Approximating the Asset Pricing Kernel

Author(s): David A. Chapman
Source: *The Journal of Finance*, Vol. 52, No. 4 (Sep., 1997), pp. 1383-1410
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2329440
Accessed: 25/11/2009 10:11

THE JOURNAL OF FINANCE • VOL. LII, NO. 4 • SEPTEMBER 1997

Approximating the Asset Pricing Kernel

DAVID A. CHAPMAN*

ABSTRACT

This article tests a simple consumption-based asset pricing model by approximating the true asset pricing kernel using low-order orthonormal polynomials based on the model's state variables. Approximated kernels based solely on next period's consumption growth are not rejected by overall measures of model fit, but they produce statistically and economically large pricing errors. Approximated kernels based on two quarters of future consumption growth and technology shocks have substantially improved overall fit. In particular, the best of these kernels are capable of eliminating the small firm effect.

A CONVENTIONAL APPROACH TO examining the implications of dynamic asset pricing models is to assume that markets are complete, the economy is in equilibrium, and that the unique state-price deflator can be interpreted as the intertemporal marginal rate of substitution of a representative consumer. Given assumptions about the form of this (unobservable) utility function and on the observability of aggregate consumption, the model can be tested by examining the conditional or unconditional covariation of estimated intertemporal marginal rates of substitution with measured asset returns.

The fact that this method of testing dynamic models requires strong assumptions is, in and of itself, not particularly troubling. However, the fact that these attempts have been almost uniformly inconsistent with the data on asset returns and consumption is problematic. For example, when the utility function is assumed to be time and state separable, the absolute and relative levels of the real returns on risk-free and risky securities (see, for example, Mehra and Prescott (1985) and Weil (1989)) are inconsistent with the model's predictions. The primary finding from a large literature testing consumption-based models is that measured consumption is too "smooth" to rationalize the observed level and variability of asset returns, for "reasonable" parameterizations of time-separable utility functions.

An alternate approach—pursued in this paper—is to approximate the state-price deflator or "asset pricing kernel" using orthonormal polynomials in a

* Finance Department, Graduate School of Business, The University of Texas at Austin. A portion of this work was completed while I was on leave at the William E. Simon Graduate School of Business Administration at the University of Rochester. I thank Tom Cooley, David Ellis, Ludger Hentschel, Burton Hollifield, John Long, Neil Pearson, Ehud Ronn, Bill Schwert, Jay Shanken, Dan Slesnick, Laura Starks, René Stulz (the editor), Stanley Zin, an anonymous referee, and participants in the Finance workshops at Texas A&M University, the University of British Columbia, the William E. Simon Graduate School of Business at the University of Rochester, and the University of Illinois at Urbana-Champaign.

small number of "state variable(s)." These state variables must be implied by a particular model of the underlying economy. As an example, this article examines state variables based on a stochastic version of the neoclassical growth model, which is used extensively in the literature on "real business cycles." This structure suggests aggregate consumption as a candidate state variable that can be used in constructing approximate kernels. Alternatively, the model implies that consumption is a time-invariant function of the level of the capital stock and transitory shocks to total factor productivity (or "technology shocks"), which suggests that technology shock growth rates *might* aid in the construction of an approximate kernel.

The linear in the polynomials structure of any of these approximations has a distinct advantage over other approximation schemes. The estimated pricing kernel is nonlinear in the underlying state variables, and therefore capable of pricing nonlinear payoffs, but it is linear in the approximating parameters. This implies that there is a closed-form expression for the approximating parameters, and the orders of the polynomials can be viewed as "factors" in a linear pricing model.

Approximating polynomials in aggregate consumption growth and a combination of consumption growth and temporary shock growth are tested using quarterly ex post real returns on Treasury bills, the first, fifth, and tenth size deciles of New York Stock Exchange/American Stock Exchange (NYSE/Amex) returns, and the returns to a portfolio of long-term corporate bonds. These returns are used both alone and in combination with a standard set of instrumental variables: dividend yields, credit spreads, and industrial production growth rates. In addition to traditional overidentifying restriction tests, the approximate kernels are examined using an asymptotically chi-squared statistic based on the mean square distance of the estimated kernel to the bounds introduced in Hansen and Jagannathan (1991). The polynomial approximations are also evaluated by testing the statistical significance of marginal polynomial orders and by examining the pricing errors on individual assets and groups of related assets, using both graphical evidence and Wald type tests.

Applying the approximate kernel approach produces the following results: kernels based on one-period ahead consumption growth alone are not typically rejected by overall measures of model fit, but they produce statistically and economically large pricing errors on individual assets. The largest improvement in the performance of the approximate kernels occurs when an additional future value of consumption growth and/or technology shock growth are added to the kernels. This is motivated by an appeal to time nonseparable preferences or the durability of consumption goods. The marginal power of shocks, in these cases, is difficult to detect directly, but the inclusion of shocks substantially improves the fit of the approximate kernels. In particular, this "nonseparable form" of the approximate kernel is the only specification capable of *consistently* reducing the average pricing errors on small market capitalization stocks to zero; i.e., there is no longer a "size effect" using this approximate kernel.

The results also provide strong evidence of the potential problems in estimating the polynomial parameters, emphasized in Cochrane (1996), associated with using the "optimal" covariance matrix for the pricing errors. If the covariance matrix is measured poorly, the resulting kernel can have undesirable properties in finite samples. In particular, the results show that a prespecified weighting matrix equal to the inverse of the second moment matrix of returns generally produces smaller pricing errors and kernels that are more nearly consistent with the Hansen-Jagannathan bounds.

This is not surprising, since the choice of this weighting matrix is motivated by the analysis in Hansen and Jagannathan (1997), which shows that parameters estimated using this weighting matrix minimize the mean-square distance between the estimated kernel and the Hansen-Jagannathan bounds. In the case of the kernels that include two future growth rates of consumption and shocks, the kernels estimated using the optimal weighting matrix and the inverse second moment matrix of returns have very similar overall performance. If the constant weighting matrix estimates are viewed as a "robustness check" on the optimal parameter estimates, only the nonseparable kernels pass the test.

The emphasis and results in this article can be contrasted with the approaches taken in related articles by Bansal and Viswanathan (1993) and Bansal, Hsieh, and Viswanathan (1993). Bansal and Viswanathan (1993) also approximate nonlinear asset pricing kernels. However, their approximate kernels are based on asset returns, rather than on the state variables of the economy. As such, their results nest the market-based capital asset pricing model (CAPM) and examine the importance of nonlinearities in returns data. The results in this study can be viewed as attempting a similar exercise for models based on macroeconomic aggregates. Bansal, Hsieh, and Viswanathan (1993) examine international pricing models using returns-based approximate kernels.

Another similarity to the Bansal and Viswanathan (1993) and the Bansal, Hsieh, and Viswanathan (1993) studies is the use of the generalized method of moments (GMM) to estimate the parameters of the approximating functions. An important difference between the current study and Bansal and Viswanathan (1993), however, is the use of orthonormal polynomials in the approximation, as opposed to neural nets. As noted earlier, there are substantial advantages, both in terms of computation and interpretation, that follow from the linearity (in the parameters) of the polynomial approximation. Bansal, Hsieh, and Viswanathan (1993) do use polynomial approximations, but they are based on returns and they are not orthogonalized or scaled, which create some numerical problems in solving for the approximation that are avoided by using orthonormal polynomials.

The remainder of the article is organized as follows: Section I describes a motivating example economy, and Section II discusses the exact nature of the approximation to the pricing kernel. The empirical methods are described in Section III, and the data used in all of the subsequent empirical analysis is contained in Section IV. Section V presents estimates of the approximations to the asset pricing kernel, and conclusions and suggestions for future research are contained in Section VI.

I. A Motivating Example

The stochastic neoclassical growth model is the standard "base case" in the literature on real business cycles.[1] In this simple, stylized structure, economic decisions are made by a continuum of identical households. In particular, there is a single consumption good which is produced using a constant-returns-to-scale production function that employs both capital and labor. The production technology is subject to a random shock to total factor productivity that evolves over time as a Markov process.

Under a standard set of assumptions, it is possible to represent the competitive equilibrium of this economy as the solution to a "social planner's problem."[2] In this case, the consumption, labor supply, and capital investment decisions can all be represented as continuous, time-invariant functions of the model's state variables. In general, it will not be possible to express these decision rules as closed-form functions of the state variables. If the households' utility functions are (additively) time-separable, then the model's state variables will be the current levels of the technology shock and the capital stock. If the utility function is *not* time-separable— due either to the direct utility effects of past consumption or to durability in consumption—then the state variables will include current and lagged values of the technology shock and the aggregate capital stock. For the current analysis, the important point is that aggregate consumption is a time-invariant function of a small number of variables.[3]

Since there are already a complete set of contingent claims markets in this economy (by assumption), the introduction of markets for financial assets does not change the equilibrium decision rules and quantity allocations. All securities are priced by assuming that a "representative household" solves a consumption-investment problem, with an exogenous endowment that is specified as the equilibrium aggregate consumption process. In other words, aggregate consumption serves as a single summary variable for all of the information about the state of the economy that is relevant for pricing financial assets.

Suppose, in particular, that there are n long-lived securities. If the utility function is additively separable over time, the familiar Euler equation for asset pricing, first presented in Leroy (1973) and Lucas (1978), will characterize the equilibrium asset prices:

$$\mathrm{E}\left[\beta \frac{u'(c_{t+1})}{u'(c_t)} R_{t,t+1} \middle| \Im_t\right] = 1 \tag{1}$$

where $u'(c_t)$ is the marginal utility of consumption at time t, β is a discount rate, $\Im_t$ represents information available at time t, and $R_{t,t+1}$ is the $n \times 1$ vector of gross returns. The (discounted) intertemporal marginal rate of sub-

[1] See Cooley and Prescott (1995) for an introduction to the real business cycle literature.

[2] See Stokey and Lucas (1989) for an extended, text book treatment of this subject.

[3] This is clearly *not* the only interesting specification of the aggregate economy! Rather, it is among the simplest specifications, and it suggests a candidate set of state variables that turn out to have interesting properties.

stitution is a familiar example of an "asset pricing kernel"; i.e., a strictly positive, adapted stochastic process $\{m_t\}_{t=0}^{\infty}$ that can be used to price all assets:

$$E[m_{t+1}R_{t+1}|\Im_t] = 1 \tag{2}$$

The same basic result holds even when the representative household's utility function is not additively separable over time. For example, the Euler equation in Ferson and Constantinides (1991) is

$$E\left[\sum_{i=1}^{\infty}\left(\frac{s_{t+i}}{s_t}\right)^{-\phi}\{b_{i-1}R_{j,t,t+1} - b_i\}\middle|\Im_t\right] = 1 \tag{3}$$

where $s_t = \sum_{i=0}^{\infty} b_\tau c_{t-i}$ accumulates the effect of past consumption expenditures on current utility, b_τ is a composite coefficient that reflects both "habit formation" effects from past consumption and the durability of prior consumption purchases, and ϕ is the household's coefficient of relative risk aversion. Given assumptions about the specific functional form for utility and the correlation between measured consumption and "true" consumption, the asset pricing equations of this model, either equation (1) or equation (3), are the standard Euler equations tested in consumption-based pricing models. Preference parameters can be estimated and the model can be evaluated using GMM, as in Hansen and Singleton (1982) or Ferson and Constantinides (1991).

II. Approximating the Pricing Kernel

As described earlier, the optimal consumption choice will be a function of the state variables in the model. The pricing implications of equation (2) are examined by approximating the kernel function as a continuous function of a set of observable variables. The Weierstrass theorem (see, for example, Section 4.2 in Hämmerlin and Hoffmann (1991)) states that an *infinite* order polynomial will provide an approximation that is arbitrarily close to a given continuous function, with a domain that is a closed, bounded subset of $\mathbb{R}^k$. A well-known problem with a polynomial approximation is that different terms in the polynomial can have strong linear relationships over relevant portions of their domain.[4] A more efficient approximation algorithm can be constructed using a set of orthonormal polynomials.

Given an inner product, denoted $\langle \cdot , \cdot \rangle$, defined on the space of continuous functions defined on some closed and bounded domain D, the orthonormal polynomials $\{P_k\}_{k=0}^{\infty}$ are defined as satisfying

$$\langle P_j, P_k\rangle = 0, \qquad \forall k \quad \text{and} \quad \forall j < k$$
$$\langle P_k, P_k\rangle = 1 \tag{4}$$

[4] For example, over the interval (0, 1) the two components of the second-order polynomial $p(x) = x + x^2$ are both increasing.

For example, if $D = [-1, 1] \subset \mathbb{R}$ and the inner product is defined by the weighting function $w(x) = 1$, then the Legendre polynomials

$$\mathrm{L}_k(x) = \frac{1}{2^k k!}\sqrt{\frac{2k+1}{2}}\,\frac{\mathrm{d}^k(x^2-1)^k}{\mathrm{d}x^k}$$

form a set of orthonormal polynomials.[5]

The important question becomes: what inner product should be used to form the set of orthonormal polynomials? A choice that will ensure that the polynomials are orthogonal with respect to the state variable that determines the kernel is to define the inner product using the (unconditional) density function of the underlying state variable. The polynomials are defined by

$$P_k(x) = \varphi_{k,1} + \varphi_{k,2}x + \cdots + \varphi_{k,k+1}x^k \tag{5}$$

where the parameters $\{\varphi_{k1}, \ldots, \varphi_{k,k+1}\}$ satisfy the following set of equations

$$\begin{array}{rcl} \varphi_{k,1} + \varphi_{k,2}\mu_1 + \cdots + \varphi_{k,k+1}\mu_k & = & 0 \\ \vdots & & \vdots \\ \varphi_{k,1}\xi_{k-1,1} + \varphi_{k,2}\xi_{k-1,2} + \cdots + \varphi_{k,k+1}\xi_{k-1,k+1} & = & 0 \\ \mathrm{E}(P_k(x)^2) & = & 1 \end{array} \tag{6}$$

where μ_k is the kth noncentral moment of x and

$$\xi_{ij} \equiv \varphi_{i-1,1}\mu_{j-1} + \cdots + \varphi_{i-1,k}\mu_{j-1+i} \tag{7}$$

These equations can be solved by defining each $\varphi_{k,j}$ as a linear function of $\varphi_{k,k+1}$ and then solving the last equation in (6) for $\varphi_{k,k+1}$. The approximation of the kernel can be constructed using the orthonormal polynomials, since any arbitrary polynomial can be defined from these basis functions.

The second practical problem in using a polynomial approximation is the appropriate choice of the maximum order of the approximating polynomial, or (equivalently) the magnitude of the approximation error in any particular application. As Judd (1992) notes, "The best choice of (polynomial order) cannot be determined a priori. Generally, the only "correct" choice is (infinity)." The question examined below is: Can a *low order* polynomial approximation provide a reasonable approximation to the kernel that prices measured returns? The goal is to see if this approach can provide a general mapping from consumption to asset prices. This issue is analogous to general specification error in choosing the form of the representative agent's utility function. The

[5] The Chebyshev polynomials, used extensively in Judd (1992), provide another example. These polynomials, defined over $x \in [-1, 1]$ by $T_k(x) = \cos k\ (\arccos x)$ are orthogonal (although not normalized to one) with respect to the inner product defined by the weighting function $w(x) = 1/\sqrt{1 - x^2}$.

formal error magnitude is not important if the approximation can consistently price assets. Approximations of the form

$$\hat{m}(x_t,\ \theta) = \sum_{j=0}^{k} \theta_j P_j(x_t) \equiv \mathbb{P}_t'\theta \tag{8}$$

will be examined, for small k, where $\mathbb{P}_t$ is the $(k+1) \times 1$ vector of orthonormal polynomials of order 0 to k, evaluated at x_t.

A substantial advantage of using a polynomial approximation of the form of equation (8) is that the approximate kernel, while nonlinear in the state variables, is linear in the parameters of the approximating polynomials. This implies that the kernel has the general ability to price nonlinear payoffs, as in Bansal and Viswanathan (1993), but the pricing equations have familiar linear interpretations.[6] Bansal, Hsieh, and Viswanathan (1993) also use a polynomial approximation to the asset pricing kernel. The form of their approximation has two important differences from the versions of the orthonormal expansion tested in the following section. Their approximation does not orthogonalize or rescale the monomials used to form the expansion. This results in problems of numerical stability of the general linear approximation; i.e., collinearity in the terms of their expansion. More fundamentally, their approximation is based on asset returns, rather than aggregate quantities or other state variables.

Consider the case of a kth order approximating polynomial in a single state variable. For an individual security j, equation (2) can be written as

$$\mathrm{E}[\hat{m}(x_{t+1},\ \theta)R_{j,t,t+1}|x_t] = 1 \tag{9}$$

where $\hat{m}(x;\ \theta)$ is defined in equation (8). Following Cochrane (1996), and letting $\hat{m}_{t+1} = \hat{m}(x;\theta) = \theta_0 + \mathbb{P}_{t+1}'\theta$ (where the mean levels of $\mathbb{P}_{t+1}'$ are reflected in θ_0) equation (9) for an individual return R_j can be written as:

$$\begin{aligned}
\mathrm{E}_t[R_{j,t,t+1}] &= (1/\mathrm{E}_t(\hat{m}_{t+1})) - (\mathrm{cov}_t(\hat{m}_{t+1},\ R_{j,t+1})/\mathrm{E}_t(\hat{m}_{t+1})) \\
&= (1/\theta_0) - (\mathrm{E}_t(\mathbb{P}_{t+1}'\theta R_{j,t+1})/\theta_0) \\
&= (1/\theta_0) - [\theta'\mathrm{E}_t(\mathbb{P}_{t+1}\mathbb{P}_{t+1}')E_t(\mathbb{P}_{t+1}\mathbb{P}_{t+1}')^{-1}\mathrm{E}_t(\mathbb{P}_{t+1}'R_{j,t+1})]/\theta_0 \\
&= (1/\theta_0) - [\theta'\mathrm{E}_t(\mathbb{P}_{t+1}\mathbb{P}_{t+1}')\beta_{t,P,R_j}]/\theta_0 \\
&= \lambda_0 + \lambda_{1t}\beta_{t,P,R_j}
\end{aligned} \tag{10}$$

where $\beta_{t,P,R_j} = \mathrm{E}_t(\mathbb{P}_{t+1}\mathbb{P}_{t+1}')^{-1}\ \mathrm{E}_t(\mathbb{P}_{t+1}'R_{j,t+1})$ and $\mathrm{E}_t[\cdot]$ is a shorthand notation for $\mathrm{E}[\ \cdot \mid x_t]$.

The simplest examples of approximating the pricing kernel are based directly on aggregate consumption. In a continuous-time economy with complete markets, Breeden (1979) and Duffie and Zame (1989) demonstrate that all assets obey a consumption capital asset pricing model (CCAPM) in which

[6] Bansal and Viswanathan (1993) base their approximations on neural networks applied in return space.

equity premiums are determined by the (instantaneous) covariance of a security's return with the (instantaneous) growth rate in consumption (i.e., its "consumption beta"). In the case of a first-order approximating polynomial in consumption growth, equation (10) simplifies to the familiar expression

$$\mathrm{E}_t[R_{j,t,t+1}] = \lambda_0 + \lambda_{1t}\beta_{t,c,R_j} \tag{11}$$

which states that a security's return is a linear function of its consumption beta. These continuous-time derivations do not carry over in an exact manner to a general discrete-time formulation of the problem. Higher order approximations in consumption growth will lead to a multi-beta version of equation (11), in which expected returns depend on betas with respect to higher order orthonormal polynomials in consumption growth. These additional terms capture the impact of the nonlinearities in the pricing kernel.

The effect of time nonseparability on the pricing kernel is modeled by adding a future value (or values) of consumption growth as an additional state variable. The orthonormal polynomials in this case are defined using the 2-fold tensor product of the one dimensional polynomials, as defined in Judd (1992):

$$\hat{m}(x_{1t}, x_{2t}) = \sum_{j=0}^{k} \sum_{l=0}^{k} \theta_{jl} P_j(x_{1t}) P_l(x_{2t}) \tag{12}$$

where x_{1t} is consumption growth from $t - 1$ to t and x_{2t} is consumption growth from t to $t + 1$. Kernels in more than two state variables can be constructed in a similar manner.

A final form of approximation to the pricing kernel—which can be combined with either the time separable or time non-separable utility specification—includes a direct measure of aggregate technology shocks. In principle, true consumption already contains all of the relevant information in shocks that can be used in pricing assets. However, if measured consumption falls short of this ideal, there may be some additional information conveyed through the direct inclusion of shocks. Of course, if the empirical measurements of the shocks are no more accurate than the measurement of consumption, this approach may have no clear practical advantage over using flawed consumption data alone.

III. Estimating Approximate Kernels

The parameters of the polynomial approximation are estimated using generalized method of moments (GMM), as in Bansal and Viswanathan (1993) and Bansal, Hsieh, and Viswanathan (1993). The orthogonality conditions are:

$$\mathrm{E}_t[(R_{t+1} \otimes Z_t) m_{t+1}] = \mathbf{1} \otimes Z_t \tag{13}$$

where Z_t is an l-dimensional vector of elements in the information set available to investors and "$\otimes$" denotes the Kronecker product operator.[7] Applying the law of iterated expectations to equation (13) yields

$$\mathrm{E}[(R_{t+1} \otimes Z_t)m_{t+1}] = \mathrm{E}[\mathbf{1} \otimes Z_t] \tag{14}$$

The sample counterpart to equation (14) for an individual security i is:

$$g_{i,T}(x, R, Z; \theta) = T^{-1} \sum_{t=1}^{T} [(R_{i,t+1} \otimes Z_t)\mathbb{P}'_{t+1}\theta - \mathbf{1} \otimes Z_t] \tag{15}$$

As Cochrane (1996) notes, equation (14) is an implication of (13), and if (14) holds for *all* choices of Z, it implies equation (13). Individual elements of $R \otimes Z$, say $R_{i,t+1} \cdot Z_{j,t}$ have the interpretation as returns to "managed portfolios," where the return is generated by investing in asset i according to the observable signal Z_j. For notational convenience, let $R_z = R \otimes Z$ and let $\mathbf{1}_z = \mathbf{1} \otimes Z$.

The objective function for the GMM estimation problem is

$$J(\theta) = g_T(x, R, Z; \theta)' W_T g_T(x, R, Z; \theta) \tag{16}$$

where $g_T(\cdot)$ is the $(nl \times 1)$-vector of sample orthogonality conditions and W_T is an $(nl \times nl)$-matrix. Hansen (1982) proves that the weighting matrix that produces the efficient GMM estimator is the inverse of the long-run covariance matrix of the model disturbances.[8] Using the linear (in the parameters) structure, the GMM estimator is:

$$\hat{\theta}_{\text{gmm}} = [D'_T W_T D_T]^{-1} D'_T W_T \left(T^{-1} \sum_{t=1}^{T} 1'_{z,t} \right) \tag{17}$$

where $D_T \equiv T^{-1} \Sigma_{t=1}^{T} [R'_{z,t+1}\ \mathbb{P}_{t+1}]$, $R_{z,t+1} = R_{t+1} \otimes Z_t$, and $1_{z,t} = 1 \otimes Z_t$. The weighting matrix, W_T, is determined by iterating on the solution to equation (16) until the parameter estimates converge, which implies that it will be a

[7] Note that the elements of the polynomial approximation are not scaled by the instrumental variables. As Shanken (1991) and Cochrane (1996) note, this would be equivalent to permitting the θ coefficients to be time-varying functions of the instruments. The reason for not scaling in the current context is to evaluate the performance of the unconditional polynomial approximations. Time-varying polynomial approximations constitute an important area for future research.

[8] Formally, let $u_t = g(x_t, Z_t, R_t; \theta)$ be the nl-vector of disturbances, then the optimal weighting matrix is:

$$W^* = \lim_{J \to \infty} \sum_{\tau=-J}^{J} \mathrm{E}(u_t u'_{t+\tau})$$

This matrix can be estimated consistently while allowing for arbitrary serial correlation in the disturbances u_t using a general "heteroskedasticity and autocorrelation consistent" estimator. The results presented below are based on the quadratic spectral kernel proposed in Andrews (1991) and Andrews and Monohan (1992). Ogaki (1993) provides a general discussion of the application of kernel estimators in this context.

function of the point estimates $\hat{\theta}_{\text{gmm}}$.[9] The estimator defined in equation (17) is a consistent estimator of the true parameters, θ^*, and it has an asymptotic normal distribution with

$$\text{var}(\hat{\theta}_{\text{gmm}}) = T^{-1}(D'_T W_T D_T)^{-1} \tag{18}$$

For a proof of this result, see Hansen (1982).

The sample orthogonality conditions (15) can be interpreted as the pricing errors that are obtained using the approximate kernel. If the model provides a good description of the data, then these pricing errors should be close to zero and the minimized value of equation (16) will be small. Hansen (1982) shows that $J_T \equiv T \cdot J\,(\hat{\theta}_{\text{gmm}})$ has a χ^2 distribution with $q - p$ degrees of freedom, where $q = nl$ and $p = \dim(\theta)$. This statistic is commonly referred to as the "*J* test" or as a test of the model's *"overidentifying restrictions."* The J_T statistic tests the magnitude of a weighted-average of all of the pricing errors. Notice, however, that there are two ways to get a small value of the J_T statistic: generate small pricing errors with a high degree of precision or generate large pricing errors with even larger standard deviations of those errors. In the empirical work presented in the following section, it will be important to distinguish between these two situations.

An alternate test of the overall adequacy of the approximate kernel can be constructed using the results in Hansen and Jagannathan (1991) and Hansen and Jagannathan (1997). Given a set of asset returns, Hansen and Jagannathan (1991) develop a bound on the mean and standard deviation that must be satisfied by any candidate pricing kernel. Hansen and Jagannathan (1994), applied to the case at hand, demonstrate that the minimum mean-square distance from any given kernel to the permissible bounds is the solution to the following problem:

$$J^{\text{HJ}}(\theta) = \min_{\theta} g_T(x, R, Z; \theta)' E[R_z R'_z]^{-1} g_T(x, R, Z; \theta) \tag{19}$$

Cochrane (1996), using results in Hansen (1982), demonstrates that the covariance matrix of the pricing errors generated by the parameters $\hat{\theta}_{\text{HJ}}$ that solve equation (19) is:

$$\Sigma_g^{\text{HJ}} = T^{-1}[I_{nl} - D_T(D'_T \Re D_T)^{-1} D'_T \Re] W_T [I_{nl} - D_T(D'_T \Re D_T)^{-1} D'_T \Re]' \tag{20}$$

where $\Re = [T^{-1} \Sigma_{t=1}^{T} R_{zt} R'_{zt}]^{-1}$. Under the null hypothesis that the pricing errors are zero (i.e., that the model is well-specified), the minimized value of $J^{\text{HJ}}(\hat{\theta}_{\text{HJ}})$ can be written as:

$$g'_{HJ}[\Sigma_g^{\text{HJ}}]^{-1} g_{HJ} \sim \chi^2_{q-p} \tag{21}$$

[9] Convergence of the parameter estimates was judged to have occurred when the largest (absolute) percentage change in any parameter was 1 percent or the number of iterations exceeded 20.

By the construction of the GMM estimator, Σ_g^{HJ} is a singular matrix, but equation (21) can be calculated using the pseudo-inverse of Σ_g^{HJ}.[10,11]

Cochrane (1996) argues that the iterated GMM estimator using the optimal weighting matrix may present two important, and related, practical problems for the evaluation of asset pricing models. First, if the covariance matrix for the iterated GMM estimator is poorly measured, then the estimator will put too much weight on moments that (spuriously) appear to be measured precisely. Second, given the high correlation of asset returns, the iterated estimator may place too much weight on portfolios that are economically uninteresting, in the sense that they are composed of extreme short and long positions in some of the assets.[12]

The covariance matrix for the pricing errors, based on GMM estimation with the optimal weighting matrix, can be used to test the significance of individual pricing errors or groups of pricing errors using Wald type test statistics. This matrix is:

$$\Sigma_g^{GMM} = T^{-1}[W_T^{-1} - D_T(D'_T W_T D_T)^{-1} D'_T] \tag{22}$$

For example, consider the case of five assets ($n = 5$) and four instruments ($l = 4$). The hypothesis:

$$H_0 \colon g_{11} = g_{12} = g_{13} = g_{14} = 0$$

tests whether or not the pricing errors associated with the first security and its managed portfolios are zero. If the matrix C is defined as

$$\underset{4\times 20}{C} = \left[I_4 \; \vdots \; \underset{4\times 16}{0} \right]$$

then H_0 can be written as $C_{g_T} = 0$ and the statistic that tests it is

$$g'_T[C\ \Sigma_g^{GMM}\ C']^{-1} g_T \sim \chi_q^2 \tag{23}$$

where, in this case, q is equal to four, which is the number of linear restrictions.

[10] See Cochrane (1996), footnote 7, for one method of constructing the pseudo-inverse based on an eigenvalue decomposition of Σ_g^{HJ}. In the section that follows, the inverse was calculated using the command INVSWP in Gauss, which computes a "generalized sweep inverse."

[11] The distribution of the J^{HJ} (θ) statistic under the null hypothesis of a correctly specified model is quite straightforward. Hansen and Jagannathan (1994) and Hansen, Heaton, and Luttmer (1995) solve the more difficult problem of the distribution of J^{HJ} (θ) under the null of *misspecification;* i.e., that the estimated kernel is actually some distance (in a mean-square sense) ε from the Hansen-Jagannathan bounds. This statistic has an asymptotic normal distribution (with a variance related to ε^{-2}). These results permit an explicit discussion of the power of the J^{HJ} (θ) test.

[12] Cochrane (1996) recommends using the results of the initial iteration of the GMM estimator (i.e., OLS estimates with a covariance matrix that is corrected for possible heteroskedasticity and autocorrelation of the residuals) as a check on the robustness of the iterated estimates. Results using the OLS estimates are available upon request.

The marginal significance of each polynomial order can also be assessed using the Wald type tests described above, combined with the covariance matrix for $\hat{\theta}_{gmm}$ in equation (18). In the case of a single state variable, these tests are asymptotically identical to t-tests of individual parameters. When the kernel is based on two state variables, these will be tests of multiple coefficient restrictions. For example, with two state variables a test of the first-order versus the second order approximations (using equation (12)) would test the joint significance of three parameters, while an increase from the second-order to the third-order would test five parameters.[13]

IV. The Data

The data consist of observations on consumption, investment, asset returns, technology shocks, and a number of instruments used as conditioning information. All series are observed at the quarterly frequency. The consumption data are real "personal consumption expenditures" on nondurable goods, taken from the Citibase directory. These series are converted to a per capita basis using "Residential Population." In order to identify the temporary technology shock, the aggregate production function is assumed to have a Cobb–Douglas form:

$$Y_t = \exp(z_t) K_t^{\omega} H_t^{1-\omega} \tag{24}$$

where z_t is the temporary shock, ω is capital's share of output, and $(1 - \omega)$ is labor's share. The temporary shock is identified as the "Solow residual," and it is calculated from the following difference relation

$$z_t - z_{t-1} = \Delta \log Y_t - (1 - \omega) \cdot \Delta \log H_t; \qquad z_0 = 0 \tag{25}$$

for a given choice of ω where Δ is the first difference operator ($\Delta x_t = x_t - x_{t-1}$). Given real gross national product and average weekly hours of production workers from the Household Survey, the Solow residual is defined by setting ω equal to 0.4, consistent with the arguments in Cooley and Prescott (1995).

Summary statistics for the consumption growth, (real) value-weighted returns, and technology shock growth series are contained in Panel A of Table I.[14] The empirical measure of the temporary technology shock is seven times as variable as consumption, and it is extremely persistent. This suggests that approximate kernels based on technology shocks should have qualitatively different properties than pure consumption-based kernels.

The asset return series are all from the Center for Research in Security Prices (CRSP). They consist of the one-month holding period return on three month Treasury bills, a portfolio of high grade corporate bonds, a portfolio of long-term government bonds (these last two returns are the Ibbotson bond data

[13] Data limitations required a reduction in the total number of parameters estimated, relative to the number of terms in equation (12). Therefore, cross-product terms were omitted in the approximations, which explains the degrees of freedom for the tests reported in the tables.

[14] Approximate kernels based on consumption expenditures on "services" are examined in Chapman (1996). The results are not substantially different from the ones reported below.

Table I

Descriptive Statistics: 1954:1 to 1991:4

This table contains descriptive statistics for potential state variables, asset returns, and instruments. Panel A presents statistics for state variables. *Consumption* is defined as the growth rate of real personal consumption expenditures (PCE) on nondurables divided by the residential population. *Temporary Shock* is the growth rate of the Solow residual. Panel B presents statistics for inflation and ex post real returns. *Inflation* is the logarithmic first difference in the Consumer Price Index (Urban Consumers, seasonally adjusted). *Tbill3M* is the ex post real one month holding period return on a three month Treasury bill. All real returns are formed by subtracting ex post inflation from the nominal return. *Govt. Bonds* is the ex post real one month holding period return from a portfolio of long-term government bonds. *Corp. Bonds* is the ex post real one month holding period return on a portfolio of high-grade corporate bonds. *Decile x* is the ex post real one month return to an equal-weighted portfolio of stocks from decile x of New York Stock Exchange/American Stock Exchange (NYSE/Amex) stocks, with *Decile 1* being the smallest decile and *Decile 10* being the largest decile. Quarterly series are defined from the monthly series using the formula $(1 + R_{t,t+3}) = \Pi_{j=1}^{3} (1 + R_{t+j-1,t+j})$. All series are expressed at a quarterly rate, not in percent. Panel C contains statistics for the instruments used in the generalized method of moments (GMM) estimation. *DyS&P* is the dividend yield on the S&P 500 Composite Index. *CrdSpd* is the yield spread on Baa rated versus Aaa rated corporate bonds. *Ipgr* is the logarithmic twelfth difference in the Federal Reserve Board's index of total industrial production. It is expressed at an annual rate, not in percentage. The quarterly series are constructed as simple averages of the months within the quarter, and all series are expressed at an annual rate, not in percentage.

Series	Mean	Std. Dev.	Autocorrelations 1	2	3	4	8	12
Panel A: State Variables								
Consumption	0.003	0.007	0.30	0.07	0.16	0.02	−0.17	0.01
Temporary shock	0.010	0.050	0.17	0.01	−0.16	−0.16	0.02	−0.05
Panel B: Inflation and Real Asset Returns								
Inflation	0.011	0.009	0.75	0.71	0.72	0.60	0.33	0.27
Tbill3M	0.004	0.007	0.61	0.60	0.59	0.50	0.28	0.14
Govt. bonds	0.004	0.054	−0.04	0.07	0.15	0.08	−0.08	−0.02
Corp. bonds	0.005	0.052	0.02	0.06	0.16	0.07	−0.06	−0.03
Decile 1	0.041	0.176	−0.02	−0.08	−0.12	0.29	0.07	0.21
Decile 3	0.032	0.136	0.03	−0.06	−0.07	0.12	−0.02	0.01
Decile 5	0.028	0.123	0.03	−0.08	−0.09	0.07	−0.05	0.08
Decile 7	0.028	0.108	0.05	−0.11	−0.07	−0.00	−0.04	0.05
Decile 10	0.018	0.080	0.14	−0.08	−0.05	−0.01	−0.06	0.01
Panel C: Instrumental Variables								
DyS&P	0.039	0.008	0.93	0.84	0.76	0.69	0.56	0.51
CrdSpd	0.084	0.038	0.92	0.84	0.76	0.68	0.46	0.31
Ipgr	0.031	0.056	0.83	0.50	0.16	−0.12	−0.26	−0.03

series), and the monthly returns on ten size portfolios of NYSE/Amex stocks, where "Decile 1" contains the smallest firms and "Decile 10" contains the largest firms. Nominal returns are converted to ex post real returns by subtracting inflation, measured as the logarithmic first difference in the Consumer Price

Index (Urban, seasonally adjusted). (Real) Return series are converted to the quarterly frequency using the formula $(1 + R_{t,t+3}) = \Pi_{j=1}^{3} (1 + R_{t+j-1,t+j})$.

Panel B of Table I presents some basic summary statistics for these return series. The average real returns on the bond series are all slightly positive and approximately equal to each other. The long-term government bonds and the corporate bonds have return variances that are almost ten times larger, on average, than the variance of one-month return on three-month Treasury bills. As expected, the real returns on small stocks are both higher and more variable than the returns on large stocks. It is worth noting the large autocorrelation coefficient in the returns of the smallest decile stocks at the annual frequency. This is a manifestation of the well-documented "January effect," which is still evident in the quarterly data.

The contemporaneous correlation (not reported) between adjacent stock decile returns and the returns on the two long-term bond returns are in excess of 0.9. This suggests that only a few assets can be used in testing the model. Including additional returns provides little new information and increases the numerical instability of inversions of the second moment matrix of returns. The analysis in the following section will be based on the (quarterly) real returns to three month Treasury bills, corporate bonds, and Decile 1, 5, and 10 stocks. The use of these assets has two motivations. They have (comparatively) low contemporaneous correlation, and they include (simultaneously) small stock returns and the returns on bills and bonds. Traditionally, this has constituted the most difficult test of any asset pricing model.[15]

The instrumental variables used in the analysis are similar to a subset of the ones used in Ferson and Constantinides (1991). They are the logarithmic twelfth difference in the Federal Reserve Board's monthly index of total industrial production (*Ipgr*), the difference in yields on Baa versus Aaa corporate debt (*CrdSpd*), and the dividend yield on the S&P 500 Composite index (*DyS&P*). All of these series are from Citibase. The three series were chosen to reflect variations in the stock market, bond market, and the real economy and they have been used in numerous empirical studies on the time series properties of asset returns. Summary statistics for the instrumental variables are given in Panel C of Table I.

V. Approximating the Asset Pricing Kernel: Evidence

The evidence in Tables II through IV and Figures 1 and 2 concerns kernels that are based on next period's growth rate in aggregate consumption. The parameter estimates for consumption growth are significantly different from zero (at the 0.10 and 0.05 confidence levels, respectively). However, the parameters associated with the higher order consumption terms are all individually indistinguishable from zero. The $J_T(i)$ statistics, for $i = 1, \ldots, 4$, all fail to reject the models' overidentifying restrictions. While examining this statis-

[15] See, for example, Hansen and Jagannathan (1991) and Fama (1992).

Table II

Iterated GMM Estimates of Approximate Kernels Based on One-Period-Ahead Consumption Growth: 1954:1 to 1991:4

This table contains the estimates from approximate kernels based on the growth rate in real per capita personal consumption expenditures on nondurables. *, **, and *** indicate that the coefficient is significant at the 0.10, 0.05, and 0.01 confidence level, respectively, using a Wald type test for the individual coefficient. $J_T(i)$ refers to the conventional test of the model's overidentifying restrictions, and $J^{HJ}(i)$ refers to the test of average pricing errors using the weighting matrix $W = T^{-1} \Sigma_{t=1}^{T} (R_{zt}R'_{zt})^{-1}$, both statistics are for the ith order approximation. Both statistics have asymptotic χ^2 distributions with degrees of freedom equal to the number of securities multiplied by the number of instruments less the number of parameters in the approximate kernel. The p-values are reported in parentheses beneath each statistic. The degrees of freedom associated with each pair of statistics are reported in the column labeled *df*.

Approx.	Polynomial Parameters					Tests of Overall Fit		
Order	θ_0	θ_1	θ_2	θ_3	θ_4	$J_T(i)$	$J^{HJ}(i)$	*df*
1	0.932***	−0.106*				20.529	19.781	18
						(0.304)	(0.345)	
2	0.895***	−0.496***	−0.307*			19.604	18.320	17
						(0.295)	(0.369)	
3	0.989***	−0.246	0.148	0.370		15.955	16.517	16
						(0.456)	(0.417)	
4	0.990***	−0.244	0.155	0.382	0.024	15.823	17.452	15
						(0.394)	(0.293)	

tic is a conventional method of model evaluation, an important question addressed below is: Exactly what information is contained in this "failure to reject?"

Figure 1 shows the Hansen-Jagannathan volatility bounds (hereafter "HJ bounds") on a valid asset pricing kernel implied by the set of five original assets, augmented by the fifteen managed portfolios, formed using the instrumental variables. It also shows the point estimates of the means and standard deviations of two sets of kernels: kernels based on the iterated GMM estimates (hereafter, "iterated kernels") and kernels based on the parameters estimated by using the inverse of the second moment matrix of returns as the weighting matrix (hereafter, "HJ kernels").

The consumption-based kernel shows substantial differences in the properties of iterated kernels and HJ kernels. In particular, the HJ kernels are generally inside the HJ bounds, and their standard deviations increase (uniformly) with the order of the polynomial approximation. In contrast, the iterated kernels are never inside the HJ bounds. These results provide strong support for the contention in Cochrane (1996) that poor estimation of the covariance matrix of the GMM errors can result in perverse behavior of the "optimal" estimator. More generally, the optimal GMM estimators are designed to minimize the sampling error of the parameter estimates of the polynomial approximation. They are not designed to minimize the pricing errors associated with the original assets.

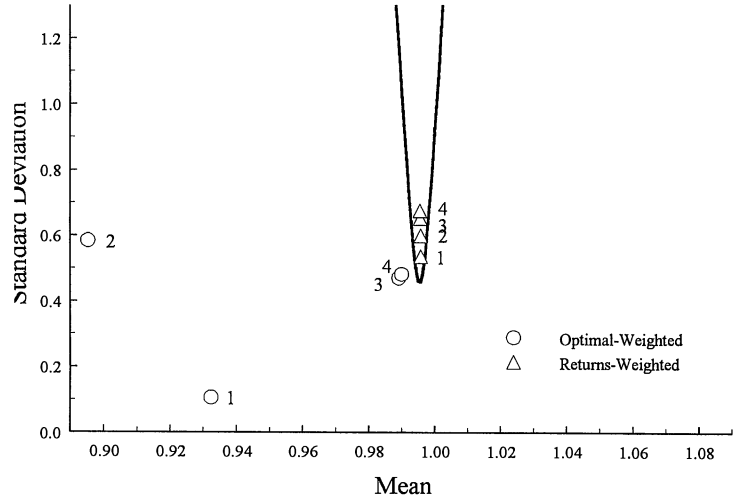

Figure 1. The Hansen-Jagannathan volatility bounds and approximate kernels based on a time-separable consumption model. This figure shows both the volatility bounds based on the real returns to Treasury bills, size-based stock portfolios, and corporate bonds; and the point estimates of the mean and standard deviation of approximate kernels based solely on real, per capita, personal consumption expenditures on nondurables. The numbers next to each of the point estimates refer to the order of the approximating polynomial. The order runs from 1 to 4. "Optimal-Weighted" refers to point estimates of the mean and standard deviation of the approximate kernels based on the parameters estimated using the optimal GMM weighting matrix. "Returns-Weighted" refers to point estimates of the mean and standard deviation of the approximate kernels based on the parameters estimated using the inverse of the sample estimate of the second moment matrix of returns as the weighting matrix in the GMM estimation.

The $J^{\mathrm{HJ}}(i)$ statistics, for $i = 1, \ldots, 4$, test the null hypotheses that the HJ kernels actually do lie inside the HJ bounds.[16] The null hypothesis is generally not rejected at conventional levels of significance, but the point estimates of the test statistics seem large relative to the impressions given by Figure 1. This suggests that the asymptotic distributions of these statistics may not provide a reasonable approximation to their small sample counterparts.

In order to gain more insight into the performance of the different kernels, Figure 2 and Tables III and IV examine the average pricing errors from the different approximate kernels based on one-period ahead consumption growth. The pricing errors reported in the figures are the average quarterly return error divided by the expected (gross) return under the model. There are twenty entries in each panel, corresponding to the original asset and the three managed portfolios associated with it. Errors from both the optimal GMM kernel and the HJ

[16] The power of these tests to detect deviations from the null of correct pricing is low. Using the (asymptotic) results in Hansen and Jagannathan (1994) and Hansen, Heaton, and Luttmer (1995), the null of a *misspecified* model is also not rejected for any of the pure consumption kernels. In other words, the point estimates of the distances shown in Figure 1 are found to be significantly different from zero under the alternative of misspecification.

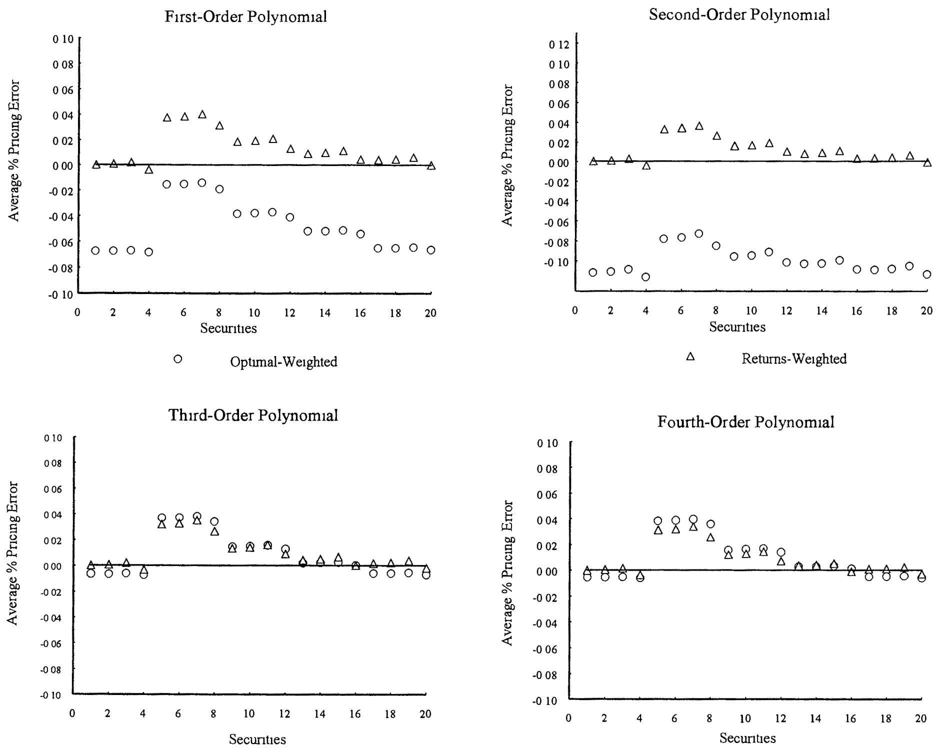

Figure 2. Average pricing errors for approximate kernels based on a time-separable consumption model. This figure shows the average pricing error from approximate kernels based solely on one-period ahead consumption growth for each of the four "managed portfolios" based on the five primary assets. The primary assets are the real returns to Treasury bills, three size-based stock portfolios, and corporate bonds. The managed portfolio returns are formed by multiplying each original asset by a constant and the other instrumental variables. The instrumental variables are the lagged dividend yield on the S&P 500, the lagged yield spread between Baa and Aaa corporate debt, and lagged industrial production growth. In all, there are twenty average pricing errors associated with each order of approximation and with each weighting method for generating parameter estimates. The first four entries are the T-bill returns and the three managed portfolios associated with T-bills, the next are for the smallest decile stocks, followed by the middle decile stocks, the largest decile stocks, and corporate bond returns. "Optimal-Weighted" refers to the pricing errors associated with the parameter estimates generated using the optimal GMM weighting matrix. "Returns-Weighted" refers to the pricing errors associated with the parameter estimates generated using the inverse of the sample estimate of the second moment matrix of returns as the weighting matrix in the GMM estimation.

kernel are reported for each of the approximations from order one to four. For each type of kernel, the first four entries are the T-bill returns and the three managed portfolios associated with T-bills, the next are for the smallest decile stocks, followed by the middle decile stocks, the largest decile stocks, and corporate bond returns. For the iterated kernels, there is at least one order of approximation that produces average pricing errors that are in excess of five percent of gross expected returns. These average errors are clearly economically significant (if the point

Table III

Testing the Statistical Significance of Pricing Errors from Kernels Based on One-Period Ahead Consumption Growth Using the Optimal GMM Weighting Matrix: 1954:1 to 1991:4

This table examines the pricing errors for kernels based on the growth rate in real per capita personal consumption expenditures on nondurables. *Tbill3M* is the ex post real one month holding period return on a three month Treasury bill. *Corp. Bonds* is the ex post real one month holding period return on a portfolio of high-grade corporate bonds. *Decile x* is the ex post real one month return to an equal-weighted portfolio of stocks from decile x of New York Stock Exchange/American Stock Exchange (NYSE/Amex) stocks, with *Decile 1* being the smallest decile and *Decile 10* being the largest decile. All real returns are formed by subtracting ex post inflation from the nominal return. Each entry in the body of the table is the Wald-type test for the null hypothesis that the pricing errors for the given asset and the managed portfolios based on the asset and the instruments are all zero. Each statistic is distributed χ^2_4, and the p-value is reported in parentheses beneath the value of the statistic.

	Approximation Order			
Asset Class	1	2	3	4
Tbill3M	17.961	5.659	0.727	2.207
	(0.001)	(0.226)	(0.947)	(0.698)
Decile 1	18.574	6.107	4.534	4.796
	(0.001)	(0.191)	(0.338)	(0.309)
Decile 5	19.097	5.196	3.719	6.190
	(0.001)	(0.268)	(0.445)	(0.185)
Decile 10	19.842	5.155	2.042	5.432
	(0.001)	(0.272)	(0.728)	(0.246)
Corp. bonds	17.348	5.436	0.786	2.248
	(0.002)	(0.245)	(0.940)	(0.690)

estimates are accurate). Generally speaking, there is no clear reduction in the magnitude of the average errors for higher order approximations.

The magnitudes of the errors in Figure 2 are consistent with the mean kernel values shown in Figure 1. For example, the second order optimal kernel has a mean that is far too small to be consistent with the HJ bounds, and the individual pricing errors are, predictably, negative. Surprisingly, the errors for the HJ kernels do *not* generally offer a substantial improvement over the iterated kernels, *despite lying in the interior of the HJ bounds.* The exception to this general result occurs in cases where the iterated kernels have means that are extremely large or small. It is also worth noting that for all of the kernels that are even close to producing small pricing errors, the errors associated with the smallest decile firms are the largest.

Table III reports the values of χ^2 tests, based on equation (23), of the magnitude of different pricing errors for the iterated kernels, where the null hypothesis is that the errors associated with a given asset and its managed portfolios are jointly zero. The null is rejected for the first-order approximations in consumption-based kernels. However, in many cases, the pricing errors are not significantly different from zero. Table IV contains the significance tests for the pricing errors for the HJ kernels. The pricing errors are not

Table IV

Testing the Statistical Significance of Pricing Errors from Kernels Based on One-Period Ahead Consumption Growth Using the Second Moment Matrix of Returns as the Weighting Matrix: 1954:1 to 1991:4

This table examines the pricing errors for kernels based on the growth rate in real per capita personal consumption expenditures on nondurables. *Tbill3M* is the ex post real one month holding period return on a three month Treasury bill. *Corp. Bonds* is the ex post real one month holding period return on a portfolio of high-grade corporate bonds. *Decile x* is the ex post real one month return to an equal-weighted portfolio of stocks from decile x of New York Stock Exchange/American Stock Exchange (NYSE/Amex) stocks, with *Decile 1* being the smallest decile and *Decile 10* being the largest decile. All real returns are formed by subtracting ex post inflation from the nominal return. Each entry in the body of the table is the Wald-type test for the null hypothesis that the pricing errors for the given asset and the managed portfolios based on the asset and the instruments are all zero. Each statistic is distributed χ^2_4, and the p-value is reported in parentheses beneath the value of the statistic.

	Approximation Order			
Asset Class	1	2	3	4
Tbill3M	2.854	2.572	1.414	1.525
	(0.583)	(0.632)	(0.842)	(0.822)
Decile 1	6.207	5.412	3.849	4.259
	(0.184)	(0.248)	(0.427)	(0.372)
Decile 5	5.233	4.569	3.034	3.377
	(0.264)	(0.334)	(0.552)	(0.497)
Decile 10	4.299	3.655	2.267	2.529
	(0.367)	(0.455)	(0.687)	(0.639)
Corp. bonds	3.273	2.943	1.681	1.741
	(0.513)	(0.567)	(0.794)	(0.783)

significantly different from zero. The general conclusions from these figures and tables are: Kernels based purely on consumption growth do not provide strong evidence in support of the importance of higher order terms in consumption- or return-based kernels. While they are not generally rejected by aggregate measures of performance, they produce point estimates for pricing errors that are sometimes quite large and imprecisely measured.[17]

Kernels motivated by the time nonseparable utility specification of equation (3) are examined in Figures 3 through 6 and Tables V through VII. For kernels based on one- and two-periods ahead consumption, the coefficients on the first- and second-order polynomial terms are all significantly different from zero. One period ahead consumption is not significantly different from zero at any polynomial order. Consumption growth two periods ahead, however, is strongly significant at all polynomial orders. When technology shock growth is

[17] Kernels based on a combination of shocks and aggregate consumption produce a modest improvement on kernels based on consumption alone. Shocks have marginal explanatory power for returns when combined with nondurables. The HJ kernels produce smaller pricing errors than the iterated GMM kernels, but the smallest size decile stocks still appear to have the largest pricing errors. These results are contained in Chapman (1996).

Table V

Tests of Significance for Approximate Kernels Based on Nonseparable Utility: 1954:1 to 1991:4

Panel A contains the results for Wald type tests of the significance of the marginal polynomial order in approximate kernels based on next period's one quarter growth rate in real per capita nondurable consumption and the one period growth rate two periods in the future *and* on the two comparable one period growth rates in the technology shock. Panel B contains tests of the marginal significance—for different approximation orders—of consumption growth 1-period ahead versus consumption growth 2-periods ahead. Panel C contains tests of the marginal significance—for different approximation orders—of consumption growth versus technology shock growth. Panel D contains the tests of the models' overall fit. $J_T(i)$ refers to the conventional test of the model's overidentifying restrictions, and $J^{\mathrm{HJ}}(i)$ refers to the test of average pricing errors using the weighting matrix $W = T^{-1}\sum_{t=1}^{T}(R_{zt}R'_{zt})^{-1}$, both statistics are for the ith order approximation. All test statistics in all panels have asymptotic χ^2 distributions. The p-values are reported in parentheses and degrees of freedom in square brackets below and to the right of the statistic.

	Polynomial Order			
Series	1	2	3	4
Panel A: Tests of Marginal Polynomial Order				
Cons.	45.718	5.909	0.594	1.483
	(0.000) [2]	(0.052) [2]	(0.743) [2]	(0.476) [2]
Cons. & shocks	25.107	3.770	3.849	2.542
	(0.000) [4]	(0.438) [4]	(0.427) [4]	(0.637) [4]
Panel B: Tests of Consumption Growth in One versus Two Periods				
1-Period cons.	0.905	1.923	1.036	0.505
	(0.342) [1]	(0.382) [2]	(0.793) [3]	(0.973) [4]
2-Period cons.	45.666	18.256	10.672	9.667
	(0.000) [1]	(0.000) [2]	(0.014) [3]	(0.046) [4]
Panel C: Tests of Consumption versus Technology Shocks				
Cons. growth	13.572	10.442	14.376	2.429
	(0.001) [2]	(0.034) [4]	(0.026) [6]	(0.965) [8]
Shock growth	2.285	2.858	4.981	2.559
	(0.319) [2]	(0.582) [4]	(0.546) [6]	(0.959) [8]
Panel D: Tests of Overall Model Fit				
Cons.: $J_T(i)$	9.678	4.960	3.454	0.199
	(0.917) [17]	(0.992) [15]	(0.996) [13]	(0.999) [11]
Cons.: $J^{\mathrm{HJ}}(i)$	13.795	12.759	8.825	6.127
	(0.682) [17]	(0.621) [15]	(0.786) [13]	(0.865) [11]
Cons. & shocks: $J_T(i)$	12.798	7.165	5.058	0.191
	(0.618) [15]	(0.786) [11]	(0.653) [7]	(0.979) [3]
Cons. & shocks: $J^{\mathrm{HJ}}(i)$	14.293	10.671	6.594	0.200
	(0.503) [15]	(0.471) [11]	(0.472) [7]	(0.978) [3]

added to the kernel, the marginal explanatory power of consumption (one and two periods ahead) is significant for polynomials of order one through three. The marginal explanatory power of shocks is not significant for any polynomial order. The combined effect of shocks and consumption is only significant for the first-order polynomial. The $J_T(i)$ fail to reject the model's overidentifying restrictions for all polynomial orders of both specifications. The values of these statistics decline dramatically in higher polynomial orders of a kernel based on

Table VI

Testing the Statistical Significance of Pricing Errors from Kernels Based on Nonseparable Utility Using the Optimal GMM Weighting Matrix: 1954:1 to 1991:4

Panel A examines the pricing errors for kernels based on next period's one quarter growth rate in real per capita nondurable consumption and the one period growth rate two periods in the future *and* on the two comparable one period growth rates in the technology shock. Panel B examines the pricing errors for kernels based on 1- and 2-period ahead consumption growth and 1- and 2-period ahead technology shock growth. *Tbill3M* is the ex post real one month holding period return on a three month Treasury bill. *Corp. Bonds* is the ex post real one month holding period return on a portfolio of high-grade corporate bonds. *Decile x* is the ex post real one month return to an equal-weighted portfolio of stocks from decile x of New York Stock Exchange/American Stock Exchange (NYSE/Amex) stocks, with *Decile 1* being the smallest decile and *Decile 10* being the largest decile. All real returns are formed by subtracting ex post inflation from the nominal return. Each entry in the body of the table is the Wald-type test for the null hypothesis that the pricing errors for the given asset and the "managed portfolios" based on the asset and the instruments are all zero. Each statistic is distributed χ_4^2, and the p-value is reported in parentheses beneath the value of the statistic.

	Approximation Order			
Asset Class	1	2	3	4
Panel A: Consumption Growth Kernels				
Tbill3M	1.151	0.755	0.765	2.302
	(0.886)	(0.944)	(0.943)	(0.680)
Decile 1	1.125	0.940	1.368	2.921
	(0.890)	(0.919)	(0.850)	(0.571)
Decile 5	1.332	0.752	0.695	2.389
	(0.856)	(0.945)	(0.952)	(0.665)
Decile 10	1.284	0.626	0.791	2.072
	(0.864)	(0.960)	(0.940)	(0.722)
Corp. bonds	1.218	0.644	0.765	1.990
	(0.875)	(0.958)	(0.943)	(0.738)
Panel B: Consumption Growth and Shock Based Kernels				
Tbill3M	0.643	0.234	3.165	0.191
	(0.958)	(0.994)	(0.531)	(0.996)
Decile 1	0.595	0.351	2.574	0.191
	(0.964)	(0.986)	(0.631)	(0.996)
Decile 5	0.767	0.272	2.751	0.191
	(0.943)	(0.992)	(0.600)	(0.996)
Decile 10	0.817	0.194	3.094	0.191
	(0.936)	(0.996)	(0.542)	(0.996)
Corp. bonds	0.710	0.202	3.107	0.192
	(0.950)	(0.995)	(0.540)	(0.996)

both shocks and consumption. In summary, in a kernel with both nondurables consumption and technology shocks, the point estimates of the parameters on the shock terms suggests that they have no marginal explanatory power with respect to asset returns.

Table VII

Testing the Statistical Significance of Pricing Errors from Kernels Based on Nonseparable Utility Using the Second Moment Matrix of Returns as the Weighting Matrix: 1954:1 to 1991:4

Panel A examines the pricing errors for kernels based on next period's one quarter growth rate in real per capita nondurable consumption and the one period growth rate two periods in the future *and* on the two comparable one period growth rates in the technology shock. Panel B examines the pricing errors for kernels based on 1- and 2-period ahead consumption growth and 1- and 2-period ahead technology shock growth. *Tbill3M* is the ex post real one month holding period return on a three month Treasury bill. *Corp. Bonds* is the ex post real one month holding period return on a portfolio of high-grade corporate bonds. *Decile x* is the ex post real one month return to an equal-weighted portfolio of stocks from decile x of New York Stock Exchange/American Stock Exchange (NYSE/Amex) stocks, with *Decile 1* being the smallest decile and *Decile 10* being the largest decile. All real returns are formed by subtracting ex post inflation from the nominal return. Each entry in the body of the table is the Wald-type test for the null hypothesis that the pricing errors for the given asset and the "managed portfolios" based on the asset and the instruments are all zero. Each statistic is distributed χ^2_4, and the p-value is reported in parentheses beneath the value of the statistic.

	Approximation Order			
Asset Class	1	2	3	4
Panel A: Consumption Growth Kernels				
Tbill3M	0.660	0.212	0.316	1.366
	(0.956)	(0.995)	(0.989)	(0.850)
Decile 1	0.887	0.254	0.332	1.408
	(0.926)	(0.993)	(0.988)	(0.843)
Decile 5	0.974	0.284	0.382	1.607
	(0.914)	(0.991)	(0.984)	(0.808)
Decile 10	0.832	0.223	0.343	1.531
	(0.934)	(0.994)	(0.987)	(0.821)
Corp. bonds	0.730	0.199	0.297	1.457
	(0.948)	(0.995)	(0.990)	(0.834)
Panel B: Consumption Growth and Shock-Based Kernels				
Tbill3M	0.824	0.356	0.521	0.027
	(0.935)	(0.986)	(0.971)	(0.999)
Decile 1	1.045	0.343	0.216	0.038
	(0.903)	(0.987)	(0.995)	(0.999)
Decile 5	1.159	0.386	0.346	0.034
	(0.885)	(0.984)	(0.987)	(0.999)
Decile 10	1.007	0.299	0.485	0.032
	(0.909)	(0.990)	(0.975)	(0.999)
Corp. bonds	0.893	0.285	0.557	0.030
	(0.926)	(0.991)	(0.968)	(0.999)

Figures 3 and 4 show the familiar HJ bounds for this set of assets, along with the point estimates of the means and standard deviations of the different kernels. In general, the point estimates of the iterated kernels are closer to the permissible region than in any of the other cases considered so far. In partic-

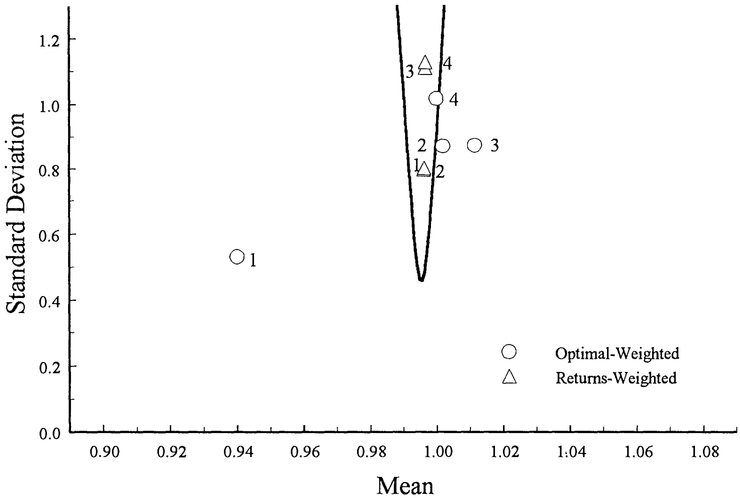

Figure 3. The Hansen-Jagannathan volatility bounds and approximate kernels based on a time-nonseparable consumption model. This figure shows both the volatility bounds based on the real returns to Treasury bills, size-based stock portfolios, and corporate bonds; and the point estimates of the mean and standard deviation of approximate kernels based on next period's one quarter growth rate in real per capita nondurable consumption and the one period growth rate two periods in the future. The numbers next to each of the point estimates refer to the order of the approximating polynomial. The order runs from 1 to 4. "Optimal-Weighted" refers to point estimates of the mean and standard deviation of the approximate kernels based on the parameters estimated using the optimal GMM weighting matrix. "Returns-Weighted" refers to point estimates of the mean and standard deviation of the approximate kernels based on the parameters estimated using the inverse of the sample estimate of the second moment matrix of returns as the weighting matrix in the GMM estimation.

ular, the fourth-order kernel based on consumption and shocks is in the interior of the HJ bounds. As usual, the HJ kernels are all inside the bounds, and the $J^{\mathrm{HJ}}(i)$ tests are consistent with this conclusion.

The individual pricing errors are examined in Figures 5 and 6. The striking results in Figure 6 show very small average pricing errors for both types of kernels for the second- and fourth-order approximations based on consumption and shocks, *including the smallest stock decile.* Furthermore, tests of the significance of the pricing errors in Tables VI and VII are all quite small, uniformly failing to reject the null that all pricing errors are zero. These strong results suggest that a combination of consumption and technology shocks does a good job as an approximation to the pricing kernel, but that the magnitude of the standard errors on individual parameter estimates makes it very difficult to assess the significance of individual polynomial terms. It should also be noted that the polynomial approximations are incapable of ensuring a strictly positive pricing kernel.[18]

[18] See Chapman (1996) for plots of the fitted values of some approximate kernels. These fitted values are not strictly positive, and they fail to reveal any consistent time series pattern in the properties of the approximate kernels.

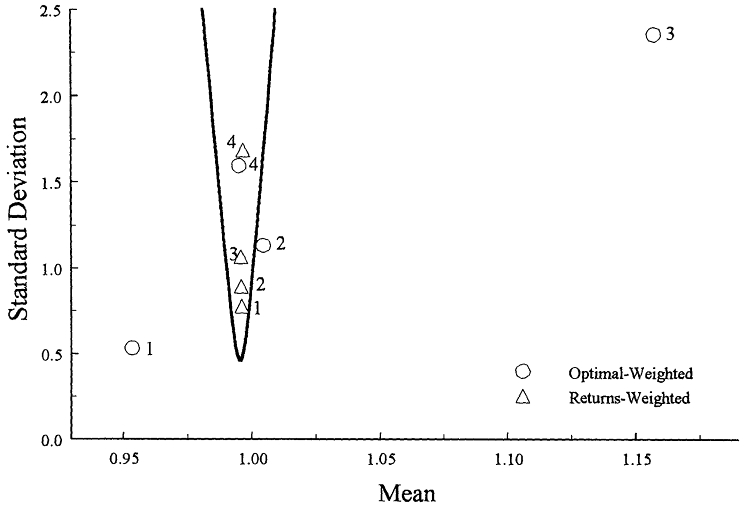

Figure 4. The Hansen-Jagannathan volatility bounds and approximate kernels based on a time-nonseparable consumption model, including technology shock growth. This figure shows both the volatility bounds based on the real returns to Treasury bills, size-based stock portfolios, and corporate bonds; and the point estimates of the mean and standard deviation of approximate kernels based on next period's one quarter growth rate in real per capita nondurable consumption and the one period growth rate two periods in the future *and* on the two comparable one period growth rates in the technology shock. The numbers next to each of the point estimates refer to the order of the approximating polynomial. The order runs from 1 to 4. "Optimal-Weighted" refers to point estimates of the mean and standard deviation of the approximate kernels based on the parameters estimated using the optimal GMM weighting matrix. "Returns-Weighted" refers to point estimates of the mean and standard deviation of the approximate kernels based on the parameters estimated using the inverse of the sample estimate of the second moment matrix of returns as the weighting matrix in the GMM estimation.

VI. Conclusions

This article has presented a new approach to examining the implications of dynamic asset pricing models. Given a particular underlying structure for the model economy, the asset pricing kernel will be a continuous function of a small number of state variables. The pricing kernel can then be approximated using orthogonal polynomials in the state variables. This approximation approach has the advantage of making the approximations linear functions of the estimated parameters while permitting the true pricing kernel to be a nonlinear function of the model's state variables. A simple version of the neoclassical growth model was used as a specific example of the approach.

Applying the approximate kernel approach to a set of quarterly (ex post) real returns to Treasury bills, size-based stock portfolios, and government bond portfolios produced the following results: Kernels based on one-period ahead consumption growth were not typically rejected by overall measures of model fit, but they produced statistically and economically large pricing errors. The largest improvement in the performance of the approximate kernels occurred when an additional future value of consumption and/or technology shocks,

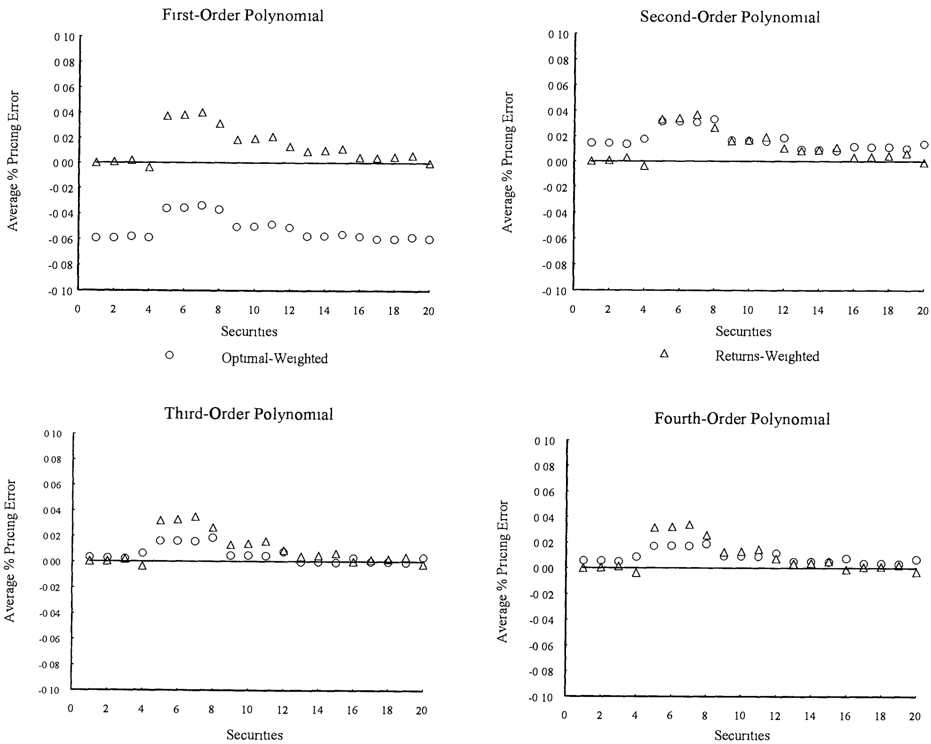

Figure 5. Average pricing errors for approximate kernels based on a time-nonseparable consumption model. This figure shows the average pricing error from approximate kernels based on next period's one quarter growth rate in real per capita nondurable consumption and the one period growth rate two periods in the future for each of the four managed portfolios based on the five primary assets. The primary assets are the real returns to Treasury bills, three size-based stock portfolios, and corporate bonds. The managed portfolio returns are formed by multiplying each original asset by a constant and the other instrumental variables. The instrumental variables are the lagged dividend yield on the S&P 500, the lagged yield spread between Baa and Aaa corporate debt, and lagged industrial production growth. In all, there are twenty average pricing errors associated with each order of approximation and with each weighting method for generating parameter estimates. The first four entries are the T-bill returns and the three managed portfolios associated with T-bills, the next are for the smallest decile stocks, followed by the middle decile stocks, the largest decile stocks, and corporate bond returns. "Optimal-Weighted" refers to the pricing errors associated with the parameter estimates generated using the optimal GMM weighting matrix. "Returns-Weighted" refers to the pricing errors associated with the parameter estimates generated using the inverse of the sample estimate of the second moment matrix of returns as the weighting matrix in the GMM estimation.

motivated by an appeal to time nonseparable preferences or durability in consumption goods, were added to the kernels. The marginal power of shocks, in these cases, is difficult to detect directly, but the inclusion of shocks substantially improved the fit of the approximate kernels. In fact, they were capable of eliminating the small firm effect. Finally, the results provided strong evidence of potential problems, emphasized in Cochrane (1996), asso-

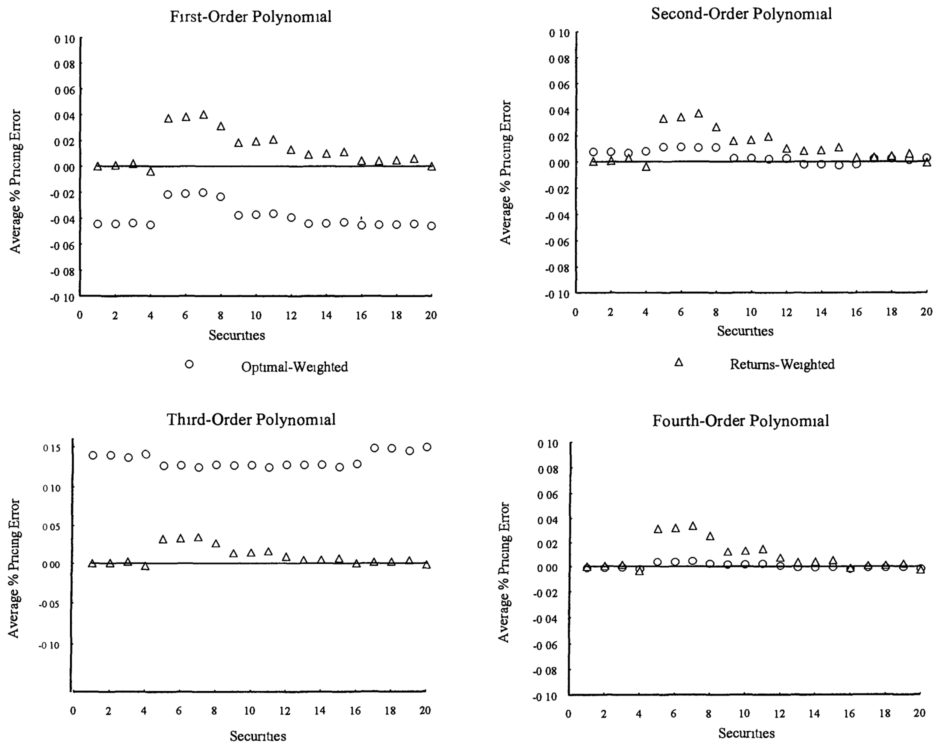

Figure 6. Average pricing errors for approximate kernels based on a time-nonseparable consumption model, including technology shock growth. This figure shows the average pricing error from approximate kernels based on next period's one quarter growth rate in real per capita nondurable consumption and the one period growth rate two periods in the future *and* on the two comparable one period growth rates in the technology shock for each of the four managed portfolios based on the five primary assets. The primary assets are the real returns to Treasury bills, three size-based stock portfolios, and corporate bonds. The managed portfolio returns are formed by multiplying each original asset with a constant and the other instrumental variables. The instrumental variables are the lagged dividend yield on the S&P 500, the lagged yield spread between Baa and Aaa corporate debt, and lagged industrial production growth. In all, there are twenty average pricing errors associated with each order of approximation and with each weighting method for generating parameter estimates. The first four entries are the T-bill returns and the three managed portfolios associated with T-bills, the next are for the smallest decile stocks, followed by the middle decile stocks, the largest decile stocks, and corporate bond returns. "Optimal-Weighted" refers to the pricing errors associated with the parameter estimates generated using the optimal GMM weighting matrix. "Returns-Weighted" refers to the pricing errors associated with the parameter estimates generated using the inverse of the sample estimate of the second moment matrix of returns as the weighting matrix in the GMM estimation.

ciated with using the "optimal" covariance matrix for the pricing errors. In particular, a prespecified weighting matrix equal to the inverse of the second moment matrix of returns generally produced smaller pricing errors, and kernels that were more nearly consistent with the HJ bounds on permissible pricing kernels.

The results in this article suggest a number of directions for future research. The parameters of the approximating kernels could be allowed to vary over time with the state variables of the model. Alternate underlying model economies could be used to suggest different specifications for the pricing kernel. An example of this would be to examine the implications of a monetary economy for the pricing of nominal securities. Finally, the risk versus return tradeoff implied by an approximate kernel could be used as a benchmark for evaluating investment performance.

REFERENCES

Andrews, Donald W. K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* 59, 817–858.

Andrews, Donald W. K., and J. Christopher Monohan, 1992, An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator, *Econometrica* 60, 953–966.

Bansal, Ravi, and S. Viswanathan, 1993, No arbitrage and arbitrage pricing: A new approach, *Journal of Finance* 48, 1231–1262.

Bansal, Ravi, David A. Hsieh, and S. Viswanathan, 1993, A new approach to international arbitrage pricing, *Journal of Finance* 48, 1719–1747.

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265–296.

Chapman, David A., 1996, Approximating the asset pricing kernel, Working Paper No. FR 96-02, Bradley Policy Research Center, William E. Simon Graduate School of Business Administration, University of Rochester.

Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572–621.

Cooley, Thomas F., and Edward C. Prescott, 1995, Economic growth and business cycles, in Thomas F. Cooley, Ed.: *Frontiers of Business Cycle Research* (Princeton University Press, Princeton).

Duffie, Darrell, and William Zame, 1989, The consumption-based capital asset pricing model, *Econometrica* 57, 1279–1297.

Fama, Eugene F., 1992, Efficient capital markets: II, *Journal of Finance* 41, 1575–1617.

Ferson, Wayne E., and George M. Constantinides, 1991, Habit persistence and durability in aggregate consumption: Empirical tests, *Journal of Financial Economics* 29, 199–240.

Hämmerlin, Günther, and Karl-Heinz Hoffmann, 1991, *Numerical Mathematics* (Springer-Verlag Inc., New York).

Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Hansen, Lars Peter, John Heaton, and Erzo Luttmer, 1995, Econometric evaluation of asset pricing models, *Review of Financial Studies* 8, 237–274.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Hansen, Lars Peter, and Ravi Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance,* Forthcoming.

Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized instrumental variable estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.

Judd, Kenneth L., 1992, Projection methods for solving aggregate growth models, *Journal of Economic Theory* 58, 410–452.

Leroy, Stephen F., 1973, Risk aversion and the martingale property of stock returns, *International Economic Review* 14, 436–446.

Lucas, Robert E., Jr., 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429–1445.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Ogaki, Masao, 1993, Generalized method of moments: Econometric applications, in C. R. Rao and H. D. Vinod, Eds.: *Handbook of Statistics, Vol. 11* (Elsevier Science Publishers, Amsterdam).

Shanken, Jay, 1991, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99–120.

Stokey, Nancy L., and Robert E. Lucas, Jr., 1989, *Recursive Methods in Economic Dynamics* (Harvard University Press, Cambridge, MA).

Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 401–421.