NBER TECHNICAL WORKING PAPER SERIES

EFFICIENT ESTIMATION OF LINEAR ASSET PRICING MODELS WITH MOVING-AVERAGE ERRORS

Lars Peter Hansen

Kenneth J. Singleton

Technical Working Paper No. 86

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1990

We are grateful to the National Science Foundation for financial support and to Kobi Boudoukh, Charles Evans, John Heaton, Robin Lumsdaine, Masao Ogaki, Karl Snow and Chi Wa Yuen for research assistance. John Cochrane and John Heaton provided helpful comments. The second author is also grateful for support from the Center for Research in Security Prices at The University of Chicago. Portions of this paper were written while the first author was a visiting scholar at the Graduate School of Business at Stanford University. This paper is part of 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 Technical Working Paper #86
March 1990

EFFICIENT ESTIMATION OF LINEAR ASSET PRICING MODELS WITH MOVING-AVERAGE ERRORS

ABSTRACT

This paper explores in depth the nature of the conditional moment restrictions implied by log-linear intertemporal capital asset pricing models (*ICAPMs*) and shows that the generalized instrumental variables (*GMM*) estimators of these models (as typically implemented in practice) are inefficient. The moment conditions in the presence of temporally aggregated consumption are derived for two log-linear *ICAPMs*. The first is a continuous time model in which agents maximize expected utility. In the context of this model, we show that there are important asymmetries between the implied moment conditions for infinitely and finitely-lived securities. The second model assumes that agents maximize non-expected utility, and leads to a very similar econometric relation for the return on the wealth portfolio. Then we describe the efficiency bound (greatest lower bound for the asymptotic variances) of the *GMM* estimators of the preference parameters in these models. In addition, we calculate the efficient *GMM* estimators that attain this bound. Finally, we assess the gains in precision from using this optimal *GMM* estimator relative to the commonly used inefficient *GMM* estimators.

Lars Peter Hansen
Department of Economics
University of Chicago
Chicago, IL 60637
(312)702-8170

Kenneth J. Singleton
Graduate School of Business
Stanford University
Stanford, CA 94305
(415)723-5753

1. INTRODUCTION

The testable implications of intertemporal asset pricing theories (*ICAPMs*) frequently take the form of conditional moment restrictions on linear econometric models with moving-average disturbances. Moving-average errors may arise, for example, when multi-period returns are examined [e.g., Hansen and Hodrick (1980,1983), Dunn and Singleton (1986), Fama and French (1988)] or in the presence of time-averaged data [e.g., Barro (1981), Grossman, Melino and Shiller (1987), Hall (1988)]. The combination of moving average disturbances and limited information about distributions[1] has led naturally to estimation of the unknown parameters in these models using the generalized method of moments (*GMM*). The parameters of multi-period forecast equations are frequently estimated by least squares [Hansen and Hodrick (1980) and Fama and French (1988)], while instrumental variables procedures have been used in estimating simultaneous equations derived from *ICAPMs* [Harvey (1988) and Hall (1988)]. For pedagogical purposes it is convenient to view both least squares and instrumental variables estimators as *GMM* estimators.

In this paper we explore in depth the nature of the conditional moment restrictions implied by log-linear *ICAPMs* and show that *GMM* estimators of these models (as typically calculated in practice) are inefficient. The moment restrictions implied by two log-linear *ICAPMs* in the presence of temporally aggregated consumption are derived in section 2. The first model is a continuous time *ICAPM* which includes the models proposed by Grossman, Melino, and Shiller (1988) and Hall (1988) as special cases. In the context of this model, we show that there are important asymmetries between the moment conditions implied by expected utility models for infinitely and finitely-lived securities. For returns on infinitely-lived securities, temporal aggregation induces autocorrelations in the disturbances that are known *a priori*, while for some finite-lived securities, the induced autocorrelations are not known *a priori*. Hence, Working's (1960) autocorrelation restriction for temporally aggregated Brownian motions need not apply to log-linear models of returns on short and intermediate term bonds. Whence, imposition of this restriction may lead to inconsistent estimators of standard errors and, in some cases, the preference parameters as well.

The second model studied is a special case of the *ICAPM* with non-expected utility proposed by Epstein and Zin (1989a,b) for the return on the wealth portfolio. Though the economic underpinnings of this model and the expected utility *ICAPM* are different, they are shown to imply strikingly similar econometric equations for this return. Consistent with Hall's (1988) analysis, the parameter on consumption growth in the non-expected utility model is the intertemporal elasticity of substitution and not the coefficient of relative risk aversion. However, these observations apply only to our expression for the return on the wealth portfolio and, in particular, do not provide a reinterpretation of log-linear models of bond returns.

There is a large (infinite-dimensional) class of *GMM* estimators for the preference parameters of these log-linear *ICAPM*s. Drawing upon the analyses in Hansen (1985) and Hansen and Singleton (1989), in section 3 we describe the efficiency bound (a greatest lower bound for the asymptotic variances) for this class of estimators, and present a *GMM* estimator that attains this bound. The *GMM* estimators typically implemented in practice for linear asset pricing models do not attain this bound because they exploit only a subset of the implied conditional moment restrictions.

The potential inefficiency of *GMM* estimation is perhaps most easily illustrated in the context of least squares estimation in the presence of moving-average errors. Ever since Fama's (1965) pioneering study of the martingale representation of stock prices, substantial attention has been given to multi-period optimal linear forecasting equations of the form:

$$(1.1) \qquad y_{t+m} = \delta_0 y_t + \delta_1 y_{t-1} + \cdots + \delta_p y_{t-p} + e_{t+m},$$

where the first element of vector y_t is a return, excess return, or difference between a forward and future spot price over m periods.[2] The δ_j's in (1.1) are square matrices that are either unrestricted or may depend on some lower-dimensional parameter vector β_o. The disturbance term e_{t+m} is an expectational error satisfying $E(e_{t+m}|I_t) = 0$, where I_t is generated by current and all past values of y_t. Under these assumptions, $\{e_t\}$ follows an MA(m-1) process.

Consistent estimators of the δ_j's (and β_o in the case of *a priori* restrictions) are commonly obtained using least squares methods that exploit the moment conditions:

$$(1.2) \qquad E(e_{t+m} y_{t-j}') = 0, \text{ for } j=0,1,\ldots,p.$$

When $\{e_t\}$ is serially correlated, least squares is in general not the most efficient estimation method in the presence of the conditional mean restriction $E(e_{t+m}|I_t) = 0$. This is because the additional moment conditions

$$(1.3) \qquad E(e_{t+m} y_{t-j}') = 0, \ j > p.$$

can be exploited to the improve the precision of the estimators.

The log-linear *ICAPMs* described in section 2 imply an analogous, infinite collection of moments conditions that can be used in estimating the preference parameters. In section 4 we calculate efficient *GMM* estimators that exploit all of the moment restrictions, and assess the gains in precision relative to the commonly used inefficient *GMM* estimators. Our calculations exploit the characterization of efficient *GMM* estimators and algorithms for calculating these estimators presented in Hansen (1985) and Hansen and Singleton (1989).

2. LOG-LINEAR, INTERTEMPORAL ASSET PRICING MODELS

In this section we investigate the conditional moment restrictions implied by two log-linear, continuous time *ICAPMs* linking consumption and asset returns. In the first of these models, consumers have state-separable preferences. This model includes the models studied by Grossman, Melino, and Shiller (1987) and Hall (1988) as special cases. The consumers in second model have preferences that are not state-separable as proposed recently by Epstein and Zin (1989a,b), Kocherlakota (1989) and Weil (1989). The models we consider depart from the assumptions of the log-linear model examined in Hansen and Singleton (1983) by replacing the assumption that the agents' decision interval coincides with the sampling interval of the data (e.g., one month) with the assumption that agents adjust their consumption and portfolios more frequently. Although the specifications of preferences in these models are different, they imply similar log-linear asset pricing relations. Therefore, a comparison of the implied relations is instructive for interpreting the conditional moment restrictions implied by log-linear *ICAPMs*. The implications of this discussion for the relative efficiency of alternative *GMM* estimators are pursued in Section 3.

2.A A Continuous Time ICAPM with State-Separable Preferences

Following Grossman, Melino, and Shiller (1987) and Hall (1988), consider a representative consumer who chooses a consumption process $\{C(t) : t \geq 0\}$ to maximize:

$$(2.1) \qquad E\int_0^\infty exp(-\delta t)U[C(t)]dt \quad ,$$

where δ is an instantaneous subjective rate of discount and the instantaneous utility function U is:

$$(2.2) \qquad U(C) = \left[\frac{C^{\gamma+1} - 1}{\gamma+1}\right], \quad \gamma < 0 \quad .$$

The marginal utility process associated with (2.2) is

$$(2.3) \qquad MU(t) = exp[\gamma c(t)] \quad ,$$

where $c(t) = log[C(t)]$.

Infinitely-Lived Securities

As in Grossman, Melino and Shiller (1987), we posit a price process for an infinitely-lived asset and deduce restrictions relating the equilibrium behavior of this price process to consumption. The price process presumes that all dividends are reinvested in the security and that the entire return is measured by the capital gain or price appreciation. Let $Q(t)$ denote the price in terms of consumption of this security, and $q(t) = log[Q(t)]$.

An implication of a large class of intertemporal asset pricing models is that in equilibrium the process $\{e^{-\delta t}MU(t)Q(t) : t\geq 0\}$ is a martingale adapted to the increasing sequence of agents' information sets $\{I(t) : t\geq 0\}$:

$$(2.4) \qquad E[e^{-\delta(t+\tau)}MU(t+\tau)Q(t+\tau)|I(t)] = e^{-\delta t}MU(t)Q(t) \quad ,$$

for all $\tau\geq 0$. Since our focus is on estimation of linear models, we impose the additional assumption that

$$(2.5) \qquad d[e^{-\delta t}MU(t)Q(t)] = e^{-\delta t}MU(t)Q(t)[\sigma\cdot dW(t)]$$

or

$$(2.6) \qquad d[MU(t)Q(t)] = \delta MU(t)Q(t)dt + MU(t)Q(t)[\sigma\cdot dW(t)] \quad ,$$

where $\{W(t) : t\geq 0\}$ is a vector of uncorrelated Brownian motions adapted to $\{I(t) : t\geq 0\}$ and the vector of real numbers σ is constant over time. Since $W(t)$ may be a vector, (2.6) allows multiple sources of uncertainty to affect $Q(t)$ and $C(t)$. Without loss of generality, we assume $E[W(t)W(t)'] = tI$. Relation (2.6) is consistent with the assumptions about the distributions of $C(t)$ and $Q(t)$ in Grossman, Melino and Shiller (1987) and Hall (1988). Using Ito's Lemma, the corresponding expression for logarithms is

$$(2.7) \qquad d(log[MU(t)Q(t)]) = \left[\delta - \frac{\sigma\cdot\sigma}{2}\right]dt + \sigma\cdot dW(t) \quad .$$

That is, $\{[\gamma c(t) + q(t)] : t\geq 1\}$ is a Brownian motion with drift $\delta-(\sigma\cdot\sigma/2)$. From (2.7) it follows that

$$(2.8) \quad [\gamma c(t+1) + q(t+1)] = [\gamma c(t) + q(t)] + \left[\delta - \frac{\sigma\cdot\sigma}{2}\right] + \sigma\cdot W(t+1) - \sigma\cdot W(t) \quad .$$

Suppose only discrete time data are available for studying (2.8). Typically, studies of (2.8) [e.g., Hansen and Singleton (1983), Ferson (1983)] have assumed that the decision interval of agents coincided with the sampling interval of consumption. However, several authors, including Grossman, Melino and Shiller (1987), and Hall (1988), have suggested that discrete time consumption data be viewed instead as a geometric average over time of the instantaneous consumption flows. Under this interpretation, a version of (2.8) expressed in terms of measured variables is obtained by averaging (2.8) backward over one unit of time:

$$\gamma[c^a(t+1) - c^a(t)] + [q^a(t+1) - q^a(t)] - \left[\delta - \frac{\sigma\cdot\sigma}{2}\right] = u^a(t+1) \quad , \tag{2.9}$$

where

$$u^a(t+1) = \int_0^1 \sigma\cdot W(t+1-\tau)d\tau - \int_0^1 \sigma\cdot W(t-\tau)d\tau \text{ and } c^a(t) = \int_0^1 c(t-\tau)d\tau , \tag{2.10}$$

etc. Notice that $[q^a(t+1) - q^a(t)]$ is a geometric average over time of real returns. Equation (2.9) is the econometric model that will be investigated empirically.

There are two implications of this model that can be tested using time series data. First,

$$E[u^a(t+2)|I(t)] = 0 \quad . \tag{2.11}$$

Second, as shown in Working (1960) the first-order autocorrelation of the temporally aggregated first difference of a Brownian motion equals .25:

$$E[.25\ u^a(t+2)^2 - u^a(t+2)u^a(t+1)|I(t)] = 0 \quad . \tag{2.12}$$

Both of these conditional moment restrictions can be tested using discrete time data without having to parameterize the continuous time law of motion for $\{[c(t),q(t)]' : t\geq 0\}$. To see this in the case of (2.11), let x(t) be any vector of variables observed by agents and the econometrician at date t, let ρ_c and ρ_q denote the vector of coefficients in the regressions of $[c^a(t+2) -$

$c^a(t+1)]$ and $[q^a(t+2) - q^a(t+1)]$ onto $x(t)$, respectively, and let ν_c and ν_q denote the coefficients on the constants in these two regressions. Then (2.11) implies that $\gamma\rho_c = \rho_q$, which identifies γ as long as $\rho_c \neq 0$, and leads to overidentifying restrictions. Once γ is identified, the discount rate δ can be identified from the variance of $u^a(t+2)$ and the coefficients ν_c and ν_q. The variance of $u^a(t+2)$ is

$$(2.13) \qquad E[u^a(t)^2] = 2\sigma\cdot\sigma/3 \quad .$$

Furthermore, (2.11) implies that $(\gamma\nu_c + \nu_q) = \left[\delta - \frac{\sigma\cdot\sigma}{2}\right]$. Therefore, given γ and $\sigma\cdot\sigma$, one can infer δ.

Pricing Discount Bonds

Relation (2.5) imposes a particular form of homoskedasticity by requiring that the σ vector be independent of time. This restriction is not plausible for all security price processes and, in particular, is not in general satisfied for the case of nominal pure discount bonds. To show this, an innovations representation for consumption and nominal price processes is posited and then the equilibrium bond prices are derived endogenously. The model used for this analysis can be viewed as a simplified version of the models investigated by Cox, Ingersoll, and Ross (1985) and Breeden (1986). We show that temporally aggregated models of bond prices do not in general lead to the same overidentifying restrictions as those deduced above for an infinitely-lived security.

Let $p(t)$ denote the logarithm of the dollar price of one unit of consumption at time t. We abstract from modeling why dollars get valued, and view $\{p(t) : t\geq 0\}$ as an exogenous process that determines value in terms of a time t numeraire. Hence, we are ignoring the distortions to real economies that might lead to valued-fiat money.[3] We assume that $\{c(t) : t\geq 0\}$ and $\{p(t) : t\geq 0\}$ have the following innovations representations

$$(2.14) \qquad c(t) = E[c(t) \mid I(0)] + \int_0^t \alpha_c(\tau)\cdot dW(t-\tau)$$

$$p(t) = E[p(t) \mid I(0)] + \int_0^t \alpha_p(\tau)\cdot dW(t-\tau) \quad ,$$

where $\{W(t):t\geq 0\}$ is defined as before and α_c and α_p are vectors of real-valued function of time. For simplicity, we assume that these functions are continuous, although weaker restrictions are permitted.

Let $\{b_r(t) : t\geq 0\}$ be the logarithm of the price of a pure discount bond at time t that pays a dollar at time t+r. This r-period bond costs $exp[b_r(t)-p(t)]$ units of consumption at time t and has a payoff of $exp[-p(t+r)]$ units of consumption at time t+r. The equilibrium bond price satisfies

$$(2.15) \qquad exp[b_r(t) - p(t) - t\delta]MU(t) = E\{exp[-p(t+r) - (t+r)\delta]MU(t+r)|I(t)\}.$$

In light of (2.14), (2.15) can be rewritten as

$$(2.16) \qquad b_r(t)-p(t)+mu(t) = E[mu(t+r)-p(t+r)-\delta r|I(t)] + Var[mu(t+r)-p(t+r)|I(t)]/2.$$

Using the fact that $mu(t) = \gamma c(t)$ and substituting (2.14) into (2.16) gives

$$(2.17) \qquad b_r(t)-p(t)+mu(t) = \gamma E[c(t+r)|I(0)] - E[p(t+r)|I(0)] - \delta r$$

$$+ \gamma\int_o^t \alpha_c(\tau+r)\cdot dW(t-\tau) - \int_o^t \alpha_p(\tau+r)\cdot dW(t-\tau) + \sigma_r/2,$$

where

$$(2.18) \qquad \sigma_r = Var[\gamma c(t+\tau) - p(t+r)|I(t)] = \int_o^r \alpha_b(\tau) \cdot \alpha_b(\tau)d\tau$$

and $\alpha_b(\tau) = \gamma\alpha_c(\tau) - \alpha_p(\tau)$.

The expectational error from forecasting [mu(t+r) - p(t+r)] is

$$(2.19) \qquad u_r(t+r) = mu(t+r) - p(t+r) - E[mu(t+r) - p(t+r)|I(t)]$$

$$= \int_o^r \alpha_b(\tau)\cdot dW(t+r-\tau) \quad .$$

Combining (2.16) - (2.19) gives

$$(2.20) \qquad \gamma c(t+r) - \gamma c(t) - \delta r - b_r(t) - p(t+r) + p(t) + \sigma_r/2 = u_r(t+r)$$

where $E[u_r(t+r)|I(t)] = 0$. Integrating (2.20) backward one unit of time gives the temporally aggregated version of (2.20):

$$(2.21) \qquad \gamma[c^a(t+r)-c^a(t)] - \delta r - b_r^a(t)-p^a(t+r) + p^a(t) + \sigma_r/2 = u_r^a(t+r).$$

The disturbance $u_r^a(t+r)$ satisfies the counterpart to the conditional moment restriction (2.11):

$$(2.22) \qquad E[u_r^a(t+r+1)|I(t)] = 0 \quad .$$

In contrast to the model for infinitely-lived securities, there are no implied restrictions on the autocorrelation function of $\{u_r^a(t+r) : t\geq 1\}$. This can be seen from the integral representation of $u_r^a(t+r)$:

$$(2.23) \qquad u_r^a(t+r) = \int_o^1 \int_o^r \alpha_b(\tau)\cdot dW(t+r-\tau-s)d\tau ds$$

$$= \int_o^r \alpha_b(\tau)\cdot[W(t+r-\tau) - W(t+r-1-\tau)]d\tau \quad .$$

In general, the function α_b cannot be identified from discrete time data. This leads to two important differences between the implications of this bond pricing model and the equity pricing model. First, with $u_r^a(t+r)$ given by (2.21) it is not possible to infer σ_r from the variance of $u_r^a(t+r)$ as in (2.13). It follows that the discount rate δ is not identifiable using discrete time data on temporally aggregated consumption and bond prices. Second, the values of the first r autocorrelations of $\{u_r^a(t+r) : t\geq 0\}$ are not known *a priori*. Hence there is no conditional moment restriction analogous to (2.12) in the case of bonds.

Hall (1988) studied a version of (2.21) for temporally aggregated returns on three-month Treasury bills (r=1) using postwar quarterly data. In constructing an IV estimator of his model, the first-order autocorrelation of the disturbance was assumed to be .25. The preceding discussion shows that this restriction is not implied by the model for this choice of returns. Thus, while Hall's (1988) parameter estimator is consistent, the standard error estimator is not. Similarly, Grossman, Melino and Shiller (1987) imposed the .25 autocorrelation restriction in a fully parameterized time series model for temporally aggregated consumption and returns. One of the assets that they used in their analysis is a similar three-month Treasury bill series. From the preceding discussion, it follows that they imposed an incorrect restriction on their time series parameterization which could render the resulting estimator of γ inconsistent.

Holding-Period Returns on Long-Term Bonds

Restrictions analogous to those deduced for infinitely-lived securities do apply approximately to holding-period returns on long-term bonds. The logarithm of the one-period return from purchasing an r period bond and selling it after one period is given by $b_{r-1}(t+1) - b_r(t)$. Differencing the versions of (2.17) for $b_r(t)$ and $b_{r-1}(t+1)$ gives:

$$(2.24)\qquad \gamma[c(t+1)-c(t)] + [b_{r-1}(t+1)-b_r(t)-p(t+1)+p(t)] - \delta - \sigma_r/2 + \sigma_{r-1}/2 = \int_{r-1}^{r} \alpha_b(\tau)\cdot dW(t+r-\tau)d\tau.$$

In addition to being continuous, suppose that $\alpha_b(\tau)$ converges to a constant $\alpha_b(\infty)$ as τ gets arbitrarily large. The right-hand side of (2.24) can be decomposed as

$$(2.25)\qquad \alpha_b(\infty)\cdot[W(t+1)-W(t)] + \int_{r-1}^{r} [\alpha_b(\tau)-\alpha_b(\infty)]\cdot dW(t+r-\tau)d\tau.$$

The convergence of $\alpha_b(\tau)$ to $\alpha_b(\infty)$ implies that the variance of the second term in (2.25) can be made arbitrarily small by choosing r to be sufficiently large. Thus, equation (2.24) is approximately of the same form as (2.8) for equities, where $\alpha_b(\infty)$ appears in the place of σ. It follows that there are counterparts to the correlation restriction (2.12) for equities which will be (approximately) satisfied by the temporally aggregated model of holding-period returns for long-term bonds. Furthermore, the subjective rate of time preference, δ, can be (approximately) inferred from the constant terms of consumption and return regressions.

Rolling Over Discount Bonds

The correlation restriction also holds approximately for one-period returns formed by rolling over bonds. Let J be an integer greater than one and let $\eta = 1/J$. Consider an investment strategy that entails purchasing an r-period bond at time t, selling it at time $t+\eta$ and repeating this strategy J times. The logarithm of the resulting return is given by

$$(2.26)\qquad v_r(t+1) = \sum_{j=1}^{J} \{b_{r-\eta}(t+\eta j) - b_r[t+\eta(j-1)]\}.$$

for bond returns. Consider the difference between the versions of (2.17) for $b_{r-\eta}(t+\eta)$ and $b_r(t)$,

$$(2.27)\qquad \gamma[c(t)-c(t+\eta)] + [b_{r-\eta}(t+\eta)-b_r(t)-p(t+\eta)+p(t)] - \eta\delta - \sigma_r/2 + \sigma_{r-\eta}/2$$
$$= \int_0^{\eta} \alpha_b(\tau+r-\eta)\cdot dW(t+\eta-\tau)d\tau.$$

Next, shifting (2.27) forward ηj units forward in time for $j=0,1,\ldots J-1$ and summing the resulting terms gives

$$(2.28)\qquad \gamma[c(t)-c(t+1)] + v_r(t+1) - \delta - J(\sigma_r/2 + \sigma_{r-\eta}/2) =$$
$$\alpha_b(r)\cdot[W(t+1)-W(t)] + \epsilon_b(t)$$

where

$$(2.29) \qquad \epsilon_b(t) = \sum_{j=1}^{J}\int_0^{\eta}[\alpha_b(\tau+r-\eta)-\alpha_b(r)]\cdot dW(t+j\eta-\tau)d\tau$$

Since α_b is continuous at r, the variance of $\epsilon_b(t)$ can be made arbitrarily small by choosing J to be sufficiently large (η to be sufficiently small). Thus, equation (2.28) is approximately of the same form as (2.8) for equities, where $\alpha_b(r)$ appears in the place of σ.

Perhaps arguments like these can be used to justify the imposition by Grossman, Melino and Shiller (1987) of autocorrelation restrictions in the models of holding period returns on long-term bonds and returns from various rollover strategies. The number of rollovers they used was quite small, however.

2.B A Log-linear ICAPM with Preferences that are not State Separable

Most *ICAPM*s have assumed that agents maximize von Neuman-Morgenstern preferences, with particular attention having been given to the HARA class of preferences. Recently, Epstein and Zin (1989a,b) proposed an *ICAPM* in which agents maximize a non-expected utility function of the type introduced by Kreps and Porteus (1979). Epstein and Zin (1989b) focused on the non-linear Euler equations associated with this model in their econometric analysis. They note that a special case of their model is a log-linear *ICAPM*. In this subsection we derive a temporally aggregated counterpart that is similar to (2.9). This derivation provides an alternative interpretation of the conditional moment restrictions and parameters in the log-linear *ICAPM*s deduced under the assumption of expected utility.

For pedagogical convenience, we adopt a discrete time formulation of the model, although a continuous formulation has been developed by Duffie and Epstein (1989). We do, however, presume that consumers make choices at a time interval that is shorter than that of an econometrician. So while an econometrician observes consumption and returns at integer points in time, the decision interval for consumers is $\eta = 1/J$ where J is an integer that is greater than one. With this in mind, we assume that a representative agent has logarithmic risk preferences and examine the following special case of the recursive utility function studied by Epstein and Zin:

$$(2.30) \qquad U_t = \Big[(1-\lambda)(C_t)^{\gamma+1} + \lambda exp[(\gamma+1)E(\log U_{t+\eta}|\mathcal{I}_t)]\Big]^{1/(\gamma+1)},$$

$t=0,\eta,2\eta,\ \ldots$ where $\gamma<0$ and λ is the subjective discount factor. In (2.30) $\mathcal{I}_t$ denotes a discrete time information set available to the consumer at date t. The parameter γ governs intertemporal substitution of consumption, with the elasticity of substitution being $-1/\gamma$. In the special case in which $\gamma = -1$, preferences are state separable with a logarithmic period utility function.

Notice that (2.30) gives U_t as a function of C_t and $U_{t+\eta}$ that is homogeneous of degree one. As a consequence the equilibrium wealth of the consumer inclusive of current consumption is proportional (conditioned on time t information) to U_t, where the proportionality factor is the reciprical of the marginal utility for time t consumption. Let ω_t denote the date t equilibrium wealth of the representative agent net of current consumption. Then

$$(2.31) \qquad \omega_t = \frac{(U_t)^{\gamma+1}(C_t)^{-\gamma}}{(1-\lambda)} - C_t.$$

Note that when $\gamma = -1$, wealth is linear in consumption which is a well known result for logarithmic preferences [e.g. see Rubinstein (1974)]. The marginal rate of substitution of consumption between dates t and $t+\eta$ is

$$(2.32) \qquad MRS_{t,t+\eta} = \lambda(C_{t+\eta}/C_t)^{\gamma} exp[(\gamma+1)E(\log U_{t+\eta}|I_t)]/(U_{t+\eta})^{\gamma+1},$$

and the logarithm of the return on the wealth portfolio over the time interval t to $t+\eta$ is

$$(2.33) \qquad v^{\omega}_{t,t+\eta} = log[(\omega_{t+\eta} + C_{t+\eta})/\omega_t].$$

Combining these observations, the standard Euler equation, $E[MRS_{t,t+\eta}\exp(v^{\omega}_{t,t+\eta})|I_t] = 1$, the return on the wealth portfolio can be expressed as

$$(2.34) \qquad E[\gamma(c_{t+\eta} - c_t) + v^{\omega}_{t,t+\eta}|I_t] - \eta\delta = 0.$$

where $\delta = -\log(\lambda)/\eta$. This is equivalent to equation (2.15) in Epstein and Zin (1989b). Note that δ can be interpreted as the continuously compounded subjective rate of time preference.

An implication of (2.34) is that the disturbance

$$(2.35) \qquad u_{t+\eta} = \gamma(c_{t+\eta} - c_t) + v^{\omega}_{t,t+\eta} - \eta\delta$$

satisfies the moment restriction $E(u_{t+\eta}|I_t) = 0$. The variance of $u_{t+\eta}$ conditional on I_t may not be constant (i.e., there may be conditional heteroskedasticity), though the process $\{u_t\}$ is assumed to be stationary.

The process $\{u_t\}$ typically will be a nondegenerate stochastic process except in the special case in which $\gamma = -1$.

Summing (2.35) over the J time intervals between t and t+1 gives

$$(2.36) \qquad \sum_{j=1}^{J} u_{t+j\eta} = \gamma(c_{t+1} - c_t) + r^{\omega}_{t,t+1} - \delta.$$

There are several important differences between (2.36) and the corresponding expression (2.8) for the continuous time, expected utility *ICAPM*. First, the parameter multiplying consumption growth in (2.33) is the inverse of the elasticity of substitution, which may be different from the coefficient of relative risk aversion. Thus, this derivation confirms Hall's (1988) interpretation of γ. Second, while the derivation of (2.8) assumed that marginal utilities and asset prices were jointly lognormal, no distributional assumptions were imposed in deriving (2.34). On the other hand, (2.8) is satisfied for any asset price that meets the distributional requirement (2.7), whereas (2.35) holds only for the return on the aggregate wealth portfolio. Finally, the constant terms in the two expressions are different due to the presence of the conditional variance in (2.8).

In spite of these differences, (2.7) and (2.32) have remarkably similar econometric implications. Equation (2.35) can be temporally aggregated to obtain

$$(2.37) \qquad \gamma[c^a_{t+1} - c^a_t] + r^{\omega a}_{t,t+1} + \delta = u^a_{t+1},$$

where the aggregation is again over the J decision intervals of length η. The disturbance term u^a_{t+1} follows an MA(1) process, $E[u^a_{t+2}|I_t] = 0$, and has a first-order autocorrelation that can be inferred *a priori*. This autocorrelation approaches .25 as η shrinks to zero. The disturbance in (2.34), however, does not in general satisfy the *conditional* correlation restriction (2.12), and may be conditionally heteroskedastic.

3. EFFICIENT GMM ESTIMATION

In this section we discuss the efficiency of *GMM* estimators of the linear asset pricing models described in Section 2. Let γ_o denote the value of the preference parameter γ for the population used in the econometric analysis. In this section we focus on methods for estimating γ_o efficiently using relation (2.9). This relation is a single equation of a simultaneous system determining consumption growth and the equity return. As such, γ_o can be estimated using a single-equation, limited information estimator or a system, full information estimation procedure that estimates γ_o along with the other parameters characterizing the joint process for consumption growth and the asset return. Initially, we focus on *GMM* estimation of the single equation (2.9). At the end of this section, we discuss the relative efficiency of limited information *GMM* estimators of γ_o in the context of (2.9) and full information estimators.

To set up this discussion, let $y_{t+1}' = [c^a(t+1)-c^a(t),\ q^a(t+1)-q^a(t)]$ and $\Delta(\gamma) = [\gamma\ 1]$, and assume that T evenly spaced observations on y at integer points in time are available for estimation. Since γ_o is the parameter of interest, we assume all variables are deviated from their means and, thereby, suppress the constant term. Extensions of this discussion to the efficient estimation of γ_o and the constant term are straightforward.

In terms of this notation, equation (2.9) can be expressed as

$$(3.1) \qquad \Delta(\gamma_o)y_t = \epsilon_t$$

where $\epsilon_t = u^a(t)$ sampled at integer points in time t. The disturbance process $\{\epsilon_t\}$ is homoskedastic, follows an MA(1) process and has first-order autocorrelation equal to .25. An implication (2.11) is that ϵ_t satisfies the conditional moment restriction

$$(3.2) \qquad E(\epsilon_{t+2}|J_t) = 0,$$

where J_t is the discrete time information set generated by y_t, y_{t-1}, ... Hence, any random variable with a finite second moment in the time t information set J_t is orthogonal to ϵ_{t+2}.

Finite Lag Efficiency

Following Hansen (1982) and Hansen and Singleton (1982), these unconditional moment restrictions can be used to construct a rich class of *GMM* estimators of γ_o. We focus on *GMM* estimators for which the orthogonal variables are linear combinations of current and past values of y_t with time invariant coefficients. Initially, we construct a family of *GMM* estimators of γ_o using the $2(lag)$ unconditional moment restrictions,

$$(3.3) \qquad E(x_t\epsilon_{t+2}) = 0, \quad x_t' = [y_t', y_{t-1}', \ldots, y_{t-lag+1}'], \; lag < \infty.$$

This is accomplished by forming linear combinations of the instrument vector x_t, i.e. $z_t = \Psi' x_t$, and selecting the *GMM* estimator of γ_o to satisfy the sample version of the scalar moment condition $E(z_t\epsilon_{t+2}) = 0$.

The asymptotic variance, $avar(z)$, of the resulting *GMM* estimator of γ_o depends on the choice of Ψ used in constructing z_t. To display this dependence, define

$$(3.4) \qquad d_{t+2} = [\partial\Delta(\gamma_o)/\partial\gamma]y_{t+2} - [c^*(t+2)-c^*(t+1)] \quad .$$

Then

$$(3.5) \qquad avar(z) = \sum_{j=-1}^{1} [E(z_t z_{t-j}) E(\epsilon_{t+2}\epsilon_{t+2-j})]/[E(d_{t+2}z_t)]^2.$$

When $E(z_t d_{t+2})$ is zero, the parameter vector γ_o cannot be identified using the unconditional moment condition $E(z_t \epsilon_{t+2}) = 0$ and we interpret $avar(z)$ as being infinite. The limits of the summation in (3.5) are dictated by the order of the MA disturbance term, 1.

Holding *lag* fixed and choosing Ψ such that the instrument $z_t = \Psi' x_t$ is

$$(3.6) \qquad z_t^o = E(x_t' d_{t+2}) \left[\sum_{j=-1}^{1} E(x_t x_{t-j}') E(\epsilon_{t+2}\epsilon_{t+2-j}) \right]^{-1} x_t.$$

gives the smallest value of (greatest lower bound for) $avar(z)$ for all possible choices of Ψ [Hansen (1982)]:

$$(3.7) \qquad inf_{lag} = \left\{ E(x_t' d_{t+2}) \left[\sum_{j=-1}^{1} E(x_t x_{t-j}') E(\epsilon_{t+2}\epsilon_{t+2-j}) \right]^{-1} E(d_{t+2}x_t) \right\}^{-1}.$$

An equivalent way of obtaining this bound is to minimize a quadratic form in the sample counterpart of the moments $E(x_t \epsilon_{t+2})$ using an optimal weighting matrix [Hansen (1982)].

A practical consideration that arises in constructing the sample counterpart to (3.6) is that this approximation is not invariant to normalization, even after accounting for the change in scale. Recall that $\Delta(\gamma) = [\gamma\ 1]$. In this case the second entry of Δ is *normalized* to be one. Alternatively, suppose that γ_o is different from zero and divide both entries of $\Delta(\gamma)$ by γ. In this case we treat $1/\gamma_o$ as the parameter to be estimated.

This changes the normalization from the second entry to the first entry as in Hall (1988). Furthermore, it preserves the conditional moment restriction (3.2) except that ϵ_{t+2} is replaced by ϵ_{t+2}/γ_o. Notice that this change in normalization alters the random variable d_{t+2} used in constructing z_t^o [see (3.6)]. Conditional moment restriction (3.2) implies that the columns of $E(x_t y_{t+2}')$ are proportional which in turns implies that the z_t^o's computed from the alternative normalizations are proportional. In practice, however, $E(x_t y_{t+2}')$ is approximated by a time series average, and the columns of this approximate matrix will typically not be proportional. As is well known in other estimation environments, the resulting *single equation* estimates may be quite sensitive to the choice of normalization [e.g. see Hillier (1990)].

We next propose an alternative estimator of γ_o that attains inf_{lag} and avoids this sensitivity to the normalization. This estimator exploits our observation in section 2 that γ_o can be identified from the restriction

$$(3.8) \qquad \gamma_o \rho_c + \rho_q = 0,$$

where ρ_c and ρ_q are the regression coefficients in the following two-period ahead forecast equation system:

$$(3.9) \qquad y_{t+2} = X_t \Pi_o + e_{t+2}; \quad X_t = \begin{bmatrix} x_t' & 0 \\ 0 & x_t' \end{bmatrix}, \quad \Pi_o = \begin{bmatrix} \rho_c \\ \rho_q \end{bmatrix},$$

$$(3.10) \qquad E(X_t' e_{t+2}) = 0.$$

The vector e_{t+2} of projection errors is, by construction, orthogonal to x_t, but not necessarily to $y_{t-lag-j}$ for $j=0,1,\ldots$.

An estimator of γ_o can be obtained by using the sample counterpart

(3.10) to estimate Π_o subject to restrictions (3.8):

$$\text{(3.11)} \qquad \underset{\Pi}{minimize}\ T\left\{\left(\tfrac{1}{T}\right)\sum_{t=1}^{T} X_t' [y_{t+2}-X_t\Pi]\right\}' W \left\{\left(\tfrac{1}{T}\right)\sum_{t=1}^{T} X_t' [y_{t+2}-X_t\Pi]\right\}$$

subject to $[\gamma I \quad I]\Pi = 0$ for some γ in $\mathbb{R}$.

In (3.11) the matrix W is a positive definite distance or *weighting* matrix which will be described shortly. As long as the solution to (3.11) is not $\gamma=0$, an identical estimator of γ_o is obtained by minimizing (3.11) subject to the constraint $[I \quad \theta I]\Pi = 0$ for some θ in $\mathbb{R}$, where it is understood that $\theta=1/\gamma$. While the minimizer to (3.11) has this invariance property, it is harder to compute than the single-equation estimator because the optimization problem (3.11) must be solved numerically.[4]

As is typically the case in *GMM* estimation, the choice of weighting matrix W has an impact on the asymptotic efficiency of the resulting estimator. To construct a W with the property that the solution to (3.11) attains the bound inf_{lag}, we first estimate Π_o without restrictions using the ordinary least squares procedures in Hansen and Hodrick (1980). This is equivalent to solving (3.11) without the constraint and with an arbitrary choice of a positive definite matrix W. Then we use the least squares residuals $\hat{e}_{t+2}$ (estimated values of e_{t+2}) to form an efficient W:

$$\text{(3.12)} \qquad \hat{W} = [\hat{C}(0) + \hat{C}(1) + \hat{C}(1)']^{-1},$$

where

$$\text{(3.13)} \qquad \hat{C}(j) = \left(\tfrac{1}{T}\right)\sum_{t=j+1}^{T} X_t' \hat{C}_e(j) X_{t-j}; \qquad \hat{C}_e(j) = \left(\tfrac{1}{T}\right)\sum_{t=j+1}^{T} \hat{e}_{t+2}\hat{e}_{t+2-j}'$$

The matrix $\hat{W}$ in (3.12) is constructed as if the vector disturbance term e_{t+2} satisfies $E(e_{t+2}|J_t) = 0$. It turns out that all that is required for this estimator of γ_o to attain the finite lag efficiency bound inf_{lag} is the weaker moment restriction $E(\epsilon_{t+2}|J_t) = 0$, where $\epsilon_{t+2} = [\gamma_o \ 1]e_{t+2}$; see the appendix.[5]

Infinite Lag Efficiency

The choice of *lag* in the previous discussion was arbitrary. Larger values of *lag* will lead to more efficient *GMM* estimators of γ_o. In this subsection we describe a *GMM* estimator that is efficient relative to all choices of *lag*. More precisely, let Z denote the family of estimators indexed by the scalar stochastic instrument process $z = \{z_t\}$, which includes the finite lag *GMM* estimators for all finite values of *lag*:

$$(3.14) \qquad Z = \{z : z_t = \sum_{j=0}^{lag-1} \psi_j' y_{t-j} \text{ for some } \psi_j \in R^2 \ j=1,2,\dots lag-1, \text{ some } lag < \infty \text{ and all } t\}.$$

(In terms of our previous notation, $\Psi' = [\psi_0', \psi_2', \ \dots \ , \psi_{lag-1}']$.) The optimal *GMM* estimator is constructed using a z^o satisfying[6]

$$(3.15) \qquad inf = avar(z^o) \le avar(z) \quad \text{for all } z \in Z \ .$$

Hansen (1985) and Hansen and Singleton (1989) provide an explicit representation of z^o for a class of linear time series models of which (3.1) is a special case. From their analyses it follows that there exists a lag polynomial $\Lambda(L) = (\lambda_0 + \lambda_1 L)^{-1}$, with $|\lambda_1|<1$ and (λ_0, λ_1) chosen to solve the equations $\lambda_0^2 + \lambda_1^2 = E(\epsilon_t^2)$ and $\lambda_0\lambda_1 = .25E(\epsilon_t^2)$ such that

$$(3.16) \qquad z_t^o = \Lambda(L)\{E[\Lambda(L^{-1})d_{t+2}|J_t]\}.$$

In general z^o is not a member of Z. The reason is that the the z's in Z are only permitted to depend on a finite number of current and lagged values of y, whereas the optimal index z^o depends on the entire past history of y.

Substituting (3.16) for z_t^o into (3.5) yields the efficiency bound for the class of *GMM* estimators Z [see Hansen (1985)]:

$$(3.17) \qquad inf = avar(z^o) = \left[E\left[E[\Lambda(L^{-1})d_{t+2}|J_t]^2\right]\right]^{-1}$$

As the number of lagged values of y_t included in x_t (*lag*) is increased, the finite-lag efficiency bound inf_{lag} given by (3.6) converges to the efficiency bound *inf* given by (3.17) [Hayashi and Sims (1983) and Hansen and Singleton (1989)]. This convergence follows from the observation that z_t^o can be approximated in mean-square arbitrarily well by finite linear combinations of current and past values of y_t.

The form of the optimal *GMM* estimator z^o has a natural interpretation. Temporal aggregation leads to an autocorrelation in the disturbance ϵ_t of .25. The coefficients in the lag polynomial $\Lambda(L)$ are chosen so that the forward filter $\Lambda(L^{-1}) = (\lambda_0 + \lambda_1 L^{-1})^{-1}$ removes the serial correlation from the process $\{\epsilon_t\}$; that is, $\{\Lambda(L^{-1})\epsilon_t\}$ is a serially uncorrelated and conditionally homoskedastic process with a unit variance. This forward filtered process satisfies $E[\Lambda(L^{-1})\epsilon_{t+2}|J_t] = 0$, because all future values of ϵ_{t+2} are mean independent of the elements of J_t. It follows that all elements of J_t with finite second moments continue to be admissible instruments for the forward-filtered disturbance. [This observation underlies the development of the forward-filtered instrumental variables estimator proposed by Hayashi and Sims (1983)]. The optimal instruments for a simultaneous equations model with serially uncorrelated and homoskedastic disturbances are obtained by taking the partial derivative of the disturbance vector with respect to the parameters to be estimated and then projecting these partial derivatives onto the past history of the predetermined variables y_t [Amemiya (1977)]. This explains the presence of $E[\Lambda(L^{-1})d_{t+2}|J_t]$ in (3.8) as $\partial[\Lambda(L^{-1})\Delta(\gamma_o)y_{t+2}]/\partial\gamma = \Lambda(L^{-1})d_{t+2}$. Algorithms for calculating (3.16) and (3.17) are described in Hansen and Singleton (1989).

To compute z_t^o requires knowledge of $\Lambda(L)$. For asymptotic efficiency it suffices to know z_t^o up to a scale factor. Hence all that is required is knowledge of the ratio λ_1/λ_0. This ratio can be inferred from the correlation of the disturbance term (.25). More generally the coefficients of $\Lambda(L)$ would have to be estimated before calculating an optimal *GMM* estimator. It turns out that this first-stage estimation has no impact on the asymptotic distribution of the resulting *GMM* estimator.

Heteroskedasticity

The optimality of z^o is relative to other *GMM* estimators that exploit orthogonality conditions implied by conditional mean restrictions of the form (3.2) for which the associated expectational errors are homoskedastic. The disturbance in (2.9) is homoskedastic by construction. However, the corresponding disturbance in (2.33) derived from the model with preferences that are not state separable may be heteroskedastic. In this case z^o is not the optimal index for estimating γ_o because there is no correction for heteroskedasticity included in (3.9).

Though z^o is not the most efficient *GMM* estimator in the presence of heteroskedasticity, the resulting estimator of γ_o is still consistent. And it seems plausible that is many circumstances the optimal adjustment for serial correlation in $\{\epsilon_t\}$ underlying the construction of z^o will result in a more efficient *GMM* estimator than simply using a small set of lagged values of y as instruments. Therefore it may be desirable to implement the estimators discussed in this paper in the presence of heteroskedastic disturbances. The asymptotic variance for the *GMM* estimator using the moment conditions $E(z_t^o \epsilon_{t+2}) = 0$ in the presence of heteroskedasticity is given by

$$(3.18) \qquad \sum_{j=-1}^{1} E(z_t^o z_{t-j}^o \epsilon_{t+2} \epsilon_{t+2-j}) / [E(d_{t+2} z_t^o)]^2 .$$

and not by *inf* in (3.17). Unlike the similar expression (3.5), (3.18) incorporates heteroskedastic-consistent estimators of the autocovariances of $\{z_t^o \epsilon_{t+2}\}$ as suggested in Hansen (1982).

Correlation Restriction

The expectational error $u_{t+2} = .25(\epsilon_{t+2})^2 - \epsilon_{t+2}\epsilon_{t+1}$ associated with the conditional second moment restriction (2.12) also follows an MA(1) process. However, because this error involves nonlinear functions of the y_{t+2} and y_{t+1}, it is not homoskedastic and does not fit into our general framework (3.1) - (3.2). Hansen, Heaton and Ogaki (1988) discussed efficient *GMM* estimation in the context of models with heteroskedastic disturbances. In general, the form of these estimators is much more complicated than the optimal estimator z^o for homoskedastic models and is correspondingly more difficult to implement.

In the case of expression (2.33) derived under the assumption of non-state separable preferences, a counterpart to the conditional correlation restriction (2.12) does not hold. Nevertheless, there is an additional efficiency gain to exploiting the unconditional correlation restriction

$$(3.19) \qquad E[.25(\epsilon_{t+2})^2 - \epsilon_{t+2}\epsilon_{t+1}] = 0$$

in addition to the restriction $E(z_t^o \epsilon_{t+2}) = 0$ in calculating the *GMM* estimator of γ_o. We pursue this observation in section 4.

System Estimation

An alternative approach to estimating the linear asset pricing model (3.1)-(3.2) is to estimate a system of equations describing the evolution of $\{y_t\}$ subject to the conditional mean restriction (3.2). This can be accomplished using either *GMM* estimation or *ML* estimation obtained by maximizing the normal likelihood function for the model (3.1). System estimation requires that an additional auxiliary equation be appended to the econometric relation (3.1) so that the model gives a complete description of the two-dimensional process $\{y_t\}$.

In section 4 we study two alternative parameterizations of the process $\{y_t\}$. First, we consider the parameterization (3.9) under the assumption that $\Pi_o x_t$ is the best predictor of y_{t+2} based on the entire history of y_t. In this case, the moment conditions (3.10) are replaced by the more stringent restrictions:

$$(3.20) \qquad E(e_{t+2} y_{t-j}') = 0, \text{ for } j=0,1,\dots .$$

An efficient system *GMM* estimator of Π_o based on (3.20) can be constructed in a manner analogous to that described for efficient single equation estimation of γ_o. This system estimator is asymptotically equivalent to the *ML* estimator of Π_o obtained by estimating both the autoregressive and moving-average parameters of the ARMA(*lag*+1,1) process (3.9) [Stoica, Soderstrum and Friedlander (1985), Hansen (1989), and Hansen and Singleton (1989)].

We used the following method in our empirical analysis. First we estimated the parameter vector Π_o using an efficient *GMM* estimator of the unconstrained equation (3.9). This estimator uses the sample counterparts of the moment conditions

$$(3.21) \qquad E(Z_t^{o\prime} e_{t+2}) = 0,$$

where Z_t^o is dimensioned 2 by 4(*lag*). As with the single-equation optimal *GMM* estimator, Z_t^o will depend in general on the infinite past history of y. The procedure described in Hansen and Singleton (1989) was implemented to construct an approximation to Z_t^o, using least squares methods to estimate Π_o and the least squares residuals to estimate $E(e_t e_t')$ and $E(e_t e_{t-1}')$. This results in a matrix time series $\hat{Z}_t^o$, t=1,2,...T. We then estimated γ_o by solving

$$(3.22) \qquad \underset{\Pi}{minimize}\; T\left[(\tfrac{1}{T})\sum_{t=1}^{T}\hat{Z}_t^{o\prime}(y_{t+2}-X_t\Pi)\right]'\hat{V}\left[(\tfrac{1}{T})\sum_{t=1}^{T}\hat{Z}_t^{o\prime}(y_{t+2}-X_t\Pi)\right]$$

$$subject\ to\ [\gamma I\ \ I]\Pi = 0 \text{ for some } \gamma \text{ in } \mathbb{R},$$

where $\hat{V}$ is the counterpart to $\hat{W}$ in (3.12)-(3.13) with $\hat{Z}_t^o$ used in place of X_t.

Since Z_t^o is the efficient instrument matrix for estimating Π_o without constraints and $\hat{V}$ is an efficient choice of weighting matrix to use for the 4(*lag*) moment conditions (3.21), it follows that the resulting estimator of γ_o is asymptotically efficient.

For comparison, we also study the following alternative two-step ahead forecasting equation for y_t:

$$(3.23) \qquad y_{t+2} = [A(L)/B(L)]w_t + e_{t+2}, \quad e_{t+2} = C_o w_{t+2} + C_1 w_{t+1}.$$

In (3.23) $\{w_t\}$ is a two-dimensional *white noise* process, the two by two matrix polynomial $A(\zeta)$ has order *lag* and the scalar polynomial $B(\zeta)$ has order *lag*+1. Thus, the two-step ahead forecasts of y_{t+2} depend on the infinite past of y_t. The counterpart to restriction (3.8) for this representation of y is

$$(3.24) \qquad [\gamma_o \ 1]A(\zeta) = 0 \text{ for all } \zeta.$$

The model can be estimated using the *ML*-based methods for estimating *exact* rational expectations models described in Hansen and Sargent (1990).

The estimators of γ_o within the systems [(3.9), (3.10)] and (3.23) have asymptotic variances that are generally smaller than the single equation efficiency bound *inf*. The efficiency gains emerge because of the assumed knowledge of the orders of the ARMA representations for these probability models of y. If these orders are in fact not known *a priori*, then parametric system estimators of γ_o must be based on approximate time series models for y. When an approximate time series model is used, *inf* is a measure of the asymptotic variance of the estimator of γ_o that accounts for the absence of prior knowledge of the orders of the ARMA representation [Hansen (1989) and Hansen and Singleton (1989)].[7] In other words, without prior knowledge of the orders of the ARMA representation, there is no asymptotic efficiency gain from using full system versus single-equation *GMM* estimators of γ_0 described previously. This result is the linear time series counterpart to the asymptotic equivalence of the limited information *ML* estimator and the two-stage least squares estimator in the classical simultaneous equations literature.

4. EMPIRICAL ANALYSIS

To re-examine the consequences of temporal aggregation for the parameter estimates of log-linear *ICAPMs*, we estimated the models described in section 2 using the methods described in section 3. The equations characterizing equity returns were estimated using monthly data for consumption of nondurables and services and the time-averaged, value-weighted return on the NYSE common stock portfolio. The nominal temporally averaged stock return was constructed for the sample period August 1962 through December 1985 using CRSP data on daily stock returns.[8] For the equations describing nominal bond returns we used quarterly data and measured the temporally averaged nominal three month return by the daily average of three month Treasury bill returns constructed by the Federal Reserve. The sample period was from the second quarter of 1947 through the fourth quarter of 1986. The nominal returns were converted to real returns appearing in (2.9) using the implicit price deflator for the monthly consumption data. The consumption, price, and bill returns are from the CITIBASE data set. We let y_t denote the two-dimensional vector of temporally aggregated logarithms of consumption growth rates (y_{1t}) and real returns-- stock or bill returns-- (y_{2t}).

Most our our analysis focuses on the lognormal model in which preferences are assumed to be state separable. As is evident from the discussion in section 2.B, some of our results can be reinterpreted as applying to a model in which preferences are not state separable, although the resulting disturbance terms may be heteroskedastic. We will comment more on this reinterpretation later in this section.

We take equation (3.9) as as starting point for our empirical analysis. If Π_o is zero, then there is no information about γ_o in the matrix of reduced-form coefficients in (3.9). Therefore, prior to estimating γ_o, we test

(4.1) $\quad H_1: \Pi_o = 0.$

H_1 is implied, for instance, by constant real interest rates as assumed by Hall (1978). For a different specification of preferences, this hypothesis has recently been tested by Christiano, Eichenbaum, and Marshall (1990) using methods similar to those discussed in this paper to account for temporal aggregation. Ignoring (3.8), Π_o is exactly identified by the least-squares normal equations (3.10). To test H_1 we make the simplifying assumption that $E(e_{t+2}|J_t) = 0$, which permits us to use the least-squares inference methods suggested by Hansen and Hodrick (1980).

The tests statistics for H_1 are reported in Table 1 for three choices of *lag*. The column labeled χ^2 gives the test statistic, which is distributed asymptotically as a chi-square with *df* degrees of freedom and probability value *Prob*. The results for stock returns indicate that there is little evidence against H_1. This finding is consistent with the large standard errors for γ_o reported in previous studies and subsequently here for stock return equations. In contrast, there is substantial against H_1 when Treasury bill returns are studied.

Next it is of interest to examine ρ_c and ρ_q separately. Recall that γ_o satisfies the restriction (3.8): $\gamma_o\rho_c = \rho_q$. When γ_o is zero, it follows that

(4.2) $\quad H_2: \rho_q = 0.$

That is, temporally aggregated returns are not predictable given x_t. Alternatively, since $\rho_c = \rho_q/\gamma_o$, when $(1/\gamma_o)$ is zero

(4.3) H_3: $\rho_c = 0$.

Equivalently, the growth rate of temporally aggregated consumption is not predictable given x_t. Hypothesis H_2 is consistent with *linear utility* while hypothesis H_3 is consistent with an intertemporal elasticity of substitution equal to zero. One of our motivations for looking at H_3 is Hall's recent finding that $(1/\gamma_0)$ is approximately zero. Hall's finding is in contrast to our earlier results [Hansen and Singleton (1982, 1983)] from a model in which the sampling interval of the data coincided with the decision interval of agents.

So far we have examined hypotheses based on two *extreme* cases of proportionality restriction (3.8). We now consider the more general hypothesis:

(4.4) H_4: $\gamma\rho_c = \rho_q$ for some γ in $\mathbb{R}$.

We test this hypothesis using the minimized value of (3.11) with $W = \hat{W}$, which is asymptotically distributed as chi square with *lag* - 1 degrees of freedom (Hansen 1982). The reduction in degrees of freedom by one relative to the test statistics for hypotheses H_2 and H_3 occurs because γ_0 is no longer specified *a priori*, but is now estimated. While the derivation in Hansen (1982) relies on the assumption that $E(e_{t+2}|J_t) = 0$, in the appendix we show that the weaker assumption $E(\epsilon_{t+2}|J_t) = 0$ suffices, where $\epsilon_{t+2} = [\gamma_0 \ 1]e_{t+2}$.

Estimates of γ_0 and the chi-square statistics for testing H_2, H_3, and H_4 using stock return data are reported in Table 2. The chi-square statistics indicate that there is little evidence against H_2, but somewhat more evidence against H_3. Since H_4 is less restrictive than H_2, there is also little evidence against H_4. The estimate of γ_0, $\hat{\gamma}$, is not very precise, and in fact has the wrong sign. The last column of the table labeled *corr* reports the estimated correlation of ϵ_{t+2} which should be .25. In all cases it is close to this magnitude.

Further insight into the sample information about γ_0 is provided in Figure 1. This figure was constructed by solving the optimization problem (3.11) repeatedly, holding γ fixed at values along the horizontal axis. In all cases W was set to $\hat{W}$. The vertical axis displays the corresponding values of the criterion functions evaluated at the solutions to these problems. Each curve represents a different choice of *lag*. The minimum value of each curve-- i.e., the statistic used to test H_4-- occurs at the $\hat{\gamma}$ reported in Table 2. For γ equal to γ_0 the minimized value of the *GMM* criterion function has an asymptotic chi square distribution with 2(*lag*) degrees of freedom. As already noted, there is a loss of one degree of freedom when γ is set to $\hat{\gamma}$.

While the point estimates of γ_0 are positive, the three curves in Figure 1 indicate that values of γ_0 less than zero are plausible. Another interesting feature of these curves is that they all peak at values of γ near -20 before asymptoting to the values at $|\gamma| = \infty$ reported in Table 2. These findings are consistent with the small values of γ_0 reported in Hansen and Singleton (1982,1983). In contrast to Hall (1988), accounting for temporal aggregation does not lead us to conclude that larger values of $|\gamma_0|$ (smaller values of $1/|\gamma_0|$) are more plausible than values of γ_0 close to zero.

In the case of Treasury bill returns, there is substantial evidence against all of the null hypotheses H_2, H_3, and H_4, especially for *lag* equal to two and three. In contrast to equity returns and consistent with our discussion in section 2, the estimated first-order autocorrelation of ϵ_{t+2} is not close to .25. The fact the estimated autocorrelations of the bond disturbances differ substantially from .25 suggests that the incorrect imposition of the restriction $\rho = .25$ may induce important biases in the estimates of parameters or standard errors. While the point estimates for γ_0 have the wrong sign, it is difficult to interpret these coefficients in light of the pervasive evidence against hypothesis H_4.

More efficient, system estimates of γ_0 can be obtained using the system *GMM* and *ML* methods described in section 3. In calculating these estimates, we focus on the stock return data because of the large values of the chi square statistics reported in Table 3 for the Treasury bill data. The rows labeled *OGMMlag*, with *lag* = 1, 2, and 3, in Table 4 display the estimates of γ_0 and the chi square statistics for the hypotheses H_2, H_3, and H_4 from fitting (3.9) subject to the restriction (3.8) using the moment conditions (3.20). Estimated standard errors and probability values of the chi-square statistics are in parentheses. The estimated autocorrelations (*corr*) of $\epsilon_t = [\gamma_0\ 1]e_t$ [the disturbance in (2.9)] are reported in the last column of the table. Comparing the *OGMM* results in Table 4 with the corresponding *GMM* results in Table 2, there are small gains in precision in using *OGMM* system estimation. Also, the test statistics are similar across the two tables.

The *ML* estimates displayed in rows *MLlag* in Table 4 were obtained using parameterization (3.23) and (3.24), where *lag* determines the orders of the lag polynomials A(L) and B(L). We used a parameterization of $B(\zeta)$ suggested in Monahan (1984) to insure that the zeros of this scalar polynomial are outside the unit circle of the complex plane. To evaluate the Gaussian likelihood function, we used the filtering methods described in Anderson and Moore (1979) and Hansen and Sargent (1990). The *ML* estimates have smaller estimated standard errors than both the *GMM* and *OGMM* estimates. Also, the associated likelihood ratio tests provide somewhat more evidence against the hypothesis $H_3 : |\gamma_o| = \infty$.

Comparing Tables 2 and 4, recall that the *GMM* estimates based on the moment conditions (3.10) (Table 2) are consistent and the inference procedures are valid whether or not the true ARMA representation for y_t is (3.9) with $E(e_{t+2}|J_t) = 0$. In contrast, the efficiency of the *OGMM* and *ML* estimates displayed in Table 4 relies on a correct specification of the time series law of motion for $\{y_t\}$. Also, the *OGMM* and *ML* estimates were obtained using different probability models, which may help to explain the apparent gain in precision in using *ML* over *OGMM*.

Up to this point, we have presumed that the correlation restriction (3.19) implied by the *ICAPMs* for stock returns in section 2 are not imposed in estimation using *GMM* or *ML* procedures. Estimator efficiency can, in general, be increased by imposing the unconditional correlation restriction (3.19) along with the moment restrictions implied by (2.11). The first three rows of Table 5 display the *GMM* estimates of γ_o based on the moment conditions (3.10) and (3.19). These estimates were calculated using the heteroskedastic-consistent weighting matrix described in Hansen (1982) in order to accommodate two potential sources of heteroskedasticity. First the original disturbance ϵ_{t+2} might be heteroskedastic as in the model non-expected utility model. Additional heteroskedasticity is introduced into the moment equations through the disturbance $u_{t+2} = .25(\epsilon_{t+2})^2 - \epsilon_{t+2}\epsilon_{t+1}$ used in (3.19) to impose the correlation restriction. The validity of this estimation for the non-expected utility model relies on the assumption that the value-weighted return is also the return on the wealth portfolio (see also Epstein and Zin 1989b). The test statistics reported are for hypothesis

H_4 with the correlation restriction added.

The most striking feature of the *GMM* results in Table 5 is that the estimated standard errors are reduced dramatically over those displayed in Table 2. At the same time, adding the correlation restriction to the set of moment conditions leads to comparable point estimates, since this restriction is almost satisfied in Table 2. It is still true that there is little evidence against the hypothesis that $\gamma_o = 0$.

The last three rows of Table 5 display *ML* results with the correlation restriction (3.19) imposed on the law of motion (3.23). In contrast to Grossman, Melino, and Shiller (1987), this restriction was imposed directly on the discrete time law of motion for $\{y_t\}$ rather than indirectly through a continuous time specification of the expected utility model. A potential advantage of focusing on the former is that we avoid the aliasing problem of identifying a continuous time probability law from discrete time data. Since homoskedasticity was imposed, these estimates relate to the expected utility model. The *ML* results are comparable to the *GMM* results: the *ML* estimates of γ_o in Tables 4 and 5 are similar but the standard errors are notably smaller in Table 5.

Finally, Table 6 displays the population efficiency bounds associated with the various estimators. The column headings refer to the six probability models in Table 4, with the parameters taken to be the point estimates obtained under H_4. For each of these probability models and each choice of *lag*, the single-equation efficiency bound, inf_{lag}, for the *GMM* estimator based on the moment conditions (3.3) was calculated. To make this bound comparable to the estimated standard errors we divided it by 278 and took square roots. The row labeled $lag = \infty$ is computed using *inf* given by (3.17). The efficiency bounds were calculated using the recursive methods described in Hansen and Singleton (1989) and GAUSS code written by John Heaton and Masao Ogaki.

In section 3 it was noted that, as $lag \to \infty$, inf_{lag} converges to the single-equation efficiency bound for the model (2.9) with the moment conditions (2.11) imposed, *inf*. From Table 6 it is seen that this convergence is very rapid for the model of stock returns with the law of motion (3.9). This is consistent with our findings in Tables 2 and 4 that there are no notable gains in precision from using *OGMM* instead of *GMM* estimation.

Convergence is slower for the parameterization (3.23). In the cases of *ML*1 and *ML*2, the estimates of the law of motion (3.23) predict much larger values of the standard errors than we found in practice in Table 2. On the other hand, *ML*3 predicts standard errors very similar to those reported in Table 2 for small values of *lag*. The $lag = \infty$ bounds reported in Table 6 for all three *ML* runs are similar to the estimated standard errors reported in Table 4 for the corresponding *ML* estimates.[9] This means that the efficiency of the system *ML* estimators can be attributed largely to their selection, at least implicitly, of efficient instruments rather than to their exploitation of the information about the order of the full ARMA model.

We draw three conclusions from these results. First, the estimated efficiency gains are sensitive to the choice of probability model for stock return and consumption data. The calculations with the *ML*3 model seem to replicate most closely our estimated standard errors for *GMM* and *ML* estimators of γ_0. Second, this probability model implies substantial gains in precision from exploiting the orthogonality conditions associated with the entire history of y. These gains are related to the location of the roots of the moving average polynomial. For the *ML* parameterizations there are complex roots near the unit circle [see Hansen and Singleton (1989)]. Third, there seems to be very little efficiency gain to implementing the efficient system estimator instead of the efficient single-equation estimator.

Finally, we investigated the impact on asymptotic efficiency of imposing the the correlation restriction. Hansen, Heaton, and Ogaki (1988) characterized *GMM* efficiency bounds for general forms of multi-period conditional moment restrictions, and Heaton and Ogaki (1988) presented an algorithm for computing an efficiency bound that incorporates a conditional moment restriction. Single-equation efficiency bounds that incorporate the conditional moment restriction are displayed in the last row of Table 6. As was anticipated by the estimates in Table 5, there is a notable efficiency gain from the imposition of this additional conditional moment restriction.

V. CONCLUSIONS

In this paper we have accomplished the following. First we have investigated the extent to which the serial correlation restriction derived by Working (1960) for temporally aggregated Brownian motions apply to log-linear asset pricing models. We find that this restriction is not as pervasive as was previously thought. For instance, while the correlation restriction may be applicable to equities returns, it is not applicable to bond returns, except in two limiting cases: long term bonds or returns constructed by rolling over a large number of short term bonds. Second, we showed that when applicable, the serial correlation restriction can improve substantially the efficiency of estimators of the preference parameter γ_o. Third we proposed two new *GMM*-type estimators. One of these estimators is asymptotically equivalent to the finite-lag efficient, single equation *GMM* estimator, but is constructed to be insensitive to the choice of normalization. The other estimator is a fully efficient system estimator for models in which the observed time series can be characterized as a multi-period counterpart to a finite-order vector autoregression. Fourth, we provide statistical evidence showing that smaller (absolute) values of the preference parameter γ_o are more plausible than larger ones for equity data. For bond data, however, we found substantial evidence against the restrictions implied by the model. This evidence is consistent with our previous empirical results, Hansen and Singleton (1982,1983), some of which did not accommodate time aggregation in consumption.

In this paper we have followed much of the theoretical and empirical literature on continuous time *ICAPM*'s by modeling preferences as time separable. In taking continuous time limits, however, it may be more plausible to introduce local durability so that consumption at near-by points in time are close substitutes. Recently, Hindy and Huang (1989) have investigated theoretical continuous time *ICAPM*'s with local durability. Such models may be better suited for analyzing the effects temporal aggregation [e.g., see Heaton (1989)].

NOTES

1. Often the conditional moment restrictions are all that is known by the econometrician about the distribution of the errors; in particular, the family of distributions from which the errors are drawn is unknown.

2. The null hypotheses that risk premia are zero or constant can be expressed as restrictions on conditional forecasts of excess holding period returns. This observation underlies many of the tests of expectations theories of the term structure of interest rates, as well as tests of whether forward exchange rates are optimal forecasts of future spot exchange rates [e.g., Hansen and Hodrick (1980)]. Similarly, recent studies of mean reversion in asset returns examine whether continuously compounded, multi-period returns have constant conditional means [e.g., Fama and French (1988)].

3. We are also ignoring the implication of a monetary economy that the prices of nominal discount bonds should be less than or equal to one.

4. For a given value of γ, optimization problem (3.11) is quadratic and can be solved by simple matrix manipulations. Hence the estimate of γ can be computed by first concentrating out all of the parameters for each γ and then doing a one-dimensional numerical search over γ.

5. The fact that this estimator of γ_0 is not sensitive to normalization is no guarantee that its finite sample behavior will dominate that of the asymptotically equivalent single equation *GMM* estimators.

6. Indices in Z could be constructed using nonlinear functions of current and

past values of y_t as well. Because of the linearity we presumed in the underlying estimation environment, there would be no efficiency gain associated with using such indices. It is still possible that linear forms of conditional moment restrictions such as (3.2) may exist even though the complete law of motion for y_t is nonlinear. In these circumstances, there may well be efficiency gains to using nonlinear functions of current and past values of y_t as indices.

7. The formal large sample justifications for using approximate time series models typically have the specification of the approximate model depend explicitly on sample size with the idea that more general models are fit with larger sample sizes. In this manner the approximation error is avoided asymptotically.

8. The monthly, temporally aggregated returns were derived by constructing monthly returns from daily data and then averaging the monthly returns over the days of the month. The monthly returns were constructed such that the monthly average involved daily returns from only the current and previous months.

9. The asymptotic variance of the *ML* estimator of γ_0 should be less than or equal to *inf*. In comparing Tables 4 and 6, we see that this ordering is actually reversed. This occurs because of the different methods used to estimate the asymptotic variance and *inf*. We used the inverse Hessian to construct an estimate of the asymptotic variance, and we used the estimated parameter values in conjunction with recursive formulas reported in Hansen and Singleton (1989) to construct an estimate of *inf*.

TABLE 1: TESTS OF H_1

lag	χ^2	*df*	*Prob*
	Stock Returns (9:62-12:85)		
GMM1	5.35	4	.252
GMM2	11.02	8	.200
GMM3	13.81	12	.313
	Treasury Bill Returns (2:47-4:86)		
GMM1	84.11	4	.234E-28
GMM2	90.73	8	.331E-15
GMM3	93.85	12	.880E-14

TABLE 2: GMM ESTIMATION FOR STOCKS (9:62-12:85)

lag	γ_0	$\chi^2[df]$ H_2	H_3	H_4	corr
1	1.73	0.26 [2]	4.54 [2]	0.07 [1]	.266
	(4.19)	(.878)	(.103)	(.792)	
2	2.14	0.73 [4]	8.85 [4]	0.38 [3]	.254
	(3.82)	(.947)	(.065	(.945)	
3	3.05	2.46 [6]	9.10 [6]	1.71 [5]	.231
	(3.68)	(.872)	(.168)	(.887)	

TABLE 3: GMM ESTIMATION FOR BONDS (2:47-4:86)

lag	γ_0	$\chi^2[df]$ H_2	H_3	H_4	corr
1	11.61	80.90 [2]	1.04 [2]	0.36 [1]	.135
	(14.25)	(.2E-17)	(.594)	(.547)	
2	11.36	72.80 [4]	18.74 [4]	18.12 [3]	.150
	(13.48)	(.5E-14)	(.001)	(.0004)	
3	8.59	73.98 [6]	21.30 [6]	20.27 [5]	.154
	(7.85)	(.6E-13)	(.002)	(.001)	

TABLE 4: EFFICIENT SYSTEM ESTIMATION (9:62-12:85)

		$\chi^2[df]$			
lag	γ_o	H_2	H_3	H_4	*corr*
OGMM1	1.01 (3.24)	0.10 [2] (.950)	6.17 [2] (.046)	0.01 [1] (.943)	.275
OGMM2	1.94 (3.62)	0.49 [4] (.974)	8.41 [4] (.077)	0.18 [3] (.981)	.258
OGMM3	2.14 (3.31)	1.42 [6] (.964)	9.48 [6] (.149)	1.01 [5] (.962)	.252
ML1	2.25 (2.26)	3.87 [4] (.424)	11.91 [6] (.018)	2.69 [3] (.442)	.253
ML2	2.40 (2.22)	4.23 [6] (.631)	13.78 [6] (.032)	2.78 [5] (.733)	.245
ML3	2.28 (2.16)	5.80 [8] (.670)	15.33 [8] (.053)	4.43 [7] (.729)	.245

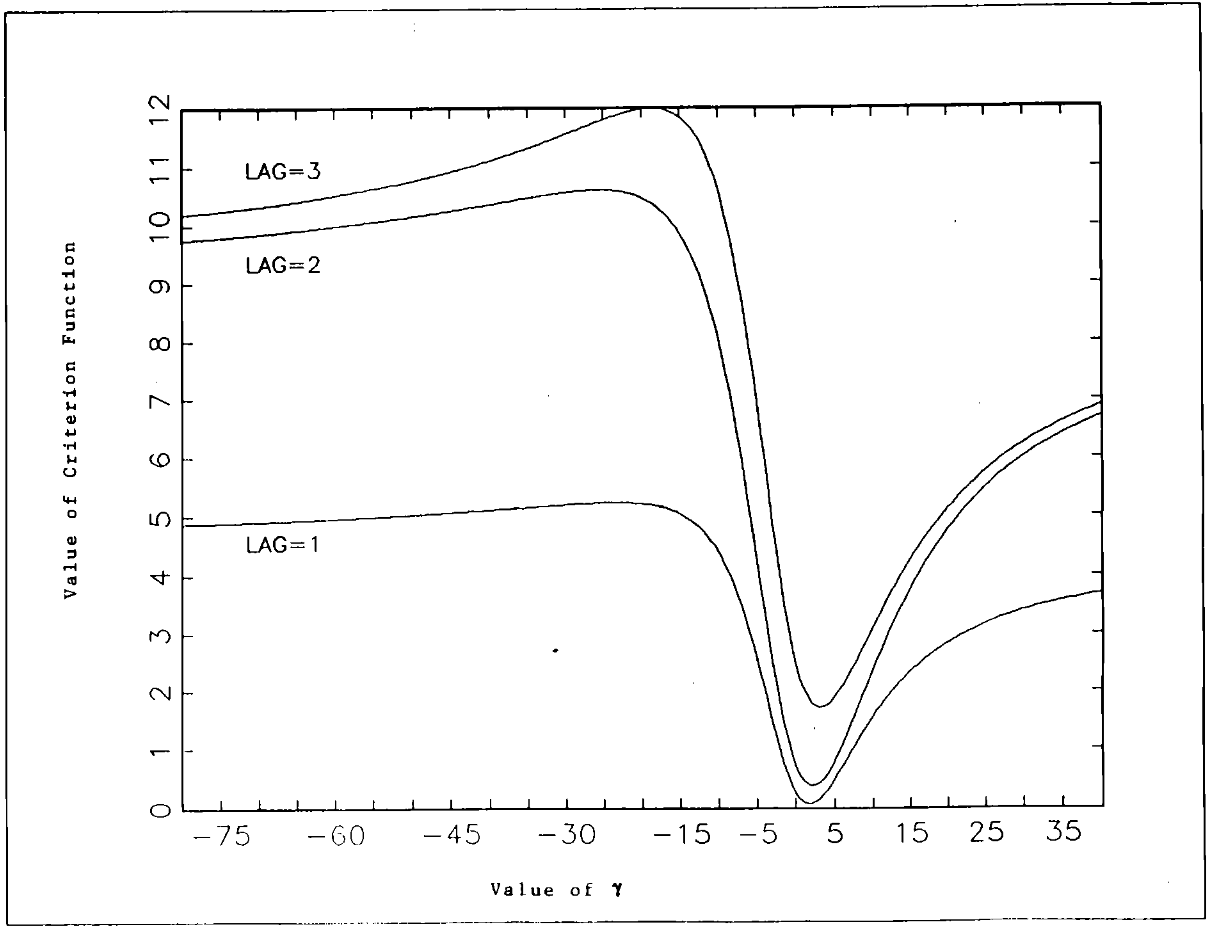
TABLE 5: ESTIMATION WITH CORRELATION RESTRICTION (9:62-12:85)

lag	γ_0	χ^2	*df*	*Prob*
*GMM*1	2.24 (1.73)	0.09	2	.954
*GMM*2	2.17 (1.73)	0.31	4	.989
*GMM*3	2.10 (2.20)	1.38	6	.967
*ML*1	2.30 (1.53)	2.69	4	.610
*ML*2	2.29 (1.42)	2.78	6	.835
*ML*3	2.17 (1.41)	4.43	8	.816

TABLE 6: EFFICIENCY BOUNDS

			Model			
	OGMM1	*OGMM2*	*OGMM3*	*ML1*	*ML2*	*ML3*
Lag						
1	3.29	4.19	4.24	12.30	6.10	4.28
2	3.16	3.63	3.83	7.01	4.35	3.81
3	3.15	3.59	3.31	5.80	4.31	3.79
4	3.15	3.59	3.28	5.08	4.19	3.68
5	3.15	3.59	3.28	4.62	3.71	3.40
10	3.15	3.59	3.28	3.57	3.07	2.89
20	3.15	3.59	3.28	2.90	2.67	2.56
40	3.15	3.59	3.28	2.49	2.37	2.29
∞	3.15	3.59	3.28	2.03	1.98	1.93
correlation restriction	2.46	1.85	1.76	1.47	1.36	1.33

Figure 1



LAG=3
LAG=2
LAG=1
0
1
2
3
4
5
6
7
8
9
10
11
12
-75
-60
-45
-30
-15
-5
5
15
25
35
Value of Criterion Function
Value of γ

REFERENCES

Amemiya, T., "The Maximum Likelihood and Nonlinear Three-stage Least Squares Estimator in the General Nonlinear Simultaneous Equations Model," *Econometrica* 45, 1977, 955-968.

Anderson, B. D.O., and J. B. Moore, *Optimal Filtering*, Prince-Hall, Inc. Englewood Cliffs, New Jersey, 1979.

Barro, R.J., "On the Predictability of Tax-Rate Changes," manuscript, University of Rochester 1981.

Breeden, D., "Consumption, Production, and Interest Rates: A Synthesis," *Journal of Financial Economics* 16, 1986, 3-39.

Christiano, L., M. Eichenbaum and D. Marshall, "*Permanent Income Model Revisited*," forthcoming in *Econometrica*, 1990.

Cox, J., J. Ingersoll and S. Ross, "A Theory of the Term Structure of Interest Rates," *Econometrica* 53, 1985, 385-408.

Duffie, D. and L. Epstein, "Stochastic Differential Utility and Asset Pricing," manuscript, Stanford University, 1989.

Dunn, K. D., and K. J. Singleton, "Modeling the Term Structure of Interest Rates Under Nonseparability of Preferences and Durability of Goods," *Journal of Financial Economics* 17, 1986, 27-55.

Epstein, L. and S. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica* 57, 1989a, 937-969.

Epstein, L. and S. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns II: An Empirical Anaysis," manuscript, University of Toronto, 1989b.

Fama, E., and K. French, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy* 96, 1988, 246-273.

Ferson, W., "Expectations of Real Interest Rates and Aggregate Consumption: Empirical Tests," *Journal of Financial and Quantitative Analysis* 18, 1983, 477-497.

Grossman, S., A. Melino and R. J. Shiller, "Estimating the Continuous Time Consumption-Based Asset Pricing Model," *Journal of Business and Economic Statistics* 5, 1987, 315-327.

Hall, R. E., "Intertemporal Substitution and Consumption," *Journal of Political Economy* 96, 1988, 339-357.

Hansen, L.P., "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50, 1982, 1029-1054.

Hansen, L.P., "A Method for Calculating Bounds on the Asymptotic Covariance Matrices of Generalized Method of Moments Estimators," *Journal of Econometrics*, 30, 1985, 203-238.

Hansen, L.P., "Semi-Parametric Efficiency Bounds for Linear Time Series Models," manuscript, University of Chicago, 1989.

Hansen, L.P., J. Heaton and M. Ogaki, "Efficiency Bounds Implied by Multiperiod Conditional Moment Restrictions," *Journal of the American Statistical Association* 83, 1988, 863-871.

Hansen, L.P., and R. J. Hodrick, "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy* 88, 1980, 829-863.

Hansen, L.P., and R. J. Hodrick, "Risk Averse Speculation in the Forward Foreign Exchange Market: an Econometric Analysis of Linear Models," in J. Frenkel ed., *Exchange Rates and International Macroeconomics*, University of Chicago Press, 1983.

Hansen, L.P. and T. J. Sargent, "Exact Linear Rational Expectations Models," forthcoming, 1990.

Hansen, L.P. and K.J. Singleton, "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica* 50, 1982, 1269-1286.

Hansen, L.P., and K.J. Singleton, "Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns," *Journal of Political Economy* 91, 1983, 249-265.

Hansen, L.P., and K.J. Singleton, "Computing Semi-parametric Efficiency Bounds for Linear Time Series Models," forthcoming in *Nonparametric and Seminonparametric Methods in Econometrics and Statistics*, eds. W. Barnett, J. Powell, and G. Tauchen, Cambridge University Press, 1989.

Harvey, Campbell, "The Real Term Structure of Consumption Growth," *Journal of Financial Economics* 22, 1988, 305-333.

Hayashi, F. and C. Sims, "Nearly Efficient Estimation of Time Series Models with Predetermined, but not Exogenous, Instruments," *Econometrica* 51, 1983, 783-798.

Heaton, J., "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications," manuscript, M.I.T., 1989.

Heaton, J. and M. Ogaki, "Efficiency Bound Calculations for a Time Series Model with Conditional Heteroskedasticity," Rochester Center for Economic Research Working Paper # 163, 1988.

Hillier, G., "On the Normalization of Structural Equations: Properties of Direction Estimators," Forthcoming in *Econometrica*, 1990.

Hindy, A. and C. Huang, "On Intemporal Preferences with a Continuous Time Dimension II: The Case of Uncertainty," manuscript, MIT, 1989.

Kocherlatoka, N., "Disintangling the Coefficient of Relative Risk Aversion from the Elasticity of Intermporal Substitution," *Journal of Finance* forthcoming, 1989.

Kreps, D., and E. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 1978, 185-200.

Monahan, J., "A Note on Enforcing Stationarity in Autoregressive-Moving Average Models," *Biometrika* 71, 1984, 403-404.

Stoica, P., T. Soderstrum, and B. Friedlander, "Optimal Instrumental Variable Estimates of the AR Parameters of an ARMA Process, *IEEE Transactions of Automatic Control*, 1985, 1066-1074.

Weil, P., "The Equity Premium Puzzle and the Risk-free Rate Puzzle," *Journal of Monetary Economics* 24, 1989, 401-421.

Working, H., "Note on the Correlation of First Differences of Averages in a Random Chain," *Econometrica* 28, 1960, 916-918.