



Generalized Method of Moments: Applications in Finance

Author(s): Ravi Jagannathan, Georgios Skoulakis, Zhenyu Wang

Reviewed work(s):

Source: *Journal of Business & Economic Statistics*, Vol. 20, No. 4, JBES Twentieth Anniversary Issue on the Generalized Method of Moments (Oct., 2002), pp. 470-481

Published by: [American Statistical Association](#)

Stable URL: <http://www.jstor.org/stable/1392417>

Accessed: 10/01/2012 04:06

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Statistical Association is collaborating with JSTOR to digitize, preserve and extend access to *Journal of Business & Economic Statistics*.

<http://www.jstor.org>

Generalized Method of Moments: Applications in Finance

Ravi JAGANNATHAN, Georgios SKOULAKIS

Kellogg School of Management, Northwestern University, Evanston, IL 60208 (*rjaganna@nwu.edu*)

Zhenyu WANG

Columbia University Business School, New York, NY 10027

We provide a brief overview of applications of generalized method of moments in finance. The models examined in the empirical finance literature, especially in the asset pricing area, often imply moment conditions that can be used in a straight forward way to estimate the model parameters without making strong assumptions regarding the stochastic properties of variables observed by the econometrician. Typically the number of moment conditions available to the econometrician would exceed the number of model parameters. This gives rise to overidentifying restrictions that can be used to test the validity of the model specifications. These advantages have led to the widespread use of the generalized method of moments in the empirical finance literature.

KEY WORDS: Asset pricing models; Bid-ask spreads; CAPM; Estimating scalar diffusions; Generalized method of moments; Linear beta pricing models; Market microstructure; Mean-variance efficiency; Mean-variance spanning; Stochastic discount factor; Term structure of interest rates.

1. INTRODUCTION

The development of the generalized method of moments (GMM) by Hansen (1982) has had a major impact on empirical research in finance, especially in the area of asset pricing. GMM has made econometric evaluation of asset-pricing models possible under more realistic assumptions regarding the nature of the stochastic process governing the temporal evolution of exogenous variables.

A substantial amount of research in finance is directed toward understanding why different financial assets earn different returns on average and why a given asset may be expected to earn different returns at different points in time. Various asset pricing models that explain how prices of financial claims are determined in financial market have been proposed in the literature to address these issues. These models differ from one another due to the nature of the assumptions that they make regarding investor characteristics, that is, preferences, endowments, and information sets; the stochastic process governing the arrival of information in financial markets; and the nature of the transactions technology for exchanging financial and real claims among different agents in the economy. Each asset-pricing model specifies what the expected return on a financial asset should be in terms of observable variables and model parameters at each point in time.

Although these models have differences, they also have similarities. Most of the models start by studying the first-order conditions to the optimal consumption, investment, and portfolio choice problem faced by a model investor. That leads to the stochastic discount factor representation of these models. The price assigned by a model to a financial asset equals the conditional expectation of its future payoff multiplied by a model-specific stochastic discount factor. In an informationally efficient market where the econometrician has less information than the model investor, it should not be possible to explain the difference between the market price of a security and the price assigned to it by a model based on information

available to the econometrician. GMM can be used to estimate the model parameters and test the set of moment conditions that arise in this natural way. For some models, it would not be convenient to work with the stochastic discount factor representation. We provide examples where even in those cases it would be natural to use the GMM.

The first two important applications of GMM in finance are those of Hansen and Hodrick (1980) and Hansen and Singleton (1982). Subsequent developments in time series econometric methods, as well as refinements and extensions to GMM, have made it a reliable and robust econometric methodology for studying dynamic asset-pricing models, allowing asset returns and the stochastic discount factor to be serially correlated, leptokurtic, and conditionally heteroscedastic. The works by Newey and West (1987), Andrews (1991), and Andrews and Monahan (1992) on estimating covariance matrices in the presence of autocorrelation and heteroscedasticity are probably the most significant among these developments.

Before the advent of GMM, the primary econometric tool in the asset-pricing area in finance was the maximum likelihood (ML) method, which is often implemented using linear or nonlinear regression methods. The ML method has several limitations. First, for each asset-pricing model, researchers have to derive a test for examining model misspecification, which is not always easy or possible. Second, linear approximation is often necessary when studying nonlinear asset-pricing models. Third, researchers must make strong distributional assumptions. To make the estimation problem tractable, the assumed distributions often must be serially uncorrelated and conditionally homoscedastic. When the distributional assumptions are not satisfied, the estimated model parameters may be biased

even in large samples. These limitations severely restrict the scope of the empirical investigations of dynamic asset-pricing models. GMM enables the econometrician to overcome these limitations. The econometrician does not have to make strong distributional assumptions—the variables of interest can be serially correlated and conditionally heteroscedastic. Further nonlinear asset-pricing models can be examined without linearization. The convenience and the generality of GMM are the two main reasons why GMM has become so popular in the finance literature.

Despite its advantages, GMM has a potential shortcoming when compared to the ML method. When the distributional assumptions are valid, the ML method provides the most efficient estimates of model parameters, whereas the GMM method may not. To apply GMM, the econometrician typically uses the moment conditions generated by the *stochastic discount factor* (SDF) representation of an asset-pricing model. The moment conditions representing the implications of an asset-pricing model chosen by the econometrician may not lead to the most efficient estimation of model parameters. It is therefore important to understand whether GMM as commonly applied in practice to examine asset-pricing models is less efficient than the ML method in the situations where the ML method can be applied. It is well known that GMM has the same estimation efficiency as the ML method when applied to the moment conditions generated by the classical beta representation of linear asset-pricing models. Jagannathan and Wang (2002) demonstrated that GMM is as efficient as the ML method when applied to the moment conditions generated by the SDF representation of linear asset-pricing models. This reinforces the importance and advantage of GMM in empirical asset-pricing applications.

In what follows we review the use of GMM in finance with emphasis on asset-pricing applications. In Section 2 we discuss using GMM for empirically evaluating the standard consumption-based asset-pricing model and some of its extensions. In Section 3 we discuss using GMM in examining factor models. In Section 4 we discuss the advantages of using alternative weighting matrices while applying GMM. We discuss the efficiency of GMM when applied to the SDF representation of asset-pricing models in Section 5. We examine the use of GMM to estimate the stochastic process for short-term interest rates in Section 6, and discuss using GMM in the market microstructure literature in Section 7. We summarize in Section 8.

2. NONLINEAR RATIONAL EXPECTATIONS MODELS

In most asset-pricing models, expectations about the future represent an important factor in the decision making process. In practice, however, it is the actions of the agents, rather than their expectations, that we observe. Empirical work on asset-pricing models relies on the hypothesis of rational expectations, which in turn implies that the errors made by the agents—namely, the differences between observed realizations and expectations—are uncorrelated with information on which the expectations are conditioned.

2.1 Consumption-Based Capital Asset-Pricing Model

One of the first applications of GMM appeared in the context of estimation of a consumption-based nonlinear rational expectations asset-pricing model of Hansen and Singleton (1982). Consider an expected utility-maximizing agent with preferences on current and future consumption streams characterized by a time-additive utility function. Let time τ refer to the present, let $c_{t+\tau}$ be the agent's consumption level during period $t + \tau$, let I_τ be the information set available to the agent at τ , let $\beta \in (0, 1)$ be the time-preference parameter, and let U be an increasing, concave utility function. The agent is then assumed to choose consumption streams to maximize

$$\sum_{t=0}^{\infty} \beta^t E[U(c_{t+\tau}) | I_\tau]. \quad (1)$$

The agent can invest in any of N available securities with gross returns ($1 +$ the rates of return) at time t , denoted by $R_{i,t}$, $i = 1, \dots, N$. Solving the agent's intertemporal consumption and portfolio choice implies the following equation, which relates security returns to the intertemporal marginal rate of substitution, $\beta \frac{U'(c_{t+1})}{U'(c_t)}$:

$$E[\beta [U'(c_{t+1})/U'(c_t)] R_{i,t+1} | I_t] = 1, \quad i = 1, \dots, N. \quad (2)$$

A standard assumption frequently used in the asset-pricing/macroeconomics literature is that the utility function U belongs to the constant relative risk-aversion (CRRA) family. The utility functions U in the CRRA family can be described by a single parameter, $\gamma > 0$, referred to as the coefficient of relative risk aversion, as

$$U(c) = \frac{c^{1-\gamma}}{1-\gamma} \quad \text{if } \gamma \neq 1 \quad \text{and} \quad U(c) = \log(c) \quad \text{if } \gamma = 1. \quad (3)$$

For such utility function U , (2) translates to

$$E[\beta (c_{t+1}/c_t)^{-\gamma} R_{i,t+1} | I_t] = 1, \quad i = 1, \dots, N, \quad (4)$$

which, by the law of iterated expectations, implies that

$$E[(1 - \beta (c_{t+1}/c_t)^{-\gamma} R_{i,t+1}) \mathbf{z}_t] = 0, \quad i = 1, \dots, N \quad (5)$$

for any random vector $\mathbf{z}_t$ measurable with respect to the information set I_t . Assume that $\mathbf{z}_t$ has dimension $K \times 1$. For inference purposes, $\mathbf{z}_t$ is chosen to be observable by the econometrician. Typically, $\mathbf{z}_t$ would include lagged rates of return and lagged consumption growth rates.

The parameter vector to be estimated is $\theta = (\beta, \gamma)'$. Let $\mathbf{1}_N$ denote the $N \times 1$ vector of 1s, let $\mathbf{0}_M$ denote the $M \times 1$ vector of 0s where $M = NK$, let $\mathbf{R}_t$ denote the vector of asset returns at time t , and let $\mathbf{y}_{t+1} = (\mathbf{R}'_t, c_{t+1}/c_t, \mathbf{z}'_t)'$ denote the data available to the econometrician at time $t + 1$. The orthogonality conditions in (5) can be written more compactly as

$$E[\mathbf{h}(\mathbf{y}_{t+1}, \theta)] = \mathbf{0}_M, \quad \text{where} \quad \mathbf{h}(\mathbf{y}_{t+1}, \theta) = [\mathbf{1}_N - \beta(1 + \mathbf{R}_{t+1})(c_{t+1}/c_t)^{-\gamma}] \otimes \mathbf{z}_t, \quad (6)$$

with $\otimes$ denoting the Kronecker product. The sample average of $\mathbf{h}(\mathbf{y}_{t+1}, \theta)$ is

$$\mathbf{g}_T(\theta) = \frac{1}{T} \sum_{t=1}^T \mathbf{h}(\mathbf{y}_{t+1}, \theta), \quad (7)$$

and the GMM estimate $\hat{\theta}_T$ is obtained by minimizing the quadratic function

$$J_T(\theta) = \mathbf{g}_T(\theta)' \mathbf{S}_T^{-1} \mathbf{g}_T(\theta), \quad (8)$$

where $\mathbf{S}_T$ is a consistent estimate of the covariance matrix $\mathbf{S}_0(\theta) = E[\mathbf{h}(\mathbf{y}_{t+1}, \theta)\mathbf{h}(\mathbf{y}_{t+1}, \theta)']$. Note that, according to (4), lags of the time series $\mathbf{h}(\mathbf{y}_{t+1}, \theta)$ are uncorrelated, and thus the spectral density matrix reduces to the covariance matrix $\mathbf{S}_0(\theta)$. A consistent estimate of $\mathbf{S}_0(\theta)$ is

$$\mathbf{S}_T = \frac{1}{T} \sum_{t=1}^T \mathbf{h}(\mathbf{y}_{t+1}, \hat{\theta}_0)\mathbf{h}(\mathbf{y}_{t+1}, \hat{\theta}_0)', \quad (9)$$

where $\hat{\theta}_0$ is an initial consistent estimate of θ . Typically, $\hat{\theta}_0$ is obtained by using the identity matrix as the weighting matrix in the quadratic form, that is, by minimizing $J_T^0(\theta) = \mathbf{g}_T(\theta)' \mathbf{g}_T(\theta)$. The test of the overidentifying restrictions is based on the statistic

$$T \mathbf{g}_T(\hat{\theta}_T)' \mathbf{S}_T^{-1} \mathbf{g}_T(\hat{\theta}_T), \quad (10)$$

which asymptotically follows a chi-squared distribution with $M - 2$ degrees of freedom, because we have M orthogonality conditions and two parameters to be estimated.

2.2 Assessment of Asset Pricing Models Using the Stochastic Discount Factor Representation

The method given in the preceding section can be used to evaluate an arbitrary asset-pricing model. For this purpose, we rewrite the asset-pricing model given in (2) as

$$E[m_{t+1} R_{i,t+1} | I_t] = 1, \quad i = 1, \dots, N, \quad (11)$$

where $m_{t+1} = \beta U'(c_{t+1})/U'(c_t)$. Any random variable m_{t+1} that satisfies (11) is referred to as a SDF. In general, a number of random variables satisfying (11) exist, and hence there is more than one SDF. An asset-pricing model designates a particular random variable as a SDF. GMM can be applied in exactly the same way as described earlier to estimate the asset-pricing model parameters and test the overidentifying restrictions implied by the asset-pricing model using its SDF representation of an asset-pricing model. In what follows, we provide some examples of SDFs. We also give an example to show that the GMM is helpful in the econometric analysis of asset-pricing models even when the SDF representation of a model is not used.

2.3 More General Utility Functions

Several researchers have modeled consumption expenditures as generating consumption services over a period of time; Notable among these are Dunn and Singleton (1986) and Eichenbaum, Hansen, and Singleton (1988). Sundaresan (1989) and Constantinides (1990) made the case for the importance of allowing for habit formation in preferences where habit depends on an agent's own past sequence of consumption. Abel (1990) and Campbell and Cochrane (1999) modeled habit in such a way that it is affected not by an agent's own decisions, but rather by the decisions of other agents in the economy. Ferson and Constantinides (1991) examined the case in which the utility function exhibits both habit persistence and durability of consumption expenditures. In these cases the utility function is no longer time separable, because the flow of consumption services in a given time period depends on consumption expenditures in the past.

Campbell, Lo, and MacKinlay (1997) considered a representative agent who maximizes the infinite-horizon utility

$$\sum_{j=0}^{\infty} \beta^j E[U(c_{t+j}) | I_t], \quad \text{where } U(c_{t+j}) = \frac{(c_{t+j}/x_{t+j})^{1-\gamma} - 1}{1-\gamma}, \quad (12)$$

c_t denotes consumption, and x_t is the variable summarizing the current state of habit in period t . When habit is external and equals the aggregate consumption c_t during the previous period, the SDF is given by

$$m_{t+1} = \beta \left(\frac{c_t}{c_{t-1}} \right)^{\gamma-1} \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma}. \quad (13)$$

When habit is internal, the SDF representation is not convenient for empirically examining the model. Instead, it is more convenient to apply GMM to the following relation between the expected return on any asset and fundamentals that must be satisfied when the model holds:

$$E_t[c_t^{\gamma-1} c_t^{-\gamma} - \beta c_t^{\gamma-1} c_{t+1}^{-\gamma} (c_{t+1}/c_t)] \\ = E_t[R_{t+1} (c_t^{\gamma-1} c_{t+1}^{-\gamma} - \beta c_{t+1}^{\gamma-1} c_{t+2}^{-\gamma} (c_{t+2}/c_{t+1}))]. \quad (14)$$

Campbell and Cochrane (1999) developed a model in which the habit variable x_t enters the infinite-horizon utility $\sum_{j=0}^{\infty} \beta^j E[U(c_{t+j}) | I_t]$ additively in the following way:

$$U(c_{t+j}) = \frac{(c_{t+j} - x_{t+j})^{1-\gamma} - 1}{1-\gamma}. \quad (15)$$

They assumed that the habit variable, x_t , is external; that is, it is determined by the consumption paths of all other agents in the economy. In particular, x_t is a weighted average of past aggregate per capita consumptions. Campbell and Cochrane showed that under some additional assumptions, the SDF is given by $m_{t+1} = \beta \left(\frac{s_{t+1} c_{t+1}}{s_t c_t} \right)^{\gamma}$, where $s_t = \frac{c_t - x_t}{c_t}$ is the surplus consumption ratio. These examples illustrate the versatility of GMM in examining a variety of asset-pricing models.

Epstein and Zin (1991) and Weil (1989) developed a utility function that breaks the tight link between the coefficient of

relative risk aversion γ and the elasticity of intertemporal substitution ψ in the time-separable CRRA case. For this class of preferences, the SDF is given by

$$m_{t+1} = \left[\beta \left(\frac{c_{t+1}}{c_t} \right)^{-1/\psi} \right]^\theta \left[\frac{1}{R_{m,t+1}} \right]^{1-\theta}, \quad \text{where } \theta = \frac{1-\gamma}{1-1/\psi}. \quad (16)$$

When the period utility function is logarithmic, the foregoing SDF reduces to the one identified by Rubinstein (1976). (For more examples, see Campbell et al. 1997 and Cochrane 2001.)

3. FACTOR MODELS

3.1 Stochastic Discount Factor Representation

Linear and nonlinear factor pricing models are popular in finance. In this section we discuss the estimation of such models by applying GMM to their SDF representation. We first consider the standard capital asset-pricing model (CAPM), which can be thought of as a linear single-factor pricing model, for econometric evaluation.

3.1.1. Standard Capital Asset-Pricing Model. Consider the standard CAPM, the most standard and widely used equilibrium model. Consider an economy with N risky assets and let R_{it} denote the gross return on the i th asset during period t . Further, let R_{mt} denote the gross return for the market portfolio of all assets in the economy over the same period t . Then the CAPM states that in equilibrium, the expected gross returns are given by

$$E[R_i] = \gamma + \lambda \beta_i, \quad \text{where } \beta_i = \frac{\text{cov}[R_i, R_m]}{\text{var}[R_m]}, \quad (17)$$

λ is the market risk premium, and γ , assumed to be nonzero, is the return of the zero-beta asset. This is equivalent to

$$E[mR_i] = 1, \quad \text{where } m = \theta_0 + \theta_1 R_m, \quad (18)$$

for some θ_0 and θ_1 . It follows that m is the SDF corresponding to the CAPM. The foregoing equivalence between the beta and SDF representations of the CAPM was first pointed out by Dybvig and Ingersoll (1982) and Hansen and Richard (1987). Once again, we can apply GMM to estimate the CAPM parameters and test the overidentifying restrictions as discussed earlier.

3.1.2. Linear Factor Models. Ross (1976) and Connor (1984) showed that when there are only K economy-wide pervasive factors, denoted by $f_1, \dots, f_K$, the expected return on any financial asset will be a linear function of the K -factor betas,

$$E[R_i] = \gamma + \sum_{k=1}^K \lambda_k \beta_{ik}, \quad (19)$$

where $(\beta_{i1}, \dots, \beta_{iK})' \equiv \beta_i = (\text{var}[\mathbf{f}])^{-1} E[(R_i - E[R_i])(\mathbf{f} - E[\mathbf{f}])]$ and $\mathbf{f} \equiv (f_1, \dots, f_K)'$. Using algebraic manipulations similar to those in the CAPM, it can be shown that the foregoing linear K -factor beta pricing model can be represented as $E[mR_i] = 1$, where the SDF m is given by

$$m = \theta_0 + \sum_{k=1}^K \theta_k f_k \quad (20)$$

for some parameters $\theta_0, \dots, \theta_K$. GMM can be used to estimate these parameters and to test the model.

3.1.3. Nonlinear Factor Models. Bansal and Viswanathan (1993) derived a nonlinear factor pricing model by assuming that the intertemporal marginal rate of substitution of the marginal investor is a constant function of a finite number of economy-wide pervasive factors. They suggested approximating the function using polynomials. In the single-factor case, this gives rise to the SDF

$$m = \beta_0 + \beta_1 r_f + \sum_{j=1,2,\dots,K} \beta_{jm} r_m^j, \quad (21)$$

where K is the order of the polynomial used and r_f and r_m denote the risk-free rate and the excess return on the market index portfolio. Bansal, Hsieh, and Viswanathan (1993) suggested approximating the function using neural nets. This gives rise to the following SDF when the excess return on the market index portfolio is the single pervasive factor:

$$m = \beta_0 + \beta_1 r_f + \sum_{j=1,2,\dots,K} \beta_{jm} g(\gamma_{oj} + \gamma_{1j} r_m), \quad (22)$$

where $g(\cdot)$ is the logistic function given by $g(\cdot) = \frac{\exp(\cdot)}{1+\exp(\cdot)}$. Again, the GMM is a natural way to estimate the model parameters and test the overidentifying restrictions.

3.1.4. Mean-Variance Spanning. Whether the returns on a subset of a given collection of financial assets are sufficient to span the unconditional mean-variance frontier of returns has received wide attention in the literature. For example, an investor may be interested in examining whether it would be possible to construct a fixed-weight portfolio of some benchmark assets that dominate a given managed portfolio in the mean-variance space. If the answer is affirmative, then the investor may not have to include the managed portfolio in the menu of opportunities. This can be examined by checking whether the set of benchmark assets spans the mean-variance frontier of returns generated by the primitive assets and the managed portfolios taken together.

Suppose that we are interested in examining whether a set of K benchmark assets spans the mean-variance efficient frontier generated by N primitive assets plus the K benchmark assets for some value for the risk-free return. Huberman and Kandel (1987) showed how to construct the statistical tests for this purpose when asset returns have an iid joint normal distribution. Note that when the mean-variance frontier of returns is spanned by K benchmark returns for some value of the risk-free return, the SDF will be an affine function only of those K benchmark returns. Bekaert and Urias (1996) showed how this property can be used to test for mean-variance spanning using GMM. The advantage to using GMM is that it allows for conditional heteroscedasticity exhibited by returns on financial assets.

3.2 Beta Representation

Applications of GMM in finance are not restricted to examining asset-pricing models using the SDF framework. In this section we present an application of GMM to empirical evaluation of linear beta pricing models using the beta representation instead of the implied SDF representation.

3.2.1. Unconditional Linear Beta Pricing Models. For notational simplicity, we assume here that there is only one economy-wide pervasive factor. The results generalize in a straightforward manner to the case of multiple factors. For expositional convenience, we assume that the single factor is the return on the market portfolio and consider the standard CAPM. Following MacKinlay and Richardson (1991), assume an economy with a risk-free asset and N risky assets with excess returns over period t , denoted by R_{it} , for $i = 1, \dots, N$. Further, let R_{pt} denote the excess return on a portfolio p over period t . Mean-variance efficiency of the portfolio p implies that

$$E[R_{it}] = \beta_i E[R_{pt}],$$

$$\text{where } \beta_i = \frac{\text{cov}[R_{it}, R_{pt}]}{\text{var}[R_{pt}]}, \quad i = 1, \dots, N. \quad (23)$$

Using vector notation will ease the exposition. Let $\mathbf{R}_t = (R_{1t}, \dots, R_{Nt})'$ denote the column vector of excess returns over period t , let $\beta = (\beta_1, \dots, \beta_N)'$ denote the column vector of betas, and define $\mathbf{u}_t = \mathbf{R}_t - E[\mathbf{R}_t] - (R_{pt} - E[R_{pt}])\beta$. This leads to the regression representation

$$\mathbf{R}_t = \alpha + R_{pt}\beta + \mathbf{u}_t, \quad \text{with } E[\mathbf{u}_t] = \mathbf{0},$$

$$E[R_{pt}\mathbf{u}_t] = \mathbf{0}, \quad \text{and} \quad \alpha = E[\mathbf{R}_t] - E[R_{pt}]\beta, \quad (24)$$

as a consequence of the definition of β . Mean-variance efficiency of p implies $\alpha = \mathbf{0}$, which imposes testable restrictions on the data.

The traditional approach to testing mean-variance efficiency relies on the assumption that excess returns are temporally iid and multivariate normal. This assumption allows use of the Wald statistic for testing the null hypothesis $\alpha = \mathbf{0}$. As Gibbons, Ross, and Shanken (1989) showed, the Wald statistic follows an $F(N, N - T - 1)$ distribution in finite samples and an asymptotic chi-squared distribution with N degrees of freedom under the null hypothesis.

As MacKinlay and Richardson (1991) pointed out, it is important to relax the iid normal assumption. This is because there is overwhelming empirical evidence suggesting that returns are heteroscedastic and have fatter tails than those implied by the normal distribution. In contrast with the traditional approach, GMM provides an econometric framework that allows for testing mean-variance efficiency without the need to make strong distributional assumptions. It is only assumed that a riskless rate of return exists and that the excess asset returns are stationary and ergodic time series with finite fourth moments. Let $\{(R_{1t}, \dots, R_{Nt}, R_{pt}) : t = 1, \dots, T\}$ be a sample of T time series observations. Using (24), define $u_{it}(\alpha_i, \beta_i) = R_{it} - \alpha_i - \beta_i R_{pt}$ for all $i = 1, \dots, N$ and $t = 1, \dots, T$. Let $\delta = (\alpha_1, \beta_1, \dots, \alpha_N, \beta_N)'$ denote the $2N \times 1$ parameter vector, and define the following $2N \times 1$ vectors:

$$\mathbf{h}_t(\delta) = [u_{1t}(\alpha_1, \beta_1) \\ \times u_{1t}(\alpha_1, \beta_1)R_{pt} \dots u_{Nt}(\alpha_N, \beta_N) \\ \times u_{Nt}(\alpha_N, \beta_N)R_{pt}]' \quad (25)$$

and $\mathbf{g}_T(\delta) = \frac{1}{T} \sum_{t=1}^T \mathbf{h}_t(\delta)$. Note that (24) implies the moment condition $E[\mathbf{h}_t(\delta)] = \mathbf{0}$, which we exploit for estimation and testing using the GMM methodology. One approach is to estimate the unrestricted system and then test the hypothesis $\alpha = \mathbf{0}$ using the unrestricted estimates. The alternative approach is to impose the restriction $\alpha = \mathbf{0}$, estimate the restricted system, and then test for overidentifying restrictions. Next we briefly describe these two approaches.

Unrestricted Case. The sample analog of $E[\mathbf{h}_t(\delta)] = \mathbf{0}$ is the system of equations

$$\frac{1}{T} \sum_{t=1}^T u_{it}(\alpha_i, \beta_i) = \frac{1}{T} \sum_{t=1}^T u_{it}(\alpha_i, \beta_i)R_{pt} = 0, \quad i = 1, \dots, N. \quad (26)$$

Hence there are $2N$ equations and $2N$ unknown parameters, and the system is identified. The foregoing equations coincide with the normal equations from ordinary least squares (OLS), and thus this version of GMM is equivalent to OLS regression for each i . It follows from the general theory of GMM as developed by Hansen (1982) that the GMM estimator $\hat{\delta}$ is asymptotically normal with mean δ and covariance matrix $(\mathbf{D}_0' \mathbf{S}_0^{-1} \mathbf{D}_0)^{-1}$, where

$$\mathbf{D}_0 = E\left[\frac{\partial \mathbf{h}_t(\delta)}{\partial \delta'}\right] \quad \text{and} \quad \mathbf{S}_0 = \sum_{r=-\infty}^{+\infty} E[\mathbf{h}_t(\delta)(\mathbf{h}_{t-r}(\delta))']. \quad (27)$$

In practice, consistent estimates $\mathbf{D}_T$ and $\mathbf{S}_T$ are used instead of the unknown population quantities $\mathbf{D}_0$ and $\mathbf{S}_0$. Let ϕ_1 denote the test statistic for the hypothesis $\alpha = \mathbf{0}$. Then, under the null hypothesis,

$$\phi_1 = T \hat{\alpha}' [\mathbf{P}(\mathbf{D}_T' \mathbf{S}_T^{-1} \mathbf{D}_T)^{-1} \mathbf{P}]^{-1} \hat{\alpha} \quad (28)$$

has an asymptotic chi-squared distribution with N degrees of freedom, where $\mathbf{P} = \mathbf{I}_N \otimes [1 \ 0]$ and $\hat{\alpha} = \mathbf{P}\hat{\delta}$.

Restricted Case. Imposing the restriction $\alpha = \mathbf{0}$ yields

$$\mathbf{g}_T(\mathbf{0}, \beta) = \frac{1}{T} \sum_{t=1}^T \mathbf{h}_t(\mathbf{0}, \beta). \quad (29)$$

Hence there are $2N$ restrictions but only N unknown parameters, the system is overidentified, and not all of the sample moments can be set equal to 0. The GMM estimate of β , say $\hat{\beta}$, is obtained by minimizing the quadratic form $\mathbf{g}_T(\mathbf{0}, \beta)' \mathbf{S}_T^{-1} \mathbf{g}_T(\mathbf{0}, \beta)$, where $\mathbf{S}_T$ is a consistent estimator of $\mathbf{S}_0(\beta) = \sum_{r=-\infty}^{+\infty} E[\mathbf{h}_t(\mathbf{0}, \beta)(\mathbf{h}_{t-r}(\mathbf{0}, \beta))']$. It follows then from the general GMM theory that $\hat{\beta}$ has an asymptotically normal distribution with mean β and covariance matrix

$$(\mathbf{D}_0(\beta)'(\mathbf{S}_0(\beta))^{-1} \mathbf{D}_0(\beta))^{-1}, \quad \text{where } \mathbf{D}_0(\beta) = E\left[\frac{\partial \mathbf{h}_t(\mathbf{0}, \beta)}{\partial \beta'}\right]. \quad (30)$$

Finally, a test of the N overidentifying restrictions can be based on the test statistic

$$\phi_2 = T \mathbf{g}_T(\hat{\beta})' \mathbf{S}_T^{-1} \mathbf{g}_T(\hat{\beta}), \quad (31)$$

which, under the null hypothesis, asymptotically follows a chi-squared distribution with $2N - N = N$ degrees of freedom.

MacKinlay and Richardson (1991) demonstrated the danger of incorrect inference under the iid multivariate normal assumption when the true return distribution is multivariate t that exhibits contemporaneous conditional heteroscedasticity given the market return. The remedy to the misspecification problem is provided by the robust GMM procedure under which the test statistic has an asymptotic chi-squared distribution regardless of the true underlying data-generating process.

3.2.2. Conditional Linear Beta Pricing Models. The lack of empirical support for linear factor models in general and the CAPM in particular, coupled with mounting evidence of predictable time variations in the joint distribution of security returns, led to the empirical evaluation of conditional linear factor pricing models. Early work examining conditional asset-pricing models includes that of Hansen and Hodrick (1983), Gibbons and Ferson (1985), and Campbell (1987). More recently, Harvey (1989), Ferson and Harvey (1993), Ferson and Korajczyk (1995), and Ghysels (1998), among others, have examined variations of the conditional version of the linear beta pricing model given in (19), allowing $E[R_i]$, γ , λ_k 's, and β_{ik} 's to vary over time. For example, Harvey (1989) examined the following conditional version of the CAPM:

$$E[r_{i,t+1}|I_t] = \beta_{it} E(r_{m,t+1}|I_t), \quad (32)$$

where $r_{i,t+1}$ and $r_{m,t+1}$ denote the date $t+1$ excess returns on security i and the market; $\beta_{it} = \frac{\text{cov}(r_{i,t+1}, r_{m,t+1}|I_t)}{\text{var}(r_{m,t+1}|I_t)}$; and $E(\cdot|I_t)$, $\text{cov}(\cdot|I_t)$, and $\text{var}(\cdot|I_t)$ denote conditional expectation, conditional covariance, and conditional variance operators based on the information set I_t . Harvey assumed that $\frac{E(r_{m,t+1}|I_t)}{\text{var}(r_{m,t+1}|I_t)} = \gamma$. With this assumption and the definitions of β_{it} and conditional covariance, $\text{cov}(\cdot|I_t)$, (32) becomes

$$E(r_{i,t+1}|I_t) = \gamma E[r_{i,t+1}(r_{m,t+1} - E(r_{m,t+1}|I_t))|I_t]. \quad (33)$$

Suppose that $E(r_{m,t+1}|I_t) = \delta'_m z_{m,t}$, where $z_{m,t}$ denotes a L_m -dimensional vector of variables in the information set I_t at date t . Substituting this expression into (33) yields

$$E[r_{i,t+1}|I_t] = \gamma E[r_{i,t+1}(r_{m,t+1} - \delta'_m z_{m,t})|I_t]. \quad (34)$$

Define

$$u_{i,t+1} = r_{i,t+1} - \gamma[r_{i,t+1}(r_{m,t+1} - \delta'_m z_{m,t})]. \quad (35)$$

Let $z_{i,t}$ denote the L_i -dimensional vector of variables in the information set, I_t , at date t . This gives the following moment conditions for each security, i :

$$E[u_{i,t+1} z_{i,t}] = 0. \quad (36)$$

GMM can be applied to these moment conditions to estimate and test the conditional CAPM. Harvey (1991) empirically examined the cross-section of stock index returns across several countries using the world CAPM and taking a related approach. Ferson and Harvey (1993) empirically examined a multifactor extension of (32) using GMM. They generated the moment conditions by assuming that conditional betas

are fixed affine functions of variables in the date t information set I_t . Ghysels and Hall (1990) showed that the standard GMM tests for overidentifying restrictions tend to not reject the model specifications even when the assumptions regarding beta dynamics are wrong. Ghysels (1998) tested for parameter stability in conditional factor models using the sup-LM test proposed by Andrews (1993) and found evidence in favor of misspecified beta dynamics.

4. ALTERNATIVE WEIGHTING MATRICES

An asset-pricing model typically implies a number of moment conditions. The number of model parameters in general will be much less than the number of moment conditions. GMM chooses a subset of the possible linear combinations of the moment conditions and picks the parameter values that make them hold exactly in the sample. The moment conditions are chosen so as to maximize asymptotic estimation efficiency. However, in some cases the econometrician may have some a priori information on which moment conditions contain relatively more information than others. In those cases it may be advisable to bring in the prior information available to the econometrician through an appropriate choice of the weighting matrix. For example, Eichenbaum et al. (1988) suggested prespecifying the subset of moment conditions to be used for estimation and testing the model using the additional moment conditions. This would correspond to choosing the weighting matrix with 0s and 1s as entries to pick out the moment conditions that should be used in the estimation process.

There is another reason for prespecifying the weighting matrix used. When making comparisons across models, it is often tempting to compare the minimized value of the criterion function, $T \mathbf{g}_T(\hat{\theta}_T)' \mathbf{S}_T^{-1} \mathbf{g}_T(\hat{\theta}_T)$, across the models. One model may do better than another not because the vector of average pricing errors, $\mathbf{g}_T$, associated with it are smaller, but rather because the inverse of the optimal weighting matrix, $\mathbf{S}_T$, associated with it is larger. To overcome this difficulty, Hansen and Jagannathan (1997) suggested examining the pricing error of the most mispriced portfolio after normalizing for the "size" of the portfolio. This corresponds to using the inverse of the second moment matrix of returns, $\mathbf{A} = (E[\mathbf{R}_t \mathbf{R}_t'])^{-1} \equiv \mathbf{G}^{-1}$, as the weighting matrix under the assumption that $\mathbf{G}$ is positive definite. Cochrane (1996) suggested using the identity matrix as the weighting matrix. Hansen and Jagannathan (1997) showed that $\text{dist}(\theta) = \sqrt{E[\mathbf{w}_t(\theta)' \mathbf{G}^{-1} \mathbf{w}_t(\theta)]}$ equals the least-squares distance between the candidate SDF and the set of all valid discount factors. Further, they showed that $\text{dist}(\theta)$ is equal to the maximum pricing error per unit norm on any portfolio of the N assets.

We now illustrate use of the nonoptimal matrix $\mathbf{G}^{-1}$ in the context of SDFs that are linear in a number of factors following Jagannathan and Wang (1996). Suppose that the candidate SDF is of the form $m_t^c(\theta) = \mathbf{F}_t' \theta$, where $\theta = (\theta_0, \dots, \theta_{K-1})'$ and $\mathbf{F}_t = (1, F_{1,t}, \dots, F_{K-1,t})'$ is the time t factor realization

vector. Let

$$\begin{aligned} \mathbf{D}_T &= \frac{1}{T} \sum_{t=1}^T \mathbf{R}_t \mathbf{F}_t' \xrightarrow{P} E[\mathbf{R}_t \mathbf{F}_t'] \equiv \mathbf{D} \quad \text{and} \\ \mathbf{G}_T &= \frac{1}{T} \sum_{t=1}^T \mathbf{R}_t \mathbf{R}_t' \xrightarrow{P} E[\mathbf{R}_t \mathbf{R}_t'] \equiv \mathbf{G}. \end{aligned} \quad (37)$$

Then $\mathbf{w}_t(\theta) = m_t^c(\theta) \mathbf{R}_t - \mathbf{1} = (\mathbf{R}_t \mathbf{F}_t') \theta - \mathbf{1}$, and so $\bar{\mathbf{w}}_T(\theta) = \mathbf{D}_T \theta - \mathbf{1}$. The GMM estimate $\hat{\theta}_T$ of θ is the solution to $\min_{\theta} \bar{\mathbf{w}}_T(\theta)' \mathbf{G}_T^{-1} \bar{\mathbf{w}}_T(\theta)$. The corresponding first-order condition is $\mathbf{D}_T' \mathbf{G}_T^{-1} \bar{\mathbf{w}}_T(\hat{\theta}_T) = \mathbf{0}$, from which we obtain $\hat{\theta}_T = (\mathbf{D}_T' \mathbf{G}_T^{-1} \mathbf{D}_T)^{-1} \mathbf{D}_T' \mathbf{G}_T^{-1} \mathbf{1}$. Because the weighting matrix $\mathbf{G}_T^{-1}$ differs from the optimal choice in the sense of Hansen (1982), the asymptotic distribution of $T \bar{\mathbf{w}}_T(\hat{\theta}_T)' \mathbf{G}_T^{-1} \bar{\mathbf{w}}_T(\hat{\theta}_T)$ will not be a chi-squared distribution. Suppose that for some θ_0 , we have $\sqrt{T} \bar{\mathbf{w}}_T(\theta_0) \xrightarrow{D} N(\mathbf{0}, \mathbf{S}_0)$, where $\mathbf{S}_0$ is a positive definite matrix, and also that the $N \times K$ matrix $\mathbf{D} = E[\mathbf{R}_t \mathbf{F}_t']$ has rank K and the matrix $\mathbf{G} = E[\mathbf{R}_t \mathbf{R}_t']$ is positive definite. Let

$$\begin{aligned} \mathbf{A} &= \mathbf{S}^{1/2} \mathbf{G}^{-1/2} [\mathbf{I}_N - (\mathbf{G}^{-1/2}) \mathbf{D} (\mathbf{D}' \mathbf{G}^{-1} \mathbf{D})^{-1} \mathbf{D}' \mathbf{G}^{-1/2}] \\ &\quad \times (\mathbf{G}^{-1/2})' (\mathbf{S}^{1/2})', \end{aligned} \quad (38)$$

where $\mathbf{S}^{1/2}$ and $\mathbf{G}^{1/2}$ are the upper triangular matrices from the Cholesky decomposition of $\mathbf{S}$ and $\mathbf{G}$, and $\mathbf{I}_N$ is the $N \times N$ identity matrix. Jagannathan and Wang (1996) showed that $\mathbf{A}$ has exactly $N - K$ nonzero eigenvalues that are positive and denoted by $\lambda_1, \dots, \lambda_{N-K}$, and the asymptotic distribution of $\text{dist}(\theta)$ is given by

$$T \bar{\mathbf{w}}_T(\hat{\theta}_T)' \mathbf{G}_T^{-1} \bar{\mathbf{w}}_T(\hat{\theta}_T) \xrightarrow{D} \sum_{j=1}^{N-K} \lambda_j v_j \quad \text{as } T \rightarrow \infty, \quad (39)$$

where $v_1, \dots, v_{N-K}$ are independent chi-squared random variables with one degree of freedom.

Using GMM with the weighting matrix suggested by Hansen and Jagannathan (1997) and the sampling theory derived by Jagannathan and Wang (1996), Hodrick and Zhang (2001) evaluated the specification errors of several empirical asset-pricing models that have been developed as potential improvements on the CAPM. On a common set of assets, they showed that all the recently proposed multifactor models can be statistically rejected.

5. EFFICIENCY OF GENERALIZED METHOD OF MOMENTS FOR STOCHASTIC DISCOUNT FACTOR MODELS

As discussed in the preceding sections, GMM is useful because it can be applied to the Euler equations of dynamic asset-pricing models, which are SDF representations. A SDF has the property that the value of a financial asset equals the expected value of the product of the payoff on the asset and the SDF. An asset-pricing model identifies a particular SDF that is a function of observable variables and model parameters. For example, a linear factor pricing model identifies a specific linear function of the factors as a SDF. The GMM-SDF method involves using the GMM to estimate the SDF representations of asset-pricing models. The GMM-SDF method has become

common in the recent empirical finance literature. It is sufficiently general so it can be used for analysis of both linear and nonlinear asset-pricing models, including pricing models for derivative securities.

Despite its wide use, there have been concerns that, compared to the classical ML-beta, the generality of the GMM-SDF method comes at the cost of efficiency in parameter estimation and power in specification tests. For this reason, researchers compare the GMM-SDF method with the GMM-beta method for linear factor pricing models. For such models, the GMM-beta method is equivalent to the ML-beta method under suitable assumptions about the statistical properties of returns and factors. On the one hand, if the GMM-SDF method turns out to be inefficient relative to the GMM-beta method for linear models, then some variation of the ML-beta method may well dominate the GMM-SDF method for nonlinear models as well, in terms of estimation efficiency. On the other hand, if the GMM-SDF method is as efficient as the beta method, then it is the preferred method because of its generality.

Kan and Zhou (1999) made the first attempt to compare the GMM-SDF and GMM-beta methods for estimating the parameters related to the factor risk premium. Unfortunately, their comparison is inappropriate. They ignored the fact that the risk premium parameters in the SDF and beta representations are not identical and directly compared the asymptotic variances of the two estimators by assuming that the risk premium parameters in the two methods take special and equal values. For that purpose, they made the simplifying assumption that the economy-wide pervasive factor can be standardized to have zero mean and unit variance. Based on their special assumption, they argued that the GMM-SDF method is far less efficient than the GMM-beta method. The sampling error in the GMM-SDF method is about 40 times as large as that in the GMM-beta method. They also concluded that the GMM-SDF method is less powerful than the GMM-beta method in specification tests.

Kan and Zhou's (1999) comparison, as well as their conclusion about the relative inefficiency of the SDF method, is inappropriate for two reasons. First, it is incorrect to ignore the fact that the risk premium measures in the two methods are not identical, even though they are equal at certain parameter values. Second, the assumption that the factor can be standardized to have zero mean and unit variance is equivalent to the assumption that the factor mean and variance are known or predetermined by the econometricians. By making that assumption, Kan and Zhou give an informational advantage to the GMM-beta method but not to the GMM-SDF method.

For a proper comparison of the two methods, it is necessary to incorporate explicitly the transformation between the risk premium parameters in the two methods and the information about the mean and variance of the factor while estimating the risk premium. When this is done, the GMM-SDF method is asymptotically as efficient as the GMM-beta method, as established by Jagannathan and Wang (2002). These authors found that asymptotically, the GMM-SDF method provides as precise an estimate of the risk premium as the GMM-beta method. Using Monte Carlo simulations, they demonstrated

that the two methods provide equally precise estimates in finite samples as well. The sampling errors in the two methods are similar even under nonnormal distribution assumptions, which allow conditional heteroscedasticity. Therefore, linearizing nonlinear asset-pricing models and estimating risk premiums using the GMM-beta or ML-beta method will not lead to increased estimation efficiency.

Jagannathan and Wang (2002) also examined the specification tests associated with the two methods. An intuitive test for model misspecification is to examine whether the model assigns the correct expected return to every asset; that is, whether the vector of pricing errors for the model is 0. They show that the sampling analog of pricing errors has smaller asymptotic variance in the GMM-beta method. However, this advantage of the beta method does not appear in the specification tests based on Hansen's (1982) J statistics. Jagannathan and Wang demonstrate that the GMM-SDF method has the same power as the GMM-beta method.

6. STOCHASTIC PROCESS FOR SHORT-TERM INTEREST RATES

Continuous-time models are widely used in finance, especially for valuing contingent claims. In such models, the contingent claim function that gives the value of the claim as a function of its characteristics depends on the parameters describing the stochastic process driving the prices of underlying securities. In view of this, a vast literature on estimating continuous-time models has evolved; refer to, for example, the work by Ait-Sahalia, Hansen, and Scheinkman (2001) for a discussion of traditional methods; Gallant and Tauchen (2001) for a discussion of the efficient method of moments; and Garcia, Ghysels, and Renault (2001) for a discussion of the literature in the context of contingent claim valuation. Although GMM may not be the preferred method for estimating model parameters, it is easy to implement and provides estimates that can be used as starting values in other methods. In what follows we provide a brief introduction to estimating continuous-time model parameters using GMM.

6.1 Discrete Time Approximations of Continuous-Time Models

Here we present an application of GMM to testing and comparing alternative continuous-time models for the short-rate interest rate as given by Chan, Karolyi, Longstaff, and Sanders (1992). Other authors who have used GMM in empirical tests of interest rate models include Gibbons and Ramaswamy (1993), who tested the Cox-Ingersoll-Ross (CIR) model; Harvey (1988); and Longstaff (1989).

The dynamics for the short-term riskless rate, denoted by r_t , as implied by a number of continuous-time term structure models, can be described by the stochastic differential equation

$$dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dW_t, \quad (40)$$

where W_t is standard Brownian motion process. The foregoing specification nests some of the most widely used models of the short-term rate. The case where $\beta = \gamma = 0$ (i.e.,

Brownian motion with drift) is the model used by Merton (1973) to derive discount bond prices. The case where $\gamma = 0$ (the so-called Ornstein-Uhlenbeck process) corresponds to the model used by Vasicek (1977) to derive an equilibrium model of discount bond prices. The case where $\gamma = 1/2$ (the so-called Feller or square-root process) is the specification that Cox, Ingersoll, and Ross (1985) used to build a single-factor general equilibrium term structure model. Other special cases of (40) as short-term rate specifications have been given by, for instance, Dothan (1978), Brennan and Schwartz (1980), Cox et al. (1980), Courtadon (1982), and Marsh and Rosenfeld (1983).

Following Brennan and Schwartz (1982) and Sanders and Unal (1988), Chan et al. (1992) used the following discrete-time approximation of (40) to facilitate the estimation of the parameters of the continuous-time model:

$$r_{t+1} - r_t = \alpha + \beta r_t + \varepsilon_{t+1}, \quad \text{with } E_t[\varepsilon_{t+1}] = 0 \\ \text{and } E_t[\varepsilon_{t+1}^2] = \sigma^2 r_t^{2\gamma}, \quad (41)$$

where $E_t[\cdot]$ denotes the conditional expectation given the information set at time t . It should be emphasized that GMM naturally provides a suitable econometric framework for testing the validity of the model (41). Application of ML techniques in this setting would be problematic, because the distribution of the increments critically depends on the value of γ . For example, if $\gamma = 0$, then the changes are normally distributed, whereas if $\gamma = 1/2$, then they follow a gamma distribution. A brief description of the GMM test follows.

Let θ denote the parameter vector comprising α, β, σ^2 , and γ . The structure of the approximating discrete-time model in (41) implies that under the null hypothesis, the moment restrictions $E[\mathbf{f}_t(\theta)] = \mathbf{0}$ hold, where

$$\mathbf{f}_t(\theta) = \begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_{t+1} r_t \\ \varepsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma} \\ (\varepsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma}) r_t \end{bmatrix}. \quad (42)$$

Given a dataset $\{r_t: t=1, \dots, T+1\}$ of $T+1$ observations, following the standard GMM methodology, let $\mathbf{g}_T(\theta) = \frac{1}{T} \sum_{t=1}^T \mathbf{f}_t(\theta)$. The GMM estimate $\hat{\theta}$ is then obtained by minimizing the quadratic form $J_T(\theta) = \mathbf{g}_T(\theta)' \mathbf{W} \mathbf{g}_T(\theta)$, where $\mathbf{W}$ is a positive definite symmetric weighting matrix, or, equivalently, by solving the system of equations $\mathbf{D}_0(\theta)' \mathbf{W} \mathbf{g}_T(\theta) = \mathbf{0}$, where $\mathbf{D}_0(\theta) = \frac{\partial \mathbf{g}_T}{\partial \theta}(\theta)$. The optimal choice of the weighting matrix in terms of minimizing the asymptotic covariance matrix of the estimator is $\mathbf{W} = (\mathbf{S}_0(\theta))^{-1}$, where $\mathbf{S}_0(\theta) = \sum_{r=-\infty}^{+\infty} E[\mathbf{f}_t(\theta)(\mathbf{f}_{t-r}(\theta))']$. The asymptotic distribution of the GMM estimator $\hat{\theta}$ is then normal with mean θ and covariance matrix consistently estimated by $(\mathbf{D}_0(\hat{\theta})' \mathbf{S}_T^{-1} \mathbf{D}_0(\hat{\theta}))^{-1}$, where $\mathbf{S}_T$ is a consistent estimator of $\mathbf{S}_0(\theta)$. In the case of the unrestricted model, we test for the significance of the individual parameters using the foregoing asymptotic distribution. If we restrict any of the four parameters, then we can test the validity of the model using the test statistic $T \mathbf{g}_T(\hat{\theta})' \mathbf{S}_T^{-1} \mathbf{g}_T(\hat{\theta})$, which is asymptotically distributed as a chi-squared random

variable with $4-k$ degrees of freedom, where k is the number of parameters to be estimated. Alternatively, we can test the restrictions imposed on the model by using the hypothesis tests developed by Newey and West (1987). Suppose that the null hypothesis representing the restrictions that we wish to test is of the form $H_0: a(\theta) = \mathbf{0}$, where $a(\cdot)$ is a vector function of dimension k . Let $J_T(\hat{\theta})$ and $J_T(\tilde{\theta})$ denote the restricted and unrestricted objective functions for the optimal GMM estimator. The test statistic $R = T[J_T(\tilde{\theta}) - J_T(\hat{\theta})]$ then follows asymptotically a chi-squared distribution with k degrees of freedom. This test provides a convenient tool for making pairwise comparisons among the several model candidates.

6.2 The Hansen–Scheinkman Test Function Method

Hansen and Scheinkman (1995) showed how to derive moment conditions for estimating continuous-time diffusion processes without using discrete-time approximations. Whereas other methods, such as the efficient method of moments, may be more efficient and may be preferred to GMM, they are computationally much more demanding than GMM. In view of this, it may be advisable to obtain initial estimates using GMM and use them as starting values in other more efficient methods to save computational effort.

In what follows, based on work of Sun (1997), we illustrate the Hansen–Scheinkman method for estimating the model parameters when the short rate evolves according to

$$dr_t = \mu(r_t)dt + \sigma(r_t)dW_t, \quad (43)$$

where W_t is the standard Brownian motion and $\mu(\cdot)$ and $\sigma(\cdot)$ are the drift and diffusion of the process assumed to satisfy the necessary regularity conditions to guarantee the existence and uniqueness of a solution process. Assume that the short rate process described by (43) has a stationary marginal distribution. The stationarity of the short rate r_t implies that for any time t , $E[\phi(r_t)] = c$, where ϕ is a smooth function and c is a constant. Taking differences of the foregoing equation between two arbitrary points t and $T > t$, and applying Ito's lemma, we obtain

$$0 = E \left[\int_t^T \left[\mu(r_s)\phi'(r_s) + \frac{1}{2}\sigma(r_s)\phi''(r_s) \right] ds \right] + E \left[\int_t^T \sigma(r_s)\phi'(r_s)dW_s \right]. \quad (44)$$

By the martingale property of the Ito integral, the second expectation vanishes at 0 as long as $f(t) = \sigma(r_t)\phi'(r_t)$ is square integrable. Applying Fubini's theorem to the first expectation in (44) and then using the stationarity property yields

$$E \left[\mu(r_t)\phi'(r_t) + \frac{1}{2}\sigma(r_t)\phi''(r_t) \right] = 0. \quad (45)$$

This is the first class of Hansen–Scheinkman moment conditions. The space of test functions ϕ for which (45) is well defined can be given a precise description (see Hansen,

Scheinkman, and Touzi (1996)). To illustrate the use of (45), consider the CIR square root process

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t, \quad a > 0, b > 0. \quad (46)$$

Letting the test function ϕ equal x^k , $k = 1, 2, 3$, we obtain from (45) that

$$\begin{aligned} E[a(b - r_t)] &= E[2a(b - r_t)r_t + \sigma^2 r_t] \\ &= E[2a(b - r_t)r_t^2 + \sigma^2 r_t^2] = 0, \end{aligned}$$

implying that the first three moments of r_t are given by b , $b^2 + \frac{\sigma^2 b}{2a}$, and $b^3 + \frac{3\sigma^2 b^2}{2a} + \frac{\sigma^4 b}{2a^2}$.

The first class of moments conditions in (45) uses only information contained in the stationary marginal distribution. Because the evolution of the short rate process is governed by the conditional distribution, (45) alone would not be expected to produce efficient estimates in finite samples, especially if there is strong persistence in the data. We now consider the second class of moment conditions, which uses the information contained in both the conditional and the marginal distributions.

Kent (1978) showed that a stationary scalar diffusion process is characterized by reversibility. This says that if the short rate process $\{r_t\}$ is modeled as a scalar stationary diffusion process, then the conditional density of r_t given r_0 is the same as that of r_{-t} given r_0 . The reversibility of r_t implies that for any two points in time t and s , $E[\psi(r_t, r_s)] = E[\psi(r_s, r_t)]$, where $\psi: \mathbb{R}^2 \rightarrow \mathbb{R}$ is a smooth function in both arguments. Using this equality twice and subtracting yields $E[\psi(r_{t+\Delta t}, r_s)] - E[\psi(r_t, r_s)] = E[\psi(r_s, r_{t+\Delta t})] - E[\psi(r_s, r_t)]$. Now applying Ito's lemma to both sides of the last equation, using the martingale property of the Ito integral and applying Fubini's theorem, we have

$$\begin{aligned} E \left[\mu(r_t)(\psi_1(r_t, r_s) - \psi_2(r_s, r_t)) \right. \\ \left. + \frac{1}{2}\sigma(r_t)(\psi_{11}(r_t, r_s) - \psi_{22}(r_s, r_t)) \right] = 0. \quad (47) \end{aligned}$$

This is the second class of moment conditions, with ψ as the test function. As an illustration, we note that for the case of the CIR square-root process and for $\psi(x, y)$ being equal to $(x - y)^3$, (47) translates to $E[a(b - r_t)(r_t - r_s)^2 + \sigma^2 r_t(r_t - r_s)] = 0$.

Appropriate choices of test functions (45) and (47) generate moment restrictions directly implied from a scalar stationary diffusion. The first class, (45), is a restriction on unconditional moments, whereas the second class, (47), is a joint restriction on conditional and unconditional moments. The two classes of moment conditions were derived by using only stationarity, reversibility, and Ito's lemma, without knowing any explicit forms of the conditional and unconditional distributions. Because times t and s in (45) and (47) are arbitrary, the two classes of moment conditions take into explicit account that we observe the short rate only at discrete intervals. This makes the Hansen–Scheinkman moment conditions natural choices in GMM estimation of the continuous-time short rate process using the discretely sampled data. GMM

estimation requires that the underlying short rate process be ergodic to approximate the time expectations with their sample counterparts. The central limit theorem is also required to assess the magnitude of these approximation errors. Hansen and Scheinkman (1995) proved that under appropriate regularity conditions, these GMM assumptions are indeed satisfied.

Although (45) and (47) provide moment conditions once test functions ϕ and ψ have been specified, the choice of appropriate test functions remains a critical issue. To illustrate the point, we consider the linear drift and constant volatility elasticity model of Chan et al. (1992), $dr_t = a(b - r_t)dt + \sigma r_t^\gamma dW_t$, where a, b, σ , and γ are the speed of adjustment, the long-run mean, the volatility coefficient, and the volatility elasticity. It is apparent that (45) and (47) can identify a and σ^2 only up to a common scale factor. This problem may be eliminated by making test functions depend on model parameters. When estimating the model of Chan et al. (1992), consider the test functions $\phi_1(y) = \int^y x^{-a} dx$ and $\phi_2(y) = \int^y x^{-\sigma} dx$ for the class (45). It can be shown that these are indeed valid test functions in the appropriate domain. In a related work, Conley, Hansen, Luttmer, and Scheinkman (1997) found that the score functions for the implied stationary density—that is, $\phi_3(y) = \int^y x^{-2\gamma} dx$ and $\phi_4(y) = \int^y x^{-2\gamma+1} dx$ —are efficient test functions for (45). They also used the cumulative function of the standard normal distribution

$$\psi(r_s, r_t) = \int^{r_t - r_s} \exp\left(-\frac{x^2}{2\delta^2}\right) dx \quad (48)$$

as the test function in (47), where δ is a scaling constant. The combination of $\phi_1, \phi_2, \phi_3, \phi_4$, and ψ yields the moment conditions $E[\mathbf{f}_{\text{HS}}(t, \theta)] = \mathbf{0}$, where

$$\mathbf{f}_{\text{HS}}(t, \theta) = \begin{bmatrix} a(b - r_{t+1})r_{t+1}^{-2\gamma} - \gamma\sigma^2 r_{t+1}^{-1} \\ a(b - r_{t+1})r_{t+1}^{-2\gamma+1} - (\gamma - \frac{1}{2})\sigma^2 \\ (b - r_{t+1})r_{t+1}^{-a} - \frac{1}{2}\sigma^2 r_{t+1}^{2\gamma-a-1} \\ a(b - r_{t+1})r_{t+1}^{-\sigma} - \frac{1}{2}\sigma^3 r_{t+1}^{2\gamma-\sigma-1} \\ \left[a(b - r_{t+1}) - \frac{r_{t+1} - r_t}{2\delta^2} \sigma^2 r_t^{2\gamma} \right] \exp\left[-\frac{(r_{t+1} - r_t)^2}{2\delta^2}\right] \end{bmatrix}.$$

Then GMM estimation of the model of Chan et al. (1992) proceeds in the standard way using these moment conditions.

7. MARKET MICROSTRUCTURE

A substantial portion of the literature in this area focuses on understanding why security prices change and why transactions prices depend on the quantity traded. Huang and Stoll (1994) developed a two-equation time series model of quote revision and transactions returns and evaluated the relative importance of different theoretical microstructure models proposed in the literature using GMM. Madhavan, Richardson, and Roomans (1997) used GMM to estimate and test a structural model of intraday price formation that allows for public information shocks and microstructure effect to understand why a U-shaped pattern is observed in intraday bid-ask spreads and volatility. Huang and Stoll (1997) estimated the different components of the bid-ask spread for 20 stocks in

the major market index using transactions data by applying GMM to a time series microstructure model.

There are several other market microstructure applications of GMM—too many to discuss in a short article of this nature. For example, Foster and Viswanathan (1993) conjectured that adverse selection would be relatively more severe on Mondays than on other days of the week. This implies that volume would be relatively low on Mondays. Because of the presence of conditional heteroscedasticity and serial correlation in the data, these authors used GMM to examine this hypothesis. The interested reader is referred to the surveys of this literature by Biais, Glosten, and Spatt (2002) and Madhavan (2000).

8. SUMMARY

GMM is one of the most widely used tools in financial applications, especially in the asset-pricing area. In this article we have provided several examples illustrating the use of GMM in the empirical asset-pricing literature in finance.

In most asset-pricing models, the value of a financial claim equals the expected discounted present value of its future payoffs. The models differ from one another in the stand that they take regarding which discount factor to use. An econometrician typically has a strictly smaller information set than investors who actively participate in financial markets. Hence the value computed using a given asset-pricing model based on the information available to the econometrician in general would not equal the market price of a financial claim. However, the difference between the observed market price and the value computed by the econometrician using an asset-pricing model should be uncorrelated with information available to the econometrician when investors have rational expectations.

By a judicious choice of instruments available in the econometrician's information set, we obtain a set of moment conditions that can be used to estimate the model parameters using GMM. The number of moment conditions in general would be greater than the number of model parameters. The overidentifying restrictions provide a natural test for model misspecification using the GMM J statistic.

Continuous-time models are used extensively in finance, especially for valuing contingent claims. These models value a contingent claim using arbitrage arguments, taking the stochastic process for the prices of the primitive assets as exogenous. This gives the value of a contingent claim as a function of the prices of underlying assets and the parameters of the stochastic process determining their evolution over time. Hence estimating the parameters of continuous-time stochastic processes describing the dynamics of primitive asset prices has received wide attention in the literature. We discussed an example showing how GMM can be used for that purpose. The GMM estimates thus obtained can be used as starting values in more efficient estimation methods.

The economic models often examined in empirical studies in finance imply moment conditions that can be used in a natural way for estimation and testing the models using GMM. This, combined with the fact that GMM does not require strong distributional assumptions, has led to its widespread use in other areas of finance as well. We cannot possibly discuss all of the numerous interesting applications of GMM in finance, here, and new applications continue to appear.

ACKNOWLEDGMENTS

The authors thank Eric Ghysels and Joel Hasbrouck for their valuable comments and suggestions.

[Received May 2002. Revised June 2002.]

REFERENCES

- Abel, A. (1990), "Asset Prices Under Habit Formation and Catching Up With the Joneses," *American Economic Review*, 80, 38–42.
- Ait-Sahalia, Y., Hansen, L. P., and Scheinkman, J. (2003), "Discretely-Sampled Diffusions," in *Handbook of Financial Econometrics*, eds. Y. Ait-Sahalia and L. P. Hansen, forthcoming.
- Andrews, D. W. K. (1991), "Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.
- (1993), "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61, 821–856.
- Andrews, D. W. K., and Monahan, J. C. (1992), "An Improved Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60, 953–966.
- Bansal, R., Hsieh, D. A., and Viswanathan, S. (1993), "A New Approach to International Arbitrage Pricing," *Journal of Finance*, 48, 1719–1747.
- Bansal, R., and Viswanathan, S. (1993), "No Arbitrage and Arbitrage Pricing: A New Approach," *Journal of Finance*, 48, 1231–1262.
- Bekaert, G., and Urias, M. S. (1996), "Diversification, Integration and Emerging Market Closed-End Funds," *Journal of Finance*, 51, 835–869.
- Biais, B., Glosten, L. R., and Spatt, C. S. (2002), "The Microstructure of Stock Markets," Discussion Paper 3288, CEPR.
- Brennan, M. J., and Schwartz, E. S. (1980), "Analyzing Convertible Bonds," *Journal of Financial and Quantitative Analysis*, 15, 907–929.
- (1982), "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency," *Journal of Financial and Quantitative Analysis*, 17, 75–100.
- Campbell, J. Y. (1987), "Stock Returns and the Term Structure," *Journal of Financial Economics*, 18, 373–399.
- Campbell, J. Y., and Cochrane, J. H. (1999), "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.
- Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.
- Chan, K. C., Karolyi, G. A., Longstaff, F. A., and Sanders, A. B. (1992), "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 1209–1227.
- Cochrane, J. H. (2001), *Asset Pricing*, Princeton, NJ: Princeton University Press.
- Cochrane, J. H. (1996), "A Cross-Sectional Test of an Investment-Based Asset Pricing Model," *Journal of Political Economy*, 104, 572–621.
- Conley, T., Hansen, L. P., Luttmer, E., and Scheinkman, J. A. (1997), "Short-Term Interest Rates as Subordinated Diffusions," *Review of Financial Studies*, 10, 525–577.
- Connor, G. (1984), "A Unified Beta Pricing Theory," *Journal of Economic Theory*, 34, 13–31.
- Constantinides, G. (1982), "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98, 519–543.
- Courtadon, G. (1982), "The Pricing of Options of Default-Free Bonds," *Journal of Financial and Quantitative Analysis*, 17, 75–100.
- Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1980), "An Analysis of Variable Rate Loan Contracts," *Journal of Finance*, 35, 389–403.
- (1985), "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–407.
- Dothan, U. L. (1978), "On the Term Structure of Interest Rates," *Journal of Financial Economics*, 6, 59–69.
- Dunn, K. B., and Singleton, K. J. (1986), "Modelling the Term Structure of Interest Rates Under Nonseparable Utility and Durability of Goods," *Journal of Financial Economics*, 17, 27–55.
- Dybvig, P. H., and Ingersoll, J. E. (1982), "Mean-Variance Theory in Complete Markets," *Journal of Business*, 55, 233–252.
- Eichenbaum, M., Hansen, L. P., and Singleton, K. J. (1988), "A Time-Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertainty," *Quarterly Journal of Economics*, 103, 51–78.
- Epstein, L., and Zin, S. (1991), "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: An Empirical Investigation," *Journal of Political Economy*, 99, 263–286.
- Ferson, W. E., and Constantinides, G. (1991), "Habit Persistence and Durability in Aggregate Consumption: Empirical Tests," *Journal of Financial Economics*, 36, 29–55.
- Ferson, W. E., and Harvey, C. R. (1993), "The Risk and Predictability of International Equity Returns," *Review of Financial Studies*, 6, 527–566.
- Ferson, W. E., and Korajczyk, R. A. (1995), "Do Arbitrage Pricing Models Explain the Predictability of Stock Returns?" *Journal of Business*, 68, 309–349.
- Foster, D., and Viswanathan, S. (1993), "Variations in Trading Volumes, Return Volatility, and Trading Costs: Evidence in Recent Price Formation Models," *Journal of Finance*, 48, 187–211.
- Gallant, R., and Tauchen, G. (2003), "Simulated Score Methods and Indirect Inference for Continuous-Time Models," in *Handbook of Financial Econometrics*, eds. Y. Ait-Sahalia and L. P. Hansen, forthcoming.
- Garcia, R., Ghysels, E., and Renault, E. (2003), "Option Pricing Models," in *Handbook of Financial Econometrics*, eds. Y. Ait-Sahalia and L. P. Hansen, forthcoming.
- Ghysels, E. (1998), "On Stable Factor Structures in the Pricing of Risk: Do Time-Varying Betas Help or Hurt?" *Journal of Finance*, 53, 549–574.
- Ghysels, E., and Hall, A. (1990), "Are Consumption-Based Intertemporal Capital Asset Pricing Models Structural?" *Journal of Econometrics*, 45, 121–139.
- Gibbons, M. R., and Ferson, W. E. (1985), "Testing Asset Pricing Models With Changing Expectations and an Unobservable Market Portfolio," *Journal of Financial Economics*, 14, 217–236.
- Gibbons, M. R., and Ramaswamy, K. (1993), "A Test of the Cox, Ingersoll and Ross Model of the Term Structure," *Review of Financial Studies*, 6, 619–658.
- Gibbons, M., Ross, S. A., and Shanken, J. (1989), "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 1121–1152.
- Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.
- Hansen, L. P., and Hodrick, R. J. (1980), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, 829–853.
- (1983), "Risk Averse Speculation in the Forward Foreign Exchange Market: An Econometric Analysis of Linear Models," in *Exchange Rates and International Macroeconomics*, ed. J. Frenkel, Chicago: University of Chicago Press.
- Hansen, L. P., and Jagannathan, R. (1997), "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Finance*, 62, 557–590.
- Hansen, L. P., and Richard, S. F. (1987), "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models," *Econometrica*, 55, 587–613.
- Hansen, L. P., and Scheinkman, J. (1995), "Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes," *Econometrica*, 63, 767–804.
- Hansen, L. P., Scheinkman, J., and Touzi, N. (1998), "Spectral Methods for Identifying Scalar Diffusions," *Journal of Econometrics*, 86, 1–32.
- Hansen, L. P., and Singleton, K. J. (1982), "Generalized Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.
- Harvey, C. R. (1988), "The Real Term Structure and Consumption Growth," *Journal of Financial Economics*, 22, 305–333.
- (1989), "Time-Varying Conditional Covariances in Tests of Asset Pricing Models," *Journal of Financial Economics*, 24, 289–318.
- (1991), "The World Price of Covariance Risk," *Journal of Financial Economics*, 46, 111–157.
- Hodrick, R., and Zhang, X. (2001), "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Financial Economics*, 62, 327–376.
- Huang, R. D., and Stoll, H. R. (1994), "Market Microstructure and Stock Return Predictions," *Review of Financial Studies*, 7, 179–213.
- (1997), "The Components of the Bid-Ask Spread: A General Approach," *Review of Financial Studies*, 10, 995–1034.
- Huberman, G., and Kandel, S. (1987), "Mean-Variance Spanning," *Journal of Finance*, 32, 339–346.
- Jagannathan, R., and Wang, Z. (1996), "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance*, 51, 3–53.
- (2002), "Empirical Evaluation of Asset Pricing Models: A Comparison of the SDF and Beta Methods," *Journal of Finance*, forthcoming.
- Kan, R., and Zhou, G. (1999), "A Critique of the Stochastic Discount Factor Methodology," *Journal of Finance*, 54, 1221–1248.
- Kent, J. (1978), "Time-Reversible Diffusions," *Advances in Applied Probability*, 10, 819–835.
- Longstaff, F. A. (1989), "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates," *Journal of Financial Economics*, 23, 195–224.
- MacKinlay, A. C., and Richardson, M. P. (1991), "Using Generalized Method of Moments to Test Mean-Variance Efficiency," *Journal of Finance*, 46, 551–527.
- Madhavan, A. (2000), "Market Microstructure: A Survey," *Journal of Financial Markets*, 3, 205–258.

- Madhavan, A., Richardson, M., and Roomans, M. (1997), "Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks," *Review of Financial Studies*, 10, 1035–1064.
- Marsh, T. A., and Rosenfeld, E. R. (1983), "Stochastic Process for Interest Rates and Equilibrium Bond Prices," *Journal of Finance*, 38, 635–646.
- Merton, R. C. (1973), "The Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4, 141–183.
- Newey, W., and West, K. (1987), "Hypothesis Testing With Efficient Method of Moments Estimation," *International Economic Review*, 28, 777–787.
- Ross, S. A. (1976), "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.
- Rubinstein, M. (1976), "The Valuation of Uncertain Income Streams and the Pricing of Options," *Bell Journal of Economics*, 7, 407–425.
- Sanders, A. B., and Unal, H. (1988), "On the Intertemporal Stability of the Short-Term Rate of Interest," *Journal of Financial and Quantitative Analysis*, 23, 417–423.
- Sun, G. (1997), "The Test Function Method for Estimating Continuous-Time Models of Short-Term Interest Rate," working paper, University of Minnesota, Carlson School of Management.
- Sundaresan, S. (1989), "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth," *Review of Financial Studies*, 2, 73–88.
- Vasicek, O. (1977), "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–188.
- Weil, P. (1989), "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.