Cross Sectional Asset Pricing Tests

Ravi Jagannathan, Ernst Schaumburg and Guofu Zhou[1]

December 29, 2009

[1]Ravi Jagannathan, Chicago Mercantile Exchange/John F. Sandner Distinguished Professor of Finance, Kellogg School of Management, Northwestern University and Research Associate, NBER; Ernst Schaumburg, Federal Reserve Bank of New York; Guofu Zhou, Professor of Foinance, Olin School of Business, Washington University St. Louis.
The views expressed in this paper are those of the authors and do not represent the views of the Federal Reserve Bank of New York or the Federal Reserve system.

1 Introduction

A major part of the research effort in finance is directed toward understanding why different financial assets earn vastly different returns on average. For example, as can be seen from Table 1, during the 82 year period 1927-2008 the U.S. stock market as a whole earned an average inflation adjusted continuously compounded real return of 5.86%, whereas U.S. Treasury bills earned only 0.72%.[1] Small value stocks earned an impressive 10.63%. In comparison, small growth stocks earned only 4.90% even though the two types of stocks had similar return volatilities and small value stocks had a substantially higher dividend growth rate. To the extent investors' beliefs are rational, realized average returns measured over a sufficiently long horizon would be close to what investors expected to get from holding those assets. The aim of asset pricing models is to explain why even rational investors would be content to hold assets that earned such vastly different returns on average.

	Returns		$\log(P_t/D_t)$		Div. growth rate		β_{Mkt}
	Mean (%)	σ (%)	Mean	σ	Mean (%)	σ (%)	
Risk free rate	0.72	4.0					
Market	5.86	20.4	3.345	0.446	1.45	11.0	1.00
Small growth	4.90	29.9	3.972	1.033	0.11	23.8	1.28
Small value	10.63	28.7	3.621	0.717	7.40	41.3	1.33

Table 1: Selected average financial asset returns, 1927-2008.[2]

According to finance theory, assets that have riskier payoffs should earn higher returns on average to compensate investors for bearing that increased risk. Not all risk matters though. When markets are perfect and frictionless, investors need only be compensated for bearing systematic risks that can not be diversified away. When there are transactions costs, other risks may matter. A variety of rational asset pricing models have been proposed in the literature, and each model takes a different stand on how systematic risks and other risks that matter should be measured. There are strengths and weaknesses associated with each model – while a given model may work better with certain asset classes, and worse in other asset classes. Differences on how systematic risk is measured come from differences in the assumptions regarding preferences of agents in the economy, their endowments, production technology, stochastic process governing the evolution of variables taken exogenous to the model, and the nature of frictions in markets for financial and real assets.

[1] Welch (2000) surveys 226 financial economists and finds a consensus expected market risk premium around 7%

[2] Sources: Risk free rate from the CRSP database; nominal returns on various stock portfolios are from Professor Kenneth French's data library; nominal returns converted to real using the CPI.

In spite of their differences, rational asset pricing models share some important commonalities: The law of one price, i.e., two securities that have the same future payoffs must have the same price; and linear pricing, i.e., the price of a portfolio of securities should equal the sum of the prices of the securities that make up that portfolio. In addition, most models impose the restriction that financial markets are in equilibrium, i.e., the first order condition for investors' (constrained) optimization problem are satisfied and security markets clear. Finally, some asset pricing models also explicitly impose a no arbitrage restriction, i.e., positive payoffs must have positive prices.[3]

The Sharpe (1964) and Lintner (1965) Capital Asset Pricing Model (CAPM), although the earliest, is still the most widely used among practitioners and in classrooms – Welch (2008) finds that about 75% of finance professors recommend using the CAPM. According to the CAPM the expected return on an asset is given by:

$$E[R_{it}] - R_{ft} = \beta_i E[R_{mt} - R_{ft}] \tag{1}$$

where R_{mt}, R_{ft}, and R_{it} are, respectively, the gross returns on the market portfolio of all assets in the economy, the risk free asset, and an arbitrary asset; $\beta_i = Cov(R_{it}, R_{mt})/Var(R_{mt})$; and $E[\cdot]$ is the expectation operator.

The common practice in empirical studies of the CAPM is to use the return on an index of all exchange traded stocks as the proxy for the unobserved return on the market portfolio of all assets in the economy. Roll (1977) pointed out that the CAPM equation (1) will continue to hold when the return on the market portfolio is replaced by the return on any other asset that is on the mean-variance frontier of all asset returns; and therefore tests of the CAPM in effect are tests of the mean variance efficiency of the stock market portfolio.

The CAPM was derived in a static one period economy. Merton (1973) showed that in a dynamic economy, the following intertemporal version of the CAPM like relation (ICAPM) will hold.

$$E_{t-1}[R_{it}] - R_{ft} = \beta_{im,t-1} E_{t-1}[R_{mt} - R_{ft}] + \sum_{s=1}^{S} \beta_{is,t-1} E_{t-1}[R_{st} - R_{ft}] \tag{2}$$

where $E_{t-1}[\cdot]$ denotes the conditional expectation operator based on the information set at date $t-1$; R_{mt} denotes the time t return on the market portfolio as before; and R_{st}, $s = 1, 2 \ldots S$, denote returns on the S basis assets that help hedge against changes in future investment opportunities; and

[3] Due to space constraints we do not cover the rich literature on behavioral finance pioneered by Benartzi and Thaler (1995), Barberis, Shleifer, and Vishny (1998) and Daniel, Hirshleifer, and Subrahmanyam (1998), among others, building on the seminal work of Kahneman and Tversky (1979).

$\beta_{m,t-1}, \beta_{s,t-1}, s = 1 \ldots S$, are the slopes in the conditional multiple regression of R_{it} on $R_{mt}, R_{st}, s = 1 \ldots S$. Ross (1976) showed that the unconditional version of the relation given in (2) will hold when returns have a linear factor structure and there are no arbitrage opportunities in the economy. Connor (1984) derived the results in Ross (1976) using equilibrium arguments. The algebra of the mean-variance frontier implies that a particular portfolio of the market and the S basis assets in the case of Merton (1973) and the returns on the primitive factor portfolios in the case of Ross (1976) will be mean-variance efficient. It is therefore not surprising that the mean-variance frontier of asset returns has received wide attention in the finance literature.

Hansen and Richard (1987) show that most asset pricing models, including the CAPM and the ICAPM, have the following conditional and unconditional stochastic discount factor (SDF) representations:

$$E_{t-1}[m_t R_{it}] = 1 \tag{3}$$

$$E[m_t R_{it}] = 1 \tag{4}$$

where R_{it} denotes the time t gross return on financial asset i as before, m_t denotes the stochastic discount factor that helps assign a price at time $t-1$ for payoffs that occur at time t, and E_{t-1} denotes the expectation operator based on information available at date $t-1$. An asset pricing model specifies what m_t is as a function of observable variables and unobserved parameters, and the collection of assets for which (3) and (4) hold. Taking the unconditional expectation of both sides of (3) gives the unconditional SDF representation in (4) that is frequently used in the empirical literature. Hansen and Jagannathan (1991) show that there is a one to one map between the mean-variance frontier of all candidate SDFs that satisfy (4) and the mean variance frontier of asset returns. Therefore, it should not come as a surprise that, once again, the mean-variance frontier of asset returns plays an important role in empirical evaluation of asset pricing models.

Hansen and Richard (1987) also showed that the SDF representation in (3) can be rearranged to get the following conditional and unconditional beta representations:

$$E_{t-1}[R_{it}] - R_{ft} = \beta_{i,m,t-1} E_{t-1}[R_t^* - R_{ft}] \tag{5}$$

$$E[R_{it}] - R_{ft} = \beta_{i,m} E[R_t^* - R_{ft}] \tag{6}$$

where R_t^* is the benchmark return on an asset that pays the SDF m_t, or the return on the portfolio whose payoff best approximates the SDF in the least squares sense when the SDF m_t is not in the linear span of the traded assets, and

$$\beta_{i,m,t-1} = Cov_{t-1}(R_{it}, R_t^*)/Var_{t-1}(R_t^*). \tag{7}$$

Further, R_t^* has the property that it is the global minimum second moment return. Hence a CAPM like relation holds for every asset pricing model, with the benchmark return replacing the market portfolio return. In an asset pricing model where the benchmark return itself is a portfolio of several other separating portfolios, the relations in equations (5) and (6) will give rise to corresponding multibeta representations of the asset pricing model.

The multi-beta representation in (2) has an intuitive appeal, especially to practitioners and risk managers, in understanding the magnitude of exposure of a security to different types of systematic risks. In view of this, we focus more on econometric methods for the beta representation of asset pricing models.

The rest of the paper is organized as follows. In section 2 we discuss various statistical methods for empirical evaluation of asset pricing models using the cross section of security returns. In section 3 we survey the empirical support for various asset pricing models. We summarize and conclude in section 4.

2 Econometric Methods

Any asset pricing model must, at least implicitly, specify a stochastic discount factor which links the cross sectional variation in observed excess returns to the conditional covariance of the underlying return with the stochastic discount factor. The stochastic discount factor representation provides a rich set of conditional moment condition based on which estimation and asset pricing tests in principle can be carried out. In order to implement (3) empirically, however, the econometrician must specify how investors form expectations as well as specifying the evolution of their information set leading to a joint inference problem. In practice, it is therefore the unconditional implications of (3) that are most often used for estimation and inference.

The stochastic discount factor model is the most general representation of pricing, requiring very few distributional assumptions beyond existence of moments and the convergence of sample averages to their population analogues. Moreover the full machinery of the Hansen (1982) GMM can be brought to bear on the estimation and inference problems. The majority of the empirical methods described in the sequel can be seen as special cases for specific choices of stochastic discount factors and/or additional distributional assumptions. By far the most important case is that of linear beta pricing in which the stochastic discount factor is assumed to be an affine function of a set of observed factors, $m_t(\delta) = \delta_0 - \delta_1' f_t$. To be consistent with the existing literature, we let δ denote the model parameters in the stochastic discount factor formulation and (β, γ) denote the factor loadings and factor risk premia in the corresponding linear beta pricing formulation, but note that a mapping between the two sets of parameters exist (conditional on the first two moments of the factors). The asset pricing tests specifically test the hypothesis that pricing errors of the chosen set of test assets is small, in a suitable metric, either in absolute terms or relative to a reference asset (e.g. the risk-free or the "market" portfolio).

2.1 The Stochastic Discount Factor Approach

The conditional moment conditions (3) can be analyzed using the Hansen (1982) GMM framework. In particular, any p dimensional vector z_t which is in the investors' information set at time $t-1$ can be used as an instrument thus transforming the conditional moment condition into a family of unconditional moment conditions indexed by the set of instruments z_t chosen by the econometrician:[4]

$$E[(R_t m_t(\delta) - 1_N) \otimes z_t'] = 0 \tag{8}$$

where R_t is the vector of gross returns on N test assets and δ is the vector of m model parameters. This of course includes (4) as a special case when $z_t \equiv 1$. Defining the pricing errors $u(R_t, \delta) \equiv (R_t m_t(\delta) - 1_N)$ we have the $l \equiv Np$ sample moments

$$g_T(\delta) = \frac{1}{T}\sum_{t=1}^{T} f(R_t, z_t, \delta) \ , \text{ where } \ f(R_t, z_t, \delta) \equiv u(R_t, \delta) \otimes z_t \tag{9}$$

Under the null that the model is correctly specified, the sample moments $g_T(\delta)$ should converge to zero when evaluated at the true parameters, δ_0 . But in general, the number of moments, l, is strictly greater than the number of parameters, m, and it will be impossible to solve the system of equations $g_T(\delta) = 0$. Instead, the GMM estimator is defined as the solution to a quadratic minimization problem:

$$\hat{\delta} = \arg\min_{\delta} g_T(\delta)' W_T g_T(\delta) \tag{10}$$

[4]In the hypothetical scenario where the econometrician's instruments span the investors information set, the conditional and unconditional moment conditions are equivalent. Otherwise the set of unconditional moments are strictly weaker.

where W_T is a positive definite weighting matrix. Hansen (1982) shows that the optimal choice of W_T is $W_T = \hat{S}_T^{-1}$, where $\hat{S}_T$ is a consistent estimator of $S = \text{ACov}\left(\sqrt{T} g_T(\delta_0)\right)$. With this choice of weighting matrix, the associated GMM estimator, $\tilde{\delta}$, is optimal, that is, its asymptotic covariance matrix, $(D' S_0^{-1} D)^{-1}$, is the smallest possible, where

$$D = E\left[\frac{\partial f}{\partial \delta}(R_t, z_t, \delta)\right] = E\left[\frac{\partial h}{\partial \delta}(R_t, \delta) \otimes z_t\right],$$

which can be estimated by its sample analogue. Since the optimal weighting matrix depends on δ, the GMM implementation is usually carried out in two steps. First, a consistent GMM estimator, $\hat{\delta}$, is obtained by choosing an arbitrary positive definite W_T (often the identity matrix) . Second, based on $\hat{\delta}$, the optimal weighting matrix can be obtained in many ways. Newey and West (1987) provide a simple estimator of S_0,

$$\hat{S}_T = \Omega_0 + \sum_{j=1}^{c_T} w(j,T)\left(\Omega_j + \Omega_j'\right), \quad w(j,T) = 1 - \frac{j}{c_T + 1},$$

where

$$\Omega_j = \sum_{t=j+1}^{T} f(R_t, z_t, \hat{\delta}) f(R_{t-j}, z_{t-j}, \hat{\delta})',$$

and c_T is a sequence of positive integers with $c_T \to +\infty$ and $c_T/T^{1/3} \to 0$ as the sample size T goes to infinity. With the optimal weighting matrix $W_T = \hat{S}_T^{-1}$, the optimal GMM estimator can then be obtained. Hansen, Heaton, and Yaron (1996) investigate the finite sample properties of the optimal GMM estimator as well as the estimator obtained by further iterating between estimation of δ and S_0.

When an asset pricing model is misspecified, it is of interest to obtain a measure of the degree of mis-specification. This is possible when the number of moments exceeds the number of parameters, $m > l$, so that there are $(m - l)$ overidentifying restrictions which can be exploited for model mis-specification tests. Hansen (1982) proposes a one-sided model specification test with an asymptotic chi-squared distribution,

$$T J_T(\tilde{\delta}) = T g_T(\tilde{\delta})'(\hat{S}_T^{-1}) g_T(\tilde{\delta}) \overset{\text{asy}}{\sim} \chi^2_{m-l} \tag{11}$$

which rejects the null that the model is correctly specified if J_T is too large. The first paper to test asset pricing model restrictions using this technique was Hansen and Singleton (1982).

While statistically optimal, the pricing error metric implied by the optimal weighting matrix has no economic interpretation. Moreover, since the optimal weighting matrix is the inverse spectral density matrix (at zero) of the moment conditions, it tends to reward more noisy pricing kernels and the resulting value of the metric cannot be compared across models to assess relative goodness of fit. In response to these concerns, Hansen and Jagannathan (1997) provide an alternative model misspecification test, known as the Hansen and Jagannathan distance or HJ-distance. Hansen and Jagannathan (1997) suggest using the matrix of second moments of returns on the test assets as the weighting matrix. They show that this choice of metric corresponds to a natural measure of misspecification as the minimum distance between the stochastic discount factor of an asset pricing model and the set of stochastic discount factors that correctly price the test assets. Define y as the stochastic discount factor associated with an asset pricing model, and $\mathcal{M}$ as the set of stochastic discount factors that price all the assets correctly. The HJ-distance is defined as

$$\boldsymbol{\delta} = \min_{m \in \mathcal{M}} \|m - y\|, \tag{12}$$

where $\|X\| = \sqrt{E[X^2]}$ is the standard L^2 norm. The HJ-distance can also be interpreted as a measure of the maximum pricing error of a portfolio that has a unit second moment. Define ξ as the random payoff of a portfolio. It can then be shown that

$$\boldsymbol{\delta} = \max_{\|\xi\|=1} |\pi(\xi) - \pi^y(\xi)|, \tag{13}$$

where $\pi(\xi)$ and $\pi^y(\xi)$ are the prices of ξ assigned by the true and the proposed asset pricing model, respectively.[5] As Jagannathan and Wang (1996) point out, the model mis-specification test using the HJ-distance is no longer asymptotically chi-squared but follows a mixed chi-squared distribution whose critical values must be simulated.

2.2 Linear Beta Pricing

Most cross-sectional asset pricing studies going back to the earliest CAPM studies, consider the case of a linear beta pricing restriction:

$$H_0: \qquad E[R_t] = \gamma_0 1_N + \gamma_1 \beta_1 + \cdots + \gamma_K \beta_K, \tag{14}$$

where $E[R_t]$ is the vector of expected returns on N test assets, $\beta_1, \ldots, \beta_K$ are N-vectors of risk exposures to K factors in the economy, and $\gamma = (\gamma_1, \ldots, \gamma_K)'$ is a vector of K risk premia. Since the linear beta pricing approach requires the separate identification of risk exposures and risk premia,

[5] Notably, the definition of $\mathcal{M}$ does not impose positivity (i.e. no-arbitrage) and Hansen and Jagannathan (1997) define a second distance measure by restricting the admissible set of stochastic discount factors to be nonnegative.

an assumption on the data-generating process must be made. For this reason, any test of H_0 is a joint test of H_0 and the associated assumption on the data-generating process. The common data-generating process is a multi-factor regression model,

$$r_{it} = \alpha_i + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + \epsilon_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{15}$$

where

r_{it} = the excess return (return in excess of the T-bill rate) on asset i in period t $(1 \leq i \leq N)$,

f_{jt} = the realization of the j-th factor in period t $(1 \leq j \leq K)$,

ϵ_{it} = the disturbances or random errors,

and T is the number of time-series observations.

To see the connection with the stochastic discount factor representation, let $m_t(\delta) = \delta_0 - \delta_1' f_t$ and consider the pricing of the i^{th} asset:

$$E[m_t(\delta) R_{i,t}] = 1 \Rightarrow E[R_{i,t}] = \gamma_0 + \gamma' \beta_i$$

where

$$\gamma_0 = \frac{1}{\delta_0 - \delta_1' E[f_t]} \ , \ \gamma = \frac{\delta_1' \mathrm{Var}(f_t)}{\delta_0 - \delta_1' E[f_t]} \ , \ \beta_i = \mathrm{Var}(f_t)^{-1} \mathrm{Cov}(f_t, R_{i,t})$$

Thus there exists a close correspondence between the parametrization of the linear stochastic discount factor model and the multi-beta pricing model, and Jagannathan and Wang (2002) show that the two approaches are indeed asymptotically equivalent.[6]

Testing of the linear beta pricing hypothesis (14) is typically carried out using a one-step (time series) or two-step (cross sectional) approach. The former focuses on measuring the size of alphas and is applicable when factors are tradable assets and $T > N$. The latter focuses on gammas and is applicable as long as factors are observable, even when $N > T$.

2.3 Time Series Approach

The time series approach to asset pricing tests takes one of two routes to approximating the true finite sample distribution. In Section 2.3.1 and 2.3.2 we survey the early literature which derived exact finite sample distributions of the tests under independence and joint normality. In Section

[6] In the special case where the factor mean and variance are known ex-ante, Kan and Zhou (1999) show that the multi-beta pricing leads to more precise estimates than the GMM estimator that does not exploit this information. Jagannathan and Wang (2002) show that, once moment conditions corresponding to the first two moments of the factors are added in the GMM estimation, the two methods are again equivalent.

2.3.3 we lay out the GMM approach which dispenses with most of the distributional assumptions at the cost of relying instead on an asymptotic approximation to the true finite sample distribution. Each approach has merits and which approach is chosen in a given study depends on whether serial correlation, heteroskedasticity and non-normality are deemed important features of the data at hand.

2.3.1 Portfolio Factors with Riskfree Asset

When a riskfree asset is assumed be available and factors are excess returns on tradable portfolios, the factor risk premia can be separately identified as the factor means.

The asset pricing restrictions are

$$H_{01}: \quad \alpha_i = 0, \qquad i = 1, \ldots, N. \tag{16}$$

which, in the single factor case with the factor equal the excess return on a market index, is the well known CAPM restriction. Standard univariate ordinary least squares (OLS) regressions of (15) provide estimates of both the alphas and betas for each test asset. Denote by $\hat{\alpha}$ the N-vector of the alphas, and by $\hat{\Sigma}$ the N by N matrix of the cross products of the estimated residuals divided by T. If H_{01} is true, $\hat{\alpha}$ should be close to zero. Under the popular normality assumption that the regression residuals are identically, and independently normally distributed, Gibbons, Ross, and Shanken (1989) provide the following test, known as Gibbons, Ross and Shanken (GRS) test,

$$\text{GRS} = \frac{(T-N-K)}{N} \frac{\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}}{1+\bar{F}'\hat{\Omega}^{-1}\bar{F}} \;\sim\; F_{N,T-N-K}(\eta), \tag{17}$$

where $\bar{F}$ and $\hat{\Omega}$ are the sample mean and sample covariance matrix of the excess returns on the K reference portfolios. The exact finite sample distribution, $F_{N,T-N-1}$, is the F distribution with degrees of freedom N and $T-N-1$ and noncentrality parameter $\eta = [T/(1+\bar{F}'\hat{\Omega}^{-1}\bar{F})]\alpha'\Sigma^{-1}\alpha$. Under the null, $\eta \equiv 0$, but in light of (17) the power of the test is easily computed under the alternative.

In the special case of a single factor model,$K = 1$, when the single factor is the excess return on a market index, r_p,

$$r_{it} = \alpha_i + \beta_i r_{pt} + \epsilon_{it}, \tag{18}$$

Gibbons, Ross, and Shanken (1989) offer an interesting economic interpretation of the GRS test,

$$\mathrm{GRS} = \frac{(T-N-1)}{N}\left(\left[\frac{\sqrt{1+\hat{\delta}_q^2}}{\sqrt{1+\hat{\delta}_p^2}}\right]^2 - 1\right), \tag{19}$$

where $\hat{\delta}_p$ is sample Sharpe ratio of r_p, and $\hat{\delta}_q$ is the sample Sharp ratio of the ex post efficient portfolio, that is, the frontier portfolio consisting of all the assets and r_p one would choose given the ex post sample mean and covariance matrix as the true parameters. Hence, the GRS test captures the relative deviations of the given portfolio r_p from the *ex post* efficient portfolio in terms of the Sharpe ratios. A higher GRS value implies r_p deviates more from ex post efficiency, and hence H_{01} will be rejected.

2.3.2 Portfolio Factors without Riskfree Asset

When the riskfree rate is unobserved or one is concerned about the difference between lending and borrowing rates, the riskfree asset assumption may be relaxed. In this case, the zero-beta CAPM replaces the CAPM. The earlier excess return regression (18), becomes

$$r_{it} = \alpha_i + \beta_i r_{pt} + \epsilon_{it}, \tag{20}$$

where all returns are the *raw* returns (i.e., without subtracting the T-bill rates). The zero-beta CAPM asset pricing restrictions are

$$E[r_{it}] = \gamma + \beta_i(E[r_{pt}] - \gamma), \tag{21}$$

where γ is the expected return on the 'zero-beta' portfolio and r_p is the return on *any* efficient portfolio (i.e. not necessarily the market portfolio). Comparing (20)-(21), the restrictions translate into the parametric restrictions,

$$H_{02}: \quad \alpha_i = \gamma(1-\beta_i), \qquad i = 1, \ldots, N \tag{22}$$

in the linear regression (20). Note that the 'zero-beta' rate γ is unknown and has to be estimated from the data, and it enters the constraints by multiplication with other parameters. Hence, the maximum likelihood estimator is not readily available from standard statistical or econometric theory.[7] However, conditional on the zero-beta rate, the solution is readily available as we saw in the previous section. Shanken (1985) therefore suggests a profile likelihood approach.

[7] Gibbons (1982) is the first to numerically solve the maximum likelihood estimator while Kandel (1984) provides an analytical solution to a restricted problem.

Consider the general case of K factors or reference portfolios, so that (20) extends to

$$R_t = \alpha + \beta F_t + \epsilon_t, \quad t = 1, \ldots, T, \tag{23}$$

where R_t is an N-vector of the asset returns, and F is a K-vector of the factors. Following Shanken (1985) we define $\alpha(\gamma) = \alpha - \gamma(1_N - \beta 1_K)$ and rewrite (23) as

$$(R_t - \gamma 1_N) = \alpha(\gamma) + \beta\,(F_t - \gamma 1_K) + \epsilon_t, \tag{24}$$

The zero-beta restriction is $\alpha(\gamma) = 0$, but conditional on γ, this is exactly the usual zero intercepts hypothesis, whose likelihood ratio test is readily available,

$$\mathrm{LRT}(\gamma) = T \log\left(1 + \frac{Q(\gamma)}{T}\right), \tag{25}$$

where

$$Q(\gamma) = \frac{T\widehat{\alpha(\gamma)}'\hat{\Sigma}^{-1}\widehat{\alpha(\gamma)}}{1 + (\bar{F} - \gamma 1_K)\hat{\Omega}^{-1}(\bar{F} - \gamma 1_K)}, \qquad \widehat{\alpha(\gamma)} = \hat{\alpha} - \gamma(1_N - \hat{\beta} 1_K),$$

where $\hat{\Omega}$ is, as before, the sample covariance matrix of the returns on the K reference portfolios. Now note that the maximum likelihood estimator of γ is obtained by maximizing the likelihood function under the zero-beta constraints, this is equivalent to minimize the LRT or $Q(\gamma)$. Since $Q(\gamma)$ is a function of only one parameter, the estimator is then obtained as

$$\tilde{\gamma} = \begin{cases} \tilde{\gamma}_1, & \text{if} \quad A > 0; \\ \tilde{\gamma}_2, & \text{otherwise.} \end{cases} \tag{26}$$

where $\tilde{\gamma}_1 \leq \tilde{\gamma}_2$ are the roots of the quadratic

$$H(\gamma) \equiv A\gamma^2 + B\gamma + C$$

with

$$\begin{aligned}
A &= (1_K'\hat{\Omega}^{-1}1_K)(Z'\hat{\Sigma}^{-1}\hat{\alpha}) - (1_K'\hat{\Omega}^{-1}\bar{F})(Z'\hat{\Sigma}^{-1}Z), \\
B &= (1 + \bar{F}'\hat{\Omega}^{-1}\bar{F})(Z'\hat{\Sigma}^{-1}Z) - (1_K'\hat{\Omega}^{-1}1_K)(\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}), \\
C &= -(1 + \bar{F}'\hat{\Omega}^{-1}\bar{F})(Z'\hat{\Sigma}^{-1}\hat{\alpha}) + (1_K'\hat{\Omega}^{-1}\bar{F})(\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}),
\end{aligned}$$

and $Z = 1_N - \hat{\beta} 1_K$.[8]

[8] When $K = 1$, Zhou (1991) re-formulate the problem via an eigenvalue approach that makes it feasible to derive the exact distribution of the likelihood ratio test. The exact distribution is unfortunately complex. Unlike the GRS test, it depends on a nuisance parameter even under the null. Zhou (1995), and Velu and Zhou (1999) analyze the case for $K > 1$ and provide simple distributional results that are useful for applications.

2.3.3 Generalized Method of Moments

The finite sample distribution of the tests in the previous sections were derived under the assumption of normality of asset returns. This assumption is, however, overwhelmingly rejected in the data (see, e.g., Zhou (1994), who also provides small sample tests under elliptical distributions). Using the framework of the Hansen (1982) GMM one can allow for general stationary and ergodic distributions of asset returns at the cost of only achieving asymptotic distribution results.

MacKinlay and Richardson (1991) provide two GMM tests of the CAPM, and Harvey and Zhou (1993) propose another such test.[9]

The factor regression model (18) implies the $2N$ moment conditions

$$E(\epsilon_{it}) = 0 \text{ and } E(\epsilon_{it} r_{pt}) = 0, \quad i = 1, \ldots, N.$$

To apply the GMM, let

$$g_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} f_t(\theta) \ , \ \text{ where } \ f_t(\theta) = \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{1t} r_{pt} \\ \vdots \\ \epsilon_{Nt} \\ \epsilon_{Nt} r_{pt} \end{pmatrix}$$

and $\theta = (\alpha_1, \beta_1, \ldots, \alpha_N, \beta_N)$. Note that there are exactly $2N$ moment conditions for the $2N$ parameters, and so the GMM estimator is independent of the weighting matrix, which is in fact the same as the OLS regression estimator. But the GMM framework provides a new asymptotic covariance matrix for $\hat{\alpha}$ that accounts for non-normality, in particular for conditional heteroskedasticity of the data. Let $\hat{\Sigma}_\alpha$ be a consistent estimator of this covariance matrix. Then the hypothesis can be tested by using the Wald test,

$$J_1 = T\hat{\alpha}'\hat{\Sigma}_\alpha^{-1}\hat{\alpha} \overset{\text{asy}}{\sim} \chi^2_N, \tag{27}$$

which is an extension of the Gibbons, Ross, and Shanken (1989) test in the nonnormality case.

The second GMM test of the CAPM is to use Hansen's model misspecification test. Imposing the null that all the alphas are zero, then the restricted moment function $g_{rT}(\beta) = g_T(0, \beta)$ has

[9] While Hansen and Singleton (1982) is the first to apply the GMM to a consumption-based asset pricing model, Harvey (1998) `[<-Guofu, Please add to bib file]` appears the first in the finance literature to apply the GMM to a version of the conditional CAPM. Since these tests are straightforward to extend to $K > 1$ factors, we consider only the one-factor case.

only N free parameters which enter into the objective function quadratically. The associated GMM estimator of β, $\tilde{\beta}$, can be solved explicitly, and so the test,

$$J_2 = T g_{rT}(\tilde{\beta})' W_T g_{rT}(\tilde{\beta}) \overset{\text{asy}}{\sim} \chi^2_N, \tag{28}$$

is easy to implement. The third GMM test of the CAPM simply compares the restricted and unrestricted model fit,

$$J_3 = T[g_{rT}(\tilde{\beta})' W_T g_{rT}(\tilde{\beta}) - g_T(\hat{\beta})' W_T g_T(\hat{\beta})] \overset{\text{asy}}{\sim} \chi^2_N, \tag{29}$$

where W_T is the weighting matrix in the unrestricted model.

Rather than using the optimal weighting matrix for estimating parameters, Zhou (1994) suggests the use of a patterned matrix of the following form,

$$W_T = W_1 \otimes W_2, \qquad W_1 : N \times N, \quad W_2 : L \times L, \tag{30}$$

where N is the number of assets and L is the number of instruments. Although this will not yield asymptotically the most efficient estimator, it often helps to solve the GMM estimation problem analytically or semi-analytically because this special weighting matrix makes it possible to apply the estimation results in the iid case. As demonstrated by Zhou (1994), this can be important in a number of asset pricing models where a direct numerical solution to the GMM estimation problem is hardly feasible. Based on the special W_T and the associate GMM estimator, Zhou (1994) provides an alternative GMM test,

$$H_z = T(M_T g_T)' V_T (M_T g_T) \overset{\text{asy}}{\sim} \chi^2_{NL-q} \tag{31}$$

where q is the number of parameters, V_T is a diagonal matrix, $V_T = \text{Diag}(1/v_1, \ldots, 1/v_q, 0, \ldots 0)$, formed by $v_1 > \ldots > v_q > 0$, the positive eigenvalues of the following $NL \times NL$ semi-definite matrix:

$$\Omega_T = \left[I - D_T(D_T' W_T D_T)^{-1} D_T' W_T\right] S_T \left[I - D_T(D_T' W_T D_T)^{-1} D_T' W_T\right]', \tag{4.2}$$

where M_T is an $NL \times NL$ matrix, of which the i-th row is the standardized eigenvector corresponding to the i-th largest eigenvalue for $i = 1, \ldots, NL$, D_T is a consistent estimator of D, and S_T is, as before, a consistent estimator of the underlying covariance structure of the model residuals. Cochrane (2001) provides an equivalent GMM test.

Applying the above methods, Velu and Zhou (1999) develop GMM tests for testing (1) when the factors are portfolio returns and when there is no riskless asset. Ghysels (1998) examines structure changes in the GMM framework and its asset pricing implications, while Hall (2005) summarizes various aspect of the GMM.

2.4 Cross-section Approach

The cross sectional two-step approach to testing asset pricing models was first developed by Fama and MacBeth (1973) and Black, Jensen, and Scholes (1972), and the methodology has since been adopted in many other areas of finance. In this section we discuss the classic two step approach as well as some subsequent extensions.

2.4.1 The Fama-MacBeth Two-pass OLS

In the first pass, estimates of the betas are obtained by applying OLS to Eq. (15) for each asset. Let $\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_K)$ be the resulting $N \times K$ matrix of OLS beta estimates. In the second pass, for each period t, one then runs a cross-sectional regression of $R_t = (R_{1t}, \ldots, R_{Nt})'$ on $\hat{X} = [1_N, \hat{\beta}]$ to get an estimator of $\Gamma = (\gamma_0, \gamma_1, \ldots, \gamma_K)' = (\gamma_0, \gamma_a')'$,

$$\hat{\Gamma}_t = (\hat{X}'\hat{X})^{-1}\hat{X}'R_t. \tag{32}$$

The average,

$$\hat{\Gamma}^{\mathrm{OLS}} = \sum_{t=1}^{T} \hat{\Gamma}_t / T = (\hat{X}'\hat{X})^{-1}\hat{X}'\bar{R}, \tag{33}$$

is taken as the final estimator of Γ, where $\bar{R}$ is the N-vector of sample means of the asset returns.[10].

The standard OLS t-ratios are often used for testing whether the j^{th} factor is priced,

$$\hat{t}_j = \frac{\hat{\gamma}_j^{\mathrm{OLS}}}{\hat{s}_j/\sqrt{T}}, \qquad j = 1, \ldots, K, \tag{34}$$

where $\hat{\gamma}_j^{\mathrm{OLS}}$ and $\hat{s}_j$ are the sample mean and standard deviation of the j-th component of the time series $\hat{\Gamma}_t, t = 1, \ldots, T$. The p-values are usually computed from a t distribution with degrees of freedom $T - 1$ or from a standard normal distribution. There are, however, a number of problems associated with the common practice of the Fama-MacBeth two-pass procedure. For example, the procedure ignores estimation error in the first pass betas, does not examine the overall pricing errors, and assumes implicitly that the pricing restrictions are true in computing the t-ratios. These issues will be examined in the following subsections.

[10]Some authors advocate using first pass betas estimated from a rolling window of past data, but the econometric analysis more cumbersome.

2.4.2 Alternative Two-Step Estimators and Their Asymptotic Distributions

In the second pass, Shanken (1985), among others, propose the use of the following GLS estimator:

$$\hat{\Gamma}^{\mathrm{GLS}} = (\hat{X}'\hat{\Sigma}^{-1}\hat{X})^{-1}\hat{X}'\hat{\Sigma}^{-1}\bar{R}, \tag{35}$$

where $\hat{\Sigma}$ is the estimator of the residual covariance matrix computed as the cross product of the fitted factor model residuals divided by T, with $T > (N + K)$ so that $\hat{\Sigma}$ is invertible.[11] A related alternative estimator is a weighted least squares,

$$\hat{\Gamma}^{\mathrm{WLS}} = (\hat{X}'\widehat{\Sigma}_d^{-1}\hat{X})^{-1}\hat{X}'\widehat{\Sigma}_d^{-1}\bar{R}, \tag{36}$$

where the weighting matrix $\widehat{\Sigma}_d$ consists of the diagonal elements of $\widehat{\Sigma}$. Litzenberger and Ramaswamy (1979) are perhaps the first to use this type of estimator in testing the CAPM. This estimator can be used without the assumption of $T > (N + K)$. All of the three two-pass estimators can be written in the following form:

$$\hat{\Gamma}_w = \hat{A}_w\bar{R}, \quad \hat{A}_w = (\hat{X}'\hat{W}\hat{X})^{-1}\hat{X}'\hat{W}, \tag{37}$$

where $\hat{W}$ is a symmetric weighting matrix. For example, the OLS estimator is obtained with $\hat{W}$ equal to the identity matrix.

Under an iid assumption, Shanken (1992) provides the asymptotic covariance matrix for $\hat{\Gamma}_w$,

$$\Upsilon_w = \mathrm{ACov}(\hat{\Gamma}_w) = (1 + c)\Omega_w + \Sigma_f^*, \tag{38}$$

where $c = \gamma_a'\Sigma_f^{-1}\gamma_a$, γ_a is Γ excluding the first element, $\Omega_w = A_w\Sigma A_w'$, A_w is the probability limit of $\hat{A}_w$, and

$$\Sigma_f^* = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_f \end{bmatrix}, \tag{39}$$

with Σ_f the population covariance matrix of the K factors. Asymptotically standard normal "t-ratios" are then obtained by dividing the estimates by their asymptotic standard errors. It follows that the Fama-MacBeth standard errors, computed as the time-series standard errors of the estimated gammas, understate the true asymptotic standard errors by the amount $c\Omega_w$ which can be important in certain applications. Moreover, the GLS estimator is asymptotically efficient.

[11] Kandel and Stambaugh (1995) show that the OLS estimator can be anything if one repackages the assets and if the asset pricing model is not exactly true. However, the GLS estimator will be robust to such repackaging.

Dispensing with the iid assumption, Jagannathan and Wang (1998) extend the asymptotic results about $\hat{\Gamma}_w$ and, under very general distributional assumptions, show that the asymptotic covariance matrix of $\hat{\Gamma}_w$ is given by

$$\Upsilon_w = D\Psi D' + D\Pi D' - D\left(\Gamma + \Gamma'\right)D', \tag{40}$$

where $D = (X'WX)^{-1}X'W$ with $X = [1_N \;\; \beta]$, and

$$\Psi = \sum_{k=-\infty}^{+\infty} E\left[h_t^1(h_{t+k}^1)'\right], \; \Gamma = \sum_{k=-\infty}^{+\infty} E\left[h_t^1(h_{t+k}^2)'\right] \text{ and } \Pi = \sum_{k=-\infty}^{+\infty} E\left[h_t^2(h_{t+k}^2)'\right]$$

with $h_t^1 = R_t - E[R_t]$, $h_t^2 = \left[(F_t - E[F_t])'\Sigma_f^{-1}\gamma_a\right]\epsilon_t$, the consistent estimators of which are easily obtained by their sample analogues.

Shanken and Zhou (2007) provide two GMM estimators. For simplicity, we review only the first one here. The model moment conditions are

$$E\left[\epsilon_t \otimes \begin{pmatrix} 1 \\ F_t \end{pmatrix}\right] = E\left[\begin{array}{c} R_t - \alpha - \beta F_t \\ (R_t - \alpha - \beta F_t) \otimes F_t \end{array}\right] = 0. \tag{41}$$

Let g_T be the sample moments:

$$g_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} \epsilon_t(\theta) \otimes z_t, \quad NL \times 1, \tag{42}$$

where $\theta = (\lambda', \beta_1', \ldots, \beta_K')'$ is the vector of parameters, $L = K + 1$ and $z_t = (1, F_t')'$. The GMM estimator requires the solution of

$$\min Q = g_T(\theta)'W_T g_T(\theta), \tag{43}$$

under the constraint (14), where W_T, $NL \times NL$, is a positive definite weighting matrix. With the choice of $W_T = \hat{S}_{\text{iid}}^{-1}$, where

$$\hat{S}_{\text{iid}} = \hat{\Sigma} \otimes \left(\frac{1}{T}\sum_{t=1}^{T} z_t z_t'\right). \tag{44}$$

the associated GMM estimator is in fact numerically identical to the maximum likelihood (ML) estimator of Shanken and Zhou (2007). The advantage of the GMM framework is that the asymptotic covariance matrix of the ML estimator is now given by the standard GMM theory as

$$\hat{\Sigma}_\theta = (D_T'W_T D_T)^{-1}D_T'W_T S_T W_T D_T(D_T'W_T D_T)^{-1}, \tag{45}$$

where D_T, $NL \times (NK + L)$, is the matrix of derivatives of $g_T(\theta)$ with respect to the parameters and S_T is a consistent estimator of the covariance matrix of the GMM moment conditions. This formula allows to obtain standard errors for the risk premium estimates for any weighting matrix W_T. Based on this, the associated "t-ratios" can then be computed and are asymptotically valid under conditional heteroskedasticity and/or serial correlation of the data.

2.4.3 Model Misspecification Tests

The Fama-MacBeth standard errors and t-ratios are only valid when the cross-sectional restrictions (14) hold exactly.[12] It is hence of interest to test whether these restrictions are true in any application.

Shanken (1985) and Shanken (1992) provide a simple specification test of (14) against a general alternative,

$$Q_c = T(\bar{R} - \hat{\gamma}_0^{\mathrm{GLS}} 1_N - \hat{\beta}\hat{\gamma}_a^{\mathrm{GLS}})' \widehat{\Sigma}^{-1} (\bar{R} - \hat{\gamma}_0^{\mathrm{GLS}} 1_N - \hat{\beta}\hat{\gamma}_a^{\mathrm{GLS}})/(1 + \hat{c}), \tag{46}$$

where $\hat{\Gamma}^{\mathrm{GLS}} = (\hat{\gamma}_0^{\mathrm{GLS}}, \hat{\gamma}_a^{\mathrm{GLS}\prime})'$ and $\hat{c}$ is a consistent estimator of c with the GLS weighting matrix. Statistically, $(T-N+1)Q_c/(T(N-K-1))$ is approximately F-distributed with degrees of freedom $N-K-1$ and $T-N+1$. Q_c summarizes the estimated pricing errors across assets weighted by $\widehat{\Sigma}^{-1}$, with an adjustment in the denominator for errors in the betas. The larger the pricing errors, the larger is the observed value for Q_c. Intuitively, if (14) is true, Q_c should not be "too far" from zero, as the errors will be random. On the other hand, if (14) does not hold then there will be systematic deviations as well, resulting in larger values of the observed test statistic. Roll (1985) provides an interesting geometric interpretation of Q_c when the factor is the return on a benchmark portfolio. An identical test statistic is obtained if the sample covariance matrix of asset returns is substituted for the residual covariance matrix in Q_c.

In the present multi-beta model, the stochastic discount factor y is a linear function of K common factors F, given by

$$y(\lambda) = \lambda_0 + F'\lambda_1 = x'\lambda, \tag{47}$$

where $x = [1,\ F']'$ and $\lambda = [\lambda_0,\ \lambda_1']'$. Given λ, the vector of pricing errors of the test assets is given by

$$g(\lambda) = q - E[px'\lambda] = q - D\lambda, \tag{48}$$

[12]Theoretically, Roll and Ross (1994), show the regression slope can be zero if the factor is not exactly on the efficient frontier and even if it is very close to it.

where $D = E[px']$ and it is assumed to be of full column rank. The squared HJ-distance then has an explicit expression

$$\delta^2 = \min_{\lambda} g(\lambda)' U^{-1} g(\lambda) = q'[U^{-1} - U^{-1} D(D'U^{-1}D)^{-1} D'U^{-1}]q, \tag{49}$$

where $U = E[pp']$ is assumed to be nonsingular.

In many empirical studies, p is chosen to be the gross return on N test assets, denoted as R_2. Let $Y = [f',\ R_2']'$ and define its mean and variance as

$$\mu = E[Y] \equiv \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \tag{50}$$

$$V = \mathrm{Var}[Y] \equiv \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}. \tag{51}$$

Using these notations, we can write $U = E[R_2 R_2'] = V_{22} + \mu_2 \mu_2'$ and $D = E[R_2 x'] = [\mu_2,\ V_{21} + \mu_2 \mu_1']$. Since the elements of R_2 are gross returns, we have $q = 1_N$ and it is easy to show that the λ that minimizes δ^2 is given by

$$\lambda^{HJ} = (D'U^{-1}D)^{-1}(D'U^{-1}1_N), \tag{52}$$

and hence the squared HJ-distance of (49) with $q = 1_N$ is

$$\delta^2 = 1_N'[U^{-1} - U^{-1}D(D'U^{-1}D)^{-1}D'U^{-1}]1_N. \tag{53}$$

Shanken and Zhou (2007) provide some additional tests. Note that all of the three two-pass estimators converge to the same limit when the null is true. However, when (14) is false, the estimators can differ even systematically. Shanken and Zhou (2007) show that the asymptotic distribution of the difference of two different two-pass estimators provides a simple chi-squared model specification test,

$$J_\gamma = T(\hat{\Gamma}^{OLS} - \hat{\Gamma}^{\mathrm{GLS}})'[(1 + \hat{c})\hat{\Pi}_a \hat{\Sigma} \hat{\Pi}_a']^{-1}(\hat{\Gamma}^{OLS} - \hat{\Gamma}^{\mathrm{GLS}}) \overset{\text{asy}}{\sim} \chi^2_{K+1}, \tag{54}$$

as $T \to \infty$, where $\hat{\Pi}_a = (\hat{X}'\hat{\Sigma}^{-1}\hat{X})^{-1}\hat{X}'\hat{\Sigma}^{-1} - (\hat{X}'\hat{X})^{-1}\hat{X}'$ is of rank $K+1$. In addition, one can use the likelihood ratio test and two GMM tests based aforementioned two GMM estimators. Shanken and Zhou (2007) provide a small sample comparison of these various tests via simulation.

2.4.4 Distribution Under Model Misspecification

Almost all studies using the two-pass procedures implicitly assumes that the asset pricing restriction (14) is true and then go ahead to obtain the standard errors of the risk premiums under the null.

What happens when this restriction is violated, as is likely to be the case in practice? It is hence of interest is to know the asymptotic distribution of the two-pass estimators under the alternative.

Whether the restriction is true or not, we can always project the expected return vector $E[R_t]$ onto $X = [1_N, \beta]$,

$$E[R_t] = X\Gamma_w + \eta_w, \tag{55}$$

where

$$\Gamma_w = (X'WX)^{-1}X'WE[R_t] \tag{56}$$

is the coefficient vector of the weighted projection and η_w is the projection residual or (true) pricing error vector.

Assuming normality, Shanken and Zhou (2007) show that

$$\sqrt{T}(\hat{\Gamma}_w - \Gamma_w) \overset{\text{asy}}{\sim} N(0, \Upsilon_w + \Upsilon_{w1} + \Upsilon'_{w1} + \Upsilon_{w2}) \tag{57}$$

where Υ_w is given in (38),

$$\Upsilon_{w1} = -(X'WX)^{-1}\begin{pmatrix} 0 \\ \Sigma_f^{-1}\gamma_{wa}\eta'_w W\Sigma \end{pmatrix} WX(X'WX)^{-1}, \tag{58}$$

which vanishes for the GLS estimator, and

$$\Upsilon_{w2} = (X'WX)^{-1}\left[(\eta'_w W\Sigma W\eta_w)(\Sigma_f^*)^{-1} + (\eta'_w \otimes X')V_w(\eta_w \otimes X)\right](X'WX)^{-1}, \tag{59}$$

where $(\Sigma_f^*)^{-1}$ denotes the following matrix

$$(\Sigma_f^*)^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_f^{-1} \end{bmatrix}. \tag{60}$$

In the GLS case, Υ_{w2} is given by

$$\Upsilon_{w2} = (\eta'_w\Sigma^{-1}\eta_w)\left[(X'\Sigma^{-1}X)^{-1}(\Sigma_f^*)^{-1}(X'\Sigma^{-1}X)^{-1} + (X'\Sigma^{-1}X)^{-1}\right]. \tag{61}$$

In the WLS case,

$$\Upsilon_{w2} = (X'\Sigma_d^{-1}X)^{-1}\left[(\eta'_w\Sigma_d^{-1}\Sigma\Sigma_d^{-1}\eta_w)(\Sigma_f^*)^{-1} + X'H_wX\right](X'\Sigma_d^{-1}X)^{-1}, \tag{62}$$

where Σ_d consists of the diagonal elements of Σ and H_w is an $N \times N$ matrix whose (i,j) element is given by $H_{ijw} = 2\sigma_{ij}^2\eta_{iw}\eta_{jw}/(\sigma_{ii}^2\sigma_{jj}^2)$, with η_{iw} the i-th element of η_w. With OLS, V_w is zero and so Υ_{w2} is given by the first term of (59) only. Joint normality is, therefore, not required in this case.

The asymptotic covariance matrix of the two-pass estimator has three new terms in addition to the usual Υ_w when the null hypothesis is violated. They relate to the product of the $\hat{A}_w$ matrix and the pricing error vector, η_w. Υ_{w2} is the covariance matrix of this product while Υ_{w1} is its covariance with the original disturbance terms that are present in the absence of misspecification. Although $\Upsilon_w + \Upsilon_{w1} + \Upsilon_{w1}' + \Upsilon_{w2}$ is obviously positive definite, it is not clear whether it is greater than Υ_w or not. Intuitively, the interaction between the pricing errors and the errors in $\hat{\beta}$ introduces additional "noise" that can reduce the precision of the risk premium estimates.

2.4.5 Traded Factors

Previous two-pass procedures treat the factors as non-traded. But in applications, it is often the case that some of the factors are traded. Suppose f_1 is a traded factor, the multi-beta pricing should imply that

$$\gamma_1 = E[f_{1t}]. \tag{63}$$

This says that we can simply estimate the factor risk premium by its sample mean, as is the case with the CAPM in which the market risk premium can be estimated by the sample mean of the market excess returns. In two-pass studies, the restrictions about the premiums on traded factors are usually ignored, or put differently, zero pricing error of factors is not impossed. As emphasized by Lewellen, Nagel, and Shanken (2009), models in studies that focus on the cross-sectional fit do not seem reasonable when they imply unrealistic risk premium estimates or zero-beta rates.

2.4.6 Latent Factors

Roll (1977) points that the true factors may be never observable, and this raises the issue of testability. Shanken (1987), and Kandel and Stambaugh (1987) consider the issue of factor proxies. The Ross (1976) arbitrage pricing theory implies multi-beta pricing restrictions even when the factors are latent or unobservable by an econometrician. Lehmann and Modest (1988) and Connor and Korajczyk (1988) are early tests of the APT based on the maximum likelihood method and asymptotical principal component analysis. Geweke (1996) provide a Bayesian treatment of the pricing errors. Based on recent advances in Bayesian Markov chain Monte Carlo (MCMC) methods, Nardari (2007) test the APT in a new class of linear factor models in which the factors are latent and the covariance matrix of excess returns follows a multivariate stochastic volatility process.

2.4.7 Characteristics

The beta pricing model (14) can be augmented to include firm characteristics,

$$E[r_{it}] = \gamma_0 + \lambda' Z^i + \gamma_a' \beta_i, \; i = 1, \ldots, N, \tag{64}$$

where λ is an M-dimensional constant vector representing the characteristics rewards, and β_i is a K vector of firm i's betas on the K factors. The use of firm characteristics can be viewed as either an extension of an imperfect asset pricing model or a device to detect misspecification errors. For example, Black, Jensen, and Scholes (1972) used residual variance and Banz (1981) used the firm size to examine the validity of the CAPM. Fama and French (1992) use the book-to-market ratio and provide evidence that this variable explains a larger fraction of the cross-sectional variation in expected returns. Other important characteristics that have received attention in the recent empirical literature are momentum and liquidity. Jagannathan and Wang (1998) provide a thorough econometric analysis of the Fama-MacBeth two-pass procedure and derive an asymptotic covariance matrix similar to (40).

3 Empirical Studies

We survey the main findings in the empirical asset pricing literature, starting with the early tests of the capital asset pricing model developed by Sharpe (1964), Lintner (1965) and Treynor (1965) building on the pioneering work by Markowitz (1959).

3.1 The CAPM and Linear Factor Models

The main testable implication of the Black (1972) CAPM is that the slope of the security market line should equal the market risk premium and the intercept should be a measure of the (shadow) risk-free rate. The main obstacle to testing this hypothesis is the noise associated with the estimated betas which leads to an errors-in-variables problem first identified by Miller and Scholes (1972). To alleviate this problem, Black, Jensen, and Scholes (1972) suggest forming portfolios by sorting on betas, a procedure which has since become standard practice in the literature: Estimate betas based on history, sort assets using historical betas, group them into portfolios for a selected number of years, and rebalance portfolios periodically. As long historical betas contain information about future betas this procedure will alleviate the measurement error problem while producing portfolios with sufficient dispersion in betas. Black, Jensen, and Scholes (1972) find support for the CAPM albeit the intercept term is too large and the slope term is less than the observed average excess return on the stock market. Fama and MacBeth (1973) using the novel two-step procedure described in Section 2.4.1, also find that the CAPM cannot be rejected.

The primary empirical challenges to the CAPM come from several well-documented anomalies: Managed portfolios constructed based on certain firm characteristics appear to earn very different returns on average from those predicted by the CAPM. Notable among the anomalies that challenge the validity of the CAPM are the findings that the average returns on stocks is related to firm size (Banz (1981)), earnings to price ratio (Basu (1985)), book-to-market value of equity (BM) (Rosenberg, Reid, and Lanstein (1985)), cash flow to price ratio, sales growth (Lakonishok, Shleifer, and Vishny (1994)), past returns (De Bondt and Thaler (1985) and Jegadeesh and Titman (1993)), and past earnings announcement surprise (Ball and Brown (1968)). These anomalies are not a isolated phenomena of US data nor of a particular choice of sample period. Numerous subsequent studies have confirmed the presence of similar patterns in different datasets and time periods, including those of international markets.[13]

Fama and French, in a series of papers, make a convincing case that CAPM fails to describe the cross-section of stock returns (Fama and French (1992, 1996, 1997, 1998, 2004 and 2006)). They apply as simple technique to convincingly refute the CAPM. First, sort stocks into portfolios based on a given stock characteristic or combination of characteristics that produces a spread in average returns and check whether the spread in returns can be explained by a spread in betas. Using firm size and book-to-market value as the characteristics, Fama and French (1992) show that the CAPM can be overwhelmingly rejected in the data. Fama and French (1993) conjecture that two additional risk factors, in addition to the stock market factor used in empirical implementations of the CAPM, are necessary to fully characterize economy-wide pervasive risk in stocks. The Fama and French (1993) three-factor model (FF3) has received wide attention and has become the standard model for computing risk-adjusted returns in the empirical finance literature, although it too has trouble explaining the return on certain types of stocks such as small value stocks. However, the interpretation of the two additional Fama and French factors as risk factors is still the subject of much debate. Jagannathan and Wang (1996) and Campbell (1996) argue that human capital is an important part of the total wealth portfolio and including it in the definition of the CAPM market

[13]Schwert (2003) argues that most of the anomalies are more apparent than real and often disappear after they have been noticed and publicized.

protfolio significantly improves the CAPM pricing performance.[14] A forceful argument is made by Ferson, Sarkissian, and Simin (1999) who argue that returns on portfolios constructed using stock attributes may appear to be useful risk factors, even when the attributes are completely unrelated to risk.

A significant strand of the literature argues that the empirical failure of factor pricing models is not due to mis-measurement or omission of risk factors but to market inefficiencies. Lakonishok, Shleifer, and Vishny (1994) argue that high book-to-market stocks earn a higher return because they are underpriced to start with, and not because they have higher exposure to systematic risk. Consistent with that point of view, Daniel and Titman (1997) find that firms' characteristics help explain the cross-section of returns. Piotroski (2000) provides evidence consistent with the presence of mispricing by showing that among high book-to-market stocks firms with better fundamentals outperform the rest. Mohanram (2005) reaches a similar conclusion for low book-to-market stocks. Ang, Hodrick, Xing, and Zhang (2009) find that idiosyncratic volatility is related to low returns even after controlling for systematic risk factors

A final source pricing error in asset pricing tests are market frictions. Brennan and Subrahmanyam (1996) and Brennan, Chordia, and Subrahmanyam (1998) find that security characteristics like liquidity measures and trading volume are important in explaining the cross section of stock returns. Similarly, Easley, O'Hara, and Hvidkjaer (2002) find that information risk (related to liquidity) is also important in explaining the cross section of stock returns.

3.2 Conditional CAPM

The CAPM of Sharpe (1964), Lintner (1965) and Black (1972) is cast is inherently static. The model predicts that the unconditional expected return on a security is a linear function of the market risk premium and the associated beta. Merton (1980) extends the CAPM to the following dynamic form:

$$E[R_{it} \,|\, I_t] = \gamma_{0t-1} + \gamma_{1t-1}\beta_{it-1}, \tag{65}$$

[14]The need for more than one factor was also emphasized by Connor and Korajczyk (1988), Connor and Korajczyk (1993), Lehmann and Modest (1988) and Lehmann and Modest (2005) who used statistical methods to extract additional factors from the cross section of stock returns.Bansal, Hsieh, and Viswanathan (1993) on the other hand find support for nonlinear factor models.

where I_t is the investors' information set at time t, γ_{0t-1} is the conditional expected return on a zero-beta portfolio, γ_{1t-1} is the conditional market risk premium, and

$$\beta_{it-1} = \text{cov}(R_{it}, R_t^m \mid I_{t-1})/\text{var}(R_t^m \mid I_{t-1}),$$

the conditional beta.

Jagannathan and Wang (1996) show that the conditional CAPM is related to an unconditional two factor model

$$E[R_{it}] = \tilde{\gamma}_0 + \tilde{\gamma}_1\beta_i + \tilde{\gamma}_2\beta i^{\gamma} \tag{66}$$

where β_i^{γ} captures the covariation of stock i's conditional beta with the market risk premium. Jagannathan and Wang (1996) argue that time variations in betas and their comovement with the market risk premium can explain part of the mispricing relative to the CAPM.

For the conditional CAPM model to have any bite, there must be significant covariation of betas and the market risk premium over time. Ferson and Harvey (1991) and Ferson and Korajczyk (1995) find that betas vary over the business cycle in a systematic stochastic way. Several empirical studies suggest that value stocks may be more risky when the risk premium going forward is high, thus offering at least a partial explanation to the value premium. For example, Petkova and Zhang (2005) find that value betas tend to covary positively, and growth betas tend to covary negatively with the expected market risk premium. Why are value stocks more risky when expected risk premium is high? Zhang (2005) provides one explanation: It is more costly for value firms to downsize their capital stocks since they are typically burdened with more unproductive capital. As a result, the value stocks' returns covary more with economic downturns when the expected risk premium is high.

More recently, a number of studies have provided empirical evidence against the conditional CAPM. Lewellen and Nagel (2006) use high frequency returns to argue that the variation in betas and the equity premium would have to be implausibly large for the conditional CAPM to explain the magnitude of the value premium. However, Chan, Hameed, and Lau (2003) demonstrate that price and return may be in part driven by factors unrelated to fundamental cash flow risk. Such factors, together with liquidity events, may contaminate the estimation of beta at higher frequencies (see Pastor and Stambaugh (2003)). Bali, Cakici, and Tang (2009) improve the cross-sectional performance of the conditional CAPM by using more efficient estimation techniques. In addition, a recent study by Kumar, Srescu, Boehme, and Danielsen (2008) demonstrate that once the estimation risk or parameter uncertainty associated with beta and risk premium are accounted for, the conditional CAPM will have significantly more explanatory power in the cross-section and may explain the value premium after all.

3.3 Real Options

Most firms have the option to turn down, undertake, or defer a new project, in addition to the option to modify or terminate an existing project. Therefore, a firm can be viewed as a collection of existing and future projects and complex options on those projects. McDonald and Siegel (1984) observe that a firm should optimally exercise these real options to maximize its total value. The resulting firm value will consist of both the NPVs of the projects and the value of associated real options which is determined by how those options are expected to be exercised by the firm. Dybvig and Ingersoll (1982) and Hansen and Richard (1987) show that, while the CAPM will assign the right expected returns to the primitive assets (projects), it will in general assign the wrong expected returns to options on those primitive assets. It is therefore plausible that the theory of real options may explain the value premium.

Berk, Green, and Naik (1999) build on that insight and present a model where the expected returns on all projects satisfy the CAPM but the expected returns on the firm's stocks do not. Gomes, Kogan, and Zhang (2003), Carlson, Fisher, and Giammarino (2004), Cooper (2006), and Zhang (2005) provide several additional insights by building on the Berk, Green, and Naik (1999) framework. Anderson and Garcia-Feijoo (2006) and Xing (2008) find that the BM effect disappears when one controls for proxies for firms' investment activities. Bernardo, Chowdhry, and Goyal (2007) highlight the importance of separating out the growth option from equity beta. Jagannathan and Wang (1996) argue that because of the nature of the real options vested with firms, the systematic risk of firms will vary depending on economic conditions, and the stock returns of such firms will exhibit option like behavior. An econometrician using standard time series methods may conclude that the CAPM does not hold for such firms, even when the returns on such firms satisfy the CAPM in a conditional sense.

3.4 Consumption Based Asset Pricing

Consumption based asset pricing is derived from the first order condition determining the representative households optimal consumption plan in general equilibrium. In absence of any frictions, the household Euler condition states that, along the optimal consumption path, for any asset i we have the relationship:

$$u'(c_t) = \beta E\left[u'(c_{t+1})\, R_{i,t+1}\, |I_t\right] \tag{67}$$

where β is the households subjective discount factor and I_t is its information set at time t. Comparing (3) and (67) we see that the Euler equation has the powerful implication that the stochastic discount factor should be a function of the expected increase in the marginal utility of consumption,

$$m_t \equiv \frac{\beta E[u'(c_{t+1})|I_t]}{u'(c_t)} \tag{68}$$

a result first applied to asset pricing by Breeden (1979) and in its full GMM implementation by Hansen and Singleton (1982). However, the low volatility of the observed aggregate consumption growth has lead to a number of empirical challenges for the naive consumption based model. To explain the observed market risk premium identified by Mehra and Prescott (1985), a very high curvature of the utility function is required implying that households are extremely risk averse which does not match well with other dimension of the data such as the low observed riskfree rate.

Several solutions have been proposed in the literature. Parker and Julliard (2005) and Yogo (2006) show that using a broader measure of consumption improves the pricing performance of the consumption based model while Jagannathan and Wang (2007) show that it measuring consumption from Q4 to Q4 is beneficial. Other authors have proposed more complicated utility functions as a possible solution. Dunn and Singleton (1986), Eichenbaum, Hansen, and Singleton (1988), Sundaresan (1989), Constantinides (1990) and Eichenbaum and Hansen (1990) investigate internal habit formation while Abel (1990) and Campbell and Cochrane (1999) consider external habit formation. More recently, Bansal and Yaron (2004) and Hansen, Heaton, and Li (2008) have shown that long run risk models, in which small but highly persistent innovations to consumption growth and the volatility of consumption growth can have large pricing implications, are able to resolve many of the shortcomings of the basic consumption based asset pricing model.

4 Summary

In this review we discussed econometric methods that are available for evaluating the extent to which a given asset pricing model is able to explain the cross section of returns on various financial assets. Given the vast nature of the literature, we had to necessarily restrict our discussion to a subset of the articles. We hope that our review will provide the necessary background for the readers to go over the other article in the literature on their own.

References

Abel, Andrew B., 1990, Asset Prices under Habit Formation and Catching Up with the Joneses, *American Economic Review* 80, 38 – 42.

Anderson, Christopher W., and Luis Garcia-Feijoo, 2006, Empirical evidence on capital investment, growth options, and security returns., *Journal of Finance* 61, 171 – 194.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2009, High idiosyncratic volatility and low returns: International and further u.s. evidence., *Journal of Financial Economics* 91, 1 – 23.

Bali, TG, N Cakici, and Y Tang, 2009, The conditional beta and the cross-section of expected returns, *Financial Management* 38, 103–137.

Ball, Ray, and Philip Brown, 1968, An empirical evaluation of accounting income numbers., *Journal of Accounting Research* 6, 159 – 178.

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

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Banz, Rolf W., 1981, The relationship between returns and market value of common stocks, *Journal of Financial Economics* 9, 3–18.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307 – 343.

Basu, S, 1985, The relationship between earnings yield, market value, and return on nyse common stocks: Further evidence, *Journal of Financial Economics* 12, 129–156.

Benartzi, Shlomo, and Richard H. Thaler, 1995, Myopic loss aversion and the equity premium puzzle, *The Quarterly Journal of Economics* 110, 73–92.

Berk, J. B., R. C. Green, and V. Naik, 1999, Optimal investment, growth optionsand security returns, *Journal of Finance* LIV, 1553–1607.

Bernardo, Antonio E., Bhagwan Chowdhry, and Amit Goyal, 2007, Growth options, beta, and the cost of capital., *Financial Management* 36, 5 – 17.

Black, Fisher, 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 45, 444–454.

———, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, in Michael C. Jensen, ed.: *Studies in the Theory of Capital Markets* . pp. 79–121 (Praeger, New York, NY).

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

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristicsand the cross section of expected returns, *Journal of Financial Economics* 49, 345–373.

Brennan, Michael J., and Avanidhar Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 441–464.

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

———, and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior., *Journal of Political Economy* 107, 205 – 251.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns., *Journal of Finance* 59, 2577 – 2603.

Chan, K, Allaudeen Hameed, and Sie Ting Lau, 2003, What if trading location is different from business location? evidence from the jardine group, *Journal of Finance* 58, 1221–1246.

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

Connor, Gregory, 1984, A unified beta pricing theory, *Journal of Economic Theory* 34, 13–31.

——— , and Robert A. Korajczyk, 1988, Risk and return in an equilibrium apt : Application of a new test methodology, *Journal of Financial Economics* 21, 255–289.

——— , 1993, A test for the number of factors in an approximate factor model, *Journal of Financial Economics* 48, 1263–1291.

Constantinides, George M., 1990, Habit formation: A resolution of the equity premium puzzle., *Journal of Political Economy* 98, 519 – 543.

Cooper, Ilan, 2006, Asset pricing implications of nonconvex adjustment costs and irreversibility of investment., *Journal of Finance* 61, 139 – 170.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *The Journal of Finance* 53, 1839–1885.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation of stock returns, *Journal of Finance* 52, 1–33.

De Bondt, Werner F. M., and Richard Thaler, 1985, Does the stock market overreact?., *Journal of Finance* 40, 793 – 805.

Dunn, K.B., and K.J. Singleton, 1986, Modeling the term structure of interest rates under non-separable utility and durability of goods, *Journal of Financial Economics* 17, 27–55.

Dybvig, Philip H., and Jonathan E. Ingersoll, 1982, Mean-variance theory in complete markets, *Journal of Business* 55, 233–251.

Easley, David, Maureen O'Hara, and Soeren Hvidkjaer, 2002, Is information risk a determinant of asset returns?, *Journal of Finance* 57, 2185–2221.

Eichenbaum, M.S., L.P. Hansen, and K.J. Singleton, 1988, A time series analysis of representative agent models of consumption and leisure choice under uncertainty, *The Quarterly Journal of Economics* 103, 51–78.

Eichenbaum, Martin, and Lars Peter Hansen, 1990, Estimating models with intertemporal substitution using aggregate time series data, *Journal of Business and Economic Statistics* 8, 53–69.

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

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

——— , 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

——— , 1997, Industry costs of equity., *Journal of Financial Economics* 43, 153 – 193.

——— , 1998, Value versus growth: The international evidence., *Journal of Finance* 53, 1975 – 1999.

——— , 2004, The capital asset pricing model: Theory and evidence., *Journal of Economic Perspectives* 18, 25 – 46.

——— , 2006, The value premium and the capm., *Journal of Finance* 61, 2163 – 2185.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, returnand equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Ferson, Wayne E., and Campbell Harvey, 1991, The variation of economic risk premiums, *Journal of Political Economy* 99, 385–415.

Ferson, Wayne E., and Robert A. Korajczyk, 1995, Do arbitrage pricing models explain the predictability of stock returns?, *Journal of Business* 68, 309–349.

Ferson, Wayne E., Sergei Sarkissian, and Timothy Simin, 1999, The alpha factor asset pricing model: A parable., *Journal of Financial Markets* 2, 49 – 68.

Geweke, Johnand Guofu Zhou, 1996, Measuring the pricing error of the arbitrage pricing theory, *Review of Financial Studies* 9, 557–587.

Ghysels, Eric, 1998, On stable factor structures in the pricing of risk: Do time-varying betas help or hurt?, *Journal of Finance* 53, 549–573.

Gibbons, Michael R., 1982, Multivariate tests of financial models: A new approach, *Journal of Financial Economics* 10, 3–27.

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

Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns., *Journal of Political Economy* 111, 693 – 732.

Hall, Alastair R., 2005, *Generalized method of moments* (Oxford University Press, New York, NY).

Hansen, Lars, John Heaton, and Nan Li, 2008, Consumption strikes back? measuring long run risk, *Journal of Political Economy* 116, 260–302.

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

Hansen, Lars Peter, John Heaton, and Amir Yaron, 1996, Finite-sample properties of some alternative gmm estimators, *Journal of Business and Economic Statistics* 14, 262–280.

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

——— , 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance* 62, 557–590.

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

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

Harvey, Campbell R., and Guofu Zhou, 1993, International asset pricing with alternative distributional specifications, *Journal of Empirical Finance* 1, 107–131.

Jagannathan, R., and Y. Wang, 2007, Lazy investors, discretionary consumption, and the cross-section of stock returns, *Journal of Finance* 62, 1623–1661.

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

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

——— , 2002, Empirical evaluation of asset-pricing models: A comparison of the sdf and beta methods, *Journal of Finance* 57, 2337–2367.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263–291.

Kan, Raymond, and Guofu Zhou, 1999, A critique of the stochastic discount factor methodology, *Journal of Finance* 54, 1021–1048.

Kandel, Shmuel, 1984, The likelihood ratio test statistic of mean-variance efficiency without a riskless asset, *Journal of Financial Economics* 13, 575–592.

——— , and Robert F. Stambaugh, 1987, On correlations and inferences about mean-variance efficiency, *Journal of Financial Economics* 18, 61–90.

——— , 1995, Portfolio inefficiency and the cross-section of expected returns, *Journal of Finance* 50, 157–184.

Kumar, P., S. Srescu, R. Boehme, and B. Danielsen, 2008, Estimation risk, information, and the conditional capm: Theory and evidence, *Review of Financial Studies* 21, 1037–1075.

Lakonishok, J., A. Shleifer, and R. Vishny, 1994, Contrarian investment, extrapolationand risk, *Journal of Finance* 49, 1541–1578.

Lehmann, Bruce N., and David M. Modest, 1988, The empirical foundations of the arbitrage pricing theory, *Journal of Financial Economics* 21, 213–254.

——— , 2005, Diversification and the optimal construction of basis portfolios, *Management Science* 51, 581–598.

Lewellen, Jonathan, and Stefan Nagel, 2006, The conditional capm does not explain asset-pricing anomalies, *Journal of Financial Economics* 82, 289–314.

——— , and Jay Shanken, 2009, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* In Press, Accepted Manuscript, –.

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

Litzenberger, Robert H., and Krishna Ramaswamy, 1979, The effect of personal taxes and dividends of capital asset prices: The theory and evidence, *Journal of Financial Economics* 7, 163–196.

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

Markowitz, Harry Max, 1959, *Portfolio selection; efficient diversification of investments* (New York, Wiley).

McDonald, Robert, and Daniel Siegel, 1984, Option pricing when the underlying asset earns a below-equilibrium rate of return: A note, *The Journal of Finance* 39, 261–265.

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

Merton, Robert C., 1973, An intertemporal asset pricing model, *Econometrica* 41, 867–888.

——— , 1980, On estimating the expected return on the market : An exploratory investigation, *Journal of Financial Economics* 8, 323 – 361.

Miller, Merton H., and Myron Scholes, 1972, Rate of return in relation to risk: A reexamination of some recent findings, in Michael C. Jensen, ed.: *Studies in the Theory of Capital Markets* . pp. 79–121 (Praeger, New York, NY).

Mohanram, Partha S., 2005, Separating winners from losers among low book-to-market stocks using financial statement analysis., *Review of Accounting Studies* 10, 133 – 170.

Nardari, Federicoand John T. Scruggs, 2007, Bayesian analysis of linear factor models with latent factors, multivariate stochastic volatilityand apt pricing restrictions, *Journal of Financial and Quantitative Analysis* 42, 957–992.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Parker, Jonathan, and Christian Julliard, 2005, Consumption risk and cross-sectional returns, *Journal of Political Economy* 113, 185–222.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns., *Journal of Political Economy* 111, 642 – 685.

Petkova, Ralitsa, and Lu Zhang, 2005, Is value riskier than growth?., *Journal of Financial Economics* 78, 187 – 202.

Piotroski, Joseph D., 2000, Value investing: The use of historical financial statement information to seperate winners from losers., *Journal of Accounting Research* 38, 1 – 41.

Roll, Richard, 1977, A critique of the asset pricing theory's tests; part i: On past and potential testability of the theory, *Journal of Financial Economics* 4, 129–176.

——— , 1985, A note on the geometry of shanken's csr t2 test for mean/variance efficiency, *Journal of Financial Economics* 14, 349 – 357.

——— , and Stephen A Ross, 1994, On the cross-sectional relation between expected returns and betas, *Journal of Finance* 49, 101–21.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9–17.

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

Schwert, G. William, 2003, *Anomalies and Market Efficiency.* . pp. 939 – 974 (Handbooks in Economics, vol. 21.: U Rochester).

Shanken, Jay, 1985, Multivariate tests of the zero-beta capm, *Journal of Financial Economics* 14, 327–348.

——— , 1987, Multivariate proxies and asset pricing relations : Living with the roll critique, *Journal of Financial Economics* 18, 91–110.

——— , 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1–33.

——— , and Guofu Zhou, 2007, Estimating and testing beta pricing models: alternative methods and their performance in simulations, *Journal of Financial Economics* 84, 40–86.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Sundaresan, SM, 1989, Intertemporally dependent preferences and the volatility of consumption and wealth, *Rev. Financ. Stud.* 2, 73–89.

Treynor, Jack, 1965, How to rate management of investment funds, *Harvard Business Review* 43, 63–75.

Velu, Raja, and Guofu Zhou, 1999, Testing multi-beta asset pricing models, *Journal of Empirical Finance* 6, 219–241.

Welch, Ivo, 2000, Views of financial economists on the equity premium and on professional controversies., *Journal of Business* 73, 501 – 537.

——— , 2008, The consensus estimate for the equity premium by academic financial economists in december 2007, *Working Paper, Brown University.*

Xing, Yuhang, 2008, Interpreting the Value Effect Through the Q-Theory: An Empirical Investigation, *Rev. Financ. Stud.* 21, 1767–1795.

Yogo, M., 2006, A consumption-based explanation of expected stock returns, *Journal of Finance* 61, 539–580.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

Zhou, Guofu, 1991, Small sample tests of portfolio efficiency, *Journal of Financial Economics* 30, 165–191.

——— , 1994, Analytical gmm tests: Asset pricing with time-varying risk premiums, *Review of Financial Studies* 7, 687–709.

——— , 1995, Small sample rank tests with applications to asset pricing, *Journal of Empirical Finance* 2, 71–93.