Efficiency of the Stochastic Discount Factor Method for Estimating Risk Premiums

Ravi Jagannathan Zhenyu Wang*

April 19, 2000†

Abstract

The stochastic discount factor (SDF) method provides an elegant and unified general framework for econometric analysis of linear and nonlinear asset pricing models including derivative pricing models. We examine whether the generality of the SDF methodology comes at a cost in estimation efficiency. For linear beta pricing models, we show that the SDF method provides estimates of factor risk premiums that are as precise as those estimates obtained using classical regression methods. In the special case where the mean and variance of the factors are known, the common practice of ignoring the implied restrictions on the moments of the factors makes the SDF method substantially less precise than the regression method. However, proper incorporation of the relevant restrictions makes the SDF method asymptotically as precise as the classical regression method.

*Jagannathan: Kellogg Graduate School of Management, Northwestern University, and the National Bureau of Economic Research. Wang: Graduate School of Business, Columbia University, and the Federal Reserve Bank of New York. We thank John Cochrane, Wayne Ferson, Bob Hodrick, Narayana Kocherlakota, Guofu Zhou and the seminar participants at the Federal Reserve Bank of New York, Columbia University and the University of Southern California for helpful comments, Adam Kolasinski for excellent research assistance, Lolotte Olkowski for editorial comments. We especially benefited from discussions with Kent Daniel and John Heaton. Views expressed are those of the authors and not necessarily those of the Federal Reserve Bank of New York or the Federal Reserve System.

†First draft: September 12, 1999.

1 Introduction

The stochastic discount factor (SDF) method is widely used in empirical evaluation of asset pricing models. Unlike the classical method that estimates parameters from linear regression equations, the SDF method is more general. It provides a coherent and unified framework for econometric analysis of both linear and nonlinear asset pricing models found in the literature, including pricing models for derivative securities. The generality of the framework provides a deeper understanding of the econometric issues involved in evaluating asset pricing models using financial markets data. As Cochrane (1999) points out, such unified framework is the primary attraction of the SDF method. Hence, it is not surprising that the SDF method is replacing the regression method for most applications.

In spite of its wide use, little is known about the estimation efficiency of the SDF method relative to the classical regression method. A question that arises is whether the generality of the SDF framework comes at a cost in estimation efficiency, especially for risk premiums. In order to provide an answer to this question, we compare the precision of estimated factor risk premiums obtained by using the SDF method and the classical regression method. If the SDF method turns out to be inefficient relative to the classical regression method for linear models, some variation of the classical regression method will dominate the SDF method for nonlinear models as well, in terms of estimation efficiency. This is because a nonlinear model can be locally approximated as a linear model.

Estimates of factor risk premiums are frequently used for calculating the cost of capital using beta pricing models while making capital budgeting decisions and monitoring regulated monopolies. Applications of factor models in capital budgeting decisions can be found in standard textbooks on corporate finance. In order to calculate the cost of capital for the regulation of New York public utility companies, Elton, Gruber and Mei (1994) estimate the risk premiums for macro economic factors, and Schink and Bower (1994) estimate the risk premiums for book-to-market and size factors. Fama and French (1997) focus on the precision of the estimated industry cost of capital using various factor models. In view of the importance given to precise estimation of factor risk premiums in the literature, we compare the precision of the risk premiums estimated using the SDF and the classical regression methods.

For expositional convenience, we consider the case where there is only one economy-wide pervasive risk factor. We prove that, when a financial economist does not know the mean and variance of the economy-wide pervasive factor, asymptotically the SDF method provides as precise an estimate of the risk premium as the classical regression method. Using Monte Carlo simulations, we show that the two methods provide equally precise estimates in finite samples as well. The sampling errors in the two methods are also similar in the presence of conditional heteroscedasticity. These results are contrary to the assertions made by Kan and Zhou (1999) that the SDF method is far less efficient than the regression method.

To simplify their analysis, Kan and Zhou (1999) assume that the pervasive risk factor has a zero mean and unit variance. This is equivalent to the assumption that a financial economist knows the mean and variance of the factor. This is not an entirely unrealistic assumption. The financial economist may have strong priors regarding the variance of the factor based on information from actively traded index options. Variances can be estimated rather precisely using high frequency data or with options prices. The estimation error in variances obtained through such methods maybe small enough to be ignored. In many empirical studies of the APT, researchers define a factor as the estimated innovation of an economic variable and ignore the associated sampling errors (e.g., Chen, Roll and Ross, 1986). This is equivalent to the assumption that the mean of the factor is known.

When the financial economist assumes that the mean and the variance of the factor are known, the common practice in this literature is to ignore the restriction that the sample first and second moments of the factor equal their population counterparts in expectation. This is probably based on our intuition from working with the regression model where ignoring the restrictions on the factor moments does not affect the asymptotic precision of the estimated risk premium. Following this standard practice, Kan and Zhou (1999) also ignore the restrictions on the factor moments and conclude that the sampling error in the SDF method is 20–30 times as high as the sampling error in the regression method. A central issue in financial economics is to understand the magnitude of the risk premium on a pervasive economy-wide risk and its relation to fundamentals. Hence, a method that provides a more precise estimate of risk premiums will be preferred. If the sampling error in the SDF method is 20–30 times as large as that in the regression method, the cost of using the SDF method will be too high. In that case, the regression method will dominate the SDF method.

We show that it is important to properly incorporate the information contained in the restrictions on the first two moments of the factor while estimating the factor risk premium using the SDF method. When this is done, the SDF method becomes asymptotically as efficient as the regression method. Therefore, when the mean and the variance of the factor are known, there is a substantial reduction in the estimation efficiency associated with the common practice of ignoring the restrictions on the factor moments.

We organize the rest of the paper as follows. In Section 2, we describe the SDF and the classical regression methods for estimating linear factor pricing models. In section 3, we first compare the asymptotic variance of the estimated risk premiums obtained by using the SDF and regression methods. We then compare the finite sample properties of the estimators using Monte Carlo simulations. We conclude in section 4.

2 Description of the SDF and Regression Methods

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

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

where $\beta = \text{cov}(r_t, f_t)/\sigma^2$ is the sensitivity of asset returns to the factor, and δ is the factor risk premium. Note that the vector β can be consistently estimated using the time-series regression: $r_t = \alpha + \beta f_t + \epsilon_t$. 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 in the above regression equation: $\alpha = (\delta - \mu)\beta$. Substituting out α in the regression, we obtain the following specification of stock returns:

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

For convenience, we refer to equation (2) as the *regression equation*.

By substituting the expression for β and rearranging the terms, the linear asset pricing model given in equation (1) can be equivalently written as the following:

$$E[r_t m_t] = 0 \ , \tag{3}$$

where $m_t = 1 - \lambda f_t$ and $\lambda = \delta/(\sigma^2 + \mu\delta)$. According to equation (3), the expected value of the excess return on an asset discounted by m_t equals zero. In view of this interpretation, any random variable, m_t, that satisfies equation (3) is referred to as a *stochastic discount factor* (SDF).

When the law of one price holds, there always exists a random variable m_t that satisfies equation (3). In general, a number of random variables satisfying equation (3) will exist and hence there will be several stochastic discount factors. As Dybvig and Ingersoll (1982) showed, the linear factor pricing model identifies the random variable, $m_t = 1 - \lambda f_t$, as a stochastic discount factor. In that case, equation (3) can be viewed as the *Euler equation* (first order condition) for the portfolio choice problem faced by an investor whose intertemporal marginal rate of substitution is a linear function of the factor. Hansen and Richard (1987) observed that every asset pricing model can be written in the form of equation (3) for some candidate stochastic discount factor m_t. However, the stochastic discount factor identified by an asset pricing model will in general be a nonlinear function of observable factors and model parameters[1]. Hence the stochastic discount factor method provides a unified way to represent linear and nonlinear asset pricing models.

2.1 Estimating the factor risk premium using the SDF method

The stochastic discount factor (SDF) method uses the Euler equation (3) to estimate the risk premium. Since the stochastic discount factor m_t is parameterized by the risk premium as well as the mean and variance of the factor, it is necessary to estimate the mean and variance of the factor along with the risk premium. Let $\theta = (\delta, \mu, \sigma^2)$ denote the vector of unknown parameters that the financial economist has to estimate. We need the following three moment restrictions for estimating θ.

$$E[r_t(1 - \frac{\delta}{\sigma^2 + \mu\delta} f_t)] = 0 \tag{4}$$

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

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

Let $e_T(\theta)$ denote the sample analog of the left-hand side of the moment restrictions given

[1]Theoretical models that lead to the stochastic discount factor representation of asset pricing models is discussed extensively in Ingersoll (1987). He does not use the term 'stochastic discount factor' – the term was coined by Hansen and Richard (1987). The use of this representation for econometric evaluation of asset pricing models is developed in Hansen and Singleton (1982). Hansen and Jagannathan (1991, 1997), discuss the general properties of stochastic discount factors. Ferson (1995), Campbell, Lo and MacKinlay (1997) and Cochrane (1999) provide a comprehensive introduction to the stochastic discount factor framework.

above:

$$e_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} g_t(\theta) \qquad \text{where} \qquad g_t(\theta) = \begin{pmatrix} r_t(1 - \frac{\delta}{\sigma^2+\mu\delta} f_t) \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix} .$$

Then the expected value of $g_t(\theta)$ is zero. We assume that the elements of the vector $(r_t', f_t)'$ follow a jointly stationary and ergodic process. Since $e_T(\theta)$ is the sample mean of a process whose population mean is zero, this implies that e_T will satisfy the Central Limit Theorem (see Hamilton, 1994). Thus, we have:

$$\lim_{T\to\infty} \sqrt{T} e_T(\theta) \sim N(0, S) ,$$

for some positive definite matrix S.

Let W_T be a consistent estimator of S^{-1}. The GMM estimator of θ is given by:

$$\hat{\theta} = \arg\min_{\theta} e_T(\theta)' W_T e_T(\theta) .$$

Under some conditions, the Law of Large Numbers implies that $\lim_{T\to\infty} E\left[\partial e_T(\theta)/\partial\theta'\right]$ exists. Let us denote it by D and assume that the rank of D equals the dimension of θ. It follows from Hansen (1982) that the asymptotic distribution of $\hat{\theta}$ is given by:

$$\lim_{T\to\infty} \sqrt{T}(\hat{\theta} - \theta) \sim N(0, (D'S^{-1}D)^{-1}) .$$

Let Avar$(\hat{\delta})$ denote the asymptotic variance of the estimator for the risk premium δ. Then Avar$(\hat{\delta})$ is the first element of the matrix $(D'S^{-1}D)^{-1}$. Financial economists often use the J-statistic to test the validity of the model.

We also consider the case where μ and σ^2 are assumed to be known. In this case, the only unknown parameter is the risk premium δ. First, following common practice, we ignore the moment restrictions given by equations (5) and (6) and use only equation (4) to estimate δ. We denote such an estimator by $\bar{\delta}_{\mu,\sigma}$. This estimator is equivalent to the one considered in Kan and Zhou (1999) for estimating the risk premium with the SDF method. Even when μ and σ^2 are assumed to be known, one can still use all the moment restrictions in equations (4), (5) and (6) to estimate the risk premium δ. The resulted estimator is denoted by $\hat{\delta}_{\mu,\sigma}$. We will show that the estimator $\hat{\delta}_{\mu,\sigma}$ distributes very differently than the estimator $\bar{\delta}_{\mu,\sigma}$.

2.2 The regression method

The regression method uses the regression equation (2) to estimate the risk premium. To implement the GMM[2], we make use of the moment restrictions implied by the regression equation. There are two moment restrictions from the regression equation — the zero mean of the residuals and the orthogonality between the residuals and the factor. In addition, there are two moment restrictions defining the mean and variance of the factor. The four restrictions are given below:

$$E[r_t - (\delta - \mu + f_t)\beta] = 0 \tag{7}$$

$$E[(r_t - (\delta - \mu + f_t)\beta)f_t] = 0 \tag{8}$$

$$E[f_t - \mu] = 0 \tag{9}$$

$$E[(f_t - \mu)^2 - \sigma^2] = 0 \; . \tag{10}$$

We use the GMM to estimate the parameter δ along with β, μ and σ. Define $\theta = (\delta, \beta', \mu, \sigma^2)'$ and

$$e_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} g_t(\theta) \; , \quad \text{where} \quad g_t(\theta) = \begin{pmatrix} r_t - (\delta - \mu + f_t)\beta \\ (r_t - (\delta - \mu + f_t)\beta)f_t \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix} \; .$$

Let $D = \lim_{T\to\infty} E[\partial e_T / \partial \theta']$ and S be the variance of $\lim_{T\to\infty} \sqrt{T} e_T$. Then the asymptotic distribution of the GMM estimator θ^* (see Hansen, 1982) is given by:

$$\lim_{T\to\infty} \sqrt{T}(\theta^* - \theta) \sim N(0, (D'S^{-1}D)^{-1}) \; .$$

The asymptotic variance of the estimator δ^* is the first element of the matrix $(D'S^{-1}D)^{-1}$.

When the mean and variance of the factor are assumed to be known, δ is the only unknown parameter. In that case, we first consider using all of the above moment restrictions to estimate δ and denote the obtained estimator as $\delta^*_{\mu,\sigma}$. We also consider estimating the risk premium by dropping equations (9) and (10) and using only equations (7) and (8). We denote the resulted estimator by $\delta^\dagger_{\mu,\sigma}$. This estimator is equivalent to the one considered in Kan and Zhou (1999) for estimating the risk premium with the regression method.

[2]The regression equation can be estimated using OLS, GLS or GMM. MacKinlay and Richardson (1991) show that GLS and GMM are equivalent when returns and factors are normally distributed jointly but the GMM estimator is to be preferred in the presence of conditional heteroscedasticity. Ferson and Harvey (1997) extend the equivalence argument to models with time varying betas that are linear functions of lagged instruments.

3 Comparing the SDF and Regression Methods

3.1 Asymptotic precision

Having described the SDF and the classical estimators, we now proceed to compare their asymptotic efficiency. For analytical convenience, we assume that time-series observations, $\{(r_t', f_t)'\}_{t=1,\cdots,T}$, are drawn from identical and independent multivariate Normal distributions. In that case, the asymptotic variances of the GMM estimators described in the earlier section can be derived analytically. This assumption gives conditional homoscedasticity, i.e., ϵ_t and f_t are independent, and ϵ_t, being conditional on f_t, has a constant covariance matrix, which is denoted by Ω. The derivation of the asymptotic variances is provided in Appendix A and B and the results are summarized in the following three theorems.

The first theorem shows that, in large samples, the two estimators for the risk premium have the same precision, when the financial economist does not know the mean and variance.

Theorem 1 *If the financial economist estimates the mean and the variance of the factor together with the risk premium, the GMM estimator for the risk premium obtained by using the Euler equation is asymptotically as efficient as that obtained by using the regression equation, i.e.,* $\text{Avar}(\hat{\delta}) = \text{Avar}(\delta^*)$.

When the factor f_t is the excess return of a portfolio of financial assets traded in the market, application of the asset pricing model to the factor implies that the risk premium equals the mean of the factor, i.e., $\delta = \mu$. It can be shown that imposing the restriction $\delta = \mu$ does not affect Theorem 1. In fact, under this restriction the asymptotic variances of $\text{Avar}(\hat{\delta})$ and $\text{Avar}(\delta^*)$ are both equal to the asymptotic variance of the sample average of f_t. Therefore, if imposing the restriction $\delta = \mu$, we only need to estimate the risk premium from the observations of the factor, and incorporation of asset returns does not improve estimation efficiency. In the rest of this paper, however, we focus on the general case in which the restriction $\delta = \mu$ is not imposed.

Now consider the case where the mean and the variance of the factor are known, and the financial economist ignores the two related moment restrictions. For this case, the next theorem shows that the SDF method provides a less precise estimate of the risk premium than the regression method.

Theorem 2 *Suppose the financial economist knows the mean and variance of the factor and ignores the two moment restrictions that define the mean and variance of the factor while estimating the risk premium. The GMM estimator for the risk premium obtained by using the Euler equation is asymptotically less efficient than that obtained by using the regression equation, i.e.,* $\mathrm{Avar}(\bar{\delta}_{\mu,\sigma}) > \mathrm{Avar}(\delta^{\dagger}_{\mu,\sigma})$.

The next theorem shows that, even when the mean and variance of the factor are assumed to be known, the two moment restrictions that define the mean and variance of the factor contain useful information for estimating the risk premium using the SDF method.

Theorem 3 *Suppose the financial economist knows the mean and variance of the factor and incorporates the two moment restrictions on the factor while estimating the risk premium. Then the GMM estimator for the risk premium obtained by using the Euler equation is asymptotically as efficient as that obtained by using the regression equation, i.e.,* $\mathrm{Avar}(\hat{\delta}_{\mu,\sigma}) = \mathrm{Avar}(\delta^{*}_{\mu,\sigma})$.

The derivation in Appendix B indicates that for the regression method, the last two elements of g_t are uncorrelated with the other elements. Therefore, omitting restrictions (9) and (10) does not affect the asymptotic efficiency of the estimator. For the SDF method, however, the last two elements of g_t are correlated with the other elements, as can be seen in Appendix A. Therefore, omission of restrictions (5) and (6) causes loss of information.

In Theorems 1, 2 and 3 we assume that the linear beta pricing model correctly prices the given set of assets. In order to examine the validity of the model, the standard practice is to use the J-statistic in Hansen (1982). In the case where the mean and variance of the factor are unknown, the we use $\hat{J}$ to denote the J-statistic in the SDF method, which is

$$\hat{J} = Te_T(\hat{\theta})'W_T e_T(\hat{\theta}) \ ,$$

where e_T and $\hat{\theta}$ are as defined in Section 2.1. In the case where the mean and variance of the factor are known and the restrictions of the factor moments are ignored, we use $\bar{J}_{\mu,\sigma}$ to denote the J-statistic in the SDF method. In the case where the mean and variance of the factor are known and the restrictions of the factor moments are incorporated, we use $\hat{J}_{\mu,\sigma}$ to denote the J-statistic in the SDF method. Similarly, we use J^{*}, $J^{\dagger}_{\mu,\sigma}$ and $J^{*}_{\mu,\sigma}$, for the three cases about the factor moments respectively, to denote the J-statistics in the regression method.

The asymptotic distributions of the above J-statistics in the SDF and regression methods are derived in Appendix C and summarized in the following theorem.

Theorem 4 *In each of the three cases about factor moments, the J-statistics in the SDF and the regression methods have the same asymptotic distribution. To be more specific, as $T \to \infty$, we have $\lim \hat{J} = \lim J^* = \chi^2(n-1)$, $\lim \bar{J}_{\mu,\sigma} = \lim J^{\dagger}_{\mu,\sigma} = \chi^2(n-1)$ and $\lim \hat{J}_{\mu,\sigma} = \lim J^*_{\mu,\sigma} = \chi^2(n+1)$.*

Therefore, the distributions of the J-statistics in the SDF and regression methods can only be different in finite samples or under misspecified models. Such difference may lead to difference in size and power of the tests.

3.2 Precision in finite samples – joint Normal distribution

The formulae for the asymptotic variance of the parameters estimated using the SDF and regression methods given in the theorems are valid only when the length of the time series of observations is *sufficiently large*. Theory does not tell us what the length of the time series of observations should be for it to be considered sufficiently large. Hence, in this section, we examine this issue using Monte Carlo simulation methods.

In this section we examine the finite sample properties of the factor risk premiums estimated using the SDF and the regression methods. The size as well as the power of the J-statistics for testing the validity of the linear factor model is examined by Cochrane (2000). He shows that the test statistics in the SDF and the regression methods have similar size and power. This result is contrary to Kan and Zhou (1999) because they assume that the econometrician knows the mean and the variance of the factor but do not exploit this knowledge in the SDF method.

In the first set of simulations, we assume that the returns and the factor are drawn from identical and independent multivariate Normal distributions, as assumed in our theorems. For purposes of a calibration, we use excess returns on the ten size-sorted portfolios provided by the Center for Research on Security Prices (CRSP). We assumed that the factor is given by the excess return on the value-weighted index of stocks traded on NYSE, AMEX and NASDAQ. We obtained data on monthly gross returns from January 1926 to December 1998 for the 10 size-sorted index portfolios, the value-weighted stock index and the one-month Treasury Bills from CRSP through Wharton Research Data Services (WRDS). The excess returns are obtained by subtracting the Treasury Bill returns from the gross returns.

For our simulation, we choose the values for the elements of the mean vector and the covariance matrix of the excess returns and the factor as follows. We set μ and σ^2 to be the sample mean and sample variance of the excess return on the value-weighted index. We set β to be the slope coefficient obtained from the ten ordinary least squares regressions (OLS) of the portfolio excess returns on the index excess return. The covariance matrix, $\Omega = E[\epsilon_t \epsilon_t' | f_t]$, is set equally to the sample covariance matrix of the residuals obtained in the ten 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 the historical betas. Table 1 provides all the parameters used in the simulation.

The calibrated risk premium is $\delta = 1.3692$ for monthly returns. This large risk premium is in fact the risk premium on firm size because the calibrated β is highly correlated with firm size in portfolios constructed by ranking firms on market capitalization. If we construct portfolios by ranking firms on both market capitalization and estimated beta, the calibrated beta will not be highly correlated with the firm size. In that case, as in Fama and French (1992) and Jagannathan and Wang (1996), the calibrated risk premium on beta will be close to zero, and the calibrated risk premium on firm size will be large.

We draw repeated samples of the excess returns and the factor from the multivariate Normal distribution with means and covariances chosen as above. We considered the following six different time horizons in our simulations: 60 months; 120 months; 360 months; 600 months; 900 and 1200 months. The first two horizons are often used for estimating time-varying risk premiums. The third horizon is similar to the horizon used by many influential studies on equity returns[3]. The fourth and the last horizon are a half and a full century respectively. The fifth horizon roughly corresponds to the length of the period for which CRSP's monthly return data is available. For each time series length, T (= 60; 120; 360; 600; 900; 1200), we drew 100 independent samples of length T and estimated the parameters 100 times. To see the effects of time horizons on the estimators, Table 2 gives the asymptotic standard deviation of the estimated risk premium for different sample sizes.

For each estimator, Table 3 gives the mean and the standard deviation of the 100 esti-

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

mated risk premiums. When the mean and variance of the factor are estimated with the risk premium, the estimator $\hat{\delta}$ in the SDF method and the estimator δ^* in the regression method have the same precision for any of the sample size T considered. Especially, 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 regression method instead of the SDF method.

If we assume that the mean and variance of the factor are known and we ignore the related moment restrictions, the SDF method is much less efficient than the regression method. This is true in both small as well as large samples. When we have 30 years of monthly observations, the standard deviation of the estimator in the SDF method is nearly 20 times as high as that in the regression method. These results are similar to the results reported by Kan and Zhou (1999). However, when all the moment restrictions are used, the relative efficiency of the SDF method improves substantially. Notice that the standard error of $\hat{\delta}_{\mu,\sigma}$ is nearly twice as large as that of $\delta^*_{\mu,\sigma}$ when thirty years of observations are used. As the sample size increases to 100 years of observations, the precision of the SDF method approaches that of the regression method. The improvement in efficiency of the SDF method, relative to the regression method, is expected because the two methods have the same asymptotic variance for estimated risk premiums. Thus, the sharp disadvantage of the SDF method to the regression method reported by Kan and Zhou (1999) is mainly due to the common practice of ignoring the information contained in moment restrictions (5) and (6).

3.3 Precision in finite samples – non-Normal distributions

In our theoretical derivation of the asymptotic distribution theory, 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, as pointed out by MacKinlay and Richardson (1991). In this section, we therefore examine the applicability of our results when returns exhibit conditional heteroscedasticity.

For this purpose, following MacKinlay and Richardson (1991), we make independent draws of the returns and the factor from a multivariate t distribution instead of a joint Normal distribution as in the previous section. 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, 1991, equation 14):

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

Notice that the dependence of the conditional covariance on the realized value of the factor increases as the degree of freedom ν decreases. However, there is a lower bound for the number of degrees of freedom. The asymptotic distribution theory for the GMM requires that returns and factor have finite fourth moments. Hence the degree of freedom has to be higher than 4. In view of this we use 5 degrees of freedom for the multivariate t distribution.

Table 4 gives the simulation results. As can be seen, the standard errors computed using the asymptotic theory are about 10 percent less than the standard deviation given by the simulation, indicating a small bias. It is also true that knowing the mean and variance causes the SDF method less efficient than the regression method for small samples. Further, ignoring the related moment restrictions makes the SDF method far less efficient than the regression method. When the mean and variance of the factor are unknown, which is a more realistic case, and have to be estimated along with the risk premium, the two methods have almost the same precision in every sample size. This indicates that our result is robust to the existence of conditional heteroscedasticity.

As an alternative to the multi-variate t distribution, we also considered 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 as the data of f. The residuals in the regression of decile returns on the index return are used as the data of ϵ. Independent samples $\{(f_t, \epsilon_t')'\}_{t=1,\cdots,T}$ are drawn from a Gaussian kernel estimate of the empirical distribution of the data. We use the method described by Taylor and Thompson (1986). 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, δ, β, μ, σ and Ω, are set to those in Table 1. Each estimator is then calculated based on the T samples to obtain a sample of the estimator. We repeat this independently for 100 times to obtain 100 independent samples of each estimator.

The simulation results are given in Table 5. Again, when the mean and variance of the factor are estimated together with the risk premium, the sampling errors for the risk premium estimated using the SDF method and the regression method are almost identical. Ignoring the moment restrictions on the factor, when μ and σ^2 are known, causes the SDF method to be far less precise than the regression method. Incorporating the moment restrictions on the factor greatly improves the precision of the SDF method.

3.4 Estimation of the parameter λ

In empirical studies that uses the SDF method to evaluate whether the linear factor pricing model is consistent with the time-series data on asset returns, financial economists often examine the following parameterization of the model:

$$E[r_t(1-\lambda f_t)] = 0 \ , \tag{11}$$

where the mean and variance of the factor do not enter as unknown parameters. Equation (11) is typically used to estimate the parameter λ, which is a transformation of the risk premium δ. However, if we use the regression method to estimate λ, we should apply the GMM to the following moment restrictions:

$$\begin{aligned}
&E\left[r_t - \left(\frac{\lambda\sigma^2}{1-\mu\lambda} - \mu + f_t\right)\beta\right] = 0 \\
&E\left[\left(r_t - \left(\frac{\lambda\sigma^2}{1-\mu\lambda} - \mu + f_t\right)\beta\right) f_t\right] = 0 \\
&E[f_t - \mu] = 0 \\
&E[(f_t - \mu)f_t - \sigma^2] = 0 \ .
\end{aligned}$$

For estimating λ, the SDF method is clearly much simpler than the regression method. Nevertheless, the SDF method is as efficient as the regression method. Under the assumption of identical and independent Normal distribution, the asymptotic variances derived in Appendix D and E show that the two methods have the same asymptotic precision for estimating λ. The Monte Carlo simulation presented in Table 6 indicates that they also have the same precision for finite samples and for distributions that give conditional heteroscedasticity.

Although the mean and variance of the factor is not involved in the above Euler equation, one could still estimate them together with λ. In this case, the moment restrictions used in the estimation is

$$E[r_t(1-\lambda f_t)] = 0 \tag{12}$$

$$E[f_t - \mu] = 0 \tag{13}$$

$$E[(f_t - \mu)f_t - \sigma^2] = 0 \ . \tag{14}$$

The asymptotic variances derived in Appendix D show that the estimator obtained from moment restrictions (12), (13) and (14) has the same asymptotic precision as the estimator obtained from the moment restriction (11). Therefore, it also has the same asymptotic precision as the estimator obtained from the regression method.

If the financial economist obtains knowledge about the mean and variance of the factor, it does not affect the SDF method that uses the above Euler equation. In this case, it is a common practice to ignore the last two moment restrictions in the regression method. This does not affect the GMM estimator in the regression method. The asymptotic variances derived in Appendix D and E show that the SDF method is, however, asymptotically less efficient than the regression method when the financial economist treats the mean and variance of the factor as known. Monte Carlo simulations presented in Table 7 show that the SDF method is substantially less efficient relative to the regression method. This is also consistent with the result reported by Kan and Zhou (1999).

When μ and σ^2 are assumed to be known, if we add the restrictions on the factor moments to the SDF method, we obtain a GMM estimator from the moment restrictions in equations (12), (13) and (14). The asymptotic variances derived in Appendix D and E show that this estimator is asymptotically as efficient as the estimator from the regression method in the case where the financial economist knows the mean and variance of the factor before estimating the risk premium. This result is rather surprising because μ and σ^2 does not appear in the Euler equation in (11). Table 8 shows that adding the moment restrictions of the factor into the SDF method greatly improves the precision of the estimated risk premium, although it is still less precise than the regression method for finite samples.

4 Conclusion

The SDF method has received wide attention in the theoretical and empirical asset pricing literature. The main attraction of the SDF method is that it provides an elegant and general framework for econometric evaluation of linear and nonlinear asset pricing models, including pricing models for derivative securities. In this study, we examine whether the generality of the SDF framework comes at a cost in estimation efficiency for risk premiums. Since a nonlinear model can often be well approximated by a linear one, it is natural to examine whether the SDF method provides a more precise estimate for risk premiums than the classical regression method when attention is limited to linear factor pricing models. If the SDF method does poorly in linear models than the regression method, it is likely that we can improve estimation efficiency of the SDF method in nonlinear models.

We show that in spite of its generality, the SDF method and the regression method have the same asymptotic precision for estimating risk premiums in linear asset pricing models. Monte Carlo simulations suggest that they provide estimates with similar precision even in finite samples. Although this result is for linear models, it indicates that transforming the SDF representation of non-linear asset pricing models into different specifications is unlikely to lead to a gain in estimation efficiency, unless we impose assumptions that are more restrictive.

In an environment where the financial economist is blessed with knowledge of the first two moments of the economy-wide pervasive risk factor, the two methods are shown in this paper to have the same asymptotic efficiency as long as the information contained in the restrictions on the first two moments of the factor is properly incorporated. The common practice of ignoring the moment restrictions on the mean and variance of the factor causes the SDF method to be far less efficient than the regression method. Therefore, it is important to take into account all the relevant moment restrictions while using the SDF method.

Appendix: Derivation of the Asymptotic Variances

Under the assumption that the observation of returns and the factor, $(r_t', f_t)'$, is a time-series process with identical and independent normal distribution, the asymptotic variances of the GMM estimators for the risk premium δ can be derived analytically. As we point out in Section 3, this assumption gives conditional homoscedasticity for ϵ_t in the regression equation. In this case, ϵ_t is independent of f_t and its covariance matrix is constant. We denote this covariance matrix as Ω. This assumption also implies that $S = E[g_t g_t']$ and $D = E[\partial g_t / \partial \theta]$ in Section 2.

A Using Euler equation to estimate δ

When we use Euler equation to estimate the risk premium, the moment restrictions are

$$\begin{aligned} &E[r_t(1 - \frac{\delta}{\sigma^2 + \mu\delta} f_t)] = 0 \\ &E[f_t - \mu] = 0 \\ &E[(f_t - \mu)^2 - \sigma^2] = 0 \ . \end{aligned}$$

Let us define a random variable g_t as

$$g_t = \begin{pmatrix} r_t(1 - \frac{\delta}{\sigma^2+\mu\delta} f_t) \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix} \ .$$

The covariance matrix of g_t is $S = E[g_t g_t']$. It follows from $r_t = (\delta + f_t - \mu)\beta + \epsilon_t$ that S is

$$\begin{pmatrix} \frac{\sigma^2}{(\sigma^2+\mu\delta)^2}\left((\sigma^4 + \delta^4)\beta\beta' + (\sigma^2 + \delta^2)\Omega\right) & \frac{\sigma^2(\sigma^2-\delta^2)}{\sigma^2+\mu\delta}\beta & -\frac{2\delta\sigma^4}{\sigma^2+\mu\delta}\beta \\ \frac{\sigma^2(\sigma^2-\delta^2)}{\sigma^2+\mu\delta}\beta' & \sigma^2 & 0 \\ -\frac{2\delta\sigma^4}{\sigma^2+\mu\delta}\beta' & 0 & 2\sigma^4 \end{pmatrix} \ .$$

Using the formula for the inverse of partitioned matrix, the inverse matrix of S is obtained as

$$S^{-1} = \begin{pmatrix} \frac{(\sigma^2+\mu\delta)^2}{\sigma^2(\sigma^2+\delta^2)}\Omega^{-1} & -\frac{(\sigma^2+\mu\delta)(\sigma^2-\delta^2)}{\sigma^2(\sigma^2+\delta^2)}\Omega^{-1}\beta & \frac{\delta(\sigma^2+\mu\delta)}{\sigma^2(\sigma^2+\delta^2)}\Omega^{-1}\beta \\ -\frac{(\sigma^2+\mu\delta)(\sigma^2-\delta^2)}{\sigma^2(\sigma^2+\delta^2)}\beta'\Omega^{-1} & \frac{1}{\sigma^2} + \frac{(\sigma^2-\delta^2)^2}{\sigma^2(\sigma^2+\delta^2)}\beta'\Omega^{-1}\beta & -\frac{\delta(\sigma^2-\delta^2)}{\sigma^2(\sigma^2+\delta^2)}\beta'\Omega^{-1}\beta \\ \frac{\delta(\sigma^2+\mu\delta)}{\sigma^2(\sigma^2+\delta^2)}\beta'\Omega^{-1} & -\frac{\delta(\sigma^2-\delta^2)}{\sigma^2(\sigma^2+\delta^2)}\beta'\Omega^{-1}\beta & \frac{1}{2\sigma^4} + \frac{\delta^2}{\sigma^2(\sigma^2+\delta^2)}\beta'\Omega^{-1}\beta \end{pmatrix} \ .$$

If μ and σ are unknown, the vector of unknown parameters is $\theta = (\delta, \mu, \sigma^2)'$. The expected value of the derivative of g_t with respect to θ is

$$D = E[\frac{\partial g_t}{\partial \theta'}] = \begin{pmatrix} -\frac{\sigma^2}{\sigma^2+\mu\delta}\beta & \frac{\delta^2}{\sigma^2+\mu\delta}\beta & \frac{\delta}{\sigma^2+\mu\delta}\beta \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \ .$$

The matrix $D'S^{-1}D$ is then

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

and the asymptotic covariance matrix of estimator for θ is

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

Then, the asymptotic variance of the estimator for δ is the first element of $(D'S^{-1}D)^{-1}$, which is

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

If μ and σ are known, the only unknown parameter is δ. The expected value of the derivative of g_t with respect to δ is

$$D = E[\frac{\partial g_t}{\partial \delta}] = \begin{pmatrix} -\frac{\sigma^2}{\sigma^2+\mu\delta}\beta \\ 0 \\ 0 \end{pmatrix} .$$

Then, the asymptotic variance of the estimator for δ is

$$\text{Avar}(\hat{\delta}_{\mu,\sigma}) = (D'S^{-1}D)^{-1} = \frac{\sigma^2+\delta^2}{\sigma^2}(\beta'\Omega^{-1}\beta)^{-1} .$$

If μ and σ are known, financial economists usually drop the two moment restrictions that define μ and σ^2 because they don not need to estimate them. Then, the moment restrictions for the risk premium is

$$E\left[r_t\left(1 - \frac{\delta}{\sigma^2+\mu\delta}f_t\right)\right] = 0 .$$

Let us define a random variable g_t as

$$g_t = r_t\left(1 - \frac{\delta}{\sigma^2+\mu\delta}f_t\right) .$$

Substituting r_t by $(\delta + f_t - \mu)\beta + \epsilon_t$, we obtain the covariance of g_t as

$$\begin{aligned} S &= E[g_t g_t'] = E\left[\left(1 - \frac{\delta}{\sigma^2+\mu\delta}f_t\right)^2 r_t r_t'\right] \\ &= E\left[\left(1 - \frac{\delta}{\sigma^2+\mu\delta}f_t\right)^2 (\delta + f_t - \mu)^2\right]\beta\beta' + E\left[\left(1 - \frac{\delta}{\sigma^2+\mu\delta}f_t\right)^2\right]\Omega \\ &= \frac{\sigma^2(\sigma^4+\delta^4)}{(\sigma^2+\mu\delta)^2}\beta\beta' + \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^2}\Omega , \end{aligned}$$

and its inverse is

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

The expected value of the derivative of g_t with respect to δ is

$$D = -\frac{\sigma^2}{\sigma^2+\mu\delta}\beta \ .$$

Then, it is straightforward to calculate the asymptotic variance of the estimator for δ:

$$\text{Avar}(\bar{\delta}_{\mu,\delta}) = (D'S^{-1}D)^{-1} = \frac{\sigma^2+\delta^2}{\sigma^2}(\beta'\Omega^{-1}\beta)^{-1} + \sigma^2 + \frac{\delta^4}{\sigma^2} \ .$$

B Using regression equation to estimate δ

When we use regression equation to estimate risk premium, the moment restrictions are

$$\begin{aligned}
&E[r_t - (\delta + f_t - \mu)\beta] = 0\\
&E[(r_t - (\delta + f_t - \mu)\beta)f_t] = 0\\
&E[f_t - \mu] = 0\\
&E[(f_t-\mu)^2 - \sigma^2] = 0 \ .
\end{aligned}$$

Let us define a random variable g_t as

$$g_t = \begin{pmatrix} r_t - (\delta + f_t - \mu)\beta \\ (r_t - (\delta + f_t - \mu)\beta)f_t \\ f_t - \mu \\ (f_t-\mu)^2 - \sigma^2 \end{pmatrix} = \begin{pmatrix} \epsilon_t \\ \epsilon_t f_t \\ f_t - \mu \\ (f_t-\mu)^2 - \sigma^2 \end{pmatrix} \ .$$

The covariance matrix of g_t is

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

and its inverse is

$$S^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} (\sigma^2+\mu^2)\Omega^{-1} & -\mu\Omega^{-1} & 0 & 0 \\ -\mu\Omega^{-1} & \Omega^{-1} & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma^2} & 0 \\ 0 & 0 & 0 & \frac{1}{2\sigma^4} \end{pmatrix} \ .$$

If μ and σ are unknown, the vector of unknown parameters is $\theta = (\delta, \beta', \mu, \sigma^2)'$. The expected value of the derivative of g_t with respect to θ is

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

It follows that

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

The asymptotic covariance matrix of the estimator for θ is $(D'S^{-1}D)^{-1}$. We find that the inverse matrix of $D'S^{-1}D$ is

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

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

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

If μ and σ are known, the vector of unknown parameters is $\theta = (\delta, \beta')'$. The expected value of the derivative of g_t with respect to θ is

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

It follows that

$$D'S^{-1}D = \begin{pmatrix} \beta'\Omega^{-1}\beta & \delta\beta'\Omega^{-1} \\ \delta\Omega^{-1}\beta & (\sigma^2+\delta^2)\Omega^{-1} \end{pmatrix} .$$

The asymptotic variance of the estimator for δ is the first element of $(D'S^{-1}D)^{-1}$. Using the formula for the inverse of partitioned matrix, we obtain it as

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

If we drop the last two moment restrictions that define μ and σ^2, the GMM estimator $\delta^\dagger_{\mu,\sigma}$ for δ has the same asymptotic variance because matrix S is block diagonal and because the last two rows of matrix D are zero.

C Asymptotic distributions of the J-statistics

It follows from Lemma 4.2 in Hansen (1982) that each of the J-Statistic has an asymptotic χ^2 distribution. The number of degrees of freedom of the χ^2 distribution equals the difference between the number of moment restrictions and the number of parameters that are estimated.

First, consider the case where the mean and variance of the factor are assumed unknown and the two moment restrictions on the factor are incorporated. There are $n + 2$ moment restrictions and 3 parameters in the SDF method, and there are $2n + 2$ moment restrictions and $n + 3$ parameters in the regression method. Thus, the asymptotic distribution of the J-statistics in the SDF method and the regression method have the same χ^2 distribution with $n - 1$ degrees of freedom.

Second, consider the case where the mean and variance of the factor are assumed known and the two moment restrictions on the factor are ignored. There are n moment restrictions and 1 parameters in the SDF method, and there are $2n$ moment restrictions and $n + 1$ parameters in the regression method. Thus, the asymptotic distribution of the J-statistics in the SDF method and the regression method have the same χ^2 distribution with $n - 1$ degrees of freedom.

Last, consider the case where the mean and variance of the factors are assumed known and the two moment restrictions on the factor are incorporated. There are $n + 2$ moment restrictions and 1 parameter in the SDF method, and there are $2n + 2$ moment restrictions and $n + 1$ parameters in the regression method. Thus, the asymptotic distribution of the J-statistics in the SDF method and the regression method have the same χ^2 distribution with $n + 1$ degrees of freedom.

D Using Euler equation to estimate λ

When we incorporate the restrictions on the mean and variance of the factor into the estimation of the parameter λ from Euler equation, the moment restrictions are

$$
\begin{aligned}
&E[r_t(1 - \lambda f_t)] = 0 \\
&E[f_t - \mu] = 0 \\
&E[f_t(f_t - \mu) - \sigma^2] = 0 \; .
\end{aligned}
$$

Let us define a random variable g_t as

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

The covariance matrix of g_t is $S = E[g_t g_t']$, which is

$$\begin{pmatrix} A_{11} & \frac{\sigma^2((1-\mu\lambda)^2-\sigma^2\lambda^2)}{1-\mu\lambda}\beta & \frac{\sigma^2(\mu(1-\mu\lambda)^2-\sigma^2\lambda(2-\mu\lambda))}{1-\mu\lambda}\beta \\ \frac{\sigma^2((1-\mu\lambda)^2-\sigma^2\lambda^2)}{1-\mu\lambda}\beta' & \sigma^2 & \mu\sigma^2 \\ \frac{\sigma^2(\mu(1-\mu\lambda)^2-\sigma^2\lambda(2-\mu\lambda))}{1-\mu\lambda}\beta' & \mu\sigma^2 & \sigma^2(\mu^2+2\sigma^2) \end{pmatrix} ,$$

where

$$A_{11} = \frac{\sigma^2(\sigma^4\lambda^4 + (1-\mu\lambda)^4)}{(1-\mu\lambda)^2}\beta\beta' + (\sigma^2\lambda^2 + (1-\mu\lambda)^2)\Omega^{-1} ,$$

and its inverse is:

$$\begin{pmatrix} \frac{1}{(1-\mu\lambda)^2+\sigma^2\lambda^2}\Omega^{-1} & \frac{(1-\mu\lambda)-\lambda\sigma^2}{(1-\mu\lambda)((1-\mu\lambda)^2+\sigma^2\lambda^2)}\Omega^{-1}\beta & \frac{\lambda}{(1-\mu\lambda)^2+\sigma^2\lambda^2}\Omega^{-1}\beta \\ \frac{(1-\mu\lambda)-\lambda\sigma^2}{(1-\mu\lambda)((1-\mu\lambda)^2+\sigma^2\lambda^2)}\beta'\Omega^{-1}\beta & B_{22} & B_{23} \\ \frac{\lambda}{(1-\mu\lambda)^2+\sigma^2\lambda^2}\Omega^{-1}\beta' & B_{32} & B_{33} \end{pmatrix} ,$$

where

$$\begin{aligned} B_{22} &= \frac{\mu^2+2\sigma^2}{2\sigma^4} + \frac{(1-\mu\lambda)-\lambda\sigma^2}{(1-\mu\lambda)((1-\mu\lambda)^2+\sigma^2\lambda^2)}\beta'\Omega^{-1} \\ B_{23} &= B_{32} = -\frac{\mu}{2\sigma^4} - \frac{\lambda(\lambda\sigma^2-(1-\mu\lambda)}{(1-\mu\lambda)((1-\mu\lambda)^2+\sigma^2\lambda^2)}\beta'\Omega^{-1}\beta \\ B_{33} &= \frac{1}{2\sigma^4} + \frac{\lambda^2}{(1-\mu\lambda)^2+\sigma^2\lambda^2}\beta'\Omega^{-1}\beta . \end{aligned}$$

If μ and σ^2 are unknown, the vector of unknown parameters is $\theta = (\lambda, \mu, \sigma^2)$. The expected value of derivative of g_t with respect to θ is

$$D = E\left[\frac{\partial g_t}{\partial \theta}\right] = \begin{pmatrix} -\frac{\sigma^2}{1-\mu\lambda}\beta & 0 & 0 \\ 0 & -1 & -\mu \\ 0 & 0 & -1 \end{pmatrix}$$

The matrix $D'S^{-1}D$ is

$$\begin{pmatrix} \frac{\sigma^4}{(1-\mu\lambda)((1-\mu\lambda)^2+\sigma^2\lambda^2)}\beta'\Omega^{-1}\beta & \frac{\sigma^2(\lambda\sigma^2-(1-\mu\lambda))}{(1-\mu\lambda)(((1-\mu\lambda)^2+\sigma^2\lambda)}\beta'\Omega^{-1}\beta & \frac{\sigma^2((\mu+\lambda)(1-\mu\lambda)-\mu\sigma^2\lambda^2)}{(1-\mu\lambda)(((1-\mu\lambda)^2+\sigma^2\lambda)}\beta'\Omega^{-1}\beta \\ \frac{\sigma^2(\lambda\sigma^2-(1-\mu\lambda))}{(1-\mu\lambda)(((1-\mu\lambda)^2+\sigma^2\lambda)}\beta'\Omega^{-1}\beta & A_{22} & A_{23} \\ \frac{\sigma^2((\mu+\lambda)(1-\mu\lambda)-\mu\sigma^2\lambda^2)}{(1-\mu\lambda)(((1-\mu\lambda)^2+\sigma^2\lambda)}\beta'\Omega^{-1}\beta & A_{32} & A_{33} \end{pmatrix}$$

where

$$
\begin{aligned}
A_{22} &= \frac{\mu^2+2\sigma^2}{2\sigma^4} + \frac{(\lambda\sigma^2-(1-\mu\lambda))^2}{(1-\mu\lambda)(((1-\mu\lambda)^2+\sigma^2\lambda)}\beta'\Omega^{-1}\beta \\
A_{23} &= A_{32} = \frac{2\sigma^2+\mu^2-1}{2\sigma^4} + \frac{\lambda\sigma^2-(1-\mu\lambda))(\mu\lambda^2\sigma^2-(\mu+\lambda)(1-\lambda))}{(1-\mu\lambda)(((1-\mu\lambda)^2+\sigma^2\lambda)}\beta'\Omega^{-1}\beta \\
A_{33} &= \frac{(1-\mu^2)^2+\sigma^2}{2\sigma^4} + \left(\frac{\lambda^2}{(1-\mu\lambda)^2+\sigma^2\lambda^2} + \frac{((\lambda-\mu)(1-\mu\lambda)+\mu\sigma^2\lambda^2)^2}{(1-\mu\lambda)(((1-\mu\lambda)^2+\sigma^2\lambda)}\right)\beta'\Omega^{-1}\beta
\end{aligned}
$$

The asymptotic covariance matrix of the estimator for θ is the inverse of $D'S^{-1}D$, which can be calculated as

$$
\begin{pmatrix}
B_{11} & B_{12} & \mu(1-\mu\lambda)^2-\lambda\sigma^2(1-\mu\lambda) \\
B_{21} & \sigma^2(1-\mu^2)^2+2\mu^2\sigma^2) & \mu\sigma^2(1-2\sigma^2-\mu^2) \\
\mu(1-\mu\lambda)^2-\lambda\sigma^2(1-\mu\lambda) & \mu\sigma^2(1-2\sigma^2-\mu^2) & \sigma^2(\mu^2+2\sigma^2)
\end{pmatrix},
$$

where

$$
\begin{aligned}
B_{11} &= \frac{(1-\mu\lambda)^4+\sigma^4\lambda^4}{\sigma^2} + \frac{(1-\mu\lambda)^2((1-\mu\lambda)^2+\sigma^2\lambda^2)}{\sigma^4}(\beta'\Omega^{-1}\beta)^{-1} \\
B_{12} &= B_{21} = (1-\mu\lambda)(1-\mu^2)-\lambda\sigma^2(\lambda-2\mu+\lambda\mu^2) .
\end{aligned}
$$

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

$$
\text{Avar}(\hat{\lambda}) = B_{11} = \frac{(1-\mu\lambda)^4+\sigma^4\lambda^4}{\sigma^2} + \frac{(1-\mu\lambda)^2((1-\mu\lambda)^2+\sigma^2\lambda^2)}{\sigma^4}(\beta'\Omega^{-1}\beta)^{-1} .
$$

If μ and σ^2 are known, the only unknown parameter is λ. The expected value of the derivative of g_t with respect to λ is

$$
D = E\left[\frac{\partial g_t}{\partial \theta}\right] = \begin{pmatrix} -\frac{\sigma^2}{1-\mu\lambda}\beta \\ 0 \\ 0 \end{pmatrix} .
$$

It follows that $D'S^{-1}D$ is

$$
\frac{\sigma^4}{(1-\mu\lambda)^2((1-\mu\lambda)^2+\sigma^2\lambda^2)}\beta'\Omega^{-1}\beta ,
$$

and the asymptotic variance of the estimator for λ is

$$
\text{Avar}(\hat{\lambda}_{\mu,\sigma}) = \frac{(1-\mu\lambda)^2((1-\mu\lambda)^2+\sigma^2\lambda^2)}{\sigma^4}(\beta'\Omega^{-1}\beta)^{-1} .
$$

No matter μ and σ^2 are known or nor, people usually drop the last two restrictions that define μ and σ^2. In this case, to estimate λ in Euler equation, the moment restriction is

$$
E[r_t(1-\lambda f_t)] = 0 .
$$

Let us define a random variable g_t as

$$g_t = r_t(1 - \lambda f_t)$$

Substituting r_t by the regression model, we obtain the covariance matrix of g_t as

$$S = E[g_t g_t'] = \frac{\sigma^2(\sigma^4\lambda^4 + (1-\mu\lambda)^4)}{(1-\mu\lambda)^2}\beta\beta' + (\sigma^2\lambda^2 + (1-\mu\lambda)^2)\Omega \ ,$$

and its inverse is

$$\begin{aligned} S^{-1} &= \frac{1}{(\sigma^2\lambda^2 + (1-\mu\lambda)^2)^2} \\ & \left(\Omega^{-1} - \left(\frac{1}{\sigma^2\lambda^2 + (1-\mu\lambda)^2}\beta'\Omega^{-1}\beta + \frac{(1-\mu\lambda)^2}{\sigma^2(\sigma^4\lambda^4 + (1-\mu\lambda)^4}\right)^{-1} \Omega^{-1}\beta\beta'\Omega^{-1}\right) \ . \end{aligned}$$

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

$$D = \frac{\sigma^2}{1-\mu\lambda}\beta \ .$$

It follows that the asymptotic variance of the estimator for λ is

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

E Using the regression equation to estimate λ

When we estimate parameter λ from regression equation, the moment restrictions are

$$\begin{aligned} &E[r_t - ((1-\lambda\mu)^{-1}\lambda\sigma^2 + f - \mu)\beta] = 0 \\ &E[(r_t - ((1-\lambda\mu)^{-1}\lambda\sigma^2\beta + f - \mu)\beta)f_t] = 0 \\ &E[f_t - \mu] = 0 \\ &E[(f_t - \mu)f_t - \sigma^2] = 0 \ . \end{aligned}$$

Let us define a random variable g_t as

$$g_t = \begin{pmatrix} r_t - ((1-\lambda\mu)^{-1}\lambda\sigma^2 + f_t - \mu)\beta \\ \{r_t - ((1-\lambda\mu)^{-1}\lambda\sigma^2 + f_t - \mu)\beta\} f_t \\ f_t - \mu \\ f_t(f_t - \mu) - \sigma^2 \end{pmatrix} = \begin{pmatrix} \epsilon_t \\ \epsilon_t f_t \\ f_t - \mu \\ f_t(f_t - \mu) - \sigma^2 \end{pmatrix} \ .$$

The covariance matrix of g_t is

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

and its inverse is

$$S^{-1} = \frac{1}{\sigma^2} \begin{pmatrix} \begin{pmatrix} \sigma^2+\mu^2 & -\mu \\ -\mu & 1 \end{pmatrix} \otimes \Omega^{-1} & 0 \\ 0 & \frac{1}{2\sigma^4} \begin{pmatrix} 2\sigma^2+\mu^2 & -\mu \\ -\mu & 1 \end{pmatrix} \end{pmatrix} .$$

The vector of unknown parameters is $\theta = (\lambda, \beta', \mu, \delta)'$. The expected value of the derivative of g_t with respect to θ is

$$D = \begin{pmatrix} -\frac{\sigma^2}{(1-\lambda\mu)^2}\beta & -\frac{\lambda\sigma^2}{1-\lambda\mu} I_n & \left(1 - \frac{\lambda^2\sigma^2}{(1-\lambda\mu)^2}\right)\beta & -\frac{\lambda}{1-\lambda\mu}\beta \\ -\frac{\mu\sigma^2}{(1-\lambda\mu)^2}\beta & -\frac{\sigma^2}{1-\lambda\mu} I_n & \left(1 - \frac{\lambda^2\sigma^2}{(1-\lambda\mu)^2}\right)\mu\beta & -\frac{\lambda\mu}{1-\lambda\mu}\beta \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -\mu & -1 \end{pmatrix} .$$

The matrix $D'S^{-1}D$ can be calculated out as

$$\begin{pmatrix} \frac{\sigma^4}{(1-\lambda\mu)^4}\beta'\Omega^{-1}\beta & \frac{\lambda\sigma^4}{(1-\lambda\mu)^3}\beta'\Omega^{-1} & A_{13} & \frac{\lambda\sigma^2}{(1-\lambda\mu)^3}\beta'\Omega^{-1}\beta \\ \frac{\lambda\sigma^4}{(1-\lambda\mu)^3}\Omega^{-1}\beta & \frac{\sigma^2[(1-\lambda\mu)^2+\lambda^2\sigma^2]}{(1-\lambda\mu)^2}\Omega^{-1} & A_{23} & \frac{\lambda^2\sigma^2}{(1-\lambda\mu)^2}\Omega^{-1}\beta \\ A_{31} & A_{32} & A_{33} & A_{34} \\ \frac{\lambda\sigma^2}{(1-\lambda\mu)^3}\beta'\Omega^{-1}\beta & \frac{\lambda^2\sigma^2}{(1-\lambda\mu)^2}\beta'\Omega^{-1} & A_{43} & A_{44} \end{pmatrix}$$

where

$$\begin{aligned} A_{13} &= A_{31} = -\frac{\sigma^2[(1-\lambda\mu)^2 - \lambda^2\sigma^2]}{(1-\lambda\mu)^4}\beta'\Omega^{-1}\beta \\ A_{23} &= A'_{32} = -\frac{\lambda\sigma^2[(1-\lambda\mu)^2 - \lambda^2\sigma^2]}{(1-\lambda\mu)^3}\Omega^{-1}\beta \\ A_{33} &= \frac{1}{\sigma^2} + \frac{[(1-\lambda\mu)^2 - \lambda^2\sigma^2]^2}{(1-\lambda\mu)^4}\beta'\Omega^{-1}\beta \\ A_{34} &= A_{43} = -\frac{\lambda[(1-\lambda\mu)^2 - \lambda^2\sigma^2]}{(1-\lambda\mu)^3}\beta'\Omega^{-1}\beta \\ A_{44} &= \frac{1}{2\sigma^4} + \frac{\lambda^2}{(1-\lambda\mu)^2}\beta'\Omega^{-1}\beta . \end{aligned}$$

We find that the inverse matrix of $D'S^{-1}D$ is

$$\begin{pmatrix} B_{11} & -\frac{\lambda(1-\lambda\mu)}{\sigma^2\beta'\Omega^{-1}\beta}\beta' & (1-\lambda\mu)^2 - \lambda^2\sigma^2 & -2\lambda\sigma^2(1-\lambda\mu) \\ -\frac{\lambda(1-\lambda\mu)}{\sigma^2\beta'\Omega^{-1}\beta}\beta & B_{22} & 0 & 0 \\ (1-\lambda\mu)^2 - \lambda^2\sigma^2 & 0 & \sigma^2 & 0 \\ -2\lambda\sigma^2(1-\lambda\mu) & 0 & 0 & 2\sigma^4 \end{pmatrix} ,$$

where

$$\begin{aligned} B_{11} &= \frac{(1-\lambda\mu)^4 + \lambda^4\sigma^4}{\sigma^2} + \frac{(1-\lambda\mu)^2[(1-\lambda\mu)^2 + \lambda^2\sigma^2]}{\sigma^4\beta'\Omega^{-1}\beta} \\ B_{22} &= \frac{(1-\lambda\mu)^2}{\sigma^2[(1-\lambda\mu)^2 + \lambda^2\sigma^2]}\Omega + \frac{\lambda^2}{[(1-\lambda\mu)^2 + \lambda^2\sigma^2]\beta'\Omega^{-1}\beta}\beta\beta' . \end{aligned}$$

It is straightforward but tedious to verify that multiplying this matrix to $D'S^{-1}D$ gives the identity matrix. The asymptotic variance of λ^* is then the first element of $(D'S^{-1}D)^{-1}$, which is

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

If μ and σ are known, the vector of unknown parameters is $\theta = (\lambda, \beta')'$. The expected value of the derivative of g_t with respect to θ is

$$D = \begin{pmatrix} -\frac{\sigma^2}{(1-\lambda\mu)^2}\beta & -\frac{\lambda\sigma^2}{1-\lambda\mu}I_n \\ -\frac{\mu\sigma^2}{(1-\lambda\mu)^2}\beta & -\frac{\sigma^2}{1-\lambda\mu}I_n \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \ .$$

The matrix $D'S^{-1}D$ can be calculated out as

$$\begin{pmatrix} \frac{\sigma^4}{(1-\lambda\mu)^4}\beta'\Omega^{-1}\beta & \frac{\lambda\sigma^4}{(1-\lambda\mu)^3}\beta'\Omega^{-1} \\ \frac{\lambda\sigma^4}{(1-\lambda\mu)^3}\Omega^{-1}\beta & \frac{\sigma^2[(1-\lambda\mu)^2+\lambda^2\sigma^2]}{(1-\lambda\mu)^2}\Omega^{-1} \end{pmatrix} \ .$$

Using the formula for the inverse of partitioned matrix, the inverse matrix is found out to be

$$\begin{pmatrix} \frac{(1-\lambda\mu)^2[(1-\lambda\mu)^2+\lambda^2\sigma^2]}{\sigma^4\beta'\Omega^{-1}\beta} & -\frac{\lambda(1-\lambda\mu)}{\sigma^2\beta'\Omega^{-1}\beta}\beta' \\ -\frac{\lambda(1-\lambda\mu)}{\sigma^2\beta'\Omega^{-1}\beta}\beta & \frac{(1-\lambda\mu)^2}{\sigma^2[(1-\lambda\mu)^2+\lambda^2\sigma^2]}\Omega + \frac{\lambda^2}{[(1-\lambda\mu)^2+\lambda^2\sigma^2]\beta'\Omega^{-1}\beta}\beta\beta' \end{pmatrix} \ .$$

The asymptotic variance of λ^* is then the first element of $(D'S^{-1}D)^{-1}$, which is

$$\text{Avar}(\lambda^*_{\mu,\sigma}) = \frac{(1-\lambda\mu)^2[(1-\lambda\mu)^2 + \lambda^2\sigma^2]}{\sigma^4}(\beta'\Omega^{-1}\beta)^{-1} \ .$$

If we drop the last two moment restrictions that define μ and σ^2, we obtain the same estimator because matrix S is block diagonal and because the last two rows of matrix D are zero.

Table 1. Parameter values used in Monte Carlo Simulations

This table presents the parameters used for our Monte Carlo simulations. The choice of the parameters are based on monthly historical observations (from January 1926 to December 1998) of returns (in excess of returns on one-month Treasury Bills) on size decile and the value-weighted index of the stocks listed in NYSE, AMEX, and NASDAQ. The data are obtained from the Center for Research on Security Prices (CRSP) via Wharton Research Data Services (WRDS). The mean (μ) and standard deviation (σ) of the factor is set to be the sample mean and standard deviation of returns on the value-weighted index. The beta β is set to be the slope in the time-series regression of the decile return on the index return. 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 return's historical average on beta. The parameter λ satisfies $\lambda = \delta/(\sigma^2 + \mu\delta)$. The expected return $E[r]$ is set to be $\delta\,\beta$.

	λ		δ		μ		σ		
	0.0436		1.3692		0.6914		5.5155		
				β					
1.45	1.39	1.30	1.26	1.24	1.22	1.18	1.12	1.09	0.94
				$E[r]$					
1.99	1.91	1.79	1.73	1.69	1.68	1.62	1.53	1.49	1.29
				Ω					
55.60	37.03	28.76	22.25	17.35	14.42	10.13	6.37	3.88	-2.81
37.03	29.75	22.56	17.82	13.92	11.83	8.27	5.46	3.17	-2.24
28.76	22.56	20.44	15.15	11.83	10.28	7.26	4.95	2.80	-1.90
22.25	17.82	15.15	13.64	10.10	8.85	6.64	4.79	2.65	-1.66
17.35	13.92	11.83	10.10	9.28	7.43	5.62	4.16	2.43	-1.39
14.42	11.83	10.28	8.85	7.43	7.32	5.11	3.88	2.37	-1.26
10.13	8.27	7.26	6.64	5.62	5.11	4.93	3.34	1.99	-0.99
6.37	5.46	4.95	4.79	4.16	3.88	3.34	3.33	1.85	-0.76
3.88	3.17	2.80	2.65	2.43	2.37	1.99	1.85	1.84	-0.50
-2.81	-2.24	-1.90	-1.66	-1.39	-1.26	-0.99	-0.76	-0.50	0.39

Table 2. Asymptotic Standard Deviations of the Estimators

The first row of numbers are the theoretical asymptotic standard deviations of various estimators. The observations of the returns and the factor are assumed to have identical and independent joint Normal distributions. The true parameters are assumed to be those given in Table 1. The asymptotic standard deviations are calculated according to the formulae derived in Appendix A and B. The numbers in the other rows are the asymptotic standard deviations divided by the square-root of T, which is given in the first column.

T	$\hat{\delta}$	δ^*	$\bar{\delta}_{\mu,\sigma}$	$\delta^{\dagger}_{\mu,\sigma}$	$\hat{\delta}_{\mu,\sigma}$	$\delta^*_{\mu,\sigma}$
	5.5248	5.5248	5.5352	0.3189	0.3189	0.3189
5×12	0.7132	0.7132	0.7146	0.0412	0.0412	0.0412
10×12	0.5043	0.5043	0.5053	0.0291	0.0291	0.0291
30×12	0.2912	0.2912	0.2917	0.0168	0.0168	0.0168
50×12	0.2255	0.2255	0.2260	0.0130	0.0130	0.0130
75×12	0.1842	0.1842	0.1845	0.0106	0.0106	0.0106
100×12	0.1595	0.1595	0.1598	0.0092	0.0092	0.0092

Table 3. Simulation Results from Normal Distribution

This table provides the results of Monte Carlo simulations on various estimators under the assumption that the returns and the factor have an Normal distribution. Independent samples of the factor f_t are drawn from the Normal distribution with mean μ and variance σ^2. Independent samples of the residual ϵ_t of the regression equation are drawn, independent of the factor, from the Normal distribution with mean zero and variance Ω. Excess returns on 10 portfolios are constructed to satisfy $r_t = (\delta + f_t - \mu)\beta + \epsilon_t$ for $t = 1, \cdots, T$. The parameters, δ, β, μ, σ and Ω, are set to those in Table 1. Each estimator is then calculated based on the T samples to obtain a sample of the estimator. We repeat this independently for 100 times to obtain 100 independent samples of the estimators. The table presents the average and standard deviation of the 100 samples of each estimator.

T		$\hat{\delta}$	δ^*	$\bar{\delta}_{\mu,\sigma}$	$\delta^\dagger_{\mu,\sigma}$	$\hat{\delta}_{\mu,\sigma}$	$\delta^*_{\mu,\sigma}$
5×12	Ave	1.7145	1.6960	1.7129	1.3739	1.3823	1.3726
	Std	0.9297	0.8946	0.9025	0.0794	0.1911	0.0763
10×12	Ave	1.5198	1.5216	1.5325	1.3626	1.3622	1.3639
	Std	0.5930	0.5874	0.6165	0.0383	0.0773	0.0371
30×12	Ave	1.3968	1.3978	1.3996	1.3670	1.3602	1.3667
	Std	0.3148	0.3147	0.3300	0.0172	0.0372	0.0172
50×12	Ave	1.3745	1.3749	1.3838	1.3666	1.3646	1.3667
	Std	0.2566	0.2568	0.2611	0.0128	0.0199	0.0129
75×12	Ave	1.3973	1.3978	1.4026	1.3705	1.3701	1.3706
	Std	0.2014	0.2016	0.2069	0.0113	0.0142	0.0112
100×12	Ave	1.3797	1.3800	1.3776	1.3683	1.3677	1.3683
	Std	0.1743	0.1742	0.1721	0.0100	0.0125	0.0101

Table 4. Simulation Results from t Distribution

This table provides the results of Monte Carlo simulations on various estimators under the assumption that the returns and the factor jointly have a multivariate t distribution with 5 degree of freedom. Independent samples $\{(f_t, \epsilon_t')'\}_{t=1,\cdots,T}$ are drawn from a multivariate t distribution with the mean as $(\mu, 0_n')'$ and the covariance matrix such that $E[(f_t-\mu)^2] = \sigma^2$, $E[(f_t-\mu)\epsilon_t] = 0_n$ and $E[\epsilon_t\epsilon_t'] = \Omega$, where 0_n is a 10-dimensional vector of zeros. Excess returns on 10 portfolios are constructed to satisfy $r_t = (\delta + f_t - \mu)\beta + \epsilon_t$ for $t = 1, \cdots, T$. The parameters, δ, β, μ, σ and Ω, are set to those in Table 1. Each estimator is then calculated based on the T samples to obtain a sample of the estimator. We repeat this independently for 100 times to obtain 100 independent samples of each estimator. The table presents the average and standard deviation of the 100 samples of each estimator.

T		$\hat{\delta}$	δ^*	$\bar{\delta}_{\mu,\sigma}$	$\delta^\dagger_{\mu,\sigma}$	$\hat{\delta}_{\mu,\sigma}$	$\delta^*_{\mu,\sigma}$
5×12	Ave	1.7726	1.7962	1.9849	1.3789	1.3995	1.3742
	Std	0.8455	0.8363	1.0475	0.0800	0.2344	0.0812
10×12	Ave	1.5595	1.5762	1.7122	1.3704	1.3601	1.3708
	Std	0.5892	0.5767	0.6456	0.0474	0.0984	0.0454
30×12	Ave	1.4218	1.4267	1.4788	1.3667	1.3463	1.3663
	Std	0.2792	0.2803	0.3309	0.0160	0.0549	0.0164
50×12	Ave	1.3797	1.3864	1.4367	1.3687	1.3665	1.3688
	Std	0.2233	0.2257	0.2354	0.0127	0.0228	0.0129
75×12	Ave	1.4196	1.4242	1.4654	1.3688	1.3609	1.3688
	Std	0.1714	0.1724	0.2082	0.0107	0.0199	0.0108
100×12	Ave	1.3784	1.3810	1.4112	1.3692	1.3659	1.3690
	Std	0.1612	0.1602	0.1720	0.0084	0.0156	0.0086

Table 5. Simulation Results from Empirical Distribution

This table provides the results of Monte Carlo simulations on various estimators under the assumption that the returns and the factor jointly have a multivariate empirical distribution. The monthly observations of return on the value-weighted index of NYSE, AMEX and NASDAQ are used as the data of f. The residuals in the regression of decile returns on the index return are used as the data of ϵ. Independent samples $\{(f_t, \epsilon_t')'\}_{t=1,\cdots,T}$ are drawn from the empirical distribution of the data. Excess returns on 10 portfolios are constructed to satisfy $r_t = (\delta + f_t - \mu)\beta + \epsilon_t$ for $t = 1, \cdots, T$. The parameters, δ, β, μ, σ and Ω, are set to those in Table 1. Each estimator is then calculated based on the T samples to obtain a sample of the estimator. We repeat this independently for 100 times to obtain 100 independent samples of each estimator. The table presents the average and standard deviation of the 100 samples of each estimator.

T		$\hat{\delta}$	δ^*	$\bar{\delta}_{\mu,\sigma}$	$\delta^\dagger_{\mu,\sigma}$	$\hat{\delta}_{\mu,\sigma}$	$\delta^*_{\mu,\sigma}$
5×12	Ave	1.6452	1.6235	1.5865	1.3890	1.3767	1.3874
	Std	0.8938	0.8604	0.8291	0.0927	0.1925	0.0930
10×12	Ave	1.5038	1.5010	1.4823	1.3743	1.3623	1.3739
	Std	0.5589	0.5537	0.5855	0.0444	0.0910	0.0412
30×12	Ave	1.3894	1.3900	1.3912	1.3714	1.3682	1.3715
	Std	0.2843	0.2840	0.2951	0.0158	0.0251	0.0157
50×12	Ave	1.3707	1.3710	1.3638	1.3713	1.3681	1.3716
	Std	0.2272	0.2272	0.2274	0.0121	0.0203	0.0123
75×12	Ave	1.3736	1.3738	1.3680	1.3709	1.3697	1.3709
	Std	0.1693	0.1693	0.1696	0.0097	0.0140	0.0098
100×12	Ave	1.3768	1.3770	1.3764	1.3697	1.3692	1.3698
	Std	0.1746	0.1746	0.1748	0.0087	0.0115	0.0086

Table 6. Simulation Results for $\hat{\lambda}$ and λ^*

In this table we assume that the financial economist does not know the mean and variance of the factor and thus estimates them together with the risk premium. The table provides the results of Monte Carlo simulations on the estimators for λ under the assumption that the returns and the factor have a joint Normal distribution, t distribution or empirical distribution. Independent samples of the factor f_t are drawn from a Normal, t or empirical distribution with mean μ and variance σ^2. Independent samples of the residual ϵ_t of the regression equation are drawn, uncorrelated with the factor, from a Normal, t or empirical distribution with mean zero and variance Ω. Excess returns on 10 portfolios are constructed to satisfy $r_t = (\lambda\sigma^2/(1-\mu\lambda) + f_t - \mu)\beta + \epsilon_t$ for $t = 1, \cdots, T$. The parameters, λ, β, μ, σ and Ω, are set to those in Table 1. The estimator $\hat{\lambda}$ is obtained by using the SDF method, and the estimator λ^* is obtained by using the regression method. The two estimators are calculated based on the T samples to obtain a sample for each of the estimators. We repeat this independently for 100 times to obtain 100 independent samples for each of the estimators. The table presents the average and standard deviation of the 100 samples for each estimator.

Sample size		Normal		Student-t		Empirical	
T		$\hat{\lambda}$	λ^*	$\hat{\lambda}$	λ^*	$\hat{\lambda}$	λ^*
5×12	Ave	0.0537	0.0533	0.0617	0.0624	0.0499	0.0495
	Std	0.0278	0.0271	0.0315	0.0312	0.0254	0.0249
10×12	Ave	0.0484	0.0485	0.0539	0.0546	0.0469	0.0469
	Std	0.0189	0.0187	0.0196	0.0194	0.0179	0.0178
30×12	Ave	0.0445	0.0445	0.0470	0.0471	0.0443	0.0443
	Std	0.0102	0.0102	0.0102	0.0102	0.0091	0.0091
50×12	Ave	0.0441	0.0441	0.0457	0.0459	0.0434	0.0435
	Std	0.0081	0.0081	0.0072	0.0074	0.0070	0.0070
75×12	Ave	0.0447	0.0447	0.0466	0.0467	0.0436	0.0436
	Std	0.0064	0.0064	0.0064	0.0064	0.0052	0.0052
100×12	Ave	0.0439	0.0439	0.0449	0.0450	0.0439	0.0439
	Std	0.0053	0.0053	0.0053	0.0054	0.0054	0.0054

Table 7. Simulation Results for $\bar{\lambda}_{\mu,\sigma}$ and $\lambda^{\dagger}_{\mu,\sigma}$

We assume that the mean and variance of the factor are known and the financial economist incorporates the restrictions on the first two moments of the factor while estimating the factor risk premium. The table provides the results of Monte Carlo simulations on the estimators for λ under the assumption that the returns and the factor have a joint Normal distribution, t distribution or empirical distribution. Independent samples of the factor f_t are drawn from a Normal, t or empirical distribution with mean μ and variance σ^2. Independent samples of the residual ϵ_t of the regression equation are drawn, uncorrelated with the factor, from a Normal, t or empirical distribution with mean zero and variance Ω. Excess returns on 10 portfolios are constructed to satisfy $r_t = (\lambda\sigma^2/(1-\mu\lambda) + f_t - \mu)\beta + \epsilon_t$ for $t = 1, \cdots, T$. The parameters, λ, β, μ, σ and Ω, are set to those in Table 1. The estimator $\bar{\lambda}_{\mu,\sigma}$ is obtained by using the SDF method, and the estimator $\lambda^{\dagger}_{\mu,\sigma}$ is obtained by using the regression method. The two estimators are calculated based on the T samples to obtain a sample for each of the estimators. We repeat this independently for 100 times to obtain 100 independent samples for each of the estimators. The table presents the average and standard deviation of the 100 samples for each estimator.

Sample size		Normal		Student-t		Empirical	
T		$\bar{\lambda}_{\mu,\sigma}$	$\lambda^{\dagger}_{\mu,\sigma}$	$\bar{\lambda}_{\mu,\sigma}$	$\lambda^{\dagger}_{\mu,\sigma}$	$\bar{\lambda}_{\mu,\sigma}$	$\lambda^{\dagger}_{\mu,\sigma}$
5×12	Ave	0.0537	0.0438	0.0617	0.0439	0.0499	0.0443
	Std	0.0278	0.0025	0.0315	0.0025	0.0254	0.0029
10×12	Ave	0.0484	0.0434	0.0539	0.0437	0.0469	0.0438
	Std	0.0189	0.0012	0.0196	0.0015	0.0179	0.0014
30×12	Ave	0.0445	0.0436	0.0470	0.0436	0.0443	0.0437
	Std	0.0102	0.0005	0.0102	0.0005	0.0091	0.0005
50×12	Ave	0.0441	0.0436	0.0457	0.0436	0.0434	0.0437
	Std	0.0081	0.0004	0.0072	0.0004	0.0070	0.0004
75×12	Ave	0.0447	0.0437	0.0466	0.0436	0.0436	0.0437
	Std	0.0064	0.0003	0.0064	0.0003	0.0052	0.0003
100×12	Ave	0.0439	0.0436	0.0449	0.0437	0.0439	0.0437
	Std	0.0053	0.0003	0.0053	0.0003	0.0054	0.0003

Table 8. Simulation Results for $\hat{\lambda}_{\mu,\sigma}$ and $\lambda^*_{\mu,\sigma}$

In this table we assume that the financial economist obtains knowledge about the mean and variance of the factor before estimating the risk premium and incorporates the restrictions on the first two moments of the factor. The table provides the results of Monte Carlo simulations on the estimators for λ under the assumption that the returns and the factor have a joint Normal distribution, t distribution or empirical distribution. Independent samples of the factor f_t are drawn from a Normal, t or empirical distribution with mean μ and variance σ^2. Independent samples of the residual ϵ_t of the regression equation are drawn, uncorrelated with the factor, from a Normal, t or empirical distribution with mean zero and variance Ω. Excess returns on 10 portfolios are constructed to satisfy $r_t = (\lambda\sigma^2/(1-\mu\lambda)+f_t-\mu)\beta+\epsilon_t$ for $t = 1, \cdots, T$. The parameters, λ, β, μ, σ and Ω, are set to those in Table 1. The estimator $\hat{\lambda}_{\mu,\sigma}$ is obtained by using the SDF method, and the estimator $\lambda^*_{\mu,\sigma}$ is obtained by using the regression method. The two estimators are calculated based on the T samples to obtain a sample for each of the estimators. We repeat this independently for 100 times to obtain 100 independent samples for each of the estimators. The table presents the average and standard deviation of the 100 samples for each estimator.

Sample size		Normal		Student-t		Empirical	
T		$\hat{\lambda}_{\mu,\sigma}$	$\lambda^*_{\mu,\sigma}$	$\hat{\lambda}_{\mu,\sigma}$	$\lambda^*_{\mu,\sigma}$	$\hat{\lambda}_{\mu,\sigma}$	$\lambda^*_{\mu,\sigma}$
5×12	Ave	0.0373	0.0438	0.0362	0.0438	0.0361	0.0442
	Std	0.0077	0.0024	0.0079	0.0025	0.0067	0.0029
10×12	Ave	0.0402	0.0435	0.0407	0.0437	0.0395	0.0438
	Std	0.0053	0.0011	0.0071	0.0014	0.0051	0.0013
30×12	Ave	0.0424	0.0436	0.0427	0.0436	0.0426	0.0437
	Std	0.0037	0.0005	0.0055	0.0005	0.0030	0.0005
50×12	Ave	0.0432	0.0436	0.0434	0.0436	0.0428	0.0437
	Std	0.0027	0.0004	0.0039	0.0004	0.0028	0.0004
75×12	Ave	0.0434	0.0437	0.0434	0.0436	0.0431	0.0437
	Std	0.0020	0.0003	0.0036	0.0003	0.0022	0.0003
100×12	Ave	0.0432	0.0436	0.0435	0.0436	0.0433	0.0437
	Std	0.0018	0.0003	0.0031	0.0003	0.0017	0.0003

References

Campbell, J., A. Lo, and A. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ.

Chen, N., R. Roll, and S. Ross, 1986, Economic forces and the stock market, *Journal of Business*, 59, 383–404.

Cochrane, J., 1999, "Asset pricing," manuscript, Graduate School of Business, University of Chicago.

Cochrane, J., 2000, "A resurrection of the stochastic discount factor/GMM methodology," manuscript, Graduate School of Business, University of Chicago.

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

Elton, E., and M. Gruber, and J. Mei, 1994, "Cost of capital using arbitrage pricing theory: A case study of nine New York utilities," *Financial markets, Institutions & Instruments*, 3, 46–73.

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

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

E. Fama and K. French, 1997, "Industry costs of equity," *Journal of Financial Economics*, 43, 153–193.

Ferson, W., 1995, "Theory and empirical testing of asset pricing models," *Handbooks in Operations Research and Management Science* edited by Jarrow, Maksimovic and Ziemba, 145–200.

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

Hamilton, J., 1994, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.

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

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

Hansen, L., and K. Singleton, 1982, "Generalized instrumental variables estimation of non-linear rational expectations models," *Econometrica*, 50, 1269–1286.

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

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

Ingersoll, J., 1987, *Theory of Financial Decision Making*, Rowman & Littlefield, Totowa, NJ.

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

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

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

Schink, G., and R. Bower, 1994, "Application of the Fama-French model to utility stocks," *Financial markets, Institutions & Instruments*, 3, 74–96.

Taylor, M., and J. Thompson, 1986, "Data based random number generation for a multi-variate distribution via stochastic simulation," *Computational Statistics & Data Analysis*, 4, 93–101.