Time-Series and Cross-Sectional Tests of Asset Pricing Models

63

Kyung-Jin Choi, Dongcheol Kim, and Soon-Ho Kim

Abstract

This article reviews the methodologies of testing asset pricing models which are dominantly used in the literature; time-series regression tests and cross-sectional regression tests. We provide some explanations for the test procedure of time-series regression tests and cross-sectional regression tests. We discuss individual t-test, the joint F-test by Gibbons, Ross, and Shanken (Econometrica 57:1121–1152, 1989) and tests based on the generalized method of moments estimation. We also explain the two-pass test methodology and discuss the errors-in-variables problem which occurs inevitably in the two-pass methodology.

Keywords

Asset pricing models • Cross-sectional tests • Time-series tests

63.1 Introduction

It is very important to determine an appropriate asset pricing model since it is essential in estimating cost of capital for firms and discount rates for project cash flows and evaluating the performance of managed portfolios, etc. Since Sharpe (1964) proposed Capital Asset Pricing Model (CAPM), many theoretical and empirical asset pricing models have been suggested in the literature. In particular, numerous empirical asset pricing models have been suggested with some arguments. It is important, therefore, to determine an asset pricing model among these that describes well the intertemporal and cross-sectional dynamics of asset returns. To do so, it is essential to apply a correct test procedure.

This article reviews the methodologies of testing asset pricing models (see also Kim 2011). We provide some explanations for two testing methodologies dominantly used in the literature; time-series regression tests and cross-sectional regression tests. Time-series tests focus on the intertemporal explanatory power of returns on given factors for returns on test assets returns and on the pricing error of the given asset pricing model. Meanwhile, the cross-sectional tests focus on the explanatory power of factor loadings for asset returns in the cross-section. We explain individual t-test and the joint F-test by Gibbons, Ross, and Shanken (1989) which are used in time-series tests. We also explain the two-pass test methodology that is used in cross-sectional test, and discuss the errors-in-variables problem which occurs inevitably in the two-pass methodology. Moreover, we provide some explanations of tests based on the generalized method of moments estimation and tests based on the Hansen-Jagannathan (1997) distance.

This article is organized as follows. Section 63.2 explains time-series regression tests, Section 63.3 discuss tests based on the generalized method of moments estimation, and Section 63.4 explains cross-sectional regression tests and discusses the errors-in-variables problem. Section 63.5 concludes.

K.-J. Choi (✉) • D. Kim • S.-H. Kim
Korea University Business School, Anam-dong, Sungbuk-ku, Seoul 136-701, South Korea
e-mail: jinisrolling@gmail.com; kimdc@korea.ac.kr; gotoworld@korea.ac.kr

C.-F. Lee and A.C. Lee (eds.), *Encyclopedia of Finance*, DOI 10.1007/978-1-4614-5360-4_63,

63.2 Time-Series Tests

63.2.1 The Factor Model

Consider the following return generating process of the N assets. The returns of an asset i are generated by the following linear equation

$$R_{it} - R_{ft} = \alpha_i + \beta_i^{'} F_t + \varepsilon_{it}, \qquad (63.1)$$

where R_{it} is the return on the asset i $(i = 1, \cdots, N)$ at time $t(t = 1, \cdots, T)$, R_{ft} is the risk-free return, ε_{it} is the residual or idiosyncratic return with mean zero and covariance matrix Σ_ε, β_i is a $(K \times 1)$ vector of the factor loadings, and F_t is a $(K \times 1)$ vector of the factor portfolio returns. If the model (63.1) is well-specified, the equilibrium pricing equation is described as

$$E(r_{i,t}) = \gamma^{'} \beta_i \qquad (63.2)$$

where γ is a $(K \times 1)$ risk premia vector (the j-th element γ_j is the risk premium of the j-th risk factor), $r_{i,t}$ is the return of asset i in excess of the riskless rate of return $(= R_{it} - R_{ft})$. The above pricing equation of (63.2) has one restriction that the intercepts should be zero. The intercept from the time-series factor model of Equation 63.1 is interpreted as the pricing error. It is also dubbed the Jensen alpha. Thus, if $\alpha_i = 0$ for all test assets, we interpret this as evidence that supports the validity of a given factor model or the mean-variance efficiency of a given portfolio (or a linear combination of given factor portfolios).

63.2.2 Individual *t*-Test

The typical examples of time-series tests using Equation 63.1 are Black, Jensen and Scholes (1972) (hereafter BJS) and Fama and French (1993). BJS estimate the time-series regression model of Equation 63.1 for each of test assets and examine whether the intercept estimate α_i is significantly different from zero to test the validity of the CAPM. They use ten beta-sorted portfolios as test assets. Fama and French (1993) also estimate Equation 63.1 with three stock factor portfolios and two bond market portfolios by using 25 size-BM sorted portfolios as test assets. These two studies and many other studies use individual t-statistic of the intercept estimates to test the validity of a given factor model.

63.2.3 Joint *F*-Test

To test whether the intercept estimates of the time-series regression 63.1 are jointly different from zero, Gibbons, Ross, and Shanken (1989) (hereafter GRS) develop an F-tests. If the given asset pricing model or factor model is well-specified, the intercept estimates should not be different from zero. Therefore GRS F-statistic is to test a joint hypothesis $H_0 : \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_N)^{'} = 0$. Since the null hypothesis indicates that some linear combination of the factor portfolios is on the mean-variance efficiency boundary, the GRS test is in fact to test whether the linear combination of the factor portfolios is on the mean-variance efficient boundary.

Assuming the error terms are jointly normally distributed with mean zero and nonsingular error covariance matrix Σ_ε, GRS suggest the following test statistic for the null hypothesis:

$$W = \left[\frac{T(T-N-K)}{N(T-K-1)}\right]\left[\frac{\widehat{\alpha}{\prime}\widehat{\sum}_\varepsilon^{-1}\widehat{\alpha}}{1+\widehat{\vartheta}^2}\right] \sim F(N, T-N-K), \qquad (63.3)$$

where $\widehat{\sum}_\varepsilon$ is the unbiased residual covariance matrix, $\widehat{\vartheta}^2 = \bar{F}'\widehat{\sum}_F^{-1}\bar{F}$, $\bar{F}$ is a $(K \times 1)$ vector of average returns of the factor portfolios $(= (1/T)\sum_{t=1}^T F_t)$, $\widehat{\sum}_F$ is the covariance matrix of the factor portfolio returns. Here, $\widehat{\vartheta}$ is the ex post maximum Sharpe ratio of the K factor portfolios. The GRS test statistic of Equation 63.3 has a nice geometric interpretation. Since $\widehat{\alpha}'\widehat{\sum}_\varepsilon^{-1}\widehat{\alpha} = \widehat{\vartheta}^{*2} - \widehat{\vartheta}^2$ where $\widehat{\vartheta}^*$ is the ex post maximum Sharpe ratio of the N test assets and the K factor portfolios, the GRS test statistic equals

$$W = \left[\frac{T(T-N-K)}{N(T-K-1)}\right]\left[\left(\frac{1+\widehat{\vartheta}^{*2}}{1+\widehat{\vartheta}^2}\right) - 1\right] \sim F(N, T-N-K). \qquad (63.3a)$$

Thus the GRS test determines whether $\left|\widehat{\vartheta}^{*2}\right|$ is statistically greater than $\left|\widehat{\vartheta}^2\right|$. As the squared Sharpe ratio of the tangency portfolio constructed from factor portfolios, $\widehat{\vartheta}^2$, decreases, the GRS test statistic increases, which is evidence against the mean-variance efficiency of some linear combination of the factor portfolios. If the given factor model is the one-factor model with the market index portfolio such as Standard and Poors 500 Index, the GRS test is to test the mean-variance efficiency of the given market index portfolio.

63.3 Tests Based on Generalized Methods of Moments (GMM)

63.3.1 Introduction of the GMM

The time-series tests based on individual t-tests and the GRS tests assume that asset returns or the error terms of Equation 63.1 are normally distributed. It is well-known, however, that asset returns are not normally distributed. In this circumstance, the generalized method of moments (GMM) developed by Lars Hansen (1982) is useful. The GMM refers to the estimation which uses the sample moment conditions from the population moment conditions. The main advantage of the GMM is that it makes empirical tests of an economic theory possible under a general framework. Specifically, the GMM makes nonlinear asset pricing models be empirically testable without linearization. The GMM is computationally convenient without strong distributional assumptions, and thus it can be used under serial correlation and conditional heteroskedasticity in making inferences. Moreover, it does not have to assume data generating processes. These advantages are the reason that the GMM is one of the most frequently applied methodologies in finance.

It should be mentioned, however, that the GMM estimators are only asymptotically valid. Ferson and Foerster (1994) find that a two-stage GMM approach tends to reject the models too often in a moderate sample size. Arellano and Bond (1991) also find that the estimated standard errors of the GMM estimators are severely downward biased in a small sample size. Ahn and Gadarowski (2004) show that asset pricing tests based on the GMM estimation reject the correct model too often. In other words, like other tests based on asymptotic estimators, tests based on the GMM estimation have too large Type I error.

63.3.2 Overview of the GMM Estimation

An economic model can be specified as a vector of population moment conditions: $E[g(x_t,\theta)] = E[g_t(\theta)] = 0$, where x_t is a vector of P variables for $t = 1,\ldots,T$, and θ is a parameter vector of K unknown parameters, $g(\cdot)$ is a vector of N functions, with the necessary order condition for identification such that $K \leq N$. Since the order condition is a necessary condition for identification, a failure of the order condition means that the model is not identified. When the model is over-identified, Hansen's (1982) J test can be implemented as using the over-identifying restrictions. The exactly identified model cannot be tested whether the model is correctly specified or not.

Let $g_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} g_t(\theta)$be the corresponding sample moments to the population moments. In the asset pricing test framework, $g_T(\theta)$ can be interpreted as a vector of pricing errors. Then, the GMM estimator $\widehat{\theta}$ for θ minimizes the following quadratic function with respect to θ.

$$Q = [g_T(\theta)]' \; W \; [g_T(\theta)], \tag{63.4}$$

where $\boldsymbol{W}$ is a weighting matrix which is the positive definite. The weighting matrix $\boldsymbol{W}$ implies the weight to be given for each moment. If $\boldsymbol{W}$ is an identity matrix, the GMM treats all moments symmetrically. This means that all assets have the same degree of importance in estimating the parameters and that the objective function Q is simply to minimize the sum of squared pricing errors.

Hansen (1982) proposes the optimal weighting matrix as follows. The optimal weighting matrix produces the estimates with the smallest asymptotic variance.

When $E[g_t(\theta)] = plim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} g_t(\theta) = p^{lim}_{T\to\infty} \; g_T(\theta)$ and the central limit theorem is applied,

$$\sqrt{T} g_T(\theta) \xrightarrow{d} N(0, V) \text{ as } T \to \infty, \tag{63.5}$$

where $V = \sum_{j=-\infty}^{\infty} E[g(x_t,\theta)\; g(x_{t+j},\theta)']$which is a nonsingular covariance matrix. $\boldsymbol{V}$ is the spectral density matrix of $g_T(\theta)$. When $\boldsymbol{W} = \boldsymbol{V^{-1}}$, Hansen (1982) finds that the GMM estimates are efficient. A lower bound for the asymptotic variance of the GMM estimates can be achieved as choosing $\boldsymbol{W} = \boldsymbol{V^{-1}}$. Hansen (1982) shows that the asymptotic distribution of the GMM estimator is given by

$$\sqrt{T(\widehat{\theta} - \theta)} \xrightarrow{d} N\left(0, \left(\mathbf{D}'\mathbf{V}^{-1}\mathbf{D}\right)^{-1}\right) \text{ as } T \to \infty, \tag{63.6}$$

where $D = E\left[\frac{\partial g_t(\theta)}{\partial \theta'}\right] = p^{lim}_{T\to\infty} \frac{\partial g_T(\theta)}{\partial \theta'}$.

To calculate the efficient GMM estimate, the consistent estimator $\widehat{V}$ is needed. Newey and West (1987) and Cochrane (2005) show the estimation of $\boldsymbol{V}$ in form of spectral density matrix. A two-stage procedure can be employed to obtain the efficient GMM estimator. In the first-stage, a consistent estimator $\widehat{V}$ is obtained by using the identity matrix as the weighting matrix. In the second-stage, the efficient GMM estimator can be produced using $\widehat{V}^{-1}$ as the weighting matrix. Cochrane (2005) and Hayashi (2000) describe in details. Tests for the model misspecification can be implemented as examining the extent of the pricing errors $g_T(\widehat{\theta})$. Hansen (1982) proposes J-statistics as follows

$$J = T g_T(\widehat{\theta}) V^{-1} g_T(\widehat{\theta}) \text{ and } J \xrightarrow{d} \chi^2(N-K) \text{ as } T \to \infty. \tag{63.7}$$

If the model is exactly identified, then $J = 0$.

Generally, the efficient GMM estimates are known to have poorer small sample properties. The model misspecification test based on the efficient GMM estimates also rejects the null hypothesis too often. Therefore, sufficiently large number of time-series return observations (T) relative to the number of assets (N) is required to obtain the consistent estimators. As the number of assets becomes large relatively to T, it is difficult to estimate the precise covariance matrix of returns. To avoid this problem, one approach is to group the assets into a small number of portfolios. Ferson and Foerster (1994) show that a GMM approach still tends to reject the models too often in moderate sample size. Other solutions are also proposed to improve the inference problem caused by the underestimation of the standard error of the GMM estimator. Ren and Shimotsu (2009) propose an improved method to compute the weighting matrix by applying the shrinkage method (Ledoit and Wolf (2003)). Windmeijer (2005) suggests the correction term to generate the more precise estimator of the standard error. Windmeijer (2005) adjusts the bias from the extra variation in small samples and reports the improved results.

63.3.3 Moments Conditions for the Asset Pricing Tests

Once the moment conditions are indentified, the estimation and the inference are straightforward. Generally, the optimization is performed by the numerical method. The closed-form solutions for the parameters are obtained only for some special cases in the GMM method.

The GMM method can be used for both the beta representation and the stochastic discount factor (SDF) representation in the asset pricing model. The asset pricing model under the beta representation is as follows:

$$E[r_t] = \gamma_1 \beta \text{ or } E[R_t] = \gamma_0 + \gamma_1 \ \beta, \tag{63.8}$$

where r_t is the excess return, R_t is the raw return, γ_0 is the return of the zero beta asset, γ_1 is the factor risk premium, and $\beta[= Cov(R_t, f_t)/\mathrm{Var}(f_t)]$ is the sensitivity of the asset return to the factor. The SDF representation is given by

$$E[r_t m_t] = 0 \text{ or } E[R_t m_t] = 1, \tag{63.9}$$

where m_t is referred to as a stochastic discount factor. For example, the stochastic discount factor is given by $m_t = \lambda_0 + \lambda_1 f_t$ for a one-factor linear pricing model. If the model is true, then $m_t \in M_t$, where M_t is the set of the true stochastic discount factors.

The SDF representation using the GMM becomes popular in testing asset pricing models, since it can be used within the general framework to analyze either linear or nonlinear pricing models. However, there are some concerns that the generality of the SDF method using the GMM comes at the cost of efficiency in two aspects; estimation of the parameters and power in the specification tests. Kan and Zhou (1999) point out that the SDF method using the GMM has two problems when asset returns are generated by a linear factor model. The first is that the risk premium estimate from the SDF method is not reliable. The second problem is that the model specification test using the SDF method has too low power in detecting the misspecified model. But Jagannathan and Wang (2002) argue that the SDF method using the GMM is asymptotically as efficient as the beta method using the GMM in estimating the risk premium, and that the model specification test based on the SDF is also as powerful as the one based on the beta method using the GMM. They argue that Kan and Zhou's (1999) conclusion is inappropriate in that they ignored the risk premium measures in the two methods not identical and that they assumed the factor is predetermined. Jagannathan and Wang (2002) consider explicitly the additional moments about the SDF method and the transformation between the risk premium parameters in order to make a fair comparison between the two methods.

Let F_t and $\boldsymbol{R}_t$ be a vector of K risk factors and a vector of N test asset returns respectively. The pricing model can be denoted as

$$E[R_t] = \gamma 1 + B\Gamma, \tag{63.10}$$

where γ is the nonzero return of the zero-beta asset, **1** is an $(N \times 1)$ vector of ones, $\boldsymbol{R}_t$ is an $(N \times 1)$ vector of raw returns for the test assets, $\boldsymbol{B}$ is the factor loadings of the following time-series regression, $R_t = \varphi + Bf_t + \varepsilon_t$, the error term ε_t has zero mean and uncorrelated with the factors f_t. Then $\varphi = \gamma 1 + B(\Gamma - \mu_f)$, where μ_f is the mean of the factors. Using this restriction, the following restrictions can be obtained,

$$R_t = (\gamma 1 + B\Gamma - B\mu f) + Bf_t + \varepsilon_t \tag{63.11}$$

$$E[\varepsilon_t] = 0 \tag{63.12}$$

$$E[\varepsilon_t f'_t] = 0. \tag{63.13}$$

Now, the following $N + NK + K$ moment conditions can be made,

$$E\left[R_t - \gamma 1 - B(\Gamma - \mu_f + f_t)\right] = 0 \tag{63.14}$$

$$E\left[\left[r_t - \gamma 1 - B(\Gamma - \mu_f + f_t)\right]f_t{}'\right] = 0 \tag{63.15}$$

$$E\left[f_t - \mu_f\right] = 0. \tag{63.16}$$

Then, the unknown parameters to be estimated are γ, Γ, ***B*,** and μ_f. When the factor is the return on a portfolio of traded assets, it implies $\Gamma = \mu_f$ and the moment conditions are reduced as follows,

$$E[R_t - \gamma 1 - Bf_t] = 0 \tag{63.17}$$

$$E[(R_t - \gamma 1 - Bf_t)f_t{}'] = 0. \tag{63.18}$$

The moment conditions for the SDF representation can be derived from the linear asset pricing model. Using $E[R_t] = \gamma 1 + B\Gamma$ and $B = \Sigma_{Rf}\Sigma_{ff}^{-1}$, the following equation can be obtained

$$E\left[R_t\left(\frac{1}{\gamma} + \frac{\Gamma'\Sigma_{ff}^{-1}\mu_f}{\gamma} - \frac{\Gamma'\Sigma_{ff}^{-1}}{\gamma}f_t\right)\right] = 1. \tag{63.19}$$

Then, the following SDF representation can be produced

$$E[R_t m_t] = 1, \tag{63.20}$$

where $m_t = \lambda' x_{\mathbf{t}}$, $\lambda = [\lambda_0 \lambda_1 \ldots \lambda_k]'$, and $x_t = [1 f_t']'$. Using these moment conditions, the parameters can be estimated, and the specification test can be implemented. The closed-form solution is known as follows.

$$\hat{\lambda} = (\hat{D}\hat{V}^{-1}\hat{D})\hat{D}\hat{V}^{-1}\bar{r}, \tag{63.21}$$

where $\hat{D} = (1/T)\sum_{t=1}^{T} R_t x_t'$, $\bar{r} = (1/T)\sum_{t=1}^{T} r_t$, and $\hat{V}$ is a consistent estimator in Equation 63.5. The moment conditions for the conditional asset pricing are provided by Harvey (1989, 1991), Ferson and Harvey (1993), Jagannathan and Wang (1996), and Cochrane (1996).

63.3.4 The Hansen-Jagannathan Distance

Hansen and Jagannathan (1997) develop a distance metric which is referred as the Hansen-Jagannathan (HJ)-distance. They propose how to measure the distance between a true stochastic discount factor and the implied stochastic discount factor. The HJ-distance can be interpreted as the normalized maximum pricing error of the model for the test assets. If the implied stochastic discount factor from the model is included in the set of true stochastic discount factor, then the HJ-distance is zero and the pricing error is none. To compute the HJ-distance, we can use simply the inverse of the second moment of the test asset returns as the weighting matrix, $E(R_t R_t')^{-1}$. They also provide the distribution of the HJ-distance.

The HJ-distance is defined as follows

$$\delta = \text{Min } ||y_t - m_t|| \text{ with respect to } m_t, \tag{63.22}$$

where $E(m_t R_t) = 1$, R_t is the raw returns, y_t is a candidate SDF and $m_t \in M_t$, where M_t is the set of true SDF's. The measure of distance is the usual norm. The minimization problem can be written as the Lagrangian minimization problem

$$\delta^2 = \underset{m \in L^2, m \geq 0}{\text{Min}} \; \underset{\tau \in R^n}{\sup} \left\{E(y_t - m_t)^2 - 2\tau'[E(m_t R_t) - 1]\right\}, \tag{63.23}$$

where δ is the minimum distance from the candidate SDF to the true pricing kernel, L^2 is the space of all random variables with finite second moments. Hansen and Jagannathan (1997) show the following relation

$$y_t - m_t^* = \tau^{*'} R_t, \tag{63.24}$$

where m_t^* and τ^* are the solutions for the Lagrangian minimization problem, and $\tau^* = E(R_t R_t')^{-1} E(y_t R_t - 1)$. Therefore, the HJ-distance is

$$\delta = ||y_t - m_t^*|| = \left[\tau^{*'} E(R_t R_t')\tau^*\right]^{\frac{1}{2}}. \tag{63.25}$$

Using $\tau^* = E(R_t R_t')^{-1} E(y_t R_t - 1)$, the following equation can be produced.

$$\delta = \left[E(y_t R_t - 1)' E(R_t R_t')^{-1} E(y_t R_t - 1)\right]^{\frac{1}{2}}, \tag{63.26}$$

where the pricing errors $g_t(\theta)$ can be defined as $E(y_t R_t - 1)$ here. Equation 63.26 is exactly the same as $\sqrt{Q}$ of Equation 63.4 with the weighting matrix $E(R_t R_t')^{-1}$ The sample analog can be as follows

$$\hat{\delta} = [g_T(\theta)]' \hat{G}^{-1} [g_T(\theta)], \tag{63.27}$$

where $\hat{G} = (1/T)\sum_{t=1}^{T} R_t T_t'$, and $\mathrm{G} = E(R_t R_t')$.

The HJ-distance means the smallest distance between the candidate SDF the true SDF. It can also be interpreted as the maximum pricing error of the considered test assets. Consider the following equations

$$|E(y_t R_{it}) - E(m_t^* R_{it})| = |E[(y_t - m_t^*)R_{it}]| \leq E[(y_t - m_t^*)^2]^{\frac{1}{2}} E[R_{it}^2]^{\frac{1}{2}}. \quad (63.28)$$

The Cauchy-Schwartz inequality is applied to the last term. Then, Equation 63.25 can be rewritten as follows

$$\delta = ||y_t - m_t^*|| \geq \frac{|E(y_t R_{it}) - E(m_t^* R_{it})|}{||R_{it}}. \quad (63.29)$$

Therefore, the HJ-distance can also be interpreted as the normalized maximum pricing error for the considered test assets.

We estimate a vector of the unknown parameters θ and conduct the model specification test using the typical GMM procedure as follows

$$\hat{\theta}^{HJ} = arg\,\min_{\theta}\, [g_T(\theta)]'\, \hat{G}^{-1}\, [g_T(\theta)]. \quad (63.30)$$

The distribution of $\hat{\theta}^{HJ}$ is the same as in Equation 63.6. However, the distribution of the model specification test statistic, which is T times the square of the HJ-distance, is not equal to that of the Hansen's J-statistic. Jagannathan and Wang (1996) suggest that the distribution of $T[g_T(\hat{\theta}^{HJ})]^2$ is a linear combination of χ^2 (63.1) variables. The distribution for the null hypothesis that the HJ-distance δ is zero is given by

$$T[\delta(\hat{\theta}^{HJ})]^2 \xrightarrow{d} \sum_{j=1}^{N-K-1} \varphi_j \pi_j \text{ as } T \to \infty, \quad (63.31)$$

where π_j is a random variable that follows χ^2 (63.1), and φ_j is the nonzero eigenvalues of the following matrix

$$\Lambda = V^{1/2} G^{-1/2} \left(I - (G^{-1/2})' D\, (D' G^{-1} D)^{-1} D' G^{-1/2} \right) (G^{-1/2})' (V^{1/2})', \quad (63.32)$$

where $\boldsymbol{V}^{1/2}$ and $\boldsymbol{G}^{1/2}$ are the upper triangular matrices in the Cholesky decompositions of $\boldsymbol{V}$ and $\boldsymbol{G}$. Jagannathan and Wang (1996) suggest that the empirical p-value can be obtained by comparing $T[\delta(\hat{\theta}^{HJ})]^2$ to $\sum_{j=1}^{N-K-1} \varphi_j \pi_j$ which is computed as random draws from a χ^2 (63.1) distribution. The empirical p-value of the HJ-distance can be computed from the following function.

$$\text{Empirical p-value} = \frac{1}{I} \sum_{i=1}^{I} f(x_i), \quad (63.33)$$

where $x_i = \left(\sum_{j=1}^{n-k-1} \varphi_j \pi_{ji} \right) - T[g_T(\hat{\theta}^{HJ})]^2$, $f(x_i)$ is the function which returns 1 if $x_i \geq 0$ and 0 otherwise. Jagannathan and Wang (1996) set I to 5000 in their paper. According to Ahn and Gadarowski (2004), tests based on this empirical p-value reject too often the correct models, since the computation of the empirical p-value depends on the eigenvalues from the estimate of Λ and $\widehat{\text{V}}$ has the poor finite sample property.

63.3.5 An Equality Test for the HJ-Distance

Kan and Robotti (2009) develop the method to test the HJ-distance equality of the two competing linear asset pricing models. They analyze the case of both the nested and non-nested models whether the HJ-distances of the two competing models are equal. This section introduces the general case when both non-nested models are misspecified and the both stochastic discount factors are not equal.

Let y_1 and y_2 be two competing models. δ_1 andδ_2 of Equation 63.22 in the two models are nonzero. All subscript means the model type. Let $d_t = q_{1t} - q_{2t}$, where $q_{1t} = 2u_{1t}y_{1t} - u_{1t}^2 + \delta_1^2, q_{2t} = 2u_{2t}y_{2t} - u_{2t}^2 + \delta_2^2$ with $u_{1t} = g_1' G^{-1} R_t$ and $u_{2t} = g_2' G^{-1} R_t$. Then

$$\sqrt{T}\left[\hat{\delta}_1^2 - \hat{\delta}_2^2 - (\delta_1^2 - \delta_2^2)\right] \xrightarrow{d} N(0, v_d), \quad (63.34)$$

where $v_d = \sum_{j=-\infty}^{\infty} E[d_t d_{t+j}]$ and $d_t = 2u_{1t}y_{1t} - u_{1t}^2 - 2u_{2t}y_{2t} + u_{2t}^2$. Under the null hypothesis $H_0 : \delta_1^2 = \delta_2^2 \neq 0$,

$$\sqrt{T}(\hat{\delta}_1^2 - \hat{\delta}_2^2) \xrightarrow{d} N(0, v_d) \quad (63.35)$$

63.4 Cross-Sectional Tests

63.4.1 A Risk Premia Estimation Through Two-Pass Regressions

Traditional asset pricing theories such as the CAPM (Sharpe, 1964; Lintner, 1965; Black, 1972), the APT (Ross, 1976, the Intertemporal CAPM (Merton, 1973), and the Consumption CAPM (Breeden, 1979) imply that the difference in expected returns across assets should be explained by their covariances with (systematic) risk factors or their factor loadings. When asset returns are generated from a K-factor model, their expected returns are represented by a linear combination of the factor loadings or betas, as in Equation 63.2,

$$E(r_{it}) = \gamma' \beta_i \quad \text{for } i = 1, 2, \ldots, N. \tag{63.2}$$

Since the factor loadings, β_i, are not directly observable, it is natural to test the cross-sectional relation between expected returns and factor ladings in two stages. In the first stage, beta estimates are obtained from time-series regressions of Equation 63.1,

$$r_{it} = R_{it} - R_{ft} = \alpha_i + \beta_i' F_t + \varepsilon_{it},\; E(\varepsilon_{i,t}|F_t) = 0. \tag{63.36}$$

In practice, beta estimates in the first stage regression are computed using the OLS and are given as follows

$$\hat{\beta}_i = \widehat{Var(F)}^{-1} \widehat{\mathrm{Cov}(F, r_i)}, \tag{63.37}$$

where $\widehat{Var(F)} = (1/T)\sum (F_t - \bar{F})'(F_t - \bar{F})$ and $\widehat{\mathrm{Cov}(F, r_i)} = (1/\mathrm{T})\sum (F_t - \bar{F})(r_{it} - \bar{r}_i)$. In the second stage, one runs a cross-sectional regression (CSR) of excess average returns of N assets on their estimated betas obtained from the first stage,

$$\bar{r} = \hat{\beta}\gamma + e, \tag{63.38}$$

where $\bar{r}$ is an $(\mathrm{N} \times 1)$ average returns, $\hat{\beta} = [\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_N]'$ is an $(\mathrm{N} \times \mathrm{K})$ beta estimates matrix, and $\boldsymbol{e}$ is an $(\mathrm{N} \times 1)$ residual returns with mean zero and covariance matrix $\mathbf{Var}(e) = \Sigma_e$.

Under the classical linear model assumptions, the OLS risk premia estimates of Equation 63.38 is given as

$$\hat{\gamma}_{\mathrm{OLS}} = (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'\bar{r}. \tag{63.39}$$

However, obtaining the correct distribution of $\hat{\gamma}_{\mathrm{OLS}}$ in finite samples is not so trivial, since the explanatory variables are measured with error and the least squares estimates of Equation 63.39 are subject to the errors-in-variables (EIV) bias. If we ignore the EIV bias and regard $\hat{\beta}$ as the true beta variable $\boldsymbol{\beta}$ for a while, then the distribution of the OLS estimator of Equation 63.39 is

$$\hat{\gamma}_{OLS} \sim N(\gamma, (\beta'\beta)^{-1}\beta' \mathrm{Var}(e)\beta(\beta'\beta)^{-1}). \tag{63.40}$$

Since $Var(e) = Var(\bar{r} - \beta\gamma) = Var(\alpha + \beta\bar{F} + \bar{\varepsilon} - \beta\gamma) = Var[\beta(\bar{F} - \gamma) + \bar{\varepsilon}]$ the error covariance of the cross-sectional regression in (63.38) is

$$Var(e) = \frac{1}{T}(\beta\Sigma_F\beta' + \Sigma_\varepsilon), \tag{63.41}$$

where Σ_F is the factor covariance and Σ_ε is the error covariance in the time-series regression. Therefore, Equation 63.40 can be rewritten as

$$\hat{\gamma}_{OLS} \sim N\left(\gamma, \frac{1}{T}[(\beta'\beta)^{-1}\beta'\Sigma_\varepsilon\beta(\beta'\beta)^{-1} + \Sigma_F]\right). \tag{63.42}$$

See also Cochran (2004). For statistical inference, each variable in Equation 63.42 can be replaced with its time-series sample counterpart.

63.4.2 The GLS Estimation of Risk Premia

The OLS risk premia estimation in the previous section assumes that the residuals are cross-sectionally uncorrelated. However, since this assumption could be too strong, risk premia can be estimated by the generalized least squares (GLS) in which residuals are weighted in a way to ease heteroskedasticity. The GLS estimate of risk premia is[1]

$$\begin{aligned} \hat{\gamma}_{GLS} &= [\beta' Var(e)^{-1}\beta]^{-1}\beta' Var(e)^{-1}\bar{r} \\ &= \left[\beta'(\beta\Sigma_F\beta' + \Sigma_\varepsilon)^{-1}\beta\right]^{-1}\beta'(\beta\Sigma_F\beta' + \Sigma_\varepsilon)^{-1}\bar{r} \\ &= (\beta'\Sigma_\varepsilon^{-1}\beta)^{-1}\beta'\Sigma_\varepsilon^{-1}\bar{r}. \end{aligned} \tag{63.43}$$

The distribution of the GLS estimator of Equation 63.43 is

$$\hat{\gamma}_{GLS} \sim N\left(\gamma, \frac{1}{T(\beta'\Sigma_\varepsilon^{-1}\beta + \Sigma_F)}\right). \tag{63.44}$$

63.4.3 The Errors-in-Variables Problem

63.4.3.1 Corrections for the Standard Errors

Previous discussions simply ignore the fact that beta estimates obtained in the first-pass are not true betas but are measured with error (Shanken, 1992). It is well known that the beta estimate in time-series regression is T-consistent (as T goes to infinity, the estimation error of beta estimates goes to zero). In other words, as the number of time-series observations (T) goes to infinity and they are stationary, the

[1] Let $\mathbf{A} = \mathbf{I} + \beta'\Sigma_{\boldsymbol{\varepsilon}}^{-1}\boldsymbol{\beta}\Sigma_{\mathbf{F}}$. Then, $\hat{\gamma}_{\mathrm{GLS}} = \left[\mathrm{A}\beta'(\beta\sum_{\mathrm{F}}\beta' + \sum_\varepsilon)^{-1}\beta\right]^{-1} \mathrm{A}\beta'(\beta\sum_{\mathrm{F}}\beta' + \sum_\varepsilon)^{-1}\bar{r}$. Since, $\mathbf{A}\beta' = \beta'\Sigma_{\boldsymbol{\varepsilon}}^{-1}(\Sigma_{\boldsymbol{\varepsilon}} + \beta\Sigma_{\mathbf{F}}\,\beta')$, $\hat{\gamma}_{\mathbf{GLS}} = (\beta'\Sigma_\varepsilon^{-1}\beta)^{-1}\boldsymbol{\beta}\Sigma_{\boldsymbol{\varepsilon}}^{-1}\bar{\mathrm{r}}$.

beta estimate converges to its true beta. As long as we obtain T-consistent beta estimates, the CSR least squares estimates of risk premia are N-consistent if the number of test assets in the CSR (N) is large. However, since return observations show nonstationarity (see Kim and Kon (1996, 1999)) and thus it is difficult to increase T in estimating betas, it would be hard to obtain T-consistent beta estimates, and accordingly, the least squares estimates of Equations 63.39 and 63.43 and their standard errors are not N-consistent and are biased.

Shanken (1992) suggests a method of correcting for the bias of the standard errors of the CSR least squares estimates in the two-pass methodology. Assuming the model given in Equation 63.36 is correct, Shanken (1992) provides the correction for the variance of the CSR risk premia estimates as follows.

$$Var(\widehat{\gamma}_{OLS}) = \frac{1}{T}\left[(1+\gamma'\Sigma_{\mathbf{F}}^{-1}\gamma)(\beta'\beta)^{-1}\beta'\Sigma_{\varepsilon}\beta(\beta'\beta)^{-1} + \Sigma_F\right] \tag{63.45}$$

$$Var(\widehat{\gamma}_{GLS}) = \frac{1}{T}\left[(1+\gamma'\Sigma_{\mathbf{F}}^{-1}\gamma)(\beta'\Sigma_{\varepsilon}^{-1}\beta)^{-1} + \Sigma_F\right]. \tag{63.46}$$

Notice that t-values of the CSR estimates tend to be overstated due to the measurement errors of the explanatory variables $\widehat{\beta}$. Since the corrected variance of (63.45) or (63.46) is increased by adding the correction term$\gamma'\Sigma_{\mathbf{F}}^{-1}\gamma$, the above correction attempts to lessen the EIV bias in the standard error of the estimates.

63.4.3.2 Corrections for the Risk Premium Estimate

The Shanken correction in (63.45) or (63.46) can be applied only to make a correction for the EIV bias of the standard errors of the CSR risk premia estimates, not of the estimates themselves (Kim, 1995). Instead of correcting the variance as in Shanken (1992), Kim (1995) provides a direct correction for the EIV bias of the CSR market risk premium estimate within the Fama and MacBeth methodology. Suppose we have the following CSR model to be estimated at time t with an estimated beta variable and a firm characteristic variable (e.g., firm size or book-to-market) measured without error at $t-1$,

$$R_t = \gamma_{0t} + \gamma_{1t}\widehat{\beta}_{t-1} + \gamma_{2t}V_{t-1} + \varepsilon_t. \tag{63.47}$$

In the above equation, the beta variable is measured with error. The beta variable is typically obtained from the time series regression (the first pass) of a firm's returns on the market returns by using return observations from $t-T$ to $t-1$. When the beta variable is measured error as in Equation 63.47, it is well known that the EIV problem is present. Thus, the EIV problem occurs in the two-pass methodology. The least squares estimation of the CSR in Equation 63.47 results in the underestimation of the role of market beta and the overestimation of the role of the firm characteristic variables.

To make correction for the EIV problem, Kim (1995) provides an EIV corrected estimators for the CSR coefficients of Equation 63.47, $\widehat{\gamma}_{0t}, \widehat{\gamma}_{1t}$, and $\widehat{\gamma}_{2t}$, which are N-consistent.

$$\widehat{\gamma}_{1t} = C_{1t}\widehat{\gamma}_{1t}^{LS} \tag{63.48}$$

$$\widehat{\gamma}_{2t} = C_{2t}\widehat{\gamma}_{2t}^{LS} \tag{63.49}$$

$$\hat{\gamma}_{0t} = \hat{\gamma}_{0t}^{LS} - (C_{1t}-1)\hat{\gamma}_{1t}^{LS}\bar{\hat{\beta}}_{t-1} - (C_{2t}-1)\hat{\gamma}_{2t}^{LS}\bar{V}_{t-1} \tag{63.50}$$

where

$$C_{1t} = \frac{1}{\left[1-\frac{1}{2}\left(\frac{m_{RR}\left(1-\hat{\rho}_{RV}^2\right)}{\delta_t m_{\hat{\beta}\hat{\beta}}\left(1-\hat{\rho}_{\hat{\beta}V}^2\right)}+1\right)\left(1-(1-Q)^{\frac{1}{2}}\right)\right]} \geq 1$$

$$Q = \frac{1-\hat{\rho}_{R\hat{\beta}\cdot V}^2}{\left[\frac{1}{2}+\frac{1}{4}\left(\frac{\delta_t m_{\hat{\beta}\hat{\beta}}\left(1-\hat{\rho}_{\hat{\beta}V}^2\right)}{m_{RR}\left(1-\hat{\rho}_{RV}^2\right)}+\frac{m_{RR}\left(1-\hat{\rho}_{RV}^2\right)}{\delta_t m_{\hat{\beta}\hat{\beta}}\left(1-\hat{\rho}_{\hat{\beta}V}^2\right)}\right)\right]}$$

$$C_{2t} = 1 - \frac{(C_{1t}-1)\left\{(\hat{\rho}_{R\hat{\beta}}\hat{\rho}_{\hat{\beta}V}/\hat{\rho}_{RV}) - \hat{\rho}_{\hat{\beta}}^2\right\}}{1-(\hat{\rho}_{R\hat{\beta}}\hat{\rho}_{\hat{\beta}V}/\hat{\rho}_{RV})}$$

and $\delta_t = T\hat{\sigma}_{m,t-1}^2$ where $\hat{\sigma}_{m,t-1}^2$ is the MLE of the market variance computed using market returns from $t-T$ to $t-1$, $m_{xy} = (1/N)(x-\iota\tilde{x})'\hat{\Sigma}_s^{-1}(y-\iota\tilde{y})$, $\tilde{x} = \Sigma_{i=1}^N\Sigma_{j=1}^N w_{ij}x_i/\Sigma_{i=1}^N\Sigma_{j=1}^N w_{ij}$, w_{ij} is the (i,j)-th element of $\hat{\sum}_{\varepsilon}^{-1}$, $\hat{\rho}_{xy} = m_{xy}/\left(m_{xx}m_{yy}\right)^{1/2}$ and $\hat{\rho}_{R\hat{\beta}\cdot V}^2 = \left(\hat{\rho}_{R\hat{\beta}} - \hat{\rho}_{RV}\right)^2/\left\{\left(1-\hat{\rho}_{\hat{\beta}V}^2\right)\left(1-\hat{\rho}_{RV}^2\right)\right\}$. Here, $\hat{\gamma}_t^{LS}$ is the least squares estimate. Note that $m_{..}$ is the cross-sectional sample second moment and $\hat{\rho}_{..}$ is thus the cross-sectional correlation coefficient.

Since the EIV-bias correction term C_{1t} is greater than one, this term makes correction for the underestimation of the least squares estimate,$\hat{\gamma}_{1t}^{LS}$. Also, since the term C_{2t} is usually less than one, this term makes correction for the overestimation of the least squares estimate of the coefficient on the firm characteristic variable such as firm size and book-to-market,$\hat{\gamma}_{2t}^{LS}$. Following the seminal works of Black, Jensen and Scholes (1972) and Fama and MacBeth (1973), portfolio grouping procedures are prevalent to lessen the EIV problem. Aside the fact that the portfolio grouping procedure does not completely remove the problem, however, it significantly reduces the sample size (N), henceforth weakens the power of the tests. In this regard, a direction correction for the EIV problem like Kim (1995) is important for the use of all individual assets as test assets.

63.5 Summary and Concluding Remarks

This article reviews the methodologies of testing asset pricing models. We provide some explanations for two testing methodologies dominantly used in the literature; time-series regression tests and cross-sectional regression tests. We explain individual t-test and the joint F-test by Gibbons, Ross, and Shanken (1989) which are used in time-series tests. We also explain the two-pass test methodology that is used in cross-sectional test, and discuss the errors-in-variables problem which occurs inevitably in the two-pass methodology. Moreover, we provide some explanations of tests based on the GMM estimation and tests based on the Hansen-Jagannathan (1997) distance.

References

Ahn, S. C. and C. Gadarowski. 2004. "Small sample properties of the GMM specification test based on the Hansen-Jagannathan distance." *Journal of Empirical Finance* 11, 109–132.

Arellano, M. and S. Bond. 1991. "Some tests of specification for panel data: monte carlo evidence and an application to employment equations." *Review of Economic Studies* 58, 277–297.

Black, F., 1972. "Capital market equilibrium with restricted borrowing." *Journal of Business* 45, 444–455.

Black, F., Jensen, M. C., and M. S. Scholes. 1972. "The capital asset pricing model: some empirical tests." *Studies in the Theory of Capital Markets,* Praeger Publishers Inc.

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

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

Cochrane, J. H. 2004. *Asset Pricing*. Princeton University Press.

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

Fama, E. F. and J. D. MacBeth. "Risk, return and equilibrium: empirical tests." *Journal of Political Economy* 81, 607–636.

Ferson, W. E. and S. R. 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, W. E. and C. R. Harvey. 1993. "The risk and predictability of international equity returns." *Review Financial Studies* 6, 527–566.

Gibbons, M. R., S. A. Ross and J. Shanken. 1989. "A test of the efficiency of a given portfolio." *Econometrica* 57, 1121–1152.

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

Hansen, L. P. and R. Jagannathan. 1997. "Assessing specification errors in stochastic discount factor models." *Journal of Finance* 62, 557–590.

Harvey, C. R. 1989. "Time-varying conditional covariances in tests of asset pricing models." *Journal of Financial Economics* 24, 289–317.

Harvey, C. R. 1991. "The world price of covariance risk." *Journal of Finance* 46, 111–157.

Hayashi, F. 2000. *Econometrics*. Princeton University Press.

Jagannathan, R., G. Skoulakis and Z. Wang. 2002. "Generalized method of moments: applications in finance." *Journal of Business & Economic Statistics* 20, 470–481.

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 C. Robotti. 2009. "Model comparison using the Hansen-Jagannathan Distance." *Review of Financial Studies* 22, 3449–3490.

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

Kim, D. 1995. "The Errors-In-Variables Problem in the cross-section of expected stock returns." *Journal of Finance* 50, 1605–1634.

Kim, D. 2011. "Asset pricing models in the Korean Stock Markets: a review for the Period of 1980 ~ 2009." *Asian Review of Financial Research* 24, 167–229.

Kim, D. and S. Kon. 1996. "Sequential parameter nonstationarity in stock market returns." *Review of Quantitative Finance and Accounting* 6, 103–131.

Kim, D. and S. Kon. 1999. "Structural change and time dependence in models of stock returns." *Journal of Empirical Finance* 6, 283–308.

Ledoit, O. and M. Wolf. 2003. "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection." *Journal of Empirical Finance* 10, 603–621.

Merton, R. C. 1973. "An intertemporal capital asset pricing model." *Econometrica* 41, 867–887.

Lintner, J. 1965. "The valuation of risky assets and the selection of risky investment in stock portfolios and capital budgets." *Review of Economics and Statistics* 47, 13–37.

Newey, W. K. and K. D. West. 1987. "Hypothesis testing with efficient method of moments estimation." *International Economic Review* 28, 777–787.

Ren, Y. and K. Shimotsu. 2009. "Improvement in finite sample properties of the Hansen-Jagannathan distance test." *Journal of Empirical Finance* 16, 483–506.

Ross, S. A. 1976. "The arbitrage theory of capital asset pricing." *Journal of Economic Theory* 13, 341–360.

Shanken, J. 1992. "On estimation of beta pricing models." *Review of Financial Studies* 5, 1–33.

Sharpe, W. 1964. "Capital asset prices: a theory of market equilibrium under conditions of risk." *Journal of Finance* 19, 425–442.

Windmeijer, F. 2005. "A finite sample correction for the variance of linear efficient two-step GMM estimators." *Journal of Econometrics* 126, 25–51.