Empirical Evaluation of Asset-Pricing Models: A Comparison of the SDF and Beta Methods
Author(s): Ravi Jagannathan and Zhenyu Wang
Source: *The Journal of Finance*, Vol. 57, No. 5 (Oct., 2002), pp. 2337-2367
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/3094514
Accessed: 24/11/2009 11:42

THE JOURNAL OF FINANCE • VOL. LVII, NO. 5 • OCTOBER 2002

Empirical Evaluation of Asset-Pricing Models: A Comparison of the SDF and Beta Methods

RAVI JAGANNATHAN and ZHENYU WANG*

ABSTRACT

The stochastic discount factor (SDF) method provides a unified general framework for econometric analysis of asset-pricing models. There have been concerns that, compared to the classical beta method, the generality of the SDF method comes at the cost of efficiency in parameter estimation and power in specification tests. We establish the correct framework for comparing the two methods and show that the SDF method is as efficient as the beta method for estimating risk premiums. Also, the specification test based on the SDF method is as powerful as the one based on the beta method.

THE USE OF THE *stochastic discount factor* (SDF) *method* for econometric evaluation of asset-pricing models has become common in the recent empirical finance literature. An SDF has the following property: 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 an SDF. The SDF method involves estimating the asset-pricing model using its SDF representation and the generalized method of moments (GMM). As Cochrane (2001) points out, the SDF method is sufficiently general that it can be used for analysis of linear as well as nonlinear asset-pricing models, including pricing models for derivative securities.

In spite of its wide use, little is known about the estimation efficiency of the SDF method relative to the classical *beta method*. The latter method estimates the parameters in a linear factor pricing model using its beta

* Jagannathan is from the Kellogg School of Management at Northwestern University and the National Bureau of Economic Research and Wang is from the Graduate School of Business at Columbia University. For helpful comments, we thank Mikhail Chernov; John Cochrane; Wayne Ferson; Bob Hodrick; Narayana Kocherlakota; René Stulz (editor); Guofu Zhou; the anonymous referee; and the seminar participants at the Federal Reserve Bank of New York, Columbia University, the University of Southern California, Washington University at St. Louis, the University of British Columbia, the NBER Conference on Asset Pricing and Portfolio Choice held in May 2000, and the Western Finance Association Meeting held in June 2000. We especially benefited from discussions with Kent Daniel, Lars Hansen, and John Heaton. Adam Kolasinski provided excellent research assistance, and both Adam Kolasinski and Lolotte Olkowski provided editorial help. An earlier draft was circulated under the title "Efficiency of the stochastic discount factor method for estimating risk premiums."

representation, in which the expected return on an asset is a linear function of its factor betas. A question that arises is whether the generality of the SDF framework comes at the costs of estimation efficiency for risk premiums and testing power for model specification.

When returns and factors are jointly normally distributed and independent over time, the classical beta method provides the most efficient unbiased estimator of factor risk premiums in linear models. If the SDF method turns out to be inefficient relative to the classical beta method for linear models under these assumptions, some variation of the beta method may well dominate the SDF method for nonlinear models as well in terms of estimation efficiency. On the other hand, if the SDF method is as efficient as the beta method, it would become the preferred method because of its generality.

We establish the correct framework for comparing the precision of the risk premium estimators in the two methods. The risk premium parameter in the SDF method is not identical to the risk premium parameter in the beta method. However, they are related to each other by a one-to-one transformation. Since the two methods use different parameters to represent the factor risk premium, it is necessary to take into account how they are related to each other before a valid comparison of the risk premium estimators can be made. We find that asymptotically, the SDF method provides as precise an estimate of the risk premium as the beta method. Using Monte Carlo simulations, we demonstrate 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 heteroskedasticity. Therefore, linearizing nonlinear asset-pricing models and estimating risk premiums using the beta method will not lead to an increase in estimation efficiency.

Kan and Zhou (1999) make an inappropriate comparison of the SDF method and the beta method for estimating the parameters related to the factor risk premium. They ignore the fact that the risk premium parameters in the two methods are not identical and directly compare 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 make 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 argue that the SDF method is far less efficient than the beta method. The sampling error in the SDF method is about 40 times as large as that in the beta method. They also conclude that the SDF method is less powerful than the 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, they give an informational advantage to the beta method but not to the SDF method. For a correct 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 the variance of the factor while estimating the risk premium. When this is done, the SDF method is asymptotically as efficient as the beta method.

We also examine 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 zero. We show that the sampling analog of pricing errors has smaller asymptotic variance in the beta method. However, this advantage of the beta method does not show up in the specification tests based on Hansen's (1982) J-statistics. Unlike Kan and Zhou (1999), we demonstrate that the SDF method has the same power as the beta method.

We organize the rest of the paper as follows. In Section I, we briefly describe the SDF and the beta methods for estimating linear factor pricing models. In Section II, we develop the analytical results for the comparison of the two methods. In Section III, we demonstrate our results by Monte Carlo simulations. We conclude in Section IV. For the proof of the theorems in the paper, we refer the readers to the Appendix.

I. Description of the Two Methods

A. The Beta Method

Let r_t be a vector of n asset returns in excess of the risk-free rate. To reduce notational complexity, we assume that there is only one economy-wide pervasive risk factor f_t. Let μ and σ^2 be the mean and variance of the factor f_t. Then the asset-pricing model under the beta representation is given by

$$\mathrm{E}[r_t] = \delta\beta, \tag{1}$$

where δ is the factor risk premium, and β, defined as $\mathrm{Cov}[r_t, f_t]/\sigma^2$, is the sensitivity of asset returns to the factor.

When the economy-wide factor f_t is the return on a portfolio of traded assets, we call it a traded factor. An example of a traded factor would be the return on the value-weighted portfolio of stocks used in empirical studies of Sharpe's (1964) Capital Asset-pricing Model (CAPM). Examples of nontraded factors can be found in Chen, Roll, and Ross (1986), who use the growth rate of industrial production and the rate of inflation, and Breeden, Gibbons and Litzenberger (1989), who use the growth rate in per capita consumption as a factor.

When the factor is the excess return on a traded asset, equation (1) implies $\mu = \delta$, that is, the risk premium is the mean of the factor. This restriction allows us to use the sample mean of the factor as an estimator of the risk premium. If the factor is not traded, this restriction does not hold, and we have to estimate the risk premium using returns on traded assets. We focus on the case where the factor is not traded, although we also consider the case where the factor is traded.

Note that the vector β can be consistently estimated using the time-series regression: $r_t = \phi + \beta f_t + \epsilon_t$. The residual ϵ_t has zero mean and is uncorrelated with the factor f_t. The asset-pricing model (1) imposes a restriction on the intercept, ϕ, that is, $\phi = (\delta - \mu)\beta$. By substituting this expression for ϕ in the regression equation, we obtain the following:

$$r_t = (\delta - \mu + f_t)\beta + \epsilon_t. \tag{2}$$

The beta method uses the beta representation (1), which gives rise to the factor model (2), to estimate the risk premium. The classic two-step cross-sectional regression approach proposed by Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) is widely used in finance literature. For example, Fama and French (1992) use this approach to show that there is no relationship between the expected stock return and beta, and Chen et al. (1986) use this approach to study a linear multifactor asset-pricing model. A shortcoming of the cross-sectional regression approach is that it ignores the sampling errors associated with the estimated betas. Shanken (1992) shows that the Fama–MacBeth method overstates the precision of the estimated parameters when returns and factors are conditionally homoskedastic and temporally independent. Jagannathan and Wang (1998) point out that this is not always the case when returns and factors exhibit conditional heteroskedasticity. Shanken and Jagannathan and Wang provide formulas for calculating the precision of the estimated parameters.

Assuming identical and independent normal distributions for returns and factors, we can apply the maximum likelihood procedure to the beta representation and thereby avoid the shortcomings associated with the two-step cross-sectional regression approach. When the factor is the excess return on a traded asset, Gibbons, Ross, and Shanken (1989) show that this approach is equivalent to estimating the parameters using standard linear multivariate time-series regression. Shanken (1992) shows that the cross-sectional regression method is equivalent to the normal maximum likelihood method if the estimation errors in betas are properly taken into account.

While the application of the maximum likelihood procedure to the beta representation provides the most efficient estimates for risk premiums, the assumption of identical and independent normal distribution is a major limitation. Returns on financial assets may exhibit conditional heteroskedasticity, serial dependence and nonnormality.[1] Hansen's (1982) generalized method of moments (GMM) has become popular because it allows for conditional heteroskedasticity, serial correlation, and nonnormal distributions. Also, the GMM reduces to the maximum likelihood estimation and standard regression approaches when the assumptions that justify those procedures are imposed (see Hamilton (1994) and Cochrane (2001)). We therefore conduct our analysis using the GMM.

We make use of the following moment restrictions implied by the factor model (2): the zero mean of the residuals, the zero covariance between the residuals and the factor, and the definition of the mean and the variance of the factor:[2]

$$\mathrm{E}[r_t - (\delta - \mu + f_t)\beta] = 0_{n\times 1} \tag{3}$$

$$\mathrm{E}[(r_t - (\delta - \mu + f_t)\beta)f_t] = 0_{n\times 1} \tag{4}$$

$$\mathrm{E}[f_t - \mu] = 0 \tag{5}$$

$$\mathrm{E}[(f_t - \mu)^2 - \sigma^2] = 0, \tag{6}$$

where $0_{n\times 1}$ is the vector of n zeros. The vector of unknown parameters is $\theta = (\delta, \beta', \mu, \sigma^2)'$. Throughout the paper, we assume that the regularity conditions mentioned in Hansen (1982) are satisfied. Let $\theta^* = (\delta^*, \beta^{*\prime}, \mu^*, \sigma^{*2})'$ denote the parameters estimated with the GMM, and $\mathrm{Avar}[\delta^*]$ denote the asymptotic variance of the estimated risk premium δ^*.

When returns and factors exhibit conditional homoskedasticity and independence over time, MacKinlay and Richardson (1991) show that the GMM estimator is equivalent to the multivariate regression estimator suggested by Gibbons et al. (1989). For large samples, the two are also equivalent to the maximum likelihood estimator. Ferson and Harvey (1997) extend the equivalence argument to models where betas are linear functions of observable variables. The advantage of the GMM estimator is that it is robust to the presence of conditional heteroskedasticity. MacKinlay and Richardson thus recommend estimating the parameters using the GMM and the beta representation. We refer to the combination of the GMM and the beta representation as the *beta method.*

[1] See French, Schwert, and Stambaugh (1987), Bollerslev, Chou, and Kroner (1992), and Glosten, Jagannathan, and Runkle (1992) for evidence.

[2] When it is not necessary to estimate the variance, σ^2, of the factor, the last moment restrictions can be ignored. If the factor is the return on a tradable asset, we have $\delta = \mu$ and the first two moment restrictions become $\mathrm{E}[r_t - f_t\beta] = 0_{n\times 1}$ and $\mathrm{E}[(r_t - f_t\beta)f_f] = 0_{n\times 1}$, which do not depend on the mean and variance of the factor. In this case, we can estimate β from these two moment restrictions and separately estimate μ and σ^2 from moment restrictions (5) and (6). The estimate of μ is also the estimate of the risk premium.

It is important to emphasize that we compare GMM estimates in the SDF and beta methods, not the maximum likelihood estimates. Although it is well known that the GMM estimates are no more efficient than the maximum likelihood estimates, the advantages of the maximum likelihood estimates vanishes if one does not know the joint distribution of the returns and the factors. If we make the wrong distribution assumption, the maximum likelihood estimates can be biased, while the GMM does not suffer from the same problem. This point is well explained by Hansen and Singleton (1982).

Hansen's J-statistic is often used as a specification test to examine whether the data are consistent with the model. When the linear factor pricing model holds, the J-statistic converges to a central χ^2 distribution as T becomes large. Another way to examine the validity of the pricing model is to test if Jensen's alpha, given by $\alpha = \mathrm{E}[r] - \delta\beta$, is zero. Jensen's alpha measures the deviation of the vector of excess returns from what the corresponding object should be according to the pricing model. The sample analog of Jensen's alpha is given by

$$\alpha^* = \bar{r} - \delta^*\beta^*, \quad \text{where} \quad \bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t. \tag{7}$$

B. The SDF Method

By substituting the expression for β into equation (1) and rearranging the terms, we obtain the following representation of the linear asset-pricing model given below:

$$\mathrm{E}[r_t m_t] = 0_{n\times 1}, \tag{8}$$

where $m_t = 1 - \lambda f_t$. Any random variable m_t that satisfies equation (8) is referred to as a *stochastic discount factor* (SDF). In general, a number of random variables satisfying equation (8) exist, and, hence, there is more than one stochastic discount factor. An asset-pricing model designates a particular random variable as a stochastic discount factor. The linear factor pricing model (8) designates the random variable $m_t = 1 - \lambda f_t$ as a stochastic discount factor. In the above expression, the parameter λ is the following nonlinear transformation of the risk premium δ:

$$\lambda = \frac{\delta}{\sigma^2 + \mu\delta} \qquad \text{or} \qquad \delta = \frac{\lambda\sigma^2}{1 - \mu\lambda}. \tag{9}$$

The SDF representation can be traced back to Dybvig and Ingersoll (1982), who derive the SDF representation for the CAPM. Ingersoll (1987) derives the SDF representation for a number of theoretical asset-pricing models.[3]

[3] However, he does not use the term "stochastic discount factor"—Hansen and Richard (1987) coined the term.

Hansen and Jagannathan (1991, 1997) develop diagnostic tests for asset-pricing models based on the SDF representation. Ferson (1995), Campbell, Lo, and MacKinlay (1997), and Cochrane (2001) provide introductions to the stochastic discount factor framework.

Farnsworth et al. (2000) find that adding the nominally risk-free security to the collection of assets increases estimation efficiency. However, there is a cost to doing so. Nominal interest rates exhibit very persistent near unit-root-like behavior at the empirically relevant frequencies. Hence, it may not be appropriate to rely on the asymptotic standard errors. For that reason, we, following the common practice in the empirical asset-pricing literature, use excess returns and ignore the restriction that the asset-pricing model should also correctly value the nominally risk-free asset. This also facilitates comparison of our results with those reported by Kan and Zhou (1999).

For the linear one-factor pricing model of excess returns, a simple moment restriction for the GMM is given by

$$\mathrm{E}[r_t(1 - \lambda f_t)] = 0_{n\times 1}. \tag{10}$$

Using moment restriction (10) and the GMM, we obtain an estimate of the parameter $\hat{\lambda}$. We denote the associated J-statistic by $\hat{J}$. We are also interested in studying the pricing error, which is defined as $\pi = \mathrm{E}[r_t] - \lambda\mathrm{E}[r_t f_t]$. Hansen and Jagannathan (1997) develop diagnostic tests based on the pricing error. The sample analog of the vector of pricing errors is given by:

$$\hat{\pi} = \bar{r} - \hat{\lambda}(\overline{rf}), \quad \text{where} \quad \bar{r} = \frac{1}{T}\left(\sum_{t=1}^{T} r_t\right) \quad \text{and} \quad \overline{rf} = \frac{1}{T}\left(\sum_{t=1}^{T} r_t f_t\right). \tag{11}$$

Classical tests of asset-pricing models examine whether the vector of pricing errors is different from zero after allowing for sampling errors. Substituting equation (9) into $\pi = \mathrm{E}[r_t] - \lambda\mathrm{E}[r_t f_t]$, we obtain

$$\pi = \left(\frac{\sigma^2}{\sigma^2 + \mu\delta}\right)\alpha \qquad \text{or} \qquad \alpha = \left(\frac{\sigma^2 + \mu\delta}{\sigma^2}\right)\pi. \tag{12}$$

Hence Jensen's alpha will be zero if and only if the pricing error is zero.

The SDF method, which is the combination of the SDF representation and the GMM, provides a convenient and general framework for analyzing linear and nonlinear asset-pricing models. The comparison of the SDF method with the beta method has important implications for the empirical evaluation of asset-pricing models. A nonlinear model can often be well approximated by a linear one. For example, Cochrane (1996) advocates that nonlinear stochastic discount factors be approximated by linear functions of macroeconomic variables. Campbell (1993, 1996) uses linear approximations for the consumption-based CAPM. If the beta method is more efficient, using it to estimate the linearized models may be more efficient than using the SDF to estimate the nonlinear models.

II. Analytical Results

A. Comparison of the Two Methods

First consider the estimates of the risk premium obtained using the two methods. The beta method gives the GMM estimate δ^* from the moment restrictions of the beta model, while the SDF method gives the GMM estimate $\hat{\lambda}$ from the moment restriction of the SDF model. We cannot compare the precision of the two estimates directly because δ and λ will not in general be equal. We therefore first transform the estimate of δ obtained with the beta method into an estimate of λ:

$$\lambda^* = \frac{\delta^*}{\sigma^{*2} + \mu^*\delta^*}, \tag{13}$$

and then compare the asymptotic variances of λ^* and $\hat{\lambda}$.

To understand why it is important to transform δ^* to λ^* for the comparison, consider the case when the factor is the consumption growth rate measured in real numbers. If we switch the measurement of the factor to percent, δ^* should be multiplied by 100, while $\hat{\lambda}$ should be divided by 100. The asymptotic variance of δ^* is then multiplied by 10,000 but the asymptotic variance of $\hat{\lambda}$ is divided by 10,000. Clearly, the scaling of the factor affects the relative magnitude of the asymptotic variances of δ^* and $\hat{\lambda}$. A correct comparison of the SDF and beta methods should be independent of the scaling of factors or returns. It is therefore incorrect to compare the asymptotic variance of δ^* and $\hat{\lambda}$ directly without proper transformation. However, transformation (13) implies that the asymptotic variance of λ^* will be scaled by 1/10,000, making it comparable to the scaled asymptotic variance of $\hat{\lambda}$.

Notice that transformation (13) requires the estimates of μ and σ^2. The estimation errors in μ^* and σ^{*2} affect the asymptotic variance of λ^*. To calculate the asymptotic variance of λ^*, we need to make use of the joint distribution of δ^*, μ^*, and σ^{*2}. Equivalently, we can substitute equation (9) into moment restrictions (3) and (4) to express them in terms of λ, and then estimate λ, β, μ, and σ^2 jointly. The standard errors for the estimate of λ will then automatically take into account the estimation errors in μ^* and σ^{*2}.

To simplify the mathematics involved, we assume that the asset returns and the factors are generated by an i.i.d. joint normal process, unless specially mentioned. Under the i.i.d. joint normal assumption, the beta-method estimator of the risk premium is equivalent to the maximum likelihood estimator. It is well known that the latter is asymptotically the most efficient consistent estimator. It is therefore important to compare the SDF method with the beta method in this special case. However, our results hold under more general distribution assumptions. This is demonstrated by the Monte Carlo simulations in Section III.

The main result of our comparison of the two methods is summarized in the next theorem. The proof of the theorem is provided in the Appendix.

THEOREM 1: *Assume that the beta representation (1) and the equivalent SDF representation (8) hold. The risk premium estimated using the SDF method has the same asymptotic variance as the risk premium estimated using the beta method, that is,* $\text{Avar}[\hat{\lambda}] = \text{Avar}[\lambda^*]$.

Next, consider pricing errors π and Jensen's α. The vector of sample pricing errors, $\hat{\pi}$, in the SDF method is calculated using equation (11). The sample analog of Jensen's alpha, α^*, in the beta method is calculated using equation (7). In general, π and α will not be equal. The pricing error π and Jensen's α are related to each other by equation (12). In view of this, we need to transform the sample analog of Jensen's α^* to get the sample pricing errors:

$$\pi^* = \left(\frac{\sigma^{*2}}{\sigma^{*2} + \mu^*\delta^*}\right)\alpha^*, \tag{14}$$

which is referred to as the sample pricing errors obtained in the beta method. We can then compare the asymptotic variances of π^* and $\hat{\pi}$. In the following theorem, we show that asymptotically, the beta method has smaller variance of the sampling pricing error.

THEOREM 2: *Assume that the beta representation (1) and the equivalent SDF representation (8) hold. The asymptotic variance of the sampling pricing error in the SDF method is at least as large as that in the beta method, that is,* $\text{Avar}[\hat{\pi}] - \text{Avar}[\pi^*]$ *is positive semidefinite.*

For the SDF method, testing $\alpha = 0$ with $\hat{\alpha}$ is algebraically equivalent to Hansen's (1982) tests of overidentification using the J-statistic. This is not the case for the beta method. The J-statistics in the two methods have the same asymptotic distribution. To be more precise, Hansen's J-statistic for the beta method as well as the SDF method have an asymptotic χ^2 distribution with $n - 1$ degrees of freedom, that is, $\hat{J} \xrightarrow{d} \chi^2_{n-1}$ and $J^* \xrightarrow{d} \chi^2_{n-1}$ as $T \to \infty$. The SDF method has $n - 1$ degrees of freedom because there are n restrictions and one parameter in equation (10). The beta method also has $n - 1$ degrees of freedom because there are $2n + 2$ restrictions in equations (3), (4), (5), and (6) and $n + 3$ parameters. Therefore, the asymptotic variances of $\hat{J}$ and J^* must be the same. The distributions of $\hat{J}$ and J^* can only be different in finite samples. Later, we use Monte Carlo simulation to address the issue of finite-sample distributions.

We can also consider the parameter δ and Jensen's α. In the beta method, the estimates δ^* and α^* are obtained as described in Section I.A. For proper comparison, we transform the estimates $\hat{\lambda}$ and $\hat{\pi}$ obtained using the SDF method as described in Section I.B to get estimates $\hat{\delta}$ and $\hat{\alpha}$ using formulas (9) and (12). It can be readily shown, by mimicking the proof of Theorem 1, that Avar$[\hat{\delta}]$ = Avar$[\delta^*]$ and that Avar$[\hat{\alpha}]$ − Avar$[\alpha^*]$ is positive semidefinite.

In Theorem 1, the factor can be either a return on a tradable asset or a nontradable macroeconomic variable. As we have discussed in the earlier section, if the factor is a tradable asset return, the asset-pricing model implies the restriction $\delta = \mu$. It can be shown that Theorem 1 continues to hold even when this restriction is imposed during estimation. In fact, under this restriction, the asymptotic variances of Avar$(\hat{\delta})$ and Avar(δ^*) are both equal to the asymptotic variance of the sample average of f_t. Therefore, imposing the restriction $\delta = \mu$, we only need to estimate the risk premium from the observations of the factor. This has been pointed out previously by Shanken (1992).

B. The Effects of Standardizing Factors

According to Theorem 1, the beta method and the SDF method provide risk premium estimates that are equally precise. This result contradicts Kan and Zhou's (1999) conclusion. In this section, we examine why this is so. When comparing the two methods, Kan and Zhou suggest comparing the asymptotic variance of the estimated δ and λ by looking at special parameter values that make δ and λ equal in numbers. Specifically, they assume that the factor has zero mean and unit variance, that is, $\mu = 0$ and $\sigma = 1$. At these special values of the parameters, they notice that $\lambda = \delta$ and directly compare the estimation errors of δ and λ.

Since this assumption applies to the SDF method as well as the beta method, one may expect that it would not give an advantage to one method over the other. In this subsection, we show that this is not the case. Predetermining the mean and the variance of the factor increases the efficiency of the estimator in the beta method but not in the SDF method. However, by adding additional moment conditions to incorporate the information brought in through the moments of the factor, the SDF method estimator can be made as efficient as the beta-method estimator.

Since the mean and the variance of the factor are predetermined to be zero and one, Kan and Zhou (1999) drop equations (5) and (6). For the beta method, they obtain the estimator of the parameter δ from the moment restriction

$$E[r_t - (\delta + f_t)\beta] = 0_{n\times 1} \tag{15}$$

$$E[(r_t - (\delta + f_t)\beta)f_t] = 0_{n\times 1}. \tag{16}$$

For the SDF method, they obtain the estimator of the parameter λ from the moment restriction

$$E[r_t(1 - \lambda f_t)] = 0_{n\times 1}, \tag{17}$$

which is equation (10). Then, they directly compare the asymptotic variance of the estimated δ with the asymptotic variance of the estimated λ.

The most important aspect of the assumption made by Kan and Zhou (1999) is that the mean and variance of the factor are predetermined without estimation. This assumption is the equivalent of ignoring the sampling errors associated with the estimates of μ and σ^2. To fully understand how this assumption affects the precision of the risk premium estimators, we assume that μ and σ^2 are known, but not necessarily zero and one.

In general, predetermining a subset of the parameters reduces the sampling error of the remaining estimates. The following lemma summarizes this effect.

LEMMA 1: *Let x_t be the observed data in period t and $g(x_t, \theta_1, \theta_2)$ be a vector function of $(x_t, \theta_1, \theta_2)$, where θ_1 is the vector of parameters of interest. Suppose that* $\mathrm{E}[g(x_t, \theta_1, \theta_2)] = 0$ *defines the GMM moment restrictions. Let $(\hat{\theta}_1', \hat{\theta}_2')'$ be the GMM estimator of $(\theta_1', \theta_2')'$. When θ_2 is predetermined, let $\tilde{\theta}_1$ be the GMM estimator of θ_1. Then,* $\mathrm{Avar}[\hat{\theta}_1] - \mathrm{Avar}[\tilde{\theta}_1]$ *is positive semi-definite. In addition, if*

$$\lim_{T\to\infty} \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^{T} \frac{\partial g(x_t, \theta_1, \theta_2)}{\partial \theta_2'}\right] = 0, \tag{18}$$

then $\mathrm{Avar}[\hat{\theta}_1] = \mathrm{Avar}[\tilde{\theta}_1]$.

When the factor moments are predetermined, the asymptotic variance of the estimated risk premium becomes smaller. Thus, Kan and Zhou (1999) understate the variances of the parameter estimates in the beta method relative to the realistic case where μ and σ must be estimated. Nonetheless, the variance of the estimates in the SDF method is not understated because condition (18) is satisfied.

Kan and Zhou (1999) do not use the two moment conditions that restrict the first and second sample moments of the factor. In general, dropping the moment restrictions in the GMM will increase the sampling error of the estimated parameters. However, when certain conditions are met, moment conditions can be dropped without affecting the sampling error of the estimated parameters. It turns out that these conditions are satisfied for the beta method. Thus, Kan and Zhou's analysis still gives the beta method an efficiency advantage after dropping the moment restrictions for μ and σ. The following lemma makes these statements precise.

LEMMA 2: *Let x_t be the observed data in period t, $g_1(x_t, \theta)$ be a vector function of (x_t, θ), where θ is the vector of parameters, and $g_2(x_t)$ be a vector function of x_t. Let $g(x_t, \theta)' = (g_1(x_t, \theta)', g_2(x_t)')'$. Suppose* $\mathrm{E}[g(x_t, \theta)] = 0$. *Let $\hat{\theta}$ be the GMM estimator from the moment restriction* $\mathrm{E}[g(x_t, \theta)] = 0$. *Let $\tilde{\theta}$ be* *the GMM estimator of* θ *from the moment restriction* $\mathrm{E}[g_1(x_t,\theta)] = 0$. *Then,* $\mathrm{Avar}[\hat{\theta}] - \mathrm{Avar}[\tilde{\theta}]$ *is negative semidefinite. In addition, if, however,*

$$\sum_{j=-\infty}^{\infty} \mathrm{E}[g_1(x_t,\theta)g_2(x_{t+j})'] = 0, \tag{19}$$

then $\mathrm{Avar}[\hat{\theta}] = \mathrm{Avar}[\tilde{\theta}]$.

To apply Lemma 2 to the beta method, let $\theta = (\delta, \beta')'$ and

$$g_1(x_t,\theta) = \begin{pmatrix} r_t - (\delta - \mu + f_t)\beta \\ (r_t - (\delta - \mu + f_t)\beta)f_t \end{pmatrix} \quad \text{and} \quad g_2(x_t) = \begin{pmatrix} f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix} \tag{20}$$

where μ and σ are predetermined. It is easy to show that g_1 and g_2 are uncorrelated. Hence, condition (19) is satisfied, and dropping the factor moment restrictions does not affect the variance of the risk premium estimator in the beta method.

Under the assumption of zero mean and unit variance of the factor, however, Kan and Zhou (1999) show that the beta method is *strictly* more efficient than the SDF method. More generally, we can show that the strict inequality holds when the first two moments of the factor are predetermined, not necessarily to be zero and one. Let $\delta^\dagger$ be the estimator of δ obtained from moment restrictions (15)–(16) and $\lambda^\dagger = \delta^\dagger/(\sigma^2 + \mu\delta^\dagger)$. Our derivation in the Appendix shows

$$\mathrm{Avar}[\lambda^\dagger] < \mathrm{Avar}[\hat{\lambda}]. \tag{21}$$

We have shown in Theorem 1 that the SDF method is as efficient as the beta method when the factor moments are not predetermined and have to be estimated. Hence, we conclude that Kan and Zhou reached a different conclusion because they ignored the estimation error associated with the first two moments of the factor and treat them as though they are known with certainty.

To summarize, the estimated risk premiums in the SDF and beta methods have the same sampling error when the factor moments are not predetermined. Predetermining the factor moments reduces the sampling error of the estimate in the beta method, even though the moment conditions corresponding to the predetermined parameters are ignored. In the SDF method, however, the variance of the estimator is not reduced when factor moments are predetermined. This explains why Kan and Zhou (1999) find that the SDF method is less efficient.

When the factor moments are predetermined, the information about the factor moments should be properly incorporated into the SDF method. From Lemma 2, we know that the restriction $\mathrm{E}[g_2(x_t)] = 0$ might affect the estimation efficiency if $g_1(x_t, \theta)$ and $g_2(x_t)$ are correlated. For this reason, we add the factor moment restrictions to the moment restriction (10) in the SDF method. Thus, we obtain the GMM estimate, denoted by $\bar{\lambda}$, of the parameter λ from the following moment restrictions:

$$\mathrm{E}[r_t(1 - \lambda f_t)] = 0 \tag{22}$$

$$\mathrm{E}[f_t - \mu] = 0 \tag{23}$$

$$\mathrm{E}[(f_t - \mu)^2 - \sigma^2] = 0, \tag{24}$$

where μ and σ are predetermined. Our derivation in the Appendix shows

$$\mathrm{Avar}[\bar{\lambda}] = \mathrm{Avar}[\lambda^{\dagger}]. \tag{25}$$

Therefore, a simple remedy to the inefficiency of the SDF method when factor moments are predetermined is to incorporate the factor-moment restrictions. This further explains why the additional information about the factor is incorporated into the beta method but not the standard SDF method when we predetermine the factor moments.

A natural question that arises is whether we can increase the estimation efficiency by incorporating the factor moment restrictions when the factor mean and variance are not predetermined. More specifically, we want to know whether the asymptotic variance of λ estimated together with μ and σ^2 from the moment restrictions

$$\mathrm{E}[r_t(1 - \lambda f_t)] = 0_{n\times 1} \tag{10$'$}$$

$$\mathrm{E}[f_t - \mu] = 0$$

$$\mathrm{E}[(f_t - \mu)^2 - \sigma^2] = 0,$$

is smaller than the asymptotic variance of the corresponding estimators that are obtained from the first moment restriction (10$'$) alone. The answer is no because the number of added moment restrictions is the same as the number of added unknown parameters. Readers can verify this easily.

III. Monte Carlo Simulations

A. Estimation Efficiency

In the first set of simulations, we assume that the returns and the factor are drawn from a multivariate normal distribution. To be consistent with the beta model and its equivalent SDF model, we generate the excess returns and the factor from the following process:

$$r_t = (\delta - \mu + f_t)\beta + \epsilon_t, \qquad \epsilon_t | f_t \sim \mathrm{N}(0_{n\times 1}, \Omega), \qquad f_t \sim \mathrm{N}(\mu, \sigma^2), \tag{26}$$

where $t = 1, \cdots, T$. For T, we consider the following four different time horizons in our simulations: 60 months, 120 months, 360 months, and 600 months. The first two horizons are often used for estimating betas and risk premiums using rolling windows. The third horizon is similar to that used in many related empirical studies.[4] The fourth is a half-century, which may be considered a "large" sample for monthly returns. The estimators are calculated based on the T samples of the factor and returns generated from the above process. We repeat this independently to obtain 1,000 draws of the estimator of λ.

We choose the values for the parameters to match the sample moments of the data. Monthly returns from January 1926 to December 1998 for the 10 size-decile portfolios; the value-weighted market index portfolio of stocks traded on the NYSE, AMEX, and Nasdaq; and the one-month Treasury bills are obtained from the Center for Research on Security Prices (CRSP). Excess returns are obtained by subtracting the Treasury bill returns. We set μ and σ^2 to equal the sample mean and the sample variance of the excess return on the market index portfolio. We set β for each size-decile portfolio to equal the slope coefficient in the OLS regressions of the size-decile excess returns on the excess market return. The covariance matrix, $\Omega = E[\epsilon_t \epsilon_t' | f_t]$, is set equal to the sample covariance matrix of the residuals obtained in the 10 OLS regressions. We set the risk premium δ to be the value of the slope coefficient obtained from the cross-sectional regression of the historical average excess returns on betas. The parameter λ satisfies $\lambda = \delta/(\sigma^2 + \mu\delta)$. Table I provides all the parameters used in the simulation.

In the population, we set risk premium $\delta = 1.3740$ for monthly returns. This implies $\lambda = 4.3790$. This large risk premium reflects the firm-size premium since β is strongly negatively correlated with firm size in the sample we use for calibration purposes.[5] We do not set the values of δ and μ to be equal. Therefore, we do not impose the restriction that the factor is the return on a portfolio of tradable assets.[6] The asymptotic standard deviations of the estimated λ implied by the parameters are reported in Panel A of Table II. For comparison with the simulation results, the asymptotic standard deviations are divided by the square root of T.

[4] For example, Fama and French (1992) and Jagannathan and Wang (1996) use 330 monthly observations. Fama and French (1993) use 342 monthly observations.

[5] If we were to construct portfolios by ranking firms by both market capitalization and estimated beta, the betas would not be as highly negatively correlated with firm size. In that case, as in Fama and French (1992) and Jagannathan and Wang (1996), the risk premium on beta would be close to zero, and the risk premium on firm size would be large in magnitude and negative.

[6] Our simulation results also hold well when we set $\delta = \mu$.

Table I

Parameter Values Used in Monte Carlo Simulations

This table presents the parameters used in our Monte Carlo simulations. The choice of the parameters is based on monthly historical observations (from January 1926 to December 1998) of returns (in excess of returns on one-month Treasury bills) on decile portfolios and the value-weighted market index of NYSE, AMEX, and Nasdaq. The data are obtained from the Center for Research on Security Prices (CRSP). The mean (μ) and standard deviation (σ) of the factor is set to be the sample mean and standard deviation of returns on the market index. The betas (β) are set to be the slopes in the time-series regression of the decile returns on the market index return. Jensen's α is set to be the intercept in the same regression. The sample covariance of the residuals in this regression is chosen to be the covariance matrix Ω. The risk premium δ is set to be the slope in the cross-sectional regression of the decile's historical average return on beta. The parameter λ satisfies $\lambda = \delta/(\sigma^2 + \mu\delta)$. The numbers reported for μ, σ, and α are multiplied by 100 while the numbers for Ω are multiplied by 10,000.

$\mu = 0.6914$ $\sigma = 5.5154$ $\delta = 1.3740$ $\lambda = 4.3790$

Decile Portfolios									
Small	2	3	4	5	6	7	8	9	Large
					β				
1.46	1.395	1.309	1.265	1.241	1.227	1.183	1.122	1.090	0.947
					α				
0.494	0.126	0.000	−0.004	0.000	−0.002	0.020	−0.042	0.023	0.003
					Ω				
55.680	37.110	28.820	22.300	17.390	14.450	10.160	6.375	3.890	−2.808
37.110	29.830	22.620	17.880	13.960	11.870	8.296	5.466	3.172	−2.243
28.820	22.620	20.480	15.180	11.860	10.300	7.265	4.932	2.784	−1.918
22.300	17.880	15.180	13.670	10.120	8.860	6.638	4.770	2.633	−1.687
17.390	13.960	11.860	10.120	9.290	7.432	5.616	4.134	2.404	−1.422
14.450	11.870	10.300	8.860	7.432	7.312	5.097	3.838	2.334	−1.301
10.160	8.296	7.265	6.638	5.616	5.097	4.901	3.289	1.944	−1.036
6.375	5.466	4.932	4.770	4.134	3.838	3.289	3.254	1.776	−0.840
3.890	3.172	2.784	2.633	2.404	2.334	1.944	1.776	1.775	−0.572
−2.808	−2.243	−1.918	−1.687	−1.422	−1.301	−1.036	−0.840	−0.572	0.312

In implementation of the GMM, we follow Ferson and Foerster (1994), who suggest iterating between the estimation of the weighting matrix and the estimation of parameters until the estimates converge. The simulation results for λ are reported in Panel B of Table II. For each estimator of λ, the table gives the standard deviation of the 1,000 estimated risk premiums. When the mean and variance of the factor are estimated along with the risk premium parameter, the estimator $\hat{\lambda}$ obtained using the SDF method and the estimator λ^* obtained using the beta method have essentially the same precision for all sample sizes considered. This is true even when the sample size is as small as 60 months. Therefore, there is no efficiency gain from the use of the beta method over the SDF method. Our simulations also show similar results for the estimated δ—we do not report them in order to save space. We refer those who are interested in the numerical results of δ to Jagannathan and Wang (1999).

Table II

Standard Errors of the Estimated λ

This table provides the asymptotic and simulated standard errors for various estimators of λ subject to the restriction $E[r_t(1 - \lambda f_t)] = 0_{n\times 1}$. The parameters, λ, δ, β, μ, σ, and Ω, are set to the values in Table I. The asymptotic standard errors are the asymptotic standard deviations divided by the square root of T. To obtain the simulated standard errors, independent samples $\{(f_t, \epsilon_t')'\}_{t=1,\cdots,T}$ are drawn from a normal, t, or empirical distribution. Excess returns on 10 portfolios are constructed to satisfy $r_t = (\delta + f_t - \mu)\beta + \epsilon_t$ for $t = 1,\cdots,T$. We consider the estimators obtained in four different approaches: (1) the beta method (λ^*), (2) the SDF method ($\hat{\lambda}$), (3) the beta method when the factor mean and variance are predetermined and the factor moment restrictions are dropped ($\lambda^\dagger$), and (4) the SDF method when the factor mean and variance are predetermined and the factor moment restrictions are added ($\bar{\lambda}$). In each approach, the estimator of λ is calculated based on the T samples. We repeat this independently to obtain 1,000 draws of the estimator of λ. The simulated standard error is the standard deviation of the random draws of the estimator.

T	λ^*	$\hat{\lambda}$	$\lambda^\dagger$	$\bar{\lambda}$
Panel A: Results Calculated from the Asymptotic Distribution				
60	2.2063	2.2063	0.0653	0.0653
120	1.5601	1.5601	0.0461	0.0461
360	0.9007	0.9007	0.0266	0.0266
600	0.6977	0.6977	0.0206	0.0206
Panel B: Results Simulated from the Normal Distribution				
60	3.1193	2.8903	0.1124	0.1360
120	1.9136	1.8703	0.0522	0.0528
360	0.9616	0.9560	0.0270	0.0270
600	0.7307	0.7279	0.0198	0.0199
Panel C: Results Simulated from the Student-t Distribution				
60	4.3511	3.4108	0.6487	0.2151
120	2.2490	2.1069	0.0712	0.0835
360	1.1101	1.1015	0.0267	0.0282
600	0.8235	0.8127	0.0210	0.0220
Panel D: Results Simulated from the Empirical Distribution				
60	3.2823	3.0287	0.1143	0.1324
120	1.8675	1.8274	0.0514	0.0542
360	0.9729	0.9674	0.0278	0.0279
600	0.7321	0.7306	0.0215	0.0215

If we predetermine the mean and variance of the factor and ignore the related moment restrictions, the SDF method is much less efficient than the beta method. This is true in both small as well as large samples. When we have 30 years of monthly observations ($T = 360$), the standard deviation of $\hat{\lambda}$ obtained using the SDF method is about 35 times as high as the standard deviation of $\lambda^\dagger$ obtained using the beta method. This result is similar to that reported by Kan and Zhou (1999). It confirms that when the factor moments are predetermined, the classical beta method substantially overstates the precision of the estimated risk premiums.

When the factor moment restrictions are added to the SDF method, the efficiency of the SDF method improves substantially. Notice that the standard error of $\bar{\lambda}$ is nearly the same as that of $\lambda^{\dagger}$. The increase in efficiency of the SDF method after adding the factor moments restrictions is expected, because in this case, the two methods have the same asymptotic variance for the estimated risk premium. Thus, the sharp disadvantage of the SDF method relative to the beta method reported by Kan and Zhou (1999) is mainly due to ignoring the estimation error in the factor moments.

In our theoretical derivation of the asymptotic variance, we assume that the variance of the returns do not depend on the realized value of the factor. This may be a rather restrictive assumption. We therefore examine the applicability of our results when returns exhibit conditional heteroskedasticity. For this purpose, following MacKinlay and Richardson (1991), we make independent draws of the returns and the factor from a multivariate t-distribution rather than a joint normal distribution. When the multivariate t-distribution has ν degrees of freedom, the conditional covariance matrix of the residuals in the regression equation, conditional on the realized value of the factor, is given by (see MacKinlay and Richardson, equation 14)

$$\mathrm{Var}[\epsilon_t | f_t] = \frac{\nu - 2 + (f_t - \mu)^2/\sigma^2}{\nu - 1}\,\Omega. \tag{27}$$

Notice that dependence of the conditional covariance on the realized value of the factor increases as the number of the degree of freedom, ν, decreases. There is, however, a lower bound on our choice of the number of degrees of freedom. The asymptotic distribution theory for the GMM requires that returns and factors have finite fourth moments. Hence, there must be more than four degrees of freedom. In view of this restriction, we use five degrees of freedom for the multivariate t-distribution, following MacKinlay and Richardson (1991).

The simulation results corresponding to the t distribution are presented in Panel C of Table II. The standard errors computed using the asymptotic theory and through simulation are about the same. It is also true that predetermining the factor mean and variance causes the estimator $\lambda^{\dagger}$ obtained with the beta method to be more efficient than the estimator $\hat{\lambda}$ obtained with the SDF method, but incorporation of the factor moment restrictions makes the estimator $\bar{\lambda}$ as efficient as the estimator $\lambda^{\dagger}$. When the mean and variance of the factor are not predetermined, which is a more realistic and correct approach, the two estimators λ^{*} and $\hat{\lambda}$ obtained with the two methods have almost the same precision in every sample size. This indicates that our results are robust to the existence of conditional heteroskedasticity.

As an alternative to the multivariate t-distribution, we also consider the joint empirical distribution of the excess returns and the factor. The monthly observations of the return on the value-weighted index of NYSE, AMEX, and Nasdaq are used for f_t. The residuals in the regression of decile returns on the index return are used for ϵ_t. Independent samples $\{(f_t, \epsilon_t')'\}_{t=1,\cdots,T}$ are drawn from an estimated empirical distribution of the data.[7] Excess returns on 10 portfolios are constructed to satisfy $r_t = (\delta - \mu + f_t)\beta + \epsilon_t$ for $t = 1,\cdots,T$. The parameters, λ, δ, β, and μ, are set to the values in Table I. Each estimator is then calculated based on the T observations to obtain a random draw of the estimator. We repeat this process independently 1,000 times to obtain 1,000 independent random draws of each estimator.

The simulation results are given in Panel D of Table II. When the mean and variance of the factor are estimated together with the risk premium, the sampling errors for the risk premium using the SDF method and the beta method are almost identical. Ignoring the estimation error in the factor moments causes the beta method to be far more precise than the SDF method. When the factor mean and variance are predetermined, incorporating the factor moment restrictions greatly improves the precision of the SDF method.

B. Specification Tests

It is common to evaluate model misspecification by examining the sample pricing errors. Our Theorem 2 states that the sample pricing errors in the beta method have a smaller covariance matrix than those in the SDF method, applied to excess returns. To assess the quantitative importance of the difference between the two methods, we set λ, δ, β, μ, σ, and Ω to the values in Table I and calculate the asymptotic standard deviations of the estimated pricing errors for the 10 size-decile portfolios. The calculations are based on the formulas given in the Appendix under the null hypothesis of $E[r] = \delta\beta$. The observations of the factor and returns are assumed to have a joint normal distribution, identical and independent across time.

In Panel A of Table III we report the asymptotic standard deviations divided by the square root of T. We only consider the general case where the factor mean and variance are not predetermined. As claimed by the theorems in this paper, the standard deviations of the pricing errors estimated using the SDF method are larger than those using the beta method. However, the differences are quite small. Such small differences would be difficult to detect using Monte Carlo methods.

When we use the same parameter values, the standard deviations of the estimated pricing errors obtained from Monte Carlo simulations are not very different for the two methods. Panel B of Table III reports the results of our Monte Carlo simulation for the pricing errors in finite samples. The design of the simulation is exactly the same as the design for Table II. To save space, we only report the results for the normal distribution. As can be seen in Table III, the standard deviations of the sample pricing errors obtained in the two methods are very similar for all the assets. This is similar to what Cochrane (2000) obtained. Based on Monte Carlo simulations, it would be

[7] We can estimate the empirical distribution either by the Bootstrap method or the method described by Taylor and Thompson (1986). The results are the same, and we only report the former.

Table III

Standard Errors of the Sample Pricing Errors

This table provides the asymptotic and simulated standard errors for sample pricing errors in the SDF and beta methods subject to the restriction $E[r_t(1 - \lambda f_t)] = 0_{n\times 1}$. The vector of sample pricing errors in the SDF method is defined in equation (11). The vector of sample pricing errors in the beta method is defined in equation (14). The true parameters, λ, δ, β, μ, σ, and Ω, are set to the values in Table I. The asymptotic standard errors are the asymptotic standard deviations divided by the square root of T. To obtain the simulated standard errors, independent samples $\{(f_t, \epsilon_t')'\}_{t=1,\cdots,T}$ are drawn from a normal distribution. Excess returns on 10 portfolios are constructed to satisfy $r_t = (\delta + f_t - \mu)\beta + \epsilon_t$ for $t = 1,\cdots,T$. The vector of sample pricing errors in each method is then calculated based on the T samples. We repeat this independently to obtain 1,000 random draws of the vector of sample pricing errors in each method. The simulated standard error is the sample standard deviation of the independent draws. All the standard errors reported in the table are multiplied by 100.

T	Method	Small	2	3	4	5	6	7	8	9	Large
		Panel A: Results Calculated from the Asymptotic Distribution									
60	Beta	0.9060	0.6629	0.5492	0.4485	0.3695	0.3276	0.2680	0.2181	0.1605	0.0653
	SDF	0.9622	0.7041	0.5832	0.4763	0.3924	0.3480	0.2846	0.2316	0.1704	0.0694
120	Beta	0.6407	0.4688	0.3883	0.3171	0.2613	0.2317	0.1895	0.1542	0.1135	0.0462
	SDF	0.6804	0.4979	0.4124	0.3368	0.2775	0.2460	0.2012	0.1638	0.1205	0.0491
360	Beta	0.3699	0.2706	0.2242	0.1831	0.1508	0.1338	0.1094	0.0890	0.0655	0.0267
	SDF	0.3928	0.2874	0.2381	0.1945	0.1602	0.1421	0.1162	0.0945	0.0696	0.0283
600	Beta	0.2865	0.2096	0.1737	0.1418	0.1168	0.1036	0.0847	0.0690	0.0507	0.0207
	SDF	0.3043	0.2226	0.1844	0.1506	0.1241	0.1100	0.0900	0.0732	0.0539	0.0219
		Panel B: Results Simulated from the Normal Distribution									
60	Beta	0.8885	0.6859	0.5898	0.5145	0.4584	0.4287	0.3874	0.3551	0.3314	0.2809
	SDF	1.0603	0.8301	0.7231	0.6409	0.5733	0.5367	0.4864	0.4455	0.4150	0.3397
120	Beta	0.6795	0.5212	0.4413	0.3637	0.3126	0.2914	0.2531	0.2164	0.1858	0.1335
	SDF	0.7644	0.5851	0.4981	0.4122	0.3558	0.3319	0.2867	0.2443	0.2111	0.1485
360	Beta	0.3863	0.2828	0.2349	0.1919	0.1593	0.1451	0.1181	0.0988	0.0791	0.0454
	SDF	0.4115	0.3007	0.2497	0.2040	0.1698	0.1541	0.1265	0.1062	0.0848	0.0491
600	Beta	0.2779	0.2097	0.1730	0.1424	0.1199	0.1081	0.0891	0.0745	0.0574	0.0293
	SDF	0.2978	0.2245	0.1853	0.1527	0.1286	0.1158	0.0956	0.0799	0.0612	0.0319

tempting to conclude that the beta and SDF methods give the same standard deviation of the sample pricing errors. Our theorem shows that such a conclusion would be wrong, and the difference can be large for some parameter values. However, for the parameter values that are likely to be encountered in most empirical studies, the differences between the two methods are often negligible.

A convenient way of examining model specification is the test based on Hansen's J-statistics. As we have pointed out in Section II.A, the J-statistics in the SDF and the beta methods have the same asymptotic distribution. Therefore, the sampling distribution of the J-statistics can be different only in finite samples or in misspecified models. In what follows, we therefore examine the size and power of the specification tests in finite samples.

To examine the test size in the two methods, we use Monte Carlo simulations to compute the rejection rates under the null hypothesis that $E[r_t(1 - \lambda f_t)] = 0_{n\times 1}$. For this purpose, independent samples $\{(f_t, \epsilon_t')'\}_{t=1,\cdots,T}$ are drawn from a normal distribution, student-t distribution, or the empirical distribution as we have described earlier. Excess returns on the 10 portfolios are then generated to satisfy $r_t = (\delta + f_t - \mu)\beta + \epsilon_t$ for $t = 1,\cdots,T$ so that the asset-pricing model holds. The parameters, λ, δ, β, μ, σ, and Ω, are set to the values in Table I. We obtain the parameter estimates and the J-statistic for each method based on a sample of T simulated observations. We repeat to obtain 10,000 independent random draws of the J-statistic. The test size for a significant level is calculated as the rejection rates. We report the test size for three significance levels: 1 percent, 5 percent and 10 percent. This requires estimating the tails of the sampling distribution of the J-statistics. Therefore, we need a large number of simulations, which is chosen to be 10,000.

The simulation results for the normal and t-distributions[8] are reported in Panels A and B of Table IV. For the numbers we considered for T, the rejection rates in the SDF method are all close to the theoretical p-values based on the χ^2 distribution. When T increases, the test size moves closer to the theoretical p-value. It is clear that the SDF method works as well as the beta method.

To examine the power of the two methods, we conduct similar Monte Carlo simulations, except that we add a nonzero Jensen's alpha to the model for generating excess returns. The nonzero Jensen's alpha makes the asset-pricing model misspecified. To be specific, excess returns are constructed to satisfy $r_t = \alpha + (\delta + f_t - \mu)\beta + \epsilon_t$. Jensen's alpha, α, for each portfolio is calibrated from the data and listed in Table I. As reported in the empirical finance literature, portfolios of small stocks are associated with large pricing errors. The power is the fraction of the simulated J-statistics that are larger than the critical value based on the sampling distribution of the J-statistics under the null hypothesis of $\alpha = 0$. The results are reported in Panels C and D of Table IV. Clearly, the SDF method is as powerful as the beta method.

Our simulation results on the power are strikingly different than those reported by Kan and Zhou (1999), who show that the beta method is much more powerful than the SDF method for specification tests. To obtain their results, they introduce misspecification by assuming that econometricians mistakenly specify some "useless" factors that are uncorrelated with asset returns in the model (see Kan and Zhou for the artificial construction of their "noisy," "unsystematic," or "useless" factors). In contrast, we introduce misspecification by adding Jensen's alpha calibrated from the data. Thus, the relative power of the two approaches may depend on the nature of the hypothesis. It is, however, reasonable to assume nonzero Jensen's alpha be-

[8] We have also drawn samples from the empirical distribution for 1,000 simulations. The results are very similar to the results obtained from the normal and student-t distributions. We do not report the results for the empirical distribution because it takes a very long time to conduct 10,000 simulations with the empirical distribution.

Table IV

Simulation Results for Specification Tests

This table provides the results of Monte Carlo simulations on the rejection rate of J-statistics. Independent samples $\{(f_t, \epsilon_t')'\}_{t=1,\cdots,T}$ are drawn from a normal or t distribution. Excess returns on 10 portfolios are generated to satisfy $r_t = \alpha + (\delta + f_t - \mu)\beta + \epsilon_t$ for $t = 1,\cdots,T$, where λ, δ, β, μ, σ, and Ω are set to the values in Table I. For the size of specification tests, α is set to a zero vector. For the power of specification tests, α is set to the values in Table I. We consider the J-statistics obtained in four different approaches: (1) the beta method (J^*), (2) the SDF method ($\hat{J}$), (3) the beta method when the factor mean and variance are predetermined and the factor moment restrictions are dropped ($J^\dagger$), and (4) the SDF method when the factor mean and variance are predetermined and the factor moment restrictions are added ($\bar{J}$). In each approach, the GMM is applied to the T samples to obtain a J-statistic. We repeat this independently to obtain 10,000 random draws of the J-statistics. The table presents the percent of the simulated J-statistics that are larger than the critical point at the significance levels of 10 percent, 5 percent, and 1 percent based on the sampling distribution of the J-statistics under the null hypothesis of $\alpha = 0$.

	10 Percent				5 Percent				1 Percent			
T	J^*	$\hat{J}$	$J^\dagger$	$\bar{J}$	J^*	$\hat{J}$	$J^\dagger$	$\bar{J}$	J^*	$\hat{J}$	$J^\dagger$	$\bar{J}$
Panel A: Size of Specification Tests Simulated from the Normal Distribution												
60	7.2	7.9	7.2	12.7	2.9	3.2	2.9	6.4	0.2	0.2	0.2	1.6
120	8.9	9.0	8.9	11.7	4.0	4.2	4.0	6.4	0.6	0.6	0.6	1.6
360	9.4	9.5	9.4	10.3	4.7	4.7	4.7	5.2	0.9	0.9	0.9	1.2
600	9.8	9.8	9.8	10.3	4.8	4.8	4.8	5.1	0.9	0.9	0.9	1.1
Panel B: Size of Specification Tests Simulated from the Student-t Distribution												
60	5.4	7.3	5.4	23.0	1.8	2.8	1.8	14.4	0.1	0.2	0.1	5.1
120	7.7	8.9	7.7	20.9	3.0	4.0	3.1	13.2	0.4	0.5	0.4	5.1
360	8.9	9.3	8.9	16.3	4.1	4.3	4.1	10.1	0.7	0.8	0.7	3.5
600	9.9	10.2	9.9	15.0	4.5	4.7	4.5	8.8	0.8	0.9	0.8	2.5
Panel C: Power of Specification Tests Simulated from the Normal Distribution												
60	11.1	12.1	11.1	16.9	4.6	5.3	4.6	9.1	0.3	0.6	0.3	2.1
120	19.7	20.2	19.7	22.1	10.4	10.9	10.4	12.5	2.2	2.4	2.2	3.8
360	50.3	50.3	50.3	47.7	36.7	36.9	36.7	34.6	15.9	16.1	15.9	14.9
600	74.5	74.5	74.5	71.1	62.3	62.3	62.3	58.9	37.6	37.7	37.6	34.6
Panel D: Power of Specification Tests Simulated from the Student-t Distribution												
60	9.6	12.6	9.6	28.8	3.9	5.6	3.9	18.4	0.4	0.6	0.4	6.9
120	18.4	20.8	18.4	31.9	9.4	11.2	9.4	21.8	1.8	2.5	1.8	9.0
360	49.8	50.6	49.8	53.6	35.9	36.9	35.9	41.4	14.7	15.6	14.7	20.6
600	73.0	73.5	73.0	72.6	60.7	61.3	60.7	61.0	34.6	35.4	34.6	37.0

cause a misspecified model by definition has nonzero pricing errors. Kan and Zhou's assumption that econometricians pick useless factors is not interesting because the prespecified factors that are typically used in empirical analysis of asset-pricing models are based on meaningful economic analysis and exhibit nonzero correlation with asset returns.

IV. Conclusion

The stochastic discount factor (SDF) method has received wide attention in the theoretical and empirical asset-pricing literature. The main attraction of the SDF method is its generality. It provides an elegant framework for econometric evaluation of both linear and nonlinear asset-pricing models, including pricing models for derivative securities. We examine whether the generality of the SDF framework comes at the cost of estimation efficiency for risk premiums and test power for model specifications.

For that purpose, we compare the classical beta method with the SDF method for linear factor-pricing models. For such models, the beta method is equivalent to the maximum likelihood method under suitable assumptions regarding the statistical properties of returns and factors. Hence the beta method has a natural advantage for such models. If the SDF method provides as precise an estimate of factor risk premiums even for linear factor-pricing models, then there would be less need for concern that the generality of the SDF method comes at a cost.

We show that in spite of its generality, the SDF method has the same asymptotic precision as the beta method for estimating risk premiums in linear factor-pricing models. Monte Carlo simulations suggest that they provide estimates with similar precision even in finite samples. We also find that the two methods have similar power of specification tests against the alternative hypothesis of nonzero pricing errors. If our findings were otherwise, there would have been some advantage to applying the beta method to nonlinear asset-pricing models through linear approximations. Our results suggest that there are no such gains.

The above results show that it is inappropriate to compare directly the risk premium parameters in the two methods at special values without proper transformation. In addition, by assuming that factors can be standardized to zero mean and unit variance, Kan and Zhou (1999) ignore the random errors associated with the sample mean and variance of the factors and treat them as known. It follows from our analysis that this practice leads to substantial overstatement of the precision with which risk premiums are estimated using the beta method. In contrast, the reported standard errors will be correct if the SDF method is used, which is an important advantage rather than a disadvantage of the SDF method.

Appendix: Proofs

Proof of Theorems 1 and 2

Let Ω be the variance of ϵ_t in equation (2). To calculate the asymptotic variance of the estimator $\hat{\lambda}$ obtained with the SDF method, let us consider the following vector of random variables $g_s(r_t, f_t, \lambda) = r_t(1 - \lambda f_t)$. Substituting r_t with equation (2) and λ by equation (9), we obtain the covariance matrix of g_s as

$$S_s = E[g_s(r_t, f_t, \lambda) g_s(r_t, f_t, \lambda)'] = \frac{\sigma^2(\sigma^4 + \delta^4)}{(\sigma^2 + \mu\delta)^2}\beta\beta' + \frac{\sigma^2(\sigma^2 + \delta^2)}{(\sigma^2 + \mu\delta)^2}\Omega, \quad \text{(A1)}$$

and its inverse is

$$S_s^{-1} = \frac{(\sigma^2 + \mu\delta)^2}{\sigma^2(\sigma^2 + \delta^2)}\Omega^{-1} - \frac{(\sigma^2 + \mu\delta)^2}{\sigma^2(\sigma^2 + \delta^2)}\left(\beta'\Omega^{-1}\beta + \frac{\sigma^4 + \delta^4}{\sigma^2 + \delta^2}\right)^{-1}\Omega^{-1}\beta\beta'\Omega^{-1}. \tag{A2}$$

The expected value of the derivative of g_s with respect to λ is

$$D_s = \mathrm{E}[\partial g_s/\partial\lambda] = -(\sigma^2 + \mu\delta)\beta. \tag{A3}$$

The asymptotic variance of the estimator of λ is $(D_s'S_s^{-1}D_s)^{-1}$, which gives

$$\mathrm{Avar}(\hat{\lambda}) = \frac{\sigma^2(\sigma^2 + \delta^2)}{(\sigma^2 + \mu\delta)^4}(\beta'\Omega^{-1}\beta)^{-1} + \frac{\sigma^2(\sigma^4 + \delta^4)}{(\sigma^2 + \mu\delta)^4}. \tag{A4}$$

To calculate the asymptotic variance of the sampling pricing error, we define

$$e_s(\hat{\lambda}) = \frac{1}{T}\left(\sum_{t=1}^{T} g_s(r_t, f_t, \hat{\lambda})\right). \tag{A5}$$

It follows from Hansen (1982) that

$$\mathrm{Avar}[e_s(\hat{\lambda})] = S_s - D_s[D_s'S_s^{-1}D_s]^{-1}D_s'. \tag{A6}$$

The sampling pricing error is $\hat{\pi} = e_s(\hat{\lambda})$, which gives

$$\mathrm{Avar}[\hat{\pi}] = \frac{\sigma^2(\sigma^2 + \delta^2)}{(\sigma^2 + \mu\delta)^2}[\Omega - (\beta'\Omega^{-1}\beta)^{-1}\beta\beta']. \tag{A7}$$

Next, let us calculate the asymptotic variance of λ^*. In the beta method, the vector of unknown parameters is $\theta = (\delta, \beta', \mu, \sigma^2)'$. Let us define a random variable $g_b(r_t, f_t, \theta)$ as

$$g_b(r_t, f_t, \theta) = \begin{pmatrix} r_t - (\delta + f_t - \mu)\beta \\ (r_t - (\delta + f_t - \mu)\beta)f_t \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix} = \begin{pmatrix} \epsilon_t \\ \epsilon_t f_t \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix}. \tag{A8}$$

The covariance matrix of g_b and its inverse are

$$S_b = \begin{pmatrix} \Omega & \mu\Omega & 0 & 0 \\ \mu\Omega & (\mu^2+\sigma^2)\Omega & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & 2\sigma^4 \end{pmatrix},$$

$$S_b^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} (\sigma^2+\mu^2)\Omega^{-1} & -\mu\Omega^{-1} & 0 & 0 \\ -\mu\Omega^{-1} & \Omega^{-1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \dfrac{1}{2\sigma^2} \end{pmatrix}. \tag{A9}$$

The expected value of the derivative of g_b with respect to θ is

$$D_b = E\left[\frac{\partial g_b}{\partial \theta'}\right] = \begin{pmatrix} -\beta & -\delta I_n & \beta & 0 \\ -\mu\beta & -(\sigma^2+\mu\delta)I_n & \mu\beta & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{A10}$$

It follows that

$$D_b' S_b^{-1} D_b = \begin{pmatrix} \beta'\Omega^{-1}\beta & \delta\beta'\Omega^{-1} & -\beta'\Omega^{-1}\beta & 0 \\ \delta\Omega^{-1}\beta & (\sigma^2+\delta^2)\Omega^{-1} & -\delta\Omega^{-1}\beta & 0 \\ -\beta'\Omega^{-1}\beta & -\delta\beta'\Omega^{-1} & \dfrac{1}{\sigma^2}+\beta'\Omega^{-1}\beta & 0 \\ 0 & 0 & 0 & \dfrac{1}{2\sigma^4} \end{pmatrix}. \tag{A11}$$

The asymptotic covariance matrix of the estimator θ^* is $(D_b' S_b^{-1} D_b)^{-1}$. We find that the inverse matrix of $D_b' S_b^{-1} D_b$ is

$$\begin{pmatrix} \frac{\sigma^2+\delta^2}{\sigma^2}(\beta'\Omega^{-1}\beta)^{-1}+\sigma^2 & -\frac{\delta}{\sigma^2}(\beta'\Omega^{-1}\beta)^{-1}\beta' & \sigma^2 & 0 \\ -\frac{\delta}{\sigma^2}(\beta'\Omega^{-1}\beta)^{-1}\beta & \frac{1}{\sigma^2+\delta^2}\Omega+\frac{\delta^2}{\sigma^2(\sigma^2+\delta^2)}(\beta'\Omega^{-1}\beta)^{-1}\beta\beta' & 0 & 0 \\ \sigma^2 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & 2\sigma^4 \end{pmatrix}. \tag{A12}$$

The asymptotic variance of the estimator for δ is the first element of the above matrix, which is then

$$\text{Avar}(\delta^*) = \frac{\sigma^2+\delta^2}{\sigma^2}(\beta'\Omega^{-1}\beta)^{-1}+\sigma^2. \tag{A13}$$

Using equation (13) and the Delta method, we find that

$$\text{Avar}(\lambda^*) = \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}(\beta'\Omega^{-1}\beta)^{-1}+\frac{\sigma^2(\sigma^4+\delta^4)}{(\sigma^2+\mu\delta)^4}. \tag{A14}$$

To calculate the asymptotic variance of the sampling pricing error obtained with the beta method, we define

$$e_b(\theta) = T^{-1}\sum_{t=1}^{T} g_b(r_t, f_t, \theta). \tag{A15}$$

It follows from Hansen (1982) that

$$\begin{aligned} \text{Avar}[e_b(\theta^*)] &= S_b - D_b(D_b' S_b^{-1} D_b)^{-1} D_b' \\ &= \begin{pmatrix} \frac{\sigma^2}{\sigma^2+\delta^2}[\Omega-(\beta'\Omega^{-1}\beta)^{-1}\beta\beta'] & A_{12} & 0_{n\times 1} & 0_{n\times 1} \\ A_{21} & A_{22} & 0_{n\times 1} & 0_{n\times 1} \\ 0_{1\times n} & 0_{1\times n} & 0 & 0 \\ 0_{1\times n} & 0_{1\times n} & 0 & 0 \end{pmatrix} \end{aligned} \tag{A16}$$

where

$$A_{12} = A_{21}' = \frac{(\mu(\sigma^2 + \delta^2) + \delta(\mu^2 + \sigma^2))}{\sigma^2 + \delta^2}[\Omega - (\beta'\Omega^{-1}\beta)^{-1}\beta\beta'] \tag{A17}$$

$$A_{22} = \frac{(\delta^2 - \mu^2)(\mu^2 + \sigma^2)}{\sigma^2 + \delta^2}\Omega - \frac{(\mu + \delta)^2(\sigma^2 + \mu\delta)^2}{\sigma^2(\sigma^2 + \delta^2)}(\beta'\Omega^{-1}\beta)^{-1}\beta\beta'. \tag{A18}$$

The sample analog of Jensen's alpha in the beta method is $\alpha^* = Q^*e(\theta^*)$, where $Q^* = [I_n, 0_{n\times n}, -\beta^*, 0_{n\times 1}]$. Thus, we have $\text{Avar}[\alpha^*] = Q[S_b - D_b(D_b'S^{-1}D_b)^{-1}D_b]Q'$, where $Q = [I_n, 0_{n\times n}, -\beta, 0_{n\times 1}]$, which gives $\text{Avar}[\alpha^*] = [\sigma^2/(\sigma^2 + \delta^2)][\Omega - (\beta'\Omega^{-1}\beta)^{-1}\beta\beta']$. It then follows from equation (14) that the asymptotic variance of the sample pricing errors is

$$\begin{aligned} \text{Avar}[\pi^*] &= \left(\frac{\sigma^2}{\sigma^2 + \mu\delta}\right)^2 \text{Avar}[\alpha^*] \\ &= \frac{\sigma^6}{(\sigma^2 + \delta^2)(\sigma^2 + \mu\delta)^2}[\Omega - (\beta'\Omega^{-1}\beta)^{-1}\beta\beta']. \end{aligned} \tag{A19}$$

Finally, the equality $\text{Avar}[\hat{\lambda}] = \text{Avar}[\lambda^*]$ follows from equations (A4) and (A14). The matrix $\text{Avar}[\hat{\pi}] - \text{Avar}[\pi^*]$ is positive semidefinite because equations (A7) and (A19) imply

$$\text{Avar}[\hat{\pi}] - \text{Avar}[\pi^*] = \frac{\sigma^2\delta^2(2\sigma^2 + \delta^2)}{(\sigma^2 + \delta^2)(\sigma^2 + \mu\delta)^2}[\Omega - (\beta'\Omega^{-1}\beta)^{-1}\beta\beta'] \tag{A20}$$

and because $\Omega - (\beta'\Omega^{-1}\beta)^{-1}\beta\beta'$ is positive semidefinite. The proof is then complete. Q.E.D.

Proof of Lemma 1: Let

$$D_1 = \lim_{T\to\infty} \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial g(x_t, \theta_1, \theta_2)}{\partial \theta_1'}\right] \quad \text{and} \quad D_2 = \lim_{T\to\infty} \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial g(x_t, \theta_1, \theta_2)}{\partial \theta_2'}\right]. \tag{A21}$$

It follows that

$$D = \lim_{T\to\infty} \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial g(x_t, \theta_1, \theta_2)}{\partial \theta'}\right] = (D_1, D_2). \tag{A22}$$

Let S be the spectral density matrix of $g(x_t, \theta_1, \theta_2)$. We thus have $\text{Avar}[\tilde{\theta}_1] = (D_1' S^{-1} D_1)^{-1}$ and $\text{Avar}[(\hat{\theta}_1', \hat{\theta}_2')'] = (D' S^{-1} D)^{-1}$. The inverse of the asymptotic variance of $(\hat{\theta}_1', \hat{\theta}_2')'$ is

$$D' S^{-1} D = \begin{pmatrix} D_1' S^{-1} D_1 & D_1' S^{-1} D_2 \\ D_2' S^{-1} D_1 & D_2' S^{-1} D_2 \end{pmatrix}. \tag{A23}$$

It follows from the formula of the inverse of partitioned matrix that

$$\begin{aligned} \text{Avar}(\hat{\theta}_1) = {} & (D_1' S^{-1} D_1)^{-1} + (D_1' S^{-1} D_1)^{-1} (D_1' S^{-1} D_2) \\ & \times [(D_2' S^{-1} D_2) - (D_2' S^{-1} D_1)(D_1' S^{-1} D_1)^{-1} (D_1' S^{-1} D_2)]^{-1} \\ & \times (D_2' S^{-1} D_1)(D_1' S^{-1} D_1)^{-1}. \end{aligned} \tag{A24}$$

The difference, $\text{Avar}[\hat{\theta}_1] - \text{Avar}[\tilde{\theta}_1]$, must be positive semidefinite because the first term on the right-hand side of equation (A24) equals $\text{Avar}[\tilde{\theta}_1]$ and the other term is positive semidefinite. In view of equation (A24), it is obvious that equation (18) implies $D_2 = 0$, and thus $\text{Avar}[\hat{\theta}_1] = (D_1 S^{-1} D_1)^{-1}$, which is the same as $\text{Avar}[\tilde{\theta}_1]$. Q.E.D.

Proof of Lemma 2: Suppose that the dimensions of $g_1(x_t, \theta)$ and $g_2(x_t)$ are m and n, respectively, and that the dimensions of θ is k. Let

$$S_{ij} = \sum_{j=-\infty}^{\infty} E[g_i(x_t, \theta) g_j(x_{t+j}, \theta)'], \quad \forall\ i, j = 1, 2. \tag{A25}$$

Let us denote the spectral density matrix of $g(x_t, \theta)$ and its inverse by

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \quad \text{and} \quad S^{-1} = \begin{pmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{pmatrix}. \tag{A26}$$

Define

$$D_1 = \lim_{T \to \infty} \text{E}\left[\frac{1}{T} \sum_{t=1}^{T} \frac{\partial g_1(x_t, \theta)}{\partial \theta'} \right] \quad \text{and} \quad D = \begin{pmatrix} D_1 \\ 0_{n \times k} \end{pmatrix}. \tag{A27}$$

The asymptotic variances of $\hat{\theta}$ and $\tilde{\theta}$ are, respectively,

$$\text{Avar}[\hat{\theta}] = (D' S^{-1} D)^{-1} \quad \text{and} \quad \text{Avar}[\tilde{\theta}] = (D_1' S_{11}^{-1} D_1)^{-1}. \tag{A28}$$

By the formula of the inverse of partitioned matrix, we have

$$S^{11} = S_{11}^{-1} + S_{11}^{-1}S_{12}(S_{22} - S_{21}S_{11}^{-1}S_{12})^{-1}S_{21}S_{11}^{-1}, \tag{A29}$$

which implies

$$D'S^{-1}D = D_1'S^{11}D_1 = D_1'S_{11}^{-1}D_1 + D_1'S_{11}^{-1}S_{12}(S_{22} - S_{21}S_{11}^{-1}S_{12})^{-1}S_{21}S_{11}^{-1}D_1. \tag{A30}$$

The inverse is then

$$\begin{aligned}(D'S^{-1}D)^{-1} = {} & (D_1'S_{11}^{-1}D_1)^{-1} - (D_1'S_{11}^{-1}D_1)^{-1}D_1'S_{11}^{-1}S_{12} \\ & \times [S_{22} - S_{21}S_{11}^{-1}S_{12} + S_{21}S_{11}^{-1}D_1(D_1'S_{11}^{-1}D_1)^{-1}D_1'S_{11}^{-1}S_{12}] \\ & \times S_{21}S_{11}^{-1}D_1(D_1'S_{11}^{-1}D_1)^{-1}.\end{aligned} \tag{A31}$$

Since both $S_{22} - S_{21}S_{11}^{-1}S_{12}$ and $S_{21}S_{11}^{-1}D_1(D_1'S_{11}^{-1}D_1)^{-1}D_1'S_{11}^{-1}S_{12}$ are positive semidefinite, it follows from equation (A31) that $(D'S^{-1}D)^{-1} - (D_1'S_{11}^{-1}D_1)^{-1}$ is negative semidefinite, which implies that Avar$[\hat{\theta}]$ – Avar$[\tilde{\theta}]$ is negative semidefinite. If equation (19) holds, then $S_{12} = 0$ and $S_{21} = 0$, which imply Avar$[\hat{\theta}]$ = Avar$[\tilde{\theta}]$ by equation (A31). Q.E.D.

Proof of Inequality (21): As with the derivation of Avar$[\lambda^*]$ and Avar$[\pi^*]$ in the proof of Theorem 1, one can obtain the asymptotic variances of $\lambda^\dagger$ and $\pi^\dagger$ as

$$\mathrm{Avar}(\lambda^\dagger) = \frac{\sigma^2(\sigma^2 + \delta^2)}{(\sigma^2 + \mu\delta)^4}(\beta'\Omega^{-1}\beta)^{-1}. \tag{A32}$$

The inequality Avar$[\hat{\lambda}] >$ Avar$[\lambda^\dagger]$ is obtained by comparing equations (A4) and (A32). Q.E.D.

Proof of Equation (25): To calculate the asymptotic variance of the estimator $\hat{\lambda}$ obtained with the SDF method, let us consider the following vector of random variables

$$g(r_t, f_t, \lambda) = \begin{pmatrix} r_t(1 - \lambda f_t) \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix}. \tag{A33}$$

Substituting r_t with equation (2), which is implied by the i.i.d.-normal assumption, and λ with equation (9), we obtain the covariance matrix of g as

$$S = \begin{pmatrix} \frac{\sigma^2}{(\sigma^2 + \mu\delta)^2}[(\sigma^4 + \delta^4)\beta\beta' + (\sigma^2 + \delta^2)\Omega] & \frac{\sigma^2(\sigma^2 - \delta^2)}{\sigma^2 + \mu\delta}\beta & -\frac{2\delta\sigma^4}{\sigma^4 + \mu\delta}\beta \\ \frac{\sigma^2(\sigma^2 - \delta^2)}{\sigma^2 + \delta\mu}\beta' & \sigma^2 & 0 \\ -\frac{2\delta\sigma^4}{\sigma^2 + \mu\delta}\beta' & 0 & 2\sigma^4 \end{pmatrix}. \tag{A34}$$

The expected value of the derivative of g with respect to λ is

$$D = \mathrm{E}\left[\frac{\partial g}{\partial \lambda}\right] = \begin{pmatrix} -(\sigma^2 + \mu\delta)\beta \\ 0 \\ 0 \end{pmatrix}. \tag{A35}$$

After some algebraic manipulation, we obtain

$$D'S^{-1}D = \frac{(\sigma^2 + \mu\delta)^4}{\sigma^2(\sigma^2 + \delta^2)}(\beta'\Omega^{-1}\beta). \tag{A36}$$

The asymptotic variance of the estimator of λ is $(D'S^{-1}D)^{-1}$, which gives

$$\mathrm{Avar}(\bar{\lambda}) = \frac{\sigma^2(\sigma^2 + \delta^2)}{(\sigma^2 + \mu\delta)^4}(\beta'\Omega^{-1}\beta)^{-1}. \tag{A37}$$

Finally, we obtain the equality $\mathrm{Avar}[\bar{\lambda}] = \mathrm{Avar}[\lambda^{\dagger}]$ by comparing equations (A37) and (A32). Q.E.D.

REFERENCES

Black, Fischer, Michael Jensen, and Myron S. Scholes, 1972, The capital asset pricing model: Some empirical tests, in Michael Jensen, ed.: *Studies in the Theory of Capital Markets* (Praeger, New York).

Bollerslev, Tim, Ray Chou, and Kenneth Kroner, 1992, ARCH modeling in finance, *Journal of Econometrics* 52, 5–59.

Breeden, Douglas, Michael Gibbons, and Robert Litzenberger, 1989, Empirical tests of the consumption-oriented CAPM, *Journal of Finance* 44, 231–262.

Campbell, John, 1993, Intertemporal asset pricing without consumption data, *American Economic Review* 83, 487–512.

Campbell, John, 1996, Understanding risk and return, *Journal of Political Economy* 104, 298–345.

Campbell, John, Andrew Lo, and Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press, Princeton, NJ).

Chen, Naifu, Richard Roll, and Stephen Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383–404.

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

Cochrane, John, 2000, A resurrection of the stochastic discount factor/GMM methodology, Manuscript, Graduate School of Business, University of Chicago.

Cochrane, John, 2001, *Asset Pricing* (Princeton University Press, Princeton, NJ).

Dybvig, Philip, and Jonathan Ingersoll, 1982, Mean-variance theory in complete markets, *Journal of Business* 55, 233–252.

Fama, Eugene, and Kenneth French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene, and Kenneth French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene, and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 71, 607–636.

Farnsworth, Heber, Wayne Ferson, David Jackson, and Steven Todd, 2000, Performance evaluation with stochastic discount factors, *Journal of Business*, forthcoming.

Ferson, Wayne, 1995, Theory and empirical testing of asset pricing models, in Robert Jarrow, Vojislav Maksimovic, and William Ziemba, ed.: *Handbooks in Operations Research and Management Science* 9, 145–200.

Ferson, Wayne, and Stephen Foerster, 1994, Finite sample properties of the generalized method of moments in tests of conditional asset pricing models, *Journal of Financial Economics* 36, 29–55.

Ferson, Wayne, and Campbell Harvey, 1997, Fundamental determinants of national equity market returns: A perspective on conditional asset pricing, *Journal of Banking and Finance* 21, 1625–1665.

French, Kenneth, William Schwert, and Robert Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 17, 5–26.

Gibbons, Michael, Stephen Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Glosten, Lawrence, Ravi Jagannathan, and David Runkle, 1992, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779–1801.

Hamilton, James, 1994, *Time Series Analysis* (Princeton University Press, Princeton, NJ).

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

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

Hansen, Lars, and Ravi Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance* 52, 557–590.

Hansen, Lars, and Scott Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 55, 587–613.

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

Ingersoll, Jonathan, 1987, *Theory of Financial Decision Making* (Rowman & Littlefield, Totowa, NJ).

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Jagannathan, Ravi, and Zhenyu Wang, 1998, An asymptotic theory for estimating beta-pricing models using cross-sectional regression, *Journal of Finance* 53, 1285–1309.

Jagannathan, Ravi, and Zhenyu Wang, 1999, Efficiency of the stochastic discount factor method for estimating risk premium, Working paper, Columbia University.

Kan, Raymond, and Guofu Zhou, 1999, A critique of the stochastic discount factor methodology, *Journal of Finance* 54, 1221–1248.

MacKinlay, Craig, and Matthew Richardson, 1991, Using generalized method of moments to test mean-variance efficiency, *Journal of Finance* 46, 511–527.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1–34.

Sharpe, William, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–422.

Taylor, Malcolm, and James Thompson, 1986, Data based random number generation for a multivariate distribution via stochastic simulation, *Computational Statistics & Data Analysis* 4, 93–101.