and independent of the true consumption process. If the flow of consumption is truly a random walk, the measured consumption series will be an ARMA(1,1) process but with an MA coefficient less than .268. If one half of the variance of the change in measured consumption is due to measurement error, the MA coefficient would be predicted to fall to just .127.

As pointed out earlier, our model predicts that the excess returns of stocks and bonds over the yield on short-term paper should be unpredictable. The time averaged excess returns should therefore have an MA(1) structure with a coefficient of about .268. These overidentifying predictions can be tested regardless of the quality of the consumption data by simply regressing the adjusted excess returns on various information sets. Moreover, there is no problem in justifying the standard procedures to test these orthogonality restrictions. The results are reported in Table 6. The likelihood ratio test statistic, λ, and the R^2 for each of the individual regressions is also reported. The individual R^2 are remarkably high and the orthogonality restrictions are rejected with very high confidence. Since yield data that are point sampled are readily available, we also tested these restrictions using the monthly point sampled yields corresponding to data sets 1-6. Because a monthly price index was not available for our longest samples, we used the log cumulated nominal returns, v^n_{it}, in the information set. These results are reported in the lower half of Table 6. Although the individual R^2 are much lower, as we would expect, the rejection of the orthogonality restrictions is even more pronounced. These results are very similiar to those reported in Hansen and Singleton[1983].

One explanation for this predictive failure is simply that the covariance matrix of the instantaneous innovations is not constant but is state dependent. This seems extremely plausible and could also account for the noted differences in the estimates of Σ from different sample periods. However, taking account of state dependent variances would make estimation and testing of the model practically impossible. Because our model imposes restrictions across the drift and diffusion parameters, making the latter state dependent would force us to abandon the linear constant coefficient model of the drift as well. We would be led to the more general stochastic process that solves

(3.9) $dy = B(t,y)dt + \Sigma^{1/2}(t,y)dZ.$

The restrictions across the drift and diffusion effectively rule out any of the convenient functional forms for $B(\cdot)$ and $\Sigma(\cdot)$, and the solution of the likelihood for even the point sampled process is difficult to implement. Computing the likelihood function for the unit averaged process that solves (3.9) seems unimaginable, with current technology.

IV Conclusions

The notion of insurance against events which impinge unfavourably on consumption choices can be used to rationalize, at least qualitatively, the systematic differences in average yields afforded by portfolios of stocks, bonds, and short-term paper. The sample means and covariances of portfolio returns and per capita consumption growth indicate that the quantitative differences in average yields can be rationalized only by implausibly high aversion to risk. Taking account of the fact that measured consumption is unit averaged substantially reduces the degree of relative risk aversion required to rationalize the data.

Nonetheless, there remains considerable evidence that casts doubt on this view of the world. In particular, it is difficult to reconcile the importance of unit averaging of the consumption flow with the fact that the measured logarithm of detrended real per capita consumption has essentially no moving average component. Also, although the model allows the average return on different portfolios to diverge due to different insurance characteristics, the particular specification that we examined requires that expected excess returns should be time invariant. This orthogonality property is forcefully rejected by the data. Addressing these particular predictive failures while taking account of unit averaging constitutes a formidable challenge for future research.

Table 1

Estimates of A Using Unconditional Means and Covariances

Data Set	ER_S^e	$Cov(R_S^e, dc/c)$	$\hat{A}_S$	ER_B^e	$Cov(R_B^e, dc/c)$	$\hat{A}_B$
1 Annual	.039	$.0^2 29$	13.76	-.013	$-.0^4 80$	159.10
3 Quarterly	.012	$.0^4 49$	139.13	-.005	$-.0^4 13$	376.41
5 Quarterly	.012	$.0^4 62$	198.56	-.002	$-.0^5 60$	398.17

R_S^e ≡ excess return on stocks over short term paper.

R_B^e ≡ excess return on bonds over short term paper.

Table 2

Important Differences Across Data Sets

(1) Data Set	(2) Sample	(3) Price Index	(4) Marginal Tax Rate	(5) Capital Gains Taxed As	(6) Averaging Method
1	Annual	Consumption Deflator	0	0	1
2	Annual	Consumption Deflator	Munic	Long-Term Gain	1
3	Quarterly	Consumption Deflator	0	0	1
4	Quarterly	Consumption Deflator	Munic	Long-Term Gain	1
5	Quarterly	CPI	0	0	2
6	Quarterly	CPI	Munic	Income	2

Notes: Column (3): To convert nominal returns to real returns, we used either the consumption deflator for non-durables and services or the consumer price index.

Column (4): A 0 indicates pretax returns. Munic indicates that the marginal personal income tax rate implicit in the spread between municipal and corporate bond yields was used to construct after-tax returns.

Column (6): Averaging method 1 takes the log monthly cumulated returns, forms a simple average for the period and subtracts the log price index. Averaging method 2 takes log monthly cumulated returns, subtracts the monthly log price index and computes the integral of the linear interpolation of the end of month values.

Table 3

Some Descriptive Statistics

	Variable	Mean	Correlation/Covariance[1]				Variable[3]	Mean	Box-Ljung Statistic to Lag:			
									6	12	18	24
Data Set 1 1890-1981 92 Observations	$\Delta \ln c$	.032	$.0^2 25$	.289	-.142	-.128	R_S^e	.039	11.17	22.59*	32.27*	43.03**
	$\Delta \ln v_1$	.058	$.0^2 25$	.029	.099	.286	R_B^e	-.013	29.97**	48.07**	55.11**	59.27**
	$\Delta \ln v_2$	.019	$-.0^3 40$	$.0^3 94$	$.0^2 31$	.692	† R_S^e	.039	9.10	19.45	30.01*	42.68*
	$\Delta \ln v_3$	$.0^2 60$	$-.0^3 48$	$.0^2 37$	$.0^2 29$	$.0^2 56$	† R_B^e	-.013	13.48*	25.03*	30.90*	35.69
Data Set 2 1890-1980 91 Observations	$\Delta \ln c$	.032	$.0^2 25$	.261	-.185	-.160	R_S^e	.035	13.02*	25.01*	36.72*	47.67**
	$\Delta \ln v_1$	.048	$.0^2 21$	.026	.159	.318	R_B^e	-.011	26.28**	44.22**	48.91**	54.10**
	$\Delta \ln v_2$	.013	$-.0^3 53$	$.0^2 15$	$.0^2 33$	.780	† R_S^e	.035	11.58	23.50*	36.69**	50.11**
	$\Delta \ln v_3$	$.0^2 18$	$-.0^3 59$	$.0^2 38$	$.0^2 33$	$.0^2 54$	† R_B^e	-.011	11.21	22.31*	26.48	31.50
Data Set 3 1953:3-1983:1 119 Observations	$\Delta \ln c$	$.0^2 48$	$.0^4 27$	.273	.020	-.067	R_S^e	.012	24.84**	31.52**	54.91**	62.76**
	$\Delta \ln v_1$	.015	$.0^4 88$	$.0^2 38$	.077	.216	R_B^e	$-.0^2 51$	10.16	25.80*	28.70	34.28
	$\Delta \ln v_2$	$.0^2 24$	$.0^6 55$	$.0^4 26$	$.0^4 29$	.247	† R_S^e	.012	7.93	15.66	31.10	37.91
	$\Delta \ln v_3$	$-.0^2 27$	$-.0^4 13$	$.0^3 51$	$.0^4 51$	$.0^2 15$	† R_B^e	$-.0^2 51$	4.42	15.50	17.15	21.78

Table 3 (continued)

	Variable	Mean	Correlation/Covariance[1]				Variable[3]	Mean	Box-Ljung Statistic to Lag:			
									6	12	18	24
Data Set 4 1953:3-1980:4 110 Observations	Δlnc	$.0^2 51$	$.0^4 28$	.282	.247	-.042	R_S^e	.013	22.13**	28.66**	50.82**	56.77**
	Δlnv_1	.011	$.0^4 80$	$.0^2 29$	.311	.142	R_B^e	$-.0^2 61$	10.25	14.17	19.42	27.55
	Δlnv_2	$-.0^2 17$	$.0^5 57$	$.0^4 73$	$.0^4 19$	.348	† R_S^e	.013	7.44	14.30	29.28*	34.39
	Δlnv_3	$-.0^2 79$	$-.0^5 67$	$.0^3 23$	$.0^4 45$	$.0^3 88$	† R_B^e	$-.0^2 61$	8.15	10.30	13.45	22.94
Data Set 5 1947:2-1981:4 139 Observations	Δlnc	$.0^2 43$	$.0^4 37$	.266	.224	.041	R_S^e	.015	27.24**	34.88**	80.04**	88.38**
	Δlnv_1	.015	$.0^4 95$	$.0^3 35$	.140	.241	R_B^e	$-.0^2 42$	19.71**	23.29*	27.62	37.08*
	Δlnv_2	$-.0^3 16$	$.0^4 10$	$.0^4 61$	$.0^4 54$	.344	† R_S^e	.015	8.03	14.67	42.66**	50.62**
	Δlnv_3	$-.0^2 43$	$.0^5 80$	$.0^3 46$	$.0^4 81$	$.0^2 10$	† R_B^e	$-.0^2 42$	10.33	12.15	15.12	24.51
Data Set 6 1947:2-1980:4 135 Observations	Δlnc	$.0^2 44$	$.0^4 37$	.263	.242	.033	R_S^e	.012	27.81**	35.38**	77.86**	85.16**
	Δlnv_1	$.0^2 93$	$.0^4 73$	$.0^2 21$	.287	.279	R_B^e	$-.0^2 24$	19.47**	24.17*	31.34*	42.06*
	Δlnv_2	$-.0^2 30$	$.0^4 11$	$.0^3 10$	$.0^4 59$	.487	† R_S^e	.012	8.83	15.24	41.86**	48.80**
	Δlnv_3	$-.0^2 54$	$.0^5 50$	$.0^3 32$	$.0^4 93$	$.0^3 62$	† R_B^e	$-.0^2 24$	14.35*	16.58	21.08	33.00

[1] Correlations are displayed above the diagonal.

[2] * Denotes significant at 5% level.
**Denotes significant at 1% level.

[3] †Denotes adjusted to remove the effects of time averaging.

Table 4

Estimates for the Fully Constrained Model[1]

Data Set	K_c	g_c	A	δ	$B_{1.}$				Σ^2			
1 (1890-1981)	.12	.03	21.23	.79	$-.0^{2}23$	$.0^{3}80$	$-.0^{3}21$	$-.0^{2}58$	$.0^{2}40$	.225	-.004	-.008
	(.06)	(.01)	(9.51)	(.37)	$(.0^{2}23)$	$(.0^{3}98)$	$(.0^{2}14)$	$(.0^{2}31)$	$.0^{2}31$	.047	.202	.374
									$-.0^{3}10$	$.0^{2}26$	$.0^{2}35$	.744
									$-.0^{3}50$	$.0^{2}75$	$.0^{2}40$	$.0^{2}86$
2 (1890-1980)	.14	.03	21.40	.77	$-.0^{2}45$	$.0^{2}11$	$-.0^{3}17$	$-.0^{2}60$	$.0^{2}40$	.214	-.003	-.008
	(.07)	(.01)	(9.86)	(.39)	$(.0^{2}30)$	$(.0^{2}12)$	$(.0^{2}14)$	$(.0^{2}33)$	$.0^{2}28$	.041	.223	.378
									$-.0^{3}13$	$.0^{2}27$	$.0^{2}35$	.790
									$-.0^{3}45$	$.0^{2}67$	$.0^{2}41$	$.0^{2}77$
3 (1953:3 - 1983:1)	-4.61	$.0^{2}6$	154.47	1.27	$-.0^{2}55$	$.0^{4}42$	$.0^{3}21$	$-.0^{3}26$	$.0^{4}41$	.234	.282	-.046
	(6.99)	$(.0^{2}2)$	(3.39)	(.15)	$(.0^{3}16)$	$(.0^{4}13)$	$(.0^{4}31)$	$(.0^{4}35)$	$.0^{3}11$	$.0^{2}52$	.199	.176
									$.0^{5}83$	$.0^{4}66$	$.0^{4}21$	.220
									$-.0^{4}14$	$.0^{3}59$	$.0^{4}47$	$.0^{2}21$
4 (1953:3 - 1980:4)	.14	$.0^{2}5$	185.38	1.23	$-.0^{3}21$	$.0^{4}42$	$.0^{4}49$	$-.0^{4}97$	$.0^{4}42$	.228	.334	-.068
	(1.60)	$(.0^{3}9)$	(29.85)	(.20)	$(.0^{3}12)$	$(.0^{4}16)$	$(.0^{4}77)$	$(.0^{4}40)$	$.0^{4}91$	$.0^{2}38$	.220	.049
									$.0^{5}82$	$.0^{4}51$	$.0^{4}14$	.302
									$-.0^{4}16$	$.0^{3}11$	$.0^{4}42$	$.0^{2}13$
5 (1947:2 - 1981:4)	-.23	$.0^{2}4$	2.12	1.01	-.039	$.0^{2}28$	$-.0^{2}61$	$-.0^{2}45$	$.0^{4}57$	.224	.111	.003
	(.71)	$(.0^{3}7)$	(.98)	$(.0^{2}4)$	(.027)	$(.0^{2}16)$	$(.0^{2}93)$	$(.0^{2}36)$	$.0^{3}12$	$.0^{2}49$	.155	.328
									$.0^{5}60$	$.0^{4}79$	$.0^{4}52$	.155
									$.0^{6}87$	$.0^{3}42$	$.0^{4}91$	$.0^{2}15$

Table 4 (continued)

Data Set	K_c	g_c	A	δ	$B_{1.}$				Σ^2			
6 (1947:2 - 1980:4)	.39	$.0^{2}36$	2.69	1.01	-.032	$.0^{2}32$	$-.0^{2}56$	$.0^{2}12$	$.0^{4}57$	.194	.130	-.024
	(.80)	$(.0^{3}92)$	(1.34)	$(.0^{2}6)$	(.022)	$(.0^{2}21)$	$(.0^{2}79)$	$(.0^{2}51)$	$.0^{4}79$	$.0^{2}29$	.213	.168
									$.0^{5}70$	$.0^{4}81$	$.0^{4}50$	.393
									$-.0^{5}52$	$.0^{3}26$	$.0^{4}80$	$.0^{3}83$

1. Figures in parentheses are the estimated standard errors.

2. Correlations are displayed above the diagonal.

Table 5

Goodness of Fit Tests

Data Set	L1 (NP=18)	L2 (NP=25)	L2' (NP=34)	L3 (NP=41)	L3' (NP=50)	Test of Model 1vs Model 2	Test of Model 1vs Model 3	Test of Model 2vs Model 3
1	822.95	838.34	849.33	865.79	877.57	30.78	85.68	54.90
2	832.53	845.04	856.74	871.85	885.13	25.02	78.64	53.62
3	1794.79	1803.07	1835.54	1823.07	1852.67	16.56	56.56	40.00
4	1724.42	1746.34	1755.34	1755.68	1776.96	43.84	62.52	18.68
5	2044.57	2056.02	2073.78	2098.20	2123.08	22.90	107.26	84.36
6	2064.79	2078.36	2102.61	2119.85	2141.57	27.14	110.12	82.98

NP denotes the number of free parameters.

$\chi^2_{.05}(7) = 14.07$ $\chi^2_{.05}(23) = 35.17$ $\chi^2_{.05}(16) = 26.30$

$\chi^2_{.01}(7) = 18.48$ $\chi^2_{.01}(23) = 41.64$ $\chi^2_{.01}(16) = 32.00$

Table 6
Predictability of Excess Returns

I. Adjusted Excess Returns (Time Averaged Data)

Information Set = $\left\{ \underline{1}, \bar{y}(t-1), \bar{t} \right\}$

(10 restrictions)

Data Set	T	R_S^2	R_B^2	λ	z
1	91	.17	.29	43.00	7.38
2	90	.17	.27	39.33	6.56
3	118	.26	.18	54.68	9.99
4	109	.22	.16	44.66	7.75
5	138	.18	.13	42.75	7.32
6	134	.18	.16	45.15	7.86

II. Monthly Returns (Point Sampled Data)

Information Set = $\left\{ \underline{1}, \ln v^n_{1,t-1}, \ln v^n_{2,t-1}, \ln v^n_{3,t-1}, t \right\}$

(8 restrictions)

Data Set	T	R_S^2	R_B^2	λ	z
1	1102	.20	.20	419.88	76.70
2	1090	.21	.21	439.28	80.37
3	355	.16	.13	121.04	20.23
4	328	.15	.13	121.90	20.39
5	418	.05	.04	62.80	9.22
6	406	.06	.04	76.67	11.84

Notes:

$T \equiv$ number of observations.

$R_S^2 \equiv R^2$ for the excess return on stocks equation.

$R_B^2 \equiv R^2$ for the excess return on bonds equation.

$\lambda \equiv$ Likelihood ratio test statistic.

$z \equiv (\lambda - q)/\sqrt{2q}$ where q is the number of restrictions under test.

References

Bergstrom,A.R. 'Continuous Time Stochastic Models and Issues of Aggregation Over Time,' in Z. Griliches and M. Intriligator (eds) Handbook of Econometrics,Amsterdam:North-Holland,(1984)

Breeden,D.T. 'An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities,' Journal of Financial Economics, 7(1979):265-296

Granger,C.W.J. and R.F. Engle 'Dynamic Model Specification with Equilibrium Constraints: Co-integration and Error Correction,' Department of Economics W.P.#85-18, University of California, San Diego, revised May 1985

Grossman,S.J. and R.J. Shiller 'Preliminary Results on the Determinants of Stock Market Variability,' unpublished mimeo, June 1980

-'On the Determinants of Stock Market Variability,' American Economic Review, 7(1981):222-227

-'Consumption Correlatedness and Risk Measurement in Economies with Non-traded Assets and Heterogeneous Information,' Journal of Financial Economics,10(1982):195-210

Hannan,E.J. Multiple Time Series, New York:Wiley,1970

Hannan,E.J., W.T.M. Dunsmuir and M. Deistler 'Estimation of Vector ARMAX Models,' Journal of Multivariate Analysis, 10(1980):275-295

Hansen,L.P. and K.J. Singleton 'Stochastic Consumption,Risk Aversion, and the Temporal Behaviour of Asset Returns,' Journal of Political Economy, 91(1983):249-265

Ibbotson,R.G. and R.A. Sinquefield Stocks, Bonds, Bills and Inflation:The Past and the Future, Charlottesville: Financial Analysts Research Foundation,1982

Melino,A. 'Estimation of Unit Averaged Diffusion Processes,' Department of Economics-Institute for Policy Analysis W.P.#8502, University of Toronto, March 1985

Merton,R.C. 'Optimum Consumption and Portfolio Rules in a Continuous Time Model,' Journal of Economic Theory, 3(1971):373-413

Osborne,D.R. 'Exact and Approximate Maximum Likelihood Estimators for Vector Moving Average Processes,' Journal of the Royal Statistical Society, Series B, 39(1977):114-118

Phillips,P.C.B. 'The Treatment of Flow Data in the Estimation of Continuous Time Systems,' in Bergstrom et.al (eds), Stability and Inflation, New York:Wiley,1978

Seltzer,L.H. The Nature and Tax Treatment of Capital Gains and Losses, New York:National Bureau of Economic Research,1951

Shiller,R.J. 'Consumption, Asset Markets and Macroeconomic Fluctuations,' Carnegie-Rochester Conference Series, 17(1982)

Wilson,G.T. 'The Factorization of Matricial Spectral Densities,' SIAM Journal of Applied Mathematics, 23(1972):420-426

— 'The Estimation of Parameters in Multivariate Time Series Models,' Journal of the Royal Statistical Society, Series B, 35(1973):76-85