

TESTS OF MULTIFACTOR PRICING MODELS, VOLATILITY BOUNDS AND PORTFOLIO PERFORMANCE

WAYNE E. FERSON*

Carroll School of Management, Boston College

Contents

Abstract	743
Keywords	743
1. Introduction	744
2. Multifactor asset-pricing models: Review and integration	746
2.1. The stochastic discount factor representation	746
2.2. Expected risk premiums	748
2.3. Return predictability	749
2.4. Consumption-based asset-pricing models	751
2.5. Multi-beta pricing models	752
2.5.1. Relation to the stochastic discount factor	752
2.5.2. Relation to mean-variance efficiency	754
2.5.3. A large-markets interpretation	755
2.6. Mean-variance efficiency with conditioning information	758
2.6.1. Conditional versus unconditional efficiency	760
2.6.2. Implications for tests	762
2.7. Choosing the factors	763
3. Modern variance bounds	766
3.1. The Hansen–Jagannathan bounds	766
3.2. Variance bounds with conditioning information	768
3.2.1. Efficient portfolio bounds	769
3.2.2. Optimal bounds	769
3.2.3. Discussion	770
3.3. The Hansen–Jagannathan distance	771
4. Methodology and tests of multifactor asset-pricing models	772
4.1. The Generalized Method of Moments approach	772

* The author acknowledges financial support from the Collins Chair in Finance at Boston College and the Pigott-PACCAR professorship at the University of Washington. He is also grateful to Geert Bekaert, John Cochrane, George Constantinides and Ludan Liu for helpful comments and suggestions.

4.2. Cross-sectional regression methods	773
4.2.1. The Fama–MacBeth approach	775
4.2.2. Interpreting the estimates	775
4.2.3. A caveat	776
4.2.4. Errors-in-betas	778
4.3. Multivariate regression and beta-pricing models	779
4.3.1. Comparing the SDF and beta-pricing approaches	782
5. Conditional performance evaluation	783
5.0.1. A numerical example	784
5.1. Stochastic discount factor formulation	785
5.1.1. Invariance to the number of funds	786
5.1.2. Additional issues	786
5.2. Beta-pricing formulation	787
5.3. Using portfolio weights	788
5.3.1. Conditional performance attribution	789
5.3.2. Interim trading bias	790
5.4. Conditional market-timing models	790
5.5. Empirical evidence on conditional performance	791
6. Conclusions	793
References	793

Abstract

Three concepts: stochastic discount factors, multi-beta pricing and mean-variance efficiency, are at the core of modern empirical asset pricing. This chapter reviews these paradigms and the relations among them, concentrating on conditional asset-pricing models where lagged variables serve as instruments for publicly available information. The different paradigms are associated with different empirical methods. We review the variance bounds of Hansen and Jagannathan (1991), concentrating on extensions for conditioning information. Hansen's (1982) Generalized Method of Moments (GMM) is briefly reviewed as an organizing principle. Then, cross-sectional regression approaches as developed by Fama and MacBeth (1973) are reviewed and used to interpret empirical factors, such as those advocated by Fama and French (1993, 1996). Finally, we review the multivariate regression approach, popularized in the finance literature by Gibbons (1982) and others. A regression approach, with a beta pricing formulation, and a GMM approach with a stochastic discount factor formulation, may be considered competing paradigms for empirical work in asset pricing. This discussion clarifies the relations between the various approaches. Finally, we bring the models and methods together, with a review of the recent conditional performance evaluation literature, concentrating on mutual funds and pension funds.

Keywords

stochastic discount factor, performance evaluation, asset pricing, portfolio efficiency, volatility bounds, predicting returns

JEL classification: A23, C1, C31, C51, D91, E20, G10, G11, G12, G14, G23

1. Introduction

The asset-pricing models of modern finance describe the prices or expected rates of return of *financial assets*, which are claims traded in financial markets. Examples of financial assets are common stocks, bonds, options, futures and other “derivatives”, so named because they derive their values from other, underlying assets. Asset-pricing models are based on two central concepts. The first is the *no arbitrage principle*, which states that market forces tend to align prices so as to eliminate arbitrage opportunities. An arbitrage opportunity arises when assets can be combined in a portfolio with zero cost, no chance of a loss and positive probability of a gain. In Chapter 10 of this volume, Dybvig and Ross describe this theory. The second central concept in asset pricing is financial market equilibrium. Investors’ desired holdings of financial assets derive from an optimization problem. In equilibrium the first-order conditions of the optimization problem must be satisfied, and asset-pricing models follow from these conditions. When the agent considers the consequences of the investment decision for more than a single period in the future, intertemporal asset pricing models result. These models are reviewed by Campbell in Chapter 13 of this volume, and by Duffie in Chapter 11.

The present chapter reviews multi-factor asset-pricing models from an empiricists’ perspective. Multi-factor models can be motivated by either the no-arbitrage principle or by an equilibrium model. Their distinguishing feature is that expected asset returns are determined by a linear combination of their covariances with variables representing the risk factors. This chapter has two main objectives. The first is to integrate the various empirical models and their tests in a self-contained discussion. The second is to review the application to the problem of measuring investment performance.

This chapter concentrates heavily on the role of *conditioning information*, in the form of lagged variables that serve as instruments for publicly available information. I think that developments in this area, *conditional asset pricing*, represent some of the most significant advances in empirical asset-pricing research in recent years.

The models described in this chapter are set in the classical world of perfectly efficient financial markets, and perfectly rational economic agents. Of course, a great deal of research is devoted to understanding asset prices under market imperfections like information and transactions costs, as several chapters in this volume amply illustrate. For asset pricing emphasizing human imperfections (behavioral finance) see Chapter 18 of this volume by Barberis and Thaler. The perfect-markets models reviewed here represent a baseline, and a starting point for understanding these more complex issues.

Work in empirical asset pricing over the last few years has provided a markedly improved understanding of the relations among the various asset-pricing models. Bits and pieces of this are scattered across a number of published papers, and some is “common” knowledge, shared by aficionados. This chapter provides an integrative discussion, refining the earlier review in Ferson (1995) to reflect what I hope is an improved understanding.

Much of our understanding of how asset-pricing models' empirical predictions are related flows from representing the models as *stochastic discount factors*. Section 2 presents the stochastic discount factor approach, briefly illustrates a few examples of stochastic discount factors, and then relates the representation to beta pricing and to mean variance efficiency. These three concepts: stochastic discount factors, beta pricing and mean variance efficiency, are at the core of modern empirical asset pricing. We show the relation among these three concepts, and a "large-markets" interpretation of these relations. The discussion then proceeds to refinements of these issues in the presence of conditioning information. Section 2 ends with a brief discussion of how the risk factors have been identified in the empirical literature, and what the empirical evidence has to say about the factors.

Section 3 begins with a fundamental empirical application of the stochastic discount factor approach – the variance bounds originally developed by Hansen and Jagannathan (1991). Unlike the case where a model identifies a particular stochastic discount factor, the question in the Hansen–Jagannathan bounds is: Given a set of asset returns, and some conditioning information, what can be said about the set of stochastic discount factors that could properly "price" the assets? By now, a number of reviews of the original Hansen–Jagannathan bounds are available in the literature. The discussion here is brief, quickly moving on to focus on less well-known refinements of the bounds to incorporate conditioning information.

Section 4 discusses empirical methods, starting with Hansen's (1982) Generalized Method of Moments (GMM). This important approach has also been the subject of several review articles and textbook chapters. We briefly review the use of the GMM to estimate stochastic discount factor models. This section is included only to make the latter parts of the chapter accessible to a reader who is not already familiar with the GMM. Section 4 then discusses two special cases that remain important in empirical asset pricing. The first is the cross-sectional regression approach, as developed by Fama and MacBeth (1973), and the second is the multivariate regression approach, popularized in the finance literature following Gibbons (1982).

Once the mainstay of empirical work on asset pricing, cross-sectional regression continues to be used and useful. Our main focus is on the economic interpretation of the estimates. The discussion attempts to shed light on recent studies that employ the empirical factors advocated by Fama and French (1993, 1996), or generalizations of that approach. The multivariate regression approach to testing portfolio efficiency can be motivated by its immunity to the errors-in-variables problem that plagues the two step, cross-sectional regression approach. The multivariate approach is also elegant, and provides a nice intuition for the statistical tests.

A regression approach, with a beta pricing formulation, and a GMM approach with a stochastic discount factor formulation, may be considered as competing paradigms for empirical work in asset pricing. However, under the same distributional assumptions, and when the same moments are estimated, the two approaches are essentially equivalent. The present discussion attempts to clarify these points, and suggests how to think about the choice of empirical method.

Section 5 brings the models and methods together, in a review of the relatively recent literature on *conditional performance evaluation*. The problem of measuring the performance of managed portfolios has been the subject of research for more than 30 years. Traditional measures use unconditional expected returns, estimated by sample averages, as the baseline. However, if expected returns and risks vary over time, this may confuse common time-variation in fund risk and market risk premiums with average performance. In this way, traditional methods can ascribe abnormal performance to an investment strategy that trades mechanically, based only on public information. Conditional performance evaluation attempts to control these biases, while delivering potentially more powerful performance measures, by using lagged instruments to control for time-varying expectations. Section 5 reviews the main models for conditional performance evaluation, and includes a summary of the empirical evidence. Finally, Section 6 of this chapter offers concluding remarks.

2. Multifactor asset-pricing models: Review and integration

2.1. The stochastic discount factor representation

Virtually all asset pricing models are special cases of the fundamental equation:

$$P_t = E_t\{m_{t+1}(P_{t+1} + D_{t+1})\}, \quad (1)$$

where P_t is the price of the asset at time t and D_{t+1} is the amount of any dividends, interest or other payments received at time $t + 1$. The market-wide random variable m_{t+1} is the stochastic discount factor (SDF)¹. The prices are obtained by “discounting” the payoffs using the SDF, or multiplying by m_{t+1} , so that the expected “present value” of the payoff is equal to the price.

The notation $E_t\{\cdot\}$ denotes the conditional expectation, given a market-wide information set, Ω_t . Since empiricists don’t get to see Ω_t , it will be convenient to consider expectations conditioned on an observable subset of instruments, Z_t . These expectations are denoted as $E(\cdot|Z_t)$. When Z_t is the null-information set, we have the unconditional expectation, denoted as $E(\cdot)$. Empirical work on asset-pricing models like Equation (1) typically relies on *rational expectations*, interpreted as the assumption that the expectation terms in the model are mathematical conditional expectations. Taking the expected values of Equation (1), rational expectations implies that versions of Equation (1) must hold for the expectations $E(\cdot|Z_t)$ and $E(\cdot)$.

¹ The random variable m_{t+1} is also known as the pricing kernel, benchmark pricing variable, or intertemporal marginal rate of substitution, depending on the context. The representation (1) goes at least back to Beja (1971), while the term “stochastic discount factor” is usually ascribed to Hansen and Richard (1987).

Assuming nonzero prices, Equation (1) is equivalent to:

$$E(m_{t+1}R_{t+1} - \underline{1} \mid \Omega_t) = 0, \quad (2)$$

where R_{t+1} is the N -vector of primitive asset gross returns and $\underline{1}$ is an N -vector of ones. The gross return $R_{i,t+1}$ is defined as $(P_{i,t+1} + D_{i,t+1})/P_{i,t}$. We say that a SDF “prices” the assets if Equations (1) and (2) are satisfied. Empirical tests of asset-pricing models often work directly with Equation (2) and the relevant definition of m_{t+1} .

Without more structure Equations (1) and (2) have no content because it is almost always possible to find a random variable m_{t+1} for which the equations hold. There will be some m_{t+1} that “works”, in this sense, as long as there are no redundant asset returns². With the restriction that m_{t+1} is a strictly positive random variable, Equation (1) becomes equivalent to the no-arbitrage principle, which says that all portfolios of assets with payoffs that can never be negative, but which are positive with positive probability, must have positive prices [Beja (1971), Rubinstein (1976), Ross (1977), Harrison and Kreps (1979), Hansen and Richard (1987)].

The no arbitrage condition does not uniquely identify m_{t+1} unless markets are complete. In that case, m_{t+1} is equal to primitive state prices divided by state probabilities. To see this write Equation (1) as $P_{i,t} = E_t\{m_{t+1}X_{i,t+1}\}$, where $X_{i,t+1} = P_{i,t+1} + D_{i,t+1}$. In a discrete-state setting, $P_{i,t} = \sum_s \pi_s X_{i,s} = \sum_s q_s (\pi_s/q_s) X_{i,s}$, where q_s is the probability that state s will occur and π_s is the state price, equal to the value at time t of one unit of the numeraire to be paid at time $t+1$ if state s occurs at time $t+1$. $X_{i,s}$ is the total payoff of the security i at time $t+1$ if state s occurs. Comparing this expression with Equation (1) shows that $m_s = \pi_s/q_s > 0$ is the value of the SDF in state s .

While the no-arbitrage principle places some restrictions on m_{t+1} , empirical work often explores the implications of equilibrium models for the SDF, based on investor optimization. Consider the Bellman equation for a representative consumer-investor’s optimization:

$$J(W_t, s_t) \equiv \text{Max } E_t \{U(C_t, \cdot) + J(W_{t+1}, s_{t+1})\}, \quad (3)$$

where $U(C_t, \cdot)$ is the direct utility of consumption expenditures at time t , and $J(\cdot)$ is the indirect utility of wealth. The notation allows the direct utility of current consumption expenditures to depend on variables such as past consumption expenditures or other state variables, s_t . The state variables are sufficient statistics, given wealth, for the utility of future wealth in an optimal consumption-investment plan. Thus, changes in the state variables represent future consumption-investment opportunity risk. The budget constraint is: $W_{t+1} = (W_t - C_t)x'R_{t+1}$, where x is the portfolio weight vector, subject to $x'\underline{1} = 1$.

² For example, take a sample of assets with a nonsingular second moment matrix and let m_{t+1} be $[\underline{1}'(E_t\{R_{t+1}R_{t+1}'\})^{-1}]R_{t+1}$.

If the allocation of resources to consumption and investment assets is optimal, it is not possible to obtain higher utility by changing the allocation. Suppose an investor considers reducing consumption at time t to purchase more of (any) asset. The expected utility cost at time t of the foregone consumption is the expected marginal utility of consumption expenditures, $U_c(C_t, \cdot) > 0$ (where a subscript denotes partial derivative), multiplied by the price $P_{i,t}$ of the asset, measured in the numeraire unit. The expected utility gain of selling the investment asset and consuming the proceeds at time $t + 1$ is $E_t\{(P_{i,t+1} + D_{i,t+1})J_w(W_{t+1}, S_{t+1})\}$. If the allocation maximizes expected utility, the following must hold: $P_{i,t}E_t\{U_c(C_t, \cdot)\} = E_t\{(P_{i,t+1} + D_{i,t+1})J_w(W_{t+1}, S_{t+1})\}$, which is equivalent to Equation (1), with

$$m_{t+1} = \frac{J_w(W_{t+1}, S_{t+1})}{E_t\{U_c(C_t, \cdot)\}}. \quad (4)$$

The m_{t+1} in Equation (4) is the *intertemporal marginal rate of substitution* (IMRS) of the consumer-investor, and Equations (2) and (4) combined are the *intertemporal Euler equation*.

Asset-pricing models typically focus on the relation of security returns to aggregate quantities. To get there, it is necessary to aggregate the Euler equations of individuals to obtain equilibrium expressions in terms of aggregate quantities. Theoretical conditions which justify the use of aggregate quantities are discussed by Wilson (1968), Rubinstein (1974) and Constantinides (1982), among others. Some recent empirical work does not assume aggregation, but relies on panels of disaggregated data. Examples include Zeldes (1989), Brav, Constantinides and Geczy (2002) and Balduzzi and Yao (2001).

Multiple-factor models for asset pricing follow when m_{t+1} can be written as a function of several factors. Equation (4) suggests that likely candidates for the factors are variables that proxy for consumer wealth, consumption expenditures or the state variables – the variables that determine the marginal utility of future wealth in an optimal consumption-investment plan.

2.2. Expected risk premiums

Typically, empirical work focuses on expressions for expected returns and excess rates of return. Expected excess returns are related to the risk factors that create variation in m_{t+1} . Consider any asset return $R_{i,t+1}$ and a reference asset return, $R_{0,t+1}$. Define the excess return of asset i , relative to the reference asset as $r_{i,t+1} = R_{i,t+1} - R_{0,t+1}$. If Equation (2) holds for both assets it implies:

$$E_t\{m_{t+1}r_{i,t+1}\} = 0 \text{ for all } i. \quad (5)$$

Use the definition of covariance to expand Equation (5) into the product of expectations plus the covariance, obtaining:

$$E_t\{r_{i,t+1}\} = \frac{\text{Cov}_t(r_{i,t+1}; -m_{t+1})}{E_t\{m_{t+1}\}}, \text{ for all } i, \quad (6)$$

where $\text{Cov}_t(\cdot; \cdot)$ is the conditional covariance. Equation (6) is a general expression for the expected excess return, from which most of the expressions in the literature can be derived. The conditional covariance of return with the SDF, m_{t+1} , is a very general measure of *systematic risk*. Asset-pricing models say that assets earn expected return premiums for their systematic risk, not their total risk (i.e., variance of return). The covariance with $-m_{t+1}$ is systematic risk because it measures the component of the return that contributes to fluctuations in the marginal utility of wealth. If we regressed the asset return on the SDF, the residual in the regression would capture the “unsystematic” risk and would not be “priced”, or command a risk premium.

If the conditional covariance with the SDF is zero for a particular asset, the expected excess return of that asset should be zero³. The more negative is the covariance with m_{t+1} the less desirable is the distribution of the random return, as the larger payoffs tend to occur when the marginal utility is low. The expected compensation for holding assets with this feature must be higher than for those with a more desirable distribution. Expected risk premiums should therefore differ across assets in proportion to their conditional covariances with $-m_{t+1}$.

2.3. Return predictability

Rational expectations implies that the difference between return realizations and the expectations in the model should be unrelated to the information that the expectations in the model are conditioned on. For example, Equation (2) says that the conditional expectation of the product of m_{t+1} and $R_{i,t+1}$ is the constant, 1.0. Therefore, $1 - m_{t+1}R_{i,t+1}$ should not be predictably different from zero using any information available at time t . If we run a regression of $1 - m_{t+1}R_{i,t+1}$ on any lagged variable, Z_t , the regression coefficients should be zero. If there is predictability in a return $R_{i,t+1}$ using instruments Z_t , the model implies that the predictability is removed when $R_{i,t+1}$ is multiplied by the correct m_{t+1} . This is the sense in which conditional asset-pricing models are asked to “explain” predictable variation in asset returns. This view generalizes the older “random walk” model of stock values, which states that stock returns should be completely unpredictable. That model is a special case which can be motivated by risk neutrality. Under risk neutrality the IMRS, m_{t+1} , is a constant. Therefore, in this case the model implies that the return $R_{i,t+1}$ should not differ predictably from a constant.

Conditional asset pricing presumes the existence of some return predictability. There should be instruments Z_t for which $E(R_{t+1}|Z_t)$ or $E(m_{t+1}|Z_t)$ vary over time, in order

³ Equation (6) is weaker than Equation (2), since Equation (6) is equivalent to $E_t\{m_{t+1}R_{i,t+1}\} = \Delta_t$, all i , where Δ_t is a constant across assets, while Equation (2) restricts $\Delta_t = 1$. Therefore, empirical tests based on Equation (6) do not exploit all of the restrictions implied by a model that may be stated in the form of Equation (2).

for the equation $E(m_{t+1}R_{t+1} - 1|Z_t) = 0$ to have empirical bite⁴. Interest in predicting security-market returns is about as old as the security markets themselves. Fama (1970) reviews the early evidence and Schwert, in Chapter 15 of this volume, reviews “anomalies” based on predictability.

One body of literature uses lagged returns to predict future stock returns, attempting to exploit serial dependence. High frequency serial dependence, such as daily or intraday patterns, are often considered to represent the effects of market microstructure, such as bid-ask spreads [e.g., Roll (1984)] and nonsynchronous trading of the stocks in an index [e.g., Scholes and Williams (1977)]. Serial dependence at longer horizons may represent predictable changes in the expected returns.

Conrad and Kaul (1989) report serial dependence in weekly returns. Jegadeesh and Titman (1993) find that relatively high return, “winner” stocks tend to repeat their performance over three to nine-month horizons. DeBondt and Thaler (1985) find that past high-return stocks perform poorly over the next five years, and Fama and French (1988) find negative serial dependence over two to five-year horizons. These serial-dependence patterns motivate a large number of studies which attempt to assess the economic magnitude and statistical robustness of the implied predictability, or to explain the predictability as an economic phenomenon. For more comprehensive reviews, see Campbell, Lo and MacKinlay (1997) or Kaul (1996). Research in this area continues, and it is fair to say that the jury is still out on the issue of predictability using lagged returns.

A second body of literature studies predictability using other lagged variables as instruments. Fama and French (1989) assemble a list of variables from studies in the early 1980s, that as of this writing remain the workhorse instruments for conditional asset pricing models. These variables include the lagged dividend yield of a stock market index, a yield spread of long-term government bonds relative to short term bonds, and a yield spread of low-grade (high default risk) corporate bonds over high-grade bonds. In addition, studies often include the level of a short term interest rate [Fama and Schwert (1977), Ferson (1989)] and the lagged excess return of a medium-term over a short-term Treasury bill [Campbell (1987), Ferson and Harvey (1991)]. Recently proposed instruments include an aggregate book-to-market ratio [Pontiff and Schall (1998)] and lagged consumption-to-wealth ratios [Lettau and Ludvigson (2001)]. Of course, many other predictor variables have been proposed and more will doubtless be proposed in the future.

Predictability using lagged instruments remains controversial, and there are some good reasons to question the predictability. Studies have identified various statistical biases in predictive regressions [e.g., Hansen and Hodrick (1980), Stambaugh (1999),

⁴ At one level this is easy. Since $E(m_{t+1}|Z_t)$ should be the inverse of a risk-free return, all we need is observable risk-free rates that vary over time. Ferson (1989) shows that the behavior of stock returns and short-term interest rates imply that conditional covariances of returns with m_{t+1} must also vary over time.

Ang and Bekaert (2001), Ferson, Sarkissian and Simin (2002)], questioned the stability of the predictive relations across economic regimes [e.g., Kim, Nelson and Startz (1991)] and raised the possibility that the lagged instruments arise solely through data mining [e.g., Lo and MacKinlay (1990), Foster, Smith and Whaley (1997)].

A reasonable response to these concerns is to see if the predictive relations hold out-of-sample. This kind of evidence is also mixed. Some studies find support for predictability in step-ahead or out-of-sample exercises [e.g., Fama and French (1989), Pesaran and Timmermann (1995)]. Similar instruments show some ability to predict returns outside the context of the USA, where they arose [e.g., Harvey (1991), Solnik (1993), Ferson and Harvey (1993, 1999)]. However, other studies conclude that predictability using the standard lagged instruments does not hold [e.g., Goyal and Welch (1999), Simin (2002)]. It seems that research on the predictability of security returns will always be interesting, and conditional asset-pricing models should be useful in framing many future investigations of these issues.

2.4. Consumption-based asset-pricing models

In these models the economic agent maximizes a lifetime utility function of consumption (including possibly a bequest to heirs). Consumption models may be derived from Equation (4) by exploiting the envelope condition, $U_c(\cdot) = J_w(\cdot)$, which states that the marginal utility of consumption must be equal to the marginal utility of wealth if the consumer has optimized the tradeoff between the amount consumed and the amount invested.

Breeden (1979) derived a consumption-based asset-pricing model in continuous time, assuming that the preferences are time-additive. The utility function for the lifetime stream of consumption is $\sum_t \beta^t U(C_t)$, where β is a time preference parameter and $U(\cdot)$ is increasing and concave in current consumption, C_t . Breeden's model is a linearization of Equation (2) which follows from the assumption that asset values and consumption follow diffusion processes [Bhattacharya (1981), Grossman and Shiller (1982)]. A discrete-time version follows Rubinstein (1976) and Lucas (1978), assuming a power utility function:

$$U(C) = \frac{C^{1-\alpha} - 1}{1-\alpha}, \quad (7)$$

where $\alpha > 0$ is the concavity parameter. This function displays constant relative risk aversion⁵ equal to α . Using Equation (7) and the envelope condition, the IMRS in Equation (4) becomes:

$$m_{t+1} = \beta (C_{t+1}/C_t)^{-\alpha}. \quad (8)$$

⁵ *Relative risk aversion* in consumption is defined as $-Cu''(C)/u'(C)$. *Absolute risk aversion* is $-u''(C)/u'(C)$. Ferson (1983) studies a consumption-based asset-pricing model with constant absolute risk aversion.

A large literature has tested the pricing Equation (1), with the SDF given by the consumption model (8), and generalizations of that model⁶.

2.5. Multi-beta pricing models

The vast majority of the empirical work on asset-pricing models involves expressions for expected returns, stated in terms of *beta coefficients* relative to one or more portfolios or factors. The beta is the regression coefficient of the asset return on the factor. Multi-beta models have more than one risk factor and more than one beta for each asset. The Arbitrage Pricing Theory (APT) leads to approximate expressions for expected returns with multiple beta coefficients. Models based on investor optimization and equilibrium lead to exact expressions⁷. Both of these approaches lead to models with the following form:

$$E_t(R_{i,t+1}) = \lambda_{0t} + \sum_{j=1}^K b_{ijt} \lambda_{jt}, \text{ for all } i. \quad (9)$$

The b_{i1t}, \dots, b_{iKt} are the time t betas of asset i relative to the K risk factors $F_{j,t+1}$, $j = 1, \dots, K$. These betas are the conditional multiple-regression coefficients of the assets on the factors. The λ_{jt} , $j = 1, \dots, K$ are the factor risk premiums, which represent increments to the expected return per unit of type- j beta. These premiums do not depend on the specific security i . λ_{0t} is the expected *zero-beta rate*. This is the expected return of any security that is uncorrelated with each of the K factors in the model (i.e., $b_{0jt} = 0, j = 1, \dots, K$). If there is a risk-free asset, then λ_{0t} is the return of this asset.

2.5.1. Relation to the stochastic discount factor

We first show how a multi-beta model can be derived as a special case of the SDF representation, when the factors capture the relevant systematic risks. We take this to mean that the error terms, $u_{i,t+1}$, in a regression of returns on the factors are not “priced”; that is, they are uncorrelated with m_{t+1} : $\text{Cov}_t(u_{i,t+1}, m_{t+1}) = 0$. We then state the general equivalence between the two representations. This equivalence was

⁶ An important generalization allows for nonseparabilities in the $U_c(C_t, \cdot)$ function in Equation (4), as may be implied by the durability of consumer goods, habit persistence in the preferences for consumption over time, or nonseparability of preferences across states of nature. Singleton (1990), Ferson (1995) and Campbell, in Chapter 13 of this volume, review this literature.

⁷ The multiple-beta equilibrium model was developed in continuous time by Merton (1973), Breeden (1979) and Cox, Ingersoll and Ross (1985). Long (1974), Sharpe (1977), Cragg and Malkiel (1982), Connor (1984), Dybvig (1983), Grinblatt and Titman (1983) and Shanken (1987) provide multibeta interpretations of equilibrium models in discrete time.

first discussed, for the case of a single-factor model, by Dybvig and Ingersoll (1982). The general, multi-factor case follows from Ferson and Jagannathan (1996).

Let $R_{0,t+1}$ be a zero-beta portfolio, and $\lambda_{0,t}$ the expected return on the zero-beta portfolio. Equation (6) implies:

$$E_t(R_{i,t+1}) = \lambda_{0,t} + \frac{\text{Cov}_t(R_{i,t+1}; -m_{t+1})}{E_t(m_{t+1})}. \quad (10)$$

Substituting the regression model $R_{i,t+1} = a_i + \sum_j b_{ijt} F_{j,t+1} + u_{i,t+1}$ into the right hand side of (10) and assuming that $\text{Cov}_t(u_{i,t+1}, m_{t+1}) = 0$ implies:

$$E_t(R_{i,t+1}) = \lambda_{0,t} + \sum_{j=1, \dots, K} b_{ijt} \frac{\text{Cov}_t\{F_{j,t+1}, -m_{t+1}\}}{E_t(m_{t+1})}. \quad (11)$$

The risk premium per unit of type- j beta is $\lambda_{j,t} = [\text{Cov}_t\{F_{j,t+1}, -m_{t+1}\}/E_t(m_{t+1})]$. In the special case where the factor $F_{j,t+1}$ is a traded asset return, Equation (11) implies that $\lambda_{j,t} = E_t(F_{j,t+1}) - \lambda_{0,t}$; the expected risk premium equals the factor portfolio's expected excess return.

Equation (11) is useful because it provides intuition about the signs and magnitudes of expected risk premiums for particular factors. The intuition is the same as in Equation (6) above. If a risk factor $F_{j,t+1}$ is negatively-correlated with m_{t+1} , the model implies that a positive risk premium is associated with that factor beta. A factor that is positively-related to marginal utility should carry a negative premium, because the big payoffs come when the value of payoffs is high. This implies a high present value and a low expected return. Expected risk premiums for a factor should also change over time if the conditional covariances of the factor with the scaled marginal utility $[m_{t+1}/E_t(m_{t+1})]$ vary over time.

The steps that take us from Equations (6) to (11) can be reversed, so the SDF and multi-beta representations are, in fact, equivalent. The formal statement is:

Lemma 1 [Ferson and Jagannathan (1996)]. *The stochastic discount factor representation (2) and the multi-beta model (9) are equivalent, where,*

$$m_{t+1} = c_{0t} + c_{1t}F_{1,t+1} + \dots + c_{Kt}F_{K,t+1}, \quad (12)$$

with

$$c_{0t} = \frac{1 + \sum_k \{\lambda_k E_t(F_{k,t+1}) / \text{Var}_t(F_{k,t+1})\}}{\lambda_{0,t}},$$

and

$$c_{jt} = -\frac{\lambda_{j,t}}{\lambda_{0,t} \text{Var}_t(F_{j,t+1})}, \quad j = 1, \dots, K.$$

For a proof, see Ferson and Jagannathan (1996).

If the factors are not traded asset returns, then it is typically necessary to estimate the expected risk premiums for the factors, $\lambda_{k,t}$. These may be identified as the conditional expected excess returns on *factor-mimicking portfolios*. A factor-mimicking portfolio is defined as a portfolio whose return can be used in place of a factor in the model. There are several ways to obtain mimicking portfolios, as described in more detail below⁸.

2.5.2. Relation to mean-variance efficiency

The concept of a *minimum-variance portfolio* is central in the asset-pricing literature. A portfolio $R_{p,t+1}$ is a minimum-variance portfolio if no portfolio with the same expected return has a smaller variance. Roll (1977) and others have shown that the portfolio $R_{p,t+1}$ is a minimum-variance portfolio if and only if a *beta-pricing model* holds⁹:

$$E_t \{R_{i,t+1} - R_{pz,t+1}\} = \beta_{ipt} E_t \{R_{p,t+1} - R_{pz,t+1}\}, \quad \text{all } i; \quad (13)$$

$$\beta_{ipt} = \frac{\text{Cov}_t(R_{i,t+1}; R_{p,t+1})}{\text{Var}_t(R_{p,t+1})}.$$

In Equation (13), β_{ipt} is the *conditional beta* of $R_{i,t+1}$ relative to $R_{p,t+1}$. $R_{pz,t+1}$ is a *zero beta* asset relative to $R_{p,t+1}$. A zero-beta asset satisfies $\text{Cov}_t(R_{pz,t+1}; R_{p,t+1}) = 0$. Equation (13) is essentially a restatement of the first-order condition for the optimization problem that defines a minimum-variance portfolio.

Equation (13) first appeared as an asset-pricing model in the famous Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965) and Black (1972). The CAPM is equivalent to the statement that the *market portfolio* $R_{m,t+1}$ is *mean-variance efficient*. The market portfolio is the portfolio of all marketed assets, weighted according to their relative total values. The portfolio is mean-variance efficient if it satisfies Equation (13), and also $E_t(R_{m,t+1} - R_{mz,t+1}) > 0$.

When the factors are traded assets like a market portfolio, or when mimicking portfolios are used, the multi-beta model in Equation (9) is equivalent to the statement that a combination of the factor portfolios is minimum-variance efficient¹⁰. Therefore,

⁸ Breeden (1979, footnote 7) derives maximum correlation mimicking portfolios. Grinblatt and Titman (1987), Shanken (1987), Lehmann and Modest (1988) and Huberman, Kandel and Stambaugh (1987) provide further characterizations of mimicking portfolios when there is no conditioning information. Ferson and Siegel (2002b) and Ferson, Siegel and Xu (2002) consider cases where there is conditioning information.

⁹ It is assumed that the portfolio $R_{p,t+1}$ is not the global minimum-variance portfolio; that is, the minimum variance over all levels of expected return. This is because the betas of all assets on the global minimum-variance portfolio are identical.

¹⁰ This result is proved by Grinblatt and Titman (1987), Shanken (1987) and Huberman, Kandel and Stambaugh (1987) and reviewed by Ferson (1995).

multiple-beta asset-pricing models like Equation (9) always imply that combinations of particular portfolios are minimum-variance efficient. This correspondence is exploited by Gibbons, Ross and Shanken (1989) and Kandel and Stambaugh (1989), among others, to develop tests of multi-beta models based on mean-variance efficiency. Such tests are discussed in Section 4 below.

Since the SDF representation in Equation (2) is equivalent to a multi-beta expression for expected returns, and a multi-beta model is equivalent to a statement about minimum-variance efficiency, it follows that the SDF representation is equivalent to a statement about minimum-variance efficiency. Let's now complete the loop.

Lemma 2. *A portfolio which maximizes squared (conditional) correlation with m_{t+1} in Equation (2) is a minimum (conditional) variance portfolio.*

Proof: Consider the conditional projection of m_{t+1} on the vector of returns R_{t+1} . The coefficient vector is $w_t \equiv (E_t(R_{t+1}R'_{t+1}))^{-1}E_t(R_{t+1}m_{t+1}) = (E_t(R_{t+1}R'_{t+1}))^{-1}\underline{1}$. Define the portfolio return $R_{p,t+1} = (w_t/w'_t\underline{1})'R_{t+1}$. The portfolio maximizes the squared conditional correlation with m_{t+1} . The regression of m_{t+1} on the vector of returns can be written as $m_{t+1} = (w'_t\underline{1})R_{p,t+1} + \varepsilon_{t+1}$. The error term ε_{t+1} is conditionally uncorrelated with $R_{i,t+1}$ for all i , and therefore with $R_{p,t+1}$ ¹¹. Substituting for m_{t+1} in Equation (6) from this regression produces:

$$E_t\{r_{i,t+1}\} = \frac{\text{Cov}_t(r_{i,t+1}; R_{p,t+1})}{\text{Cov}_t(r_{p,t+1}; R_{p,t+1})} E_t\{r_{p,t+1}\}, \quad \text{all } i, \quad (14)$$

where $r_{i,t+1}$ and $r_{p,t+1}$ are excess returns. If the reference asset for the excess returns is taken to be $R_{pz,t+1}$, a zero-beta asset for $R_{p,t+1}$, then Equation (13) follows directly from Equation (14). Equation (13) implies that $R_{p,t+1}$ is a conditional minimum-variance portfolio. \square

We have seen that exact multi-beta pricing is equivalent to the statement that $E(mR_i) = 1$ for all i , under the assumption that m is a linear function of the factors, and also equivalent to the statement that a portfolio of the factors is a minimum-variance-efficient portfolio. Thus, we have equivalence among the three paradigms: Exact multi-beta pricing, stochastic discount factors, and mean-variance efficiency.

2.5.3. A large-markets interpretation

This section describes how the three paradigms of empirical asset pricing work in the large markets of the Arbitrage Pricing Theory (APT) of Ross (1976), as refined

¹¹ The fitted values of the regression will have the same pricing implications as m_{t+1} . That is, $m_{t+1}^* = (w'_t\underline{1})R_{p,t+1}$ can replace m_{t+1} in Equation (1). Note that when the covariance matrix of asset returns is nonsingular, m_{t+1}^* is the unique SDF (i.e., satisfies Equation (1)) which is also an asset return. An SDF which satisfies Equation (1) is not in general an asset return, nor is it unique, unless markets are complete. If markets are complete, m_{t+1}^* is perfectly correlated with m_{t+1} [Hansen and Richard (1987)].

by Chamberlain (1983) and Chamberlain and Rothschild (1983). For this purpose, we ignore the existence of any “conditioning information”, and suppress the time subscripts and related notation. [For arbitrage pricing relations with conditioning information, see Stambaugh (1983)].

Assume that the following data-generating model describes equity returns in excess of a risk-free asset:

$$r_i = E(r_i) + \beta_i' f + e_i, \quad (15)$$

where $E(f) = 0 = E(e_i f)$, all i , and $f_i = F_t - E(F_t)$ are the unexpected factor returns. We can normalize the factors to have the identity as their covariance matrix; the β_i absorb the normalization. The $N \times N$ covariance matrix of the asset return residuals can then be expressed as:

$$\text{Cov}(R) \equiv \Sigma = BB' + V, \quad (16)$$

where V is the covariance matrix of the residual vector, e , B is the $N \times K$ matrix of the β_i , and Σ is assumed to be nonsingular for all N . The factor model assumes that the eigenvalues of V are bounded as $N \rightarrow \infty$, while the K nonzero eigenvalues of BB' become infinite as $N \rightarrow \infty$. Thus, the covariance matrix Σ has K unbounded and $N - K$ bounded eigenvalues as N becomes large. This is called an “approximate factor structure”, to distinguish it from an “exact” factor structure, where V is assumed to be diagonal.

The factor model in Equation (16) decomposes the variances of returns into “pervasive” and “nonsystematic” risks. If x is an N -vector of portfolio weights, the portfolio variance is $x' \Sigma x$, where $\lambda_{\max}(\Sigma) x' x \geq x' \Sigma x \geq \lambda_{\min}(\Sigma) x' x$, $\lambda_{\min}(\Sigma)$ being the smallest eigenvalue of Σ and $\lambda_{\max}(\Sigma)$ being the largest. Following Chamberlain (1983), a portfolio is “well diversified” iff $x' x \rightarrow 0$ as N grows without bound. For example, an equally weighted portfolio is well diversified; in this case $x' x = (1/N) \rightarrow 0$. The bounded eigenvalues of V imply that V captures the component of portfolio risk that is not pervasive or systematic, in the sense that this part of the variance vanishes in a well-diversified portfolio. The exploding eigenvalues of BB' imply that the common factor risks are pervasive, in the sense that they remain in a large, well-diversified portfolio.

The arbitrage pricing theory of Ross (1976) asserts that $\alpha' \alpha < \infty$ as N grows without bound, where α is the N vector of “alphas,” or expected abnormal returns, measured as the differences between the left- and right-hand sides of Equation (9), using the APT factors in the multi-beta model. The alphas are the differences between the assets’ expected returns and the returns predicted by the multi-beta model, sometimes called the “pricing errors”. The Ross APT implies that the multi-beta model’s pricing errors are “small”, on average, in a large market. If $\alpha' \alpha < \infty$ as N grows, then the cross-asset average of the squared pricing errors, $(\alpha' \alpha)/N$, must go to zero as N grows.

To see how the approximate beta pricing of the APT relates to the other paradigms of empirical asset pricing, we first describe how the pricing errors in the beta-pricing model are related to those of a stochastic discount factor representation. If we define $\alpha_m = E(mR - 1)$, where m is linear in the APT factors, then it follows from Equations (10) and (11) that $\alpha_m = E(m)\alpha$; the beta-pricing and stochastic discount factor alphas are proportional, where the risk-free rate determines the constant of proportionality. Provided that the risk-free rate is bounded above -100% , then $E(m)$ is bounded, and $\alpha'\alpha$ is bounded above if and only if $\alpha'_m\alpha_m$ is bounded above. Thus, the Ross APT has the same implications for the pricing errors in the stochastic discount factor and beta-pricing paradigms.

The third paradigm is mean-variance efficiency. We know that a combination of the APT factors is minimum-variance efficient, if and only if $\alpha = 0$. Thus, under the Ross APT a combination of the factors is not minimum-variance efficient. However, an upper bound on $\alpha'\alpha$ implies a lower bound on the correlation between a minimum-variance combination of the factors and a minimum-variance-efficient portfolio.

To see how the Ross APT restricts the correlation between the factors and a minimum-variance-efficient portfolio, we need two facts. The first is the “law of conservation of squared Sharpe ratios”, developed as Equation (52) below (p. 781). Here we state the law as $S^2(r) = \alpha'\Sigma^{-1}\alpha + S^2(f)$, where $S^2(f)$ is the maximum squared Sharpe ratio that can be obtained by a portfolio of the factors¹², $S^2(r)$ is the squared Sharpe ratio of a minimum-variance-efficient portfolio using all of the assets, and Σ is the covariance matrix of the assets’s excess returns. The second fact, which follows from Equation (13), describes the correlation, ρ , between the minimum-variance-efficient portfolio of the factors and the minimum-variance-efficient portfolios that uses all of the assets: $\rho = S(f)/S(r)$. Combining these results, $S^2(r) - S^2(f) = \alpha'\Sigma^{-1}\alpha \leq \alpha'\alpha\lambda_{\max}(\Sigma^{-1}) = \alpha'\alpha/\lambda_{\min}(\Sigma)$. Substituting for $S(r)$ in terms of ρ and $S(f)$, we arrive at: $[1/\rho^2 - 1] \leq \alpha'\alpha/[\lambda_{\min}(\Sigma)S^2(f)]$. Thus, an upper bound on $\alpha'\alpha$ places a lower bound on the squared correlation between the minimum-variance factor portfolio and a minimum-variance-efficient portfolio of the assets.

The “exact” version of the APT asserts that $\alpha'\alpha \rightarrow 0$ as N grows without bound, thus, the pricing errors of all assets go to zero as the market gets large. This version of the model requires stronger economic assumptions, as described by Connor (1984). Chamberlain (1983) shows that the exact APT is equivalent to the statement that all minimum-variance portfolios are well-diversified, and are thus combinations of the APT factors. In this case, we are essentially back to the original equivalence between the three paradigms holding as N gets large. That is, we have $E(mR - 1) = 0$ if and only if $\alpha = 0$ when m is linear in the APT factors, and equivalently, a combination of the factors is a minimum-variance-efficient portfolio in the large market.

¹² The Sharpe ratio is the expected excess return divided by the standard deviation.

2.6. Mean-variance efficiency with conditioning information

Most asset-pricing models are stated in terms of expected asset returns, covariances and betas, conditional on the available public information at time t . However, empirical tests traditionally examine unconditional expected returns and betas, or use instruments that are a subset of the available public information. Given the equivalence between mean variance efficiency and the other asset pricing representations, it follows that all tests of asset pricing models using portfolios, have examined whether particular portfolios are either unconditionally minimum variance, or minimum variance conditional on a subset of the information. To understand how such tests are related to the theories we need to examine different concepts of efficiency when there is conditioning information.

When there is conditioning information, minimum-variance efficiency may be defined in terms of the conditional means and variances (*conditionally efficient*), or in terms of unconditional moments. When the objective is to minimize the unconditional variance for a given unconditional mean, but where portfolio strategies may be functions of the information, we have (unconditional) *minimum-variance efficiency with respect to the conditioning information*.

Unconditional efficiency with respect to conditioning information may seem confusing because conditioning information may be employed by the portfolio, but unconditional expectations about that portfolio's returns are used to define efficiency. However, this information structure is actually quite common. Often the agent conducting a portfolio optimization uses more information than is available to the observer of the outcomes. If the observer does not have the conditioning information, he or she can only form unconditional, or less informed, expectations.

Dybvig and Ross (1985) provide an example. Consider a portfolio manager who is evaluated based on the unconditional mean and variance of the portfolio return. The manager may use conditioning information about future returns in forming the portfolio. They show that the manager's conditionally-efficient portfolio will typically not appear efficient to the uninformed investor. The portfolio that maximizes the manager's measured performance in this setting is the unconditionally-efficient portfolio with respect to the information.

Ferson and Siegel (2001) derive efficient-portfolio strategies with respect to conditioning information and illustrate their properties. Consider an example with two assets: a riskless asset (with gross rate of return R_f) and a risky asset with gross return, R . The risky asset's return is written as:

$$R = R_f + \mu(Z) + \varepsilon, \quad (17)$$

where $\mu(Z) = E(R - R_f \mid Z)$. The conditional variance of the return given Z , is $\Sigma(Z)$. The problem to be solved is:

$$\text{Min}_{x(Z)} \text{Var} \{x(Z)'R\} \quad \text{subject to :} \quad \mu_p = R_f + E \{x(Z)' [R - R_f]\}. \quad (18)$$

The weight function $x(Z)$ specifies the fraction invested in the risky asset, as a function of the conditioning information, Z . Here we provide a constructive derivation¹³. Let $\Lambda(Z)^{-1} = E\{(R - R_f)^2 \mid Z\} = \Sigma(Z) + \mu(Z)^2$ and write the Lagrangian:

$$L(x \mid Z) = E \left\{ x(Z)' \Lambda(Z)^{-1} x(Z) - 2\lambda [x(Z)' \mu(Z) - (\mu_p - R_f)] \right\}. \quad (19)$$

Let $\hat{w}(Z) = x(Z) + ay(Z)$, where $y(Z)$ is any function and $x(Z)$ is the optimal solution. If we consider $L(\hat{w}(Z) \mid Z) = L(x(Z) + ay(Z) \mid Z)$, then optimality of $x(Z)$ requires $\partial L / \partial a \big|_{a=0} = 0 = E[\{x(Z)' \Lambda(Z)^{-1} - \lambda \mu(Z)\} y(Z)] = 0$. Since this must hold for all functions $y(Z)$, it implies $x(Z)' \Lambda(Z)^{-1} - \lambda \mu(Z) = 0$. Solving for $x(Z)$ and evaluating λ by substituting the solution back into the constraint that $E\{x(Z)' \mu(Z)\} = (\mu_p - R_f)$ gives the solution for the unconditionally mean-variance efficient strategy with respect to the information, Z :

$$x(Z) = \frac{\xi^{-1} (\mu_p - R_f) \mu(Z)}{[\mu(Z)]^2 + \Sigma(Z)}, \quad (20)$$

where

$$\xi = E \left\{ \frac{\mu(Z)^2}{[\mu(Z)]^2 + \Sigma(Z)} \right\}. \quad (21)$$

The minimized variance implied by this solution is:

$$\sigma_p^2 = \frac{(\mu_p - R_f)^2 (1 - \xi)}{\xi} \quad (22)$$

Figure 1 gives an empirical example of the optimal weight as a function of the conditional expected excess return $\mu(Z)$, for a given unconditional mean μ_p , equal to 11.1% per year. This figure matches the Standard and Poors 500 stock index return for the 1963–1994 sample period. The example assumes homoskedasticity, where $\Sigma(Z)$ is a constant. The weight is shown for several values of R^2 , defined as the ratio of the variance of the conditional mean to the variance of the stock index return. As R^2 approaches zero the weight becomes a constant function.

When the conditional expected excess return of the risky asset is zero, the weight in the risky asset is zero. For conditional expected excess returns near zero, the efficient weight appears monotone and nearly linear in $\mu(Z)$. This is similar to other utility-maximizing strategies. For example, assuming a normal distribution the strategy that maximizes an exponential utility is linear in the conditional expected return. Kim and Omberg (1996) and Campbell and Viceira (1999) solve intertemporal portfolio

¹³ Thanks to Ludan Liu.

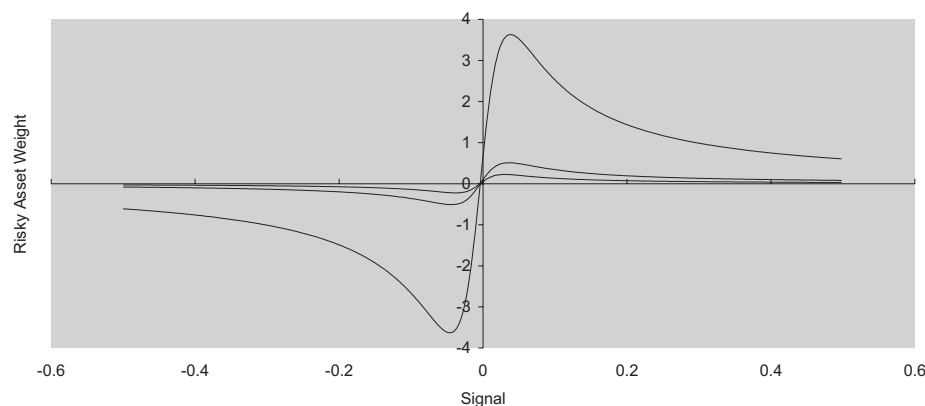


Fig. 1. Optimal weight versus signal. R^2 are 0.0045, 0.105 and 0.355.

problems and find that the portfolio weights are approximately linear in the state variables. Thus, traditional solutions to the portfolio optimization problem imply portfolio weights that are highly sensitive to extreme values of the signal. For example, if the signal is normally distributed a linear portfolio weight is unbounded.

The weight in Equation (20) satisfies $x(Z) \rightarrow 0$ as $\mu(Z) \rightarrow \pm \infty$. After a certain point, even an optimistic extreme signal leads to purchasing *less* of the risky asset, when the objective is to attain a given unconditional mean return with the smallest unconditional variance. Intuitively, an extremely high expected return presents an opportunity to reduce risk by taking a small position in the risky asset this period, without compromising the average portfolio performance¹⁴.

The “conservative” nature of the solution implies an interesting “agency” problem in a portfolio management context. The portfolio manager who is evaluated, as is common in practice, on the basis of unconditional mean return relative to unconditional return volatility may be induced to adopt a conservative response to extreme signals in order to maximize the measured performance.

2.6.1. Conditional versus unconditional efficiency

Hansen and Richard (1987) show that in the set of returns that can be generated using conditioning information, an unconditionally efficient strategy with respect to

¹⁴ The precise shape of the curve depends on the homoskedasticity assumption used in Figure 1. However, according to Equation (20), if there is heteroskedasticity, where an extreme value of the signal is associated with a large conditional variance, the conservative behavior of the strategy is reinforced. Ferson and Siegel (2001) show that the solution for an n -asset example also implies the portfolio weight is a bounded function of the signal. They also note that the graph of the unconditionally efficient portfolio weight is similar to the redescending influence curves used in robust statistics [e.g., Hampel (1974), Goodall (1983) and Carroll (1989)]. This suggests that the unconditionally efficient portfolios may be empirically robust.

the information must be conditionally efficient, but the reverse is not true. The relation between conditional and unconditional efficiency with respect to conditioning information may be understood in terms of the utility functions for which the solutions are optimal.

Ferson and Siegel (2001) show that an unconditionally-efficient portfolio with respect to the conditioning information maximizes the conditional expectation of a quadratic utility function in a single-period problem. Since quadratic-utility agents choose mean-variance-efficient portfolios, this implies that an unconditionally-efficient portfolio must be a conditionally mean-variance-efficient portfolio. However, other utility functions also lead to conditional mean-variance-efficient portfolios. One example is the exponential utility function previously mentioned, when returns are normally distributed conditional on the information. Ferson and Siegel show that the solution for the unconditionally efficient portfolio with respect to the information is unique, so the exponential utility agent chooses a conditionally mean-variance-efficient portfolio that is not unconditionally efficient with respect to the information. Thus, conditional efficiency does not imply unconditional efficiency with respect to the information. The unconditional efficient portfolios with respect to given conditioning information are a subset of the conditionally efficient portfolios with respect to the same information.

The relation between conditional and unconditional efficient portfolios with respect to given conditioning information can be represented as in Figure 2, with the unconditional mean return on the y axis and unconditional standard deviation on the x axis. The usual “fixed weight” mean-standard deviation boundary, which ignores the conditioning information, is the curve farthest to the right in the lower portion of the figure. There are an infinite number of conditionally efficient portfolio strategies, some examples of which are depicted by the other curves¹⁵. Some of the conditionally efficient strategies can plot inside the fixed-weight strategy that ignores the conditioning information, as shown by Dybvig and Ross (1985) and illustrated by one of the examples in the figure. The unconditional efficient strategy with respect to the information Z is the outer envelope of all the conditionally efficient strategies. This is shown as the left-most curve in Figure 2. Hansen and Richard (1987) provide a formal characterization and prove that the outer envelope in Figure 2 has the familiar properties associated with mean-standard deviation boundaries when there is no conditioning information, e.g., two-fund separation. See Ingersoll (1987).

¹⁵ To see that there are an infinite number, note that a conditionally minimum-variance efficient strategy solves:

$$\text{Min}_{x(Z)} \text{Var} [x(Z)'R \mid Z] \text{ s.t. } E [x(Z)'R \mid Z] = T(Z),$$

where $T(Z)$ is the target conditional mean return. Each of the infinite possible specifications for the function $T(Z)$ implies a conditionally efficient strategy.

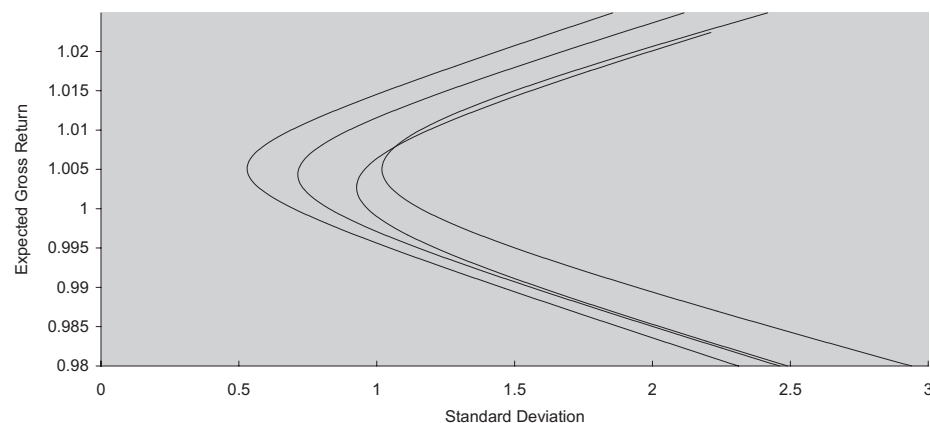


Fig. 2. Minimum variance boundaries.

2.6.2. Implications for tests

These concepts of minimum-variance efficiency have important implications for tests of asset-pricing models. In principle, we can devise tests to reject the hypothesis that a portfolio is unconditional efficient or efficient conditional on some observed instruments, but we can not tell if a portfolio is efficient given *all* the public information, Ω . If we interpret asset-pricing models as identifying which portfolios are conditionally efficient given Ω , we have a problem. The collection of minimum-variance portfolios, conditional on the market information set, Ω , is larger than the set of minimum-variance portfolios conditional on an observable subset of instruments, Z . Thus, even if we reject that a portfolio is efficient given Z , we can not infer that it is inefficient given Ω . This is similar to the Roll (1977) critique of tests of the CAPM. Roll pointed out that since the market portfolio of the CAPM can not be measured, the CAPM can not be tested without making assumptions about the unobserved market return. The problem here is that we can not test the conditional CAPM because the full information set Ω is not observed, unless we make assumptions about the unobserved information set. This problem is present even if the true market portfolio return could be measured¹⁶.

There is an important exception to this conundrum. When tests are based on Equation (2) it is possible to test the model without observing the complete information set, when m_{t+1} depends only on observable data and model parameters. Equation (2) implies that $E(m_{t+1}R_{t+1}|Z_t) = 1$, so tests may proceed using the observed instruments Z_t . This is the case, for example, in versions of the consumption-based asset-pricing model, when the relevant consumption can be measured. Given a model

¹⁶ See Wheatley (1989) for a critique of the earliest conditional asset-pricing studies based on similar logic.

in which m_{t+1} is a function of observed data and parameters, it is also possible to use the concept of unconditional efficiency with respect to given information, Z , to conduct tests of the model. This approach is developed in Ferson and Siegel (2002b).

2.7. Choosing the factors

A beta pricing model has no empirical content until the factors are specified, since there will always be a minimum-variance portfolio which satisfies Equation (13). The minimum-variance portfolio can serve as a single factor in Equation (12). Therefore, the empirical content of the model is the discipline imposed in selecting the factors.

There have been three main approaches to specifying empirical factors for multiple-beta asset-pricing models. One approach is to use statistical factor analytic or principal components methods. This approach is motivated by the APT, where the “right” factors are the ones that capture all the pervasive (unbounded eigenvalue) risk, leaving only nonsystematic (bounded eigenvalue) risk in the residuals. That approach is pursued by Roll and Ross (1980) and Connor and Korajczyk (1986), among others. The advantage of the Connor–Korajczyk approach is that the factor extraction is conducted under essentially the same large-markets assumptions that lead to the APT. This lends some rigor to the tests. The disadvantage is that purely statistical factors provide little economic intuition. Burmeister and McElroy (1988) augment statistical factors with a market portfolio and illustrate how to “rotate” the factors, to interpret them relative to more intuitive economic variables.

In a second approach the risk factors are explicitly chosen economic variables or portfolios, chosen based on economic intuition [e.g., Chen, Roll and Ross (1986), Ferson and Harvey (1991), Campbell (1993) and Cochrane (1996)]. Here is where Equation (4) should come in. According to that equation, the factors should be related to consumer wealth, consumption expenditures, and the sufficient statistics for the marginal utility of future wealth in an optimal consumption-investment plan.

A third approach for choosing factors uses the cross-sectional empirical relation of stock returns to firm attributes. For example, portfolios are formed by ranking stocks on firm characteristics that are observed to be correlated with the cross-section of average returns. Perhaps the most famous current example is the three-factor model of Fama and French (1993, 1996). Fama and French group common stocks according to their “size” (market value of equity) and their ratios of book value to market value of equity per share. Previous studies such as Keim (1983) and Reinganum (1981) found that stock returns are related to these attributes. Fama and French use the returns of small stocks in excess of large stocks, and the returns of high book-to-market in excess of low book-to-market stocks, as two “factors”. This approach is critiqued on methodological grounds in Section 4.2.

The empirical literature which examines multiple-beta pricing models is vast. Fama (1991), Connor and Korajczyk (1995) and Harvey and Kirby (1996) provide selective reviews. Studies typically focus on particular factors, and may mix the three approaches to factor selection. There is scant empirical evidence that focuses directly

on the general question: which of the three methods of factor selection is superior? The answer to this question depends on the application to which the multiple-beta model is put.

In their role as empirical models for security returns, multiple-beta models are used for essentially three things. First, they are used to explain the cross-section of average returns on different securities. This relates to Equation (6), where expected returns differ according to the return covariances with m_{t+1} . Second, the models are used to explain predictable patterns in security returns over time. This is the main goal of conditional asset pricing, as discussed in Section 2.1. Finally, multiple-beta models are used to explain the *contemporaneous* variance of security returns, through the variation of the risk factors. This relates more to multiple regression models, that are often associated with multiple-beta expected return models like Equation (9).

The cross section of expected returns is central for a number of applications. The models' fitted expected returns serve as estimates of "required" returns, in relation to risk. They are used, among other things, for the cost of equity capital, an important input in corporate project selection problems (see the surveys of Bruner et al. (1998) and Graham and Harvey (2001), and for portfolio construction and performance evaluation (see Section 5). There are problems in evaluating the three approaches to factor selection for this purpose. First, the results depend crucially on the "test assets", or portfolios for which the models are evaluated. If portfolios are formed to emphasize cross sectional variation in a particular dimension, thus de-emphasizing others¹⁷, then a model that "explains" that particular dimension will look good. For example, the Fama–French (1993) three-factor model emphasizes size and book-to-market. Fama and French (1996) find that it captures the cross section of average returns pretty well in size and book-to-market sorted portfolios. However, when confronted with industry returns [Fama and French (1997)] or with cross-sectional variation in average returns, related to the momentum effect of Jegadeesh and Titman (1993), the model performs poorly. This issue of portfolio formation has muddled some attempts in the literature to distinguish between the explanation of power of security characteristics versus betas on related factors, for the cross-section of average returns [e.g., Daniel and Titman (1997)]. See Berk (2000) for an analysis and critique.

The second problem in evaluating the methods of factor selection relates to the discussion of conditional and unconditional efficiency in Section 2.6.1. A model may identify a conditionally efficient portfolio, but the portfolio is unconditionally inefficient. In other words, conditional covariances with a portfolio return could provide an exact description of the cross section of conditional expected returns, while at the same time average returns are not explained by their unconditional covariances with the same portfolio return. To see this algebraically, take the unconditional expectation of Equation (10), and recall that the expectation of the conditional covariance differs

¹⁷ The total sum of squares in any sample must equal the across-group sum of squares plus the within-group sum of squares.

from the unconditional covariance by the covariance of the conditional means. If the cross section of assets i differ in their values of $\text{Cov}\{E_t(R_{i,t+1}), E_t(m_{t+1})\}$, we have a problem. Recall that $E_t(m_{t+1})$ is the inverse of the risk-free return. Evidence from Fama and Schwert (1977) and Ferson (1989) shows that different assets' expected returns have different sensitivities to measures of risk-free interest rates, so this problem may be a serious one.

Although these caveats make it difficult to interpret the evidence in relation to theory, it is still interesting to know which models provide a good empirical description for the cross section of average returns. A few studies compare the alternative approaches to factor selection from this perspective. Lehmann and Modest (1987) take no stand on which approach is superior, but observe that predicted expected returns for a sample of mutual funds can be very sensitive to using the CAPM versus the APT, as well as to different approaches for implementing the APT. Farnsworth et al. (2002) compare a collection of SDF models, including: (1) three-factor models based on the asymptotic principal components of Connor and Korajczyk (1986); (2) three traded economic factors relating to the stock market, government bonds and low-grade corporate bonds; and (3) the three-factor model of Fama and French (1993, 1996). They estimate the models in a common sample of nine primitive "assets", portfolios emphasizing variation in equity size, book-to-market and momentum, as well as bond market returns. They find that the principal components-based model is the worst-performing model in this group for explaining the cross section of average returns, as summarized by the Hansen–Jagannathan distance measure described in the next section.

Explaining predictability in security returns is another important and controversial application for multi-beta asset-pricing models. Much of the controversy relates to the interpretation. Fama (1970, 1991) emphasizes that evidence relating to market efficiency involves a "joint hypothesis". A model of equilibrium (essentially, a specification for the SDF) is jointly tested with the hypothesis that markets are informationally efficient with respect to particular information. If the tests reject, then logically the market could be inefficient or the SDF model could be wrong. From this perspective, predictability that cannot be explained using any of the standard asset-pricing models suggests market inefficiency; or alternatively, the need to move beyond the standard models.

A few studies have compared the alternative approaches to factor selection, for the purpose of explaining return predictability. Ferson and Korajczyk (1995) compare economic factors similar to those chosen by Chen, Roll and Ross (1986) with the asymptotic principal components of Connor and Korajczyk (1986). They study predictability in one-month to two-year returns based on a list of "standard" lagged instruments discussed above, estimating the fraction of the predictable variance of return that is captured by the models. They find that single-factor models can capture about 60% of the predictable variance in a sample of industry returns, while five-factor models capture about 80%. These results are not highly sensitive to the return horizon. The performance of a five-principal components model and a five prespecified-factor model are broadly similar for capturing predictability in returns for

all of the horizons. Farnsworth et al. (2002) find that, among the three-factor models, the Fama–French model performs the worst for explaining predictability in their study. Additional evidence that this model performs poorly for capturing return predictability is presented by Kirby (1998) and Ferson and Harvey (1999).

Factors with good contemporaneous explanatory power for security returns are useful for risk modeling and for controlling systematic variance in some research contexts. A regression of security returns on a selection of factors does not impose an asset pricing model unless the regression coefficients are restricted: examples are given in Section 4. But it is easier to draw general conclusions about the empirical performance of the methods of factor selection in this setting. In a given sample, a factor analytic approach constructs factors to be highly correlated with the asset returns. If in-sample, contemporaneous correlation is the goal, this approach almost has to be the most effective. Choosing economic variables is likely to be the worst approach, because security returns, and stock returns in particular, are only weakly correlated with most economic data [e.g., Roll (1988)]. Indeed, this low contemporaneous correlation is one motivation for the use of mimicking portfolios, described in Section 4, to replace factors based on economic data in empirical models.

3. Modern variance bounds

3.1. The Hansen–Jagannathan bounds

Hansen and Jagannathan (HJ, 1991) showed how the fundamental asset pricing Equation (1) places restrictions on the mean and variance of m_{t+1} . These restrictions depend only on the sample of assets, and thus provide a diagnostic tool for comparing different models of m_{t+1} . If a candidate for m_{t+1} , corresponding to a particular theory, fails to satisfy the HJ bounds, then it can not satisfy Equation (1). Recent papers refine and extend the HJ bounds in several directions, and a number of papers and textbooks provide basic reviews [see Ferson (1995)]¹⁸. We briefly review the case where there is no conditioning information, then move on to extensions with conditioning information.

Assume that the random column n -vector R of the assets' gross returns has mean $E(R) = \mu$ and covariance matrix Σ . When there is no conditioning information

¹⁸ Snow (1991) considers selected higher moments of the returns distribution. Bansal and Lehmann (1997) derive restrictions on $E[\ln(m)]$ that involve all higher moments of m and reduce to the HJ bounds if returns are lognormally distributed. Balduzzi and Kallal (1997) incorporate the implications for the risk premium on an economic variable. Cochrane and Hansen (1992) state restrictions in terms of the correlation between the stochastic discount factor and returns, while Cochrane and Saa'-Requejo (2000) consider bounds on the Sharpe ratios of assets' pricing errors. Hansen, Heaton and Luttmer (1995) develop asymptotic distribution theory for specification errors on stochastic discount factors, where the HJ bounds are a special case, and Ferson and Siegel (2002a) evaluate these standard errors by simulation.

a stochastic discount factor is defined as any random variable m such that $E(mR) = \underline{1}$. Hansen and Jagannathan (1991) show that the stochastic discount factor with minimum variance for its expectation $E(m)$ is given by:

$$m^* = E(m) + [\underline{1}' - E(m) E(R)'] \Sigma^{-1} [R - \mu], \quad (23)$$

and the minimum variance for a SDF is the variance of m^* :

$$\text{Var}(m) \geq [\underline{1}' - E(m) E(R)'] \Sigma^{-1} [\underline{1} - E(m) E(R)]. \quad (24)$$

The proof is instructive. Consider a regression of any m satisfying $E(mR) = \underline{1}$ on the asset returns, R . The fitted value is $m^* = E(m) + \text{Cov}(m, R) \Sigma^{-1} [R - \mu]$, and $m = m^* + \varepsilon$, where ε is the regression error satisfying $E(\varepsilon) = 0 = E(\varepsilon R)$. Since m^* is a linear function of R , it follows that $E(\varepsilon m^*) = 0$. Thus, $\text{Var}(m) = \text{Var}(m^*) + \text{Var}(\varepsilon) \geq \text{Var}(m^*)$. Finally, expanding $E(mR) = 1 = E(m) \mu + \text{Cov}(m, R)$ and substituting for $\text{Cov}(m, R)$, we arrive at Equation (23). The right-hand side of Equation (24) is just the variance of the m^* in Equation (23).

The HJ bound is related to the maximum Sharpe ratio that can be obtained by a portfolio of the assets. The Sharpe ratio is defined as the ratio of the expected excess return to the standard deviation of the portfolio return. If the vector of assets' expected excess returns is $\mu - E(m)^{-1}$ and Σ is the covariance matrix, the square of the maximum Sharpe ratio is $[\mu - E(m)^{-1}]' \Sigma^{-1} [\mu - E(m)^{-1}]$. Thus, from Equation (24) the lower bound on the variance of stochastic discount factors is the maximum squared Sharpe ratio multiplied by $E(m)^2$. The larger is the maximum squared Sharpe ratio for a given $E(m)$, the tighter is the bound on $\text{Var}(m)$ and the more potential SDFs can be ruled out.

The Hansen and Jagannathan (1991) region for $\{E(m), \sigma(m)\}$ is given by the square root of Equation (24). The boundary of this region is the minimum value of the standard deviation, $\sigma(m)$, for each value of $E(m)$. Some empirical examples are illustrated in Figure 3, corresponding to the different versions of the bounds described below. The bounds are drawn for quarterly data similar to Hansen and Jagannathan, consisting of 3-, 6-, 9- and 12-month Treasury bill returns for the 1964–1986 sample period. For a given hyperbola in Figure 3, as we vary $E(m)$ we move around the $\{E(m), \sigma(m)\}$ boundary. In order for an SDF to satisfy $E(mR) = \underline{1}$, its mean and standard deviation must plot above the boundary, “inside the cup”. The points shown by the “×” symbols are the sample means and standard deviations of the m_{t+1} of Equation (8), using quarterly total consumption data per capita in the USA over the same sample period, and various values of the relative risk aversion, α . Note that the SDF does not plot inside even the lowest cup for many values of α . In fact, the SDF just touches the boundary of Figure 3 when $\alpha = 71$. The SDF does not enter the highest cups for any value of risk aversion. The simple consumption model does not produce SDF's that are volatile enough. This is a version of the equity premium puzzle

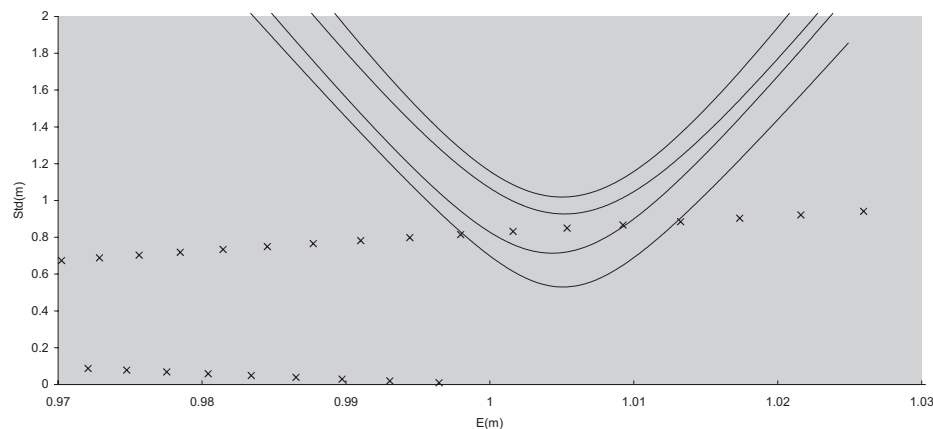


Fig. 3. Hansen–Jagannathan bounds.

of Mehra and Prescott (1985), reviewed by Mehra and Prescott in Chapter 14 of this volume.

The bound in Equation (24) is not the sharpest lower bound on $\sigma(m)$ that can be derived. Hansen and Jagannathan (1991) show how imposing that m_{t+1} is a strictly positive random variable can sharpen the bound. Computing the bounds imposing positivity requires a numerical search procedure. Another way to sharpen the bounds is through the use of conditioning information.

3.2. Variance bounds with conditioning information

The preceding analysis is based on the unconditional moments. With conditioning information Z , we may consider a stochastic discount factor for (R, Z) to be any random variable m such that $E(mR \mid Z) = \underline{1}$ for all realizations of Z . In principle, everything above could be stated for conditional means and variances, an approach pursued by Gallant, Hansen and Tauchen (1990). This would complicate Figure 3, because we would have to show a new one for each realization of Z_t . Alternatively, Hansen and Jagannathan (1991) describe a clever way to extend the analysis to partially exploit the information in a set of lagged instruments, while using unconditional moments to describe the bound. Equation (5) implies that for any set of instruments $E\{m_{t+1}r_{i,t+1} \otimes Z_t \mid Z_t\} = 0$, and therefore $E\{m_{t+1}r_{i,t+1} \otimes Z_t\} = 0$, where \otimes is the Kronecker product. If we view $\{r_{i,t+1} \otimes Z_t\}$ as the excess returns to a set of “dynamic” trading strategies, the preceding analysis goes through essentially unchanged. (The trading rule holds at time t , Z_t units of the asset i long and Z_t units of the zero-th asset short).

The approach of Hansen and Jagannathan (1991) is just one way to implement HJ bounds that use conditioning information. To understand how alternative ap-

proaches to conditioning information can refine the bounds let r_{t+1} be the vector of excess returns. In this case, Equation (5) implies the following:

$$E \{m_{t+1}r_{t+1}f(Z_t)\} = 0 \quad \text{for all functions } f(\cdot), \quad (25)$$

where the unconditional expectation is assumed to exist. In other words, if we consider $r_{t+1}f(Z_t)$ to represent a possible dynamic trading strategy, then the presence of conditioning information Z_t says that m_{t+1} should price all the dynamic trading strategies, not just $r_{t+1} \otimes Z_t$. The larger is the set of strategies for which the Equation (25) is required to hold, the smaller is the set of m_{t+1} 's that can satisfy the condition, and the tighter are the bounds. This is the motivation for extending HJs original approach, in order to use the information efficiently.

Three versions of HJ bounds with conditioning information have appeared in the literature. These may be understood through Equation (25). First are the *multiplicative bounds* of Hansen and Jagannathan (1991), who choose $f(\cdot)$ to be the linear function, $I \otimes Z_t$. Second are the *efficient portfolio bounds* of Ferson and Siegel (2002a), where $f(\cdot)$ is the set of portfolio weights that may depend on Z_t and sum to 1. Finally, the *optimal bounds* of Gallant, Hansen and Tauchen (1990) require Equation (25) to hold for all functions $f(\cdot)$.

3.2.1. Efficient portfolio bounds

Efficient portfolio bounds are based on the unconditionally efficient portfolios with respect to the information Z , derived by Ferson and Siegel (2001) and discussed in Section 2.6. Since these portfolios maximize the Sharpe ratio, over all dynamic strategies $x(Z)$ whose weights sum to 1.0, they efficiently use the information in Z to tighten the bounds. For given (R, Z) , the solutions describe an unconditional mean-standard deviation boundary, as depicted in Figure 2. Fixed-weight combinations of any two portfolios on an unconditional mean-standard-deviation boundary can describe the entire boundary [Hansen and Richard (1987)]. Thus, efficient portfolio bounds can be formed from two "arbitrary" portfolios from the boundary.

3.2.2. Optimal bounds

Gallant, Hansen and Tauchen (1990) derive optimal bounds that do not restrict to portfolio functions, with weights that sum to 1.0. The solution for the optimal bounds is presented in Ferson and Siegel (2002a) as follows. First, define the following conditional portfolio constants, which are analogous to the efficient-set constants used in the traditional mean-variance analysis [see, e.g., Ingersoll (1987)]:

$$\begin{aligned} \alpha(Z) &= \mathbf{1}'\Sigma(Z)^{-1}\mathbf{1}, \\ \beta(Z) &= \mathbf{1}'\Sigma(Z)^{-1}\mu(Z), \quad \text{and} \\ \gamma(Z) &= \mu(Z)'\Sigma(Z)^{-1}\mu(Z), \end{aligned} \quad (26)$$

where $\mu(Z)$ and $\Sigma(Z)$ are the conditional mean and variance functions. The stochastic

discount factor m for (R, Z) with minimum variance for its expectation $E(m)$ is given by

$$m^*(Z) = \zeta(Z) + [\mathbf{1} - \kappa(Z) \mu(Z)]' \Sigma(Z)^{-1} [R - \mu(Z)], \quad (27)$$

where $\zeta(Z)$ is the conditional mean of m given Z , defined as:

$$\begin{aligned} \zeta(Z) &= E(m \mid Z) \\ &= \{1 + \gamma(Z)\}^{-1} \left\{ \beta(Z) + E \left[\{1 + \gamma(Z)\}^{-1} \right] \left[E(m) - E \left(\frac{\beta(Z)}{\{1 + \gamma(Z)\}} \right) \right] \right\}. \end{aligned} \quad (28)$$

The unconditional variance of $m^*(Z)$ is:

$$\begin{aligned} \text{Var}(m^*(Z)) &= E \left[\{1 + \gamma(Z)\}^{-1} \right]^2 \left[E(m) - E \left(\frac{\beta(Z)}{\{1 + \gamma(Z)\}} \right) \right]^2 \\ &\quad + E[\alpha(Z)] - [E(m)]^2 - E \left(\frac{\beta(Z)^2}{\{1 + \gamma(Z)\}} \right). \end{aligned} \quad (29)$$

Equation (29) may be used directly to compute the optimal HJ bounds. To implement the bound, it is necessary to specify the conditional mean function $\mu(Z)$ and the conditional variance function. Then, the four unconditional expectations that appear in Equation (29) may be estimated from the corresponding sample means, independent of the value of $E(m)$.

3.2.3. Discussion

For given conditioning information, Z , the optimal bounds provide the greatest lower bound on stochastic discount factors and thus the highest, most restrictive cup. The efficient portfolio bounds incorporate an additional restriction to functions that are portfolio weights, which sum to 1.0 at each date. This reduces the flexibility of the efficient portfolio bounds to exploit the conditioning information, and thus they do not attain the greatest lower bound. Intuitively, suppose there was only one asset. Then the restricted weight could not respond at all to the conditioning information. The multiplicative bound of Hansen and Jagannathan (1991) does not restrict portfolio weights, but neither does it attempt to use the conditioning information efficiently. Bekaert and Liu (2002) and Basu and Stremme (2003) further discuss the relations among the HJ bounds with conditioning information.

Ferson and Siegel (2002a) conduct a simulation study of HJ bounds with conditioning information. They find that sample values of the bounds are upwardly biased, the bias becoming substantial when the number of assets is large relative to the number of time-series observations. This means that studies using the biased bounds run a risk of rejecting too many models for the stochastic discount factor. They derive a finite-sample adjustment for the bounds and show that it helps control the bias.

Sarkissian (2002) uses the adjustments and finds that they change the inferences about international consumption-based asset-pricing models.

The evidence to date leads to several conclusions. First, the Multiplicative bounds of HJ can be terribly biased in realistic samples. It is important to use the finite-sample adjustment to their location. Second, the Optimal bounds of Gallant, Hansen and Tauchen (1990) are more difficult to implement than the multiplicative bounds, requiring the specification of conditional means and variances, but they are the tightest bounds. While not as biased as the multiplicative bounds, they are significantly biased in some finite samples. The finite-sample adjustment should be used and improves their accuracy. Third, the efficient-portfolio bounds are similar in complexity to the optimal bounds, also requiring the specification of the conditional moments. However, unlike the optimal bounds, they remain valid (but inefficient) when these moments are not specified correctly. They have the smallest sampling error variances of all the bounds with conditioning information. They are not as biased as either the optimal or multiplicative bounds, but finite-sample adjustment is still useful.

3.3. The Hansen–Jagannathan distance

Hansen and Jagannathan (1997) develop measures of misspecification for models of the stochastic discount factor. They consider $m(\varphi)$, a “candidate” stochastic discount factor, as may be proposed by an asset-pricing model. Since the candidate SDF is misspecified, then $E(m(\varphi)R - 1|Z) \neq 0$. They propose a measure for how “close” is $m(\varphi)$ to a stochastic discount factor that “works”. We first consider the case where there is no conditioning information.

It is easy to show that a particular SDF, formed from the asset returns, $m^* = [1'E(RR')^{-1}]R$, is one that “works” for pricing R . Hansen and Jagannathan measure how close $m(\varphi)$ is to m^* . They do this by first projecting $m(\varphi)$ on the returns R to get the fitted value $\hat{m} = [E(m(\varphi)R')E(RR')^{-1}]R$. They then measure the mean square distance between the fitted values and m^* . This is the HJ Distance Measure:

$$\text{HJD} = E \left\{ (\hat{m} - m^*)^2 \right\}, \quad (30)$$

where the sample averages are used in practice to estimate the expectations.

Note that $\hat{m} - m^* = [E(m(\varphi)R) - 1]E(RR')^{-1}R$, so we may write HJD as: $[E(m(\varphi)R) - 1]E(RR')^{-1}[E(m(\varphi)R) - 1]$. This leads to a couple of nice interpretations. First, if we let $g = E(m(\varphi)R) - 1$, $W = E(RR')^{-1}$, then $\text{HJD} = g'Wg$ is Hansen’s J -test with a particular W , described in the next section. Second, by analogy with the T^2 test, HJD measures the “most mispriced” return. To see this interpretation, recall that $g = \alpha_m = E\{m(\varphi)R - 1\}$ is a measure of expected pricing error using $m(\varphi)$. The alpha of a portfolio with weight vector x is the scalar $\alpha_p = x'\alpha_m$. Consider the problem of finding the absolutely most mispriced portfolio, relative to its second moment:

$$\text{Max}_x 2x'\alpha + \lambda \{x'E(RR')x - E(r_p^2)\},$$

where λ is a Lagrange multiplier. The maximized value of $\alpha_p^2/E(r_p^2)$ is $\alpha'E(RR')^{-1}\alpha$, equivalent to the HJD measure.

When there is conditioning information in the form of lagged instruments, Z , then a correctly specified SDF has $E(m(\varphi)R - 1|Z) = 0$, which implies $E[m(\varphi)(R \otimes Z) - (1 \otimes Z)] = E(0 \otimes Z) = 0$. Let $\dot{z} \equiv Z./E(Z)$, where $./$ denotes element-by-element division. The previous equation holds only if $E[m(\varphi)(R \otimes \dot{z}) - (1 \otimes \dot{z})] = 0$, or $E[m(\varphi)(R \otimes \dot{z})] = \underline{1}$. If we define $\hat{R} \equiv R \otimes \dot{z}$, we have $E[m(\varphi)\hat{R}] = \underline{1}$, and we can proceed as before using \hat{R} instead of R .

4. Methodology and tests of multifactor asset-pricing models

The method of moments is briefly reviewed as a general way to test models based on Equation (2). This general framework is then specialized to discuss various tests of asset-pricing models. The special cases include cross-sectional regressions and multivariate regressions.

4.1. The Generalized Method of Moments approach

Let x_{t+1} be a vector of observable variables. Given a model which specifies $m_{t+1} = m(\theta, x_{t+1})$, estimation of the parameters θ and tests of the model can proceed under weak assumptions, using the Generalized Method of Moments (GMM), as developed by Hansen (1982). Define the model error term:

$$u_{i,t+1} = m(\theta, x_{t+1}) R_{i,t+1} - 1. \quad (31)$$

Suppose that we have a sample of N assets and T time periods. Combine the error terms from Equation (31) into a $T \times N$ matrix \mathbf{u} , with typical row u'_{t+1} . Equation (2) and the model for m_{t+1} imply that $E(u_{i,t+1} | Z_t) = 0$ for all i and t , and therefore $E(u_{t+1} \otimes Z_t) = 0$ for all t . The condition $E(u_{t+1} \otimes Z_t) = 0$ says that u_{t+1} is orthogonal to Z_t , and is therefore called an *orthogonality condition*. Define an $N \times L$ matrix of sample mean orthogonality conditions: $\text{vec}(\mathbf{Z}'\mathbf{u}/T)$, where Z is a $T \times L$ matrix of observed instruments with typical row Z'_t , a subset of the available information at time t ¹⁹.

Hansen's (1982) GMM estimates of θ are obtained as follows. Search for parameter values that make g close to zero by minimizing a quadratic form $g'Wg$, where W is a fixed $NL \times NL$ weighting matrix. Hansen (1982) shows that the estimators of θ that minimize $g'Wg$ are consistent and asymptotically normal, for any fixed W . If W is chosen to be the inverse of a consistent estimate of the covariance matrix of the orthogonality conditions, g , the estimators are asymptotically efficient in the class of

¹⁹ The $\text{vec}(\cdot)$ operator stacks the columns of a matrix. We assume that the same instruments are used for each of the asset equations. In general, each asset equation could use a different set of instruments, which complicates the notation.

estimators that minimize $g'Wg$ for fixed W 's. The asymptotic variance matrix of the GMM estimator of the parameter vector is then:

$$\widehat{\text{Cov}}(\theta) \approx \left[T \left(\frac{\partial g}{\partial \theta} \right)' W \frac{\partial g}{\partial \theta} \right]^{-1}. \quad (32)$$

where $\partial g/\partial \theta$ is an $NL \times \dim(\theta)$ matrix of derivatives. Hansen (1982) also shows that $J = Tg'Wg$ is asymptotically chi-square distributed, with degrees of freedom equal to the difference between the number of orthogonality conditions NL and the number of parameters, $\dim(\theta)$. This is Hansen's J -statistic, mentioned in the last section, which serves as a goodness-of-fit statistic for the model. Several choices for the weighting matrix W are available. A simple version of the optimal choice, where $W = \widehat{\text{Cov}}(g)^{-1}$, is:

$$\widehat{\text{Cov}}(g) = \frac{1}{T} \sum_t (g_t g_t') = \left[\frac{1}{T} \sum_t (u_{t+1} u_{t+1}') \otimes (Z_t Z_t') \right], \quad (33)$$

and \otimes denotes the Kronecker product. This case applies when the error terms u_t , and therefore the moment conditions g_t are serially uncorrelated. More general cases, and more detailed reviews are available in Hamilton (1994), Ferson (1995), Harvey and Kirby (1996) and Cochrane (2001), among others.

4.2. Cross-sectional regression methods

Much of the early empirical work on asset pricing used cross-sectional regressions of returns on estimates of market betas [e.g., Lintner, reported in Douglas (1969)]. The approach remains popular. Multiple-beta models, in particular, are often studied using this technique [e.g., Chen, Roll and Ross (1986), Ferson and Harvey (1991), Fama and French (1993), Lettau and Ludvigson (2001)]. Cross-sectional regression is appealing because it is an intuitive approach. Taking the simple CAPM as an example, we hypothesize: $E(R_i) = R_f + \beta_i E(R_m - R_f)$, $i = 1, \dots, N$. The model implies that the cross-sectional relation between mean returns and betas has a slope equal to the expected excess return of the market. The intercept should be a risk-free return, or a zero-beta portfolio expected return. Let's start our discussion of cross-sectional regressions with a classical two-step approach, similar to that of Black, Jensen and Scholes (1972) or Fama and MacBeth (1973).

For the first step, suppose that market betas are constant over time. The betas come from:

$$r_{it} = a_i + r_{mt}' \beta_i + \varepsilon_{it}, \quad t = 1, \dots, T \quad \text{for each } i, \quad (34)$$

For now we ignore the estimation error in the time series estimates of beta. This will be discussed later. The second step is a cross-sectional regression for each month:

$$r_{it} = \lambda_{0t} + \lambda_{1t}' \beta_i + u_{it}, \quad i = 1, \dots, N. \quad (35)$$

There could be $K > 1$ betas if we are testing a multi-factor asset-pricing model, then r_{mt} is a vector of K excess returns, and λ_{1t} is a K -vector of slope coefficients.

It is instructive to consider the GMM solution to the cross-sectional regression estimator. Define $g_t' = (1/N) \sum_i (r_{it} - \lambda_{0t} - \lambda_{1t}' \beta_i) \otimes (1, \beta_i')$. Choose the parameters to minimize $g_t' W g_t$ for some weighting matrix, W . Here the model is exactly identified, with the same number of parameters as moment conditions, so the GMM solution may be obtained by setting $g_t = 0$. This results in:

$$\begin{aligned} \hat{\lambda}_{0t} &= \frac{1}{N} \sum_i r_{it} - \hat{\lambda}_{1t}' \frac{1}{N} \sum_i \beta_i \\ \hat{\lambda}_{1t} &= \sum_i (\beta_i \beta_i')^{-1} \sum_i \beta_i (r_{it} - \hat{\lambda}_{0t})', \end{aligned} \quad (36)$$

Iteratively solving Equation (36) yields estimates of the premium for market beta, λ_{1t} , and the zero-beta return, λ_{0t} , similar to Black, Jensen and Scholes (1972). If the risk-free rate is known, we have a cross-sectional regression of excess returns on betas.

Of course, we want to use the cross-sectional regression to test hypotheses on the coefficients. For example, the hypothesis that $E(\lambda_{1t}) = 0$ says that the expected market risk premium is zero, or that beta has no cross-sectional explanatory power for returns. Alternatively, we may hypothesize that $E(\lambda_{1t} - r_{mt}) = 0$, which says that the premium is the market excess return. Standard errors for the coefficients may be obtained from Equation (36), which implies that

$$\hat{\lambda}_{1t} - \lambda_{1t} = \sum_i (\beta_i \beta_i')^{-1} \sum_i (\beta_i' u_{it}),$$

so that

$$\begin{aligned} \text{Var}(\hat{\lambda}_{1t} - \lambda_{1t}) &= \sum_i (\beta_i \beta_i')^{-1} \text{Var}(\sum_i \beta_i u_{it}) \sum_i (\beta_i \beta_i')^{-1} \\ &= (B'B)^{-1} [B' \text{Cov}(u_t) B] (B'B)^{-1}, \end{aligned} \quad (37)$$

where B is the $N \times K$ matrix of betas. Note that the variance of the estimators given by Equation (37) is not the same as the OLS solution, $\sigma_{ut}(B'B)^{-1}$, where σ_{ut} is a scalar, that one would obtain using a standard regression package to run a cross-sectional regression. Only in the special case where $\text{Cov}(u_t) = \sigma_{ut} I_n$, are the OLS standard errors correct. This would occur if the cross-sectional regression errors were uncorrelated across assets and homoskedastic across assets – a very unlikely scenario for stock market return data.

4.2.1. The Fama–MacBeth approach

Fama and MacBeth (1973) devise a simple and clever way to get estimates of the standard errors, while accounting for cross-sectional dependence. They suggest using the time-series of the estimators from a sequence of cross-sectional regressions, one for each month in the sample, to compute the standard error of the mean

coefficient. In testing the hypothesis that $\lambda_{1t} = 0$, they propose a simple t -ratio: $(1/T) \sum_t \hat{\lambda}_{1t} / \text{se}((1/T) \sum_t \hat{\lambda}_{1t})$, where the standard error is estimated by:

$$\text{se} \left(\frac{1}{T} \sum_t \hat{\lambda}_{1t} \right) \approx \frac{1}{\sqrt{T}} \left[\frac{1}{T} \sum_t \left[\hat{\lambda}_{1t} - \frac{1}{T} \sum_t \hat{\lambda}_{1t} \right]^2 \right]^{1/2}. \quad (38)$$

We can evaluate this using the previous equations. First, the sample variance of $\hat{\lambda}_{1t}$ is examined under the null hypothesis that λ_{1t} is zero.

$$\begin{aligned} E \frac{1}{T} \sum_t \left[\hat{\lambda}_{1t} - \frac{1}{T} \sum_t \hat{\lambda}_{1t} \right]^2 &= E \frac{1}{T} \sum_t \left[\left\{ \lambda_{1t} + (B' B)^{-1} B' u_t \right\} - \frac{1}{T} \sum_t \left\{ \lambda_{1t} + (B' B)^{-1} B' u_t \right\} \right]^2 \\ &= E \frac{1}{T} \sum_t \left[(B' B)^{-1} B' \left(u_t - \frac{1}{T} \sum_t u_t \right) \right]^2, \quad \text{if } \lambda_{1t} = 0, \\ &= E (B' B)^{-1} B' \left[\frac{1}{T} \sum_t \left(u_t - \frac{1}{T} \sum_t u_t \right) \left(u_t - \frac{1}{T} \sum_t u_t \right)' \right] B (B' B)^{-1} \\ &= (B' B)^{-1} B' [\text{Cov}(u)] B (B' B)^{-1}, \end{aligned}$$

which is the same as Equation (37). If we assume the $\hat{\lambda}_{1t}$ are uncorrelated over time, with a constant variance, we have $\text{Var}\{(1/T) \sum_t \hat{\lambda}_{1t}\} = (1/T)^2 \sum_t \text{Var}\{\hat{\lambda}_{1t}\}$. Estimating $\text{Var}\{\hat{\lambda}_1\}$ with the sample variance of the time series of the $\hat{\lambda}_{1t}$, as in Equation (38), produces the correct result. Thus, if we ignore estimation error in the betas, and assume that stock returns are serially uncorrelated, then under the null hypothesis that $\lambda_{1t} = 0$, the Fama–MacBeth approach delivers the correct standard errors.

4.2.2. Interpreting the estimates

Fama (1976) provides an intuitive interpretation of the cross-sectional regression estimators as portfolio returns. To fix things, start with the cross-sectional regression $R_{it} = \lambda_{0t} + \lambda_{1t} \beta_i + u_{it}$, $i = 1, \dots, N$. Let there be a single beta ($K = 1$). The CAPM implies that $E(\lambda_{1t}) = E(R_{mt} - R_{ft})$ and $E(\lambda_{0t} - R_{ft}) = 0$. Under the standard assumptions that make OLS best linear unbiased, the cross-sectional estimator solves:

$$\begin{aligned} \hat{\lambda}_{1t} \text{ solves } \quad & \text{Min}_{\{\hat{\lambda}_{1t}\}} \sum_i \hat{u}_{it}^2, \quad \text{subject to :} \\ & \text{Unbiased : } E(\hat{\lambda}_{1t}) = \lambda_{1t} \\ & \text{Linear : } \hat{\lambda}_{1t} = \sum_i w_i R_{it}. \end{aligned} \quad (39)$$

We can use these conditions to characterize $\hat{\lambda}_{1t}$ as a portfolio²⁰. In particular:

$$E(\hat{\lambda}_{1t}) = E(\sum_i w_i R_{it}) = E(\sum_i w_i [\lambda_{0t} + \lambda_{1t} \beta_i + u_{it}]) = \lambda_{1t}$$

²⁰ Since u_{it} is likely to be correlated across assets, GLS is better in theory. This amounts to a transformation of the asset returns and their betas into a different set of portfolios, then running OLS on the new portfolios. Therefore, the intuition here will translate.

implies $(\sum_i w_i) = 0$, and $(\sum_i w_i \beta_i) = 1$. This shows that the portfolio has weights, $\{w_i\}$ on the assets which sum to zero, and has a beta equal to one²¹. The first condition, $(\sum_i w_i) = 0$, says that the return is an excess return. The second condition, $(\sum_i w_i \beta_i) = 1$, says that the portfolio beta must equal 1.0. In order for the weights to sum to zero while the beta is positive, the portfolio must be “long” (positive weights) in high beta securities, and also “short” (negative weights) in low beta securities.

A similar analysis restricts the intercept estimator, implying that $(\sum_i w_i) = 1$ and $(\sum_i w_i \beta_i) = 0$. Thus, the intercept is a fully invested portfolio with no “systematic”, or factor-related risk. Its expected return should therefore be the zero-beta rate. If there is a risk-free security, this should be the risk-free rate.

The Fama–MacBeth cross-sectional regression coefficients represent one way to obtain the excess returns on mimicking portfolios for the risk factors. This is especially useful if the factors are not traded excess returns. Regression betas on nontraded variables, such as consumption, GNP growth or inflation, can be used. In this case, the Fama–MacBeth coefficients deliver excess returns, whose expected values are risk premiums for the factors. Indeed, the preceding analysis goes through if the cross-sectional regressors are not betas. For example, studies have used attributes such as firm size, dividend yield, or book-to-market ratio in place of beta coefficients.

4.2.3. A caveat

The Fama–MacBeth procedure constructs a “factor-mimicking” portfolio, for anything that we put on the right-hand side of the regression. This raises a potentially serious caveat. If a firm attribute is used that represents an anomaly, even if completely unrelated to risk, the procedure can deliver a mimicking portfolio return that may appear to work as a risk factor. This caveat is explored by Ferson, Sarkissian and Simin (1999). For a simple illustration, consider the following hypothetical regression:

$$R_{it} = \lambda_{0t} + \lambda_{1t} \alpha_i + u_{it}, \quad i = 1, \dots, N, \quad (40)$$

where α_i is an anomaly in the average return of asset i . Let $A_i \equiv \alpha_i - (1/N) \sum_i \alpha_i$, then the OLS Fama–Macbeth slope estimator constructs the portfolio:

$$R_{pt} = \hat{\lambda}_{1t} = \sum_i w_i R_{it}, \quad w_i = \frac{A_i}{\sum_i A_i^2}. \quad (41)$$

Suppose we used R_{pt} as a “factor” in an asset-pricing model. Would it appear to “price”, i.e., would returns be linear in covariances with R_{pt} ?

$$\text{Cov}(R_t, R_{pt}) = \text{Cov}(R) w = \frac{\text{Cov}(R) A}{\sum_i A_i^2}, \quad (42)$$

²¹ This condition would also apply in a multi-beta context, in which case the coefficient for a particular beta is a portfolio return with unit beta on the particular factor. Unbiasedness would also imply that the beta on the other factors equal zero, so the portfolio targets only the risk as represented by the factor in question.

where \underline{A} is the N -vector of the A_i 's. If $\text{Cov}(R)\underline{A} \propto \underline{A}$, then the vector of covariances with $R_{pt} = \hat{\lambda}_{1t}$ will appear to “explain” the cross-section of expected asset returns. For example: suppose $\text{Cov}(R) = I$. Then, returns are independent and there is no systematic risk. But, $\text{Cov}(R)\underline{A} \propto \underline{A}$, so the “factor” R_{pt} , formed by the Fama–MacBeth approach, will appear to work perfectly, in the sense that covariances with the factor return will exactly explain the cross-section of expected returns!

Similar results should be obtained when “spread” portfolios replace the FM coefficients, as in Fama and French (1993, 1996). Spread portfolios are formed as the difference between a high-attribute portfolio and a low-attribute portfolio return. Thus, they are long the high-attribute stocks, short the low-attribute stocks, and their weights sum to zero. A cross-sectional regression coefficient for stock returns on the attribute is also a linear combination of the returns, with weights that sum to zero. The portfolio is long in high-attribute stocks and short in low-attribute stocks. If a multiple regression is used, it has zero exposure to the other regressors. Subject to these conditions, it has minimum variance. A spread portfolio has a similar property if multiple independent sorts are used, as in Fama and French (1996), to control for the other attributes. While a spread portfolio does not explicitly minimize variance subject to these conditions, it avoids estimation error. Ferson, Sarkissian and Simin (1999) provide an example where Fama–MacBeth coefficients and spread portfolios, similar to Fama and French (1996), produce similar results in the face of an anomaly in asset returns. Their example shows that an arbitrary attribute, bearing an anomalous relation to returns, can be repackaged as a spurious risk factor.

Recent studies employing the approach of Fama and French (1996) do not use arbitrary anomalous attributes. Some of the most empirically powerful characteristics for the cross-sectional prediction of returns are ratios, with market price per share in the denominator. Berk (1995) emphasizes that the price of any stock is the value of its future cash flows discounted by future returns, so an anomalous pattern in the cross-section of returns would produce a corresponding pattern in book-to-market ratios or other proxies of cash-flow-to-price. A cross-sectional regression of returns on these ratios will pick out the anomalous patterns. Thus, the use of valuation ratios such as book-to-market as a sorting criterion increases the risk of creating a spurious risk factor.

In the real world, empirically measured attributes may be correlated with systematic risk and also with anomalous patterns in return. The net result of the two effects, risk versus anomaly, is complicated and model specific. However, equity market databases are inherently unbalanced panels, with more stocks than quarters or months. As new data on equity attributes becomes widely accessible, more studies will sort securities according to their attributes. The important caveat is that sorting procedures are subtle and easily abused. More work is needed to improve our understanding of the properties of such approaches.

4.2.4. Errors-in-betas

When the cross-sectional regression uses betas that are measured with error, two main issues arise. First, the cross-sectional regression coefficients suffer from a classical “attenuation bias”. Second, the standard errors are biased. Early studies that used cross-sectional regression also used portfolio grouping procedures, attempting to minimize errors in the betas, and to ensure that the remaining errors were uncorrelated with the other error terms in the model. More recently, empirical studies have taken to sorting stocks in order to accentuate some anomaly in the data, such as firm size, book-to-market, etc., in order to “challenge” the asset pricing model more forcefully. Thus, concerns about errors in the betas remain relevant.

Consider first the cross-sectional regression model with no errors in the betas:

$$R_t = \lambda_{0t} + B\lambda_t + u_t, \quad (43)$$

where $\text{Cov}(u_t, B) = 0$ and R_t is an N -vector of returns. Assume that we don't get to see the true B , instead we have B^* , where:

$$B^* = B + v = \text{true} + \text{“noise”}, \quad \text{Cov}(v, B) = 0. \quad (44)$$

Using the first stage time-series or GMM estimation, we can get an estimate of $\text{Cov}(v)$, the cross-sectional covariance matrix of the errors-in-betas. If we run the cross-sectional regression on the noisy betas:

$$R_t = \lambda_{0t} + B^*\lambda_t + e_t, \quad (45)$$

then $\hat{\lambda}_1 \xrightarrow{p} \text{Cov}(B^*)^{-1} \text{Cov}(B^*, R_t) = [\text{Cov}(B) + \text{Cov}(v)]^{-1} \text{Cov}(B + v, B\lambda_t + u_t) = [\text{Cov}(B) + \text{Cov}(v)]^{-1} \{\text{Cov}(B)\lambda_t\}$. Theil (1971) proposes an adjusted estimator to control the bias:

$$\lambda_t^* \equiv [\text{Cov}(B^*) - \text{Cov}(v)]^{-1} [\text{Cov}(B^*)] \hat{\lambda}_1 \xrightarrow{p} \lambda_t. \quad (46)$$

This estimator is used by Black and Scholes (1974), and Litzenberger and Ramaswamy (1979, 1982).

Most of the preceding analysis assumes that the same betas are used in each cross-sectional month. Under this simplifying assumption, errors in betas imply that the cross-sectional regression coefficients are not independent over time, because the same beta (with error) is used in each month. Shanken (1992) shows how to correct Fama–MacBeth standard errors for this fact.

In principle, the cleanest way to deal with errors-in-betas is to estimate the time-series model of betas and the cross-sectional regression simultaneously, thus accounting for the estimation error. This is the subject of the next section.

4.3. Multivariate regression and beta-pricing models

Tests of portfolio efficiency using multivariate regression analysis and maximum likelihood became popular in empirical finance following the work of Jobson and Korkie (1982, 1985), Gibbons (1982) and Stambaugh (1982). The traditional focus of this literature is tests of the CAPM, which implies that the market portfolio is mean-variance efficient. However, given the discussion in Section 2.3, multibeta models and stochastic discount factor models also imply that some portfolio is minimum variance efficient, and the techniques reviewed here can be applied. Following the traditional literature, we ignore the presence of conditioning information in this section. [For tests of efficiency with conditioning information, see Ferson and Siegel (2002b)]. For simplicity, consider the case where a risk-free asset exists.

Let $r_t = \{R_{it} - RF_t\}_i$ be an N -vector of excess returns, and let $r_{pt} = R_{pt} - RF_t$, be the excess return of the particular portfolio to be tested. The null hypothesis to be tested is that R_{pt} is a minimum-variance portfolio, which from Equation (13) is equivalent to:

$$E(r_t) = \beta E(r_{pt}); \quad \beta \equiv \text{Cov}(r_t; r_{pt}) \text{Var}(r_{pt})^{-1}, \quad (47)$$

when there is a given riskfree rate. The tests are based on a regression model:

$$r_t = \alpha + \beta r_{pt} + \epsilon_t, \quad \epsilon_t \sim \text{iid}(0, \Omega), \quad (48)$$

where the null hypothesis implies that the vector of alphas or intercepts, $\alpha = \underline{0}$.

MacKinlay and Richardson (1991) illustrate that it is easy to use the GMM to implement the tests of portfolio efficiency. To do so, one can form the moment conditions:

$$\begin{aligned} \epsilon_t &= r_t - \alpha - \beta r_{pt}, \\ Z_t' &= (1, r_{pt}), \\ g &= \frac{\sum_t (\epsilon_t \otimes Z_t)}{T}. \end{aligned} \quad (49)$$

The parameters are $\varphi = (\alpha', \beta')$. Choosing the parameters to $\text{Min}_\varphi g' W g$, we have the GMM estimators, which are the same as seemingly-unrelated OLS. These are consistent and asymptotically normal, even without the assumptions that the error terms are independent and identically distributed over time. It is assumed that the data are stationary, that $E(\epsilon_t) = 0 = E(\epsilon_t r_{pt})$, and other technical conditions given by Hansen (1982). If the assumptions that justify OLS as best linear unbiased are imposed, GMM delivers the OLS standard errors as well.

If the GMM uses the “optimal” weighting matrix, $W = (1/T)[\text{Cov}(g)]^{-1}$, then the asymptotic variance of the parameters is given by Equation (32). Imposing that $\epsilon_t \sim \text{iid}(0, \Omega)$, the GLS standard errors fall out as a special case. Several tests for the hypothesis that $\alpha = 0$ are available using the GMM. [See, e.g.,

Newey and West (1987)]. One example is the Wald test, which may be formed as $T\alpha'ACov(\alpha)^{-1}\alpha$, where $ACov(\cdot)$ denotes the asymptotic covariance²². The Wald statistic is asymptotically distributed as a Chi-squared variable, with degrees of freedom equal to the dimension of α .

Much of the literature works in a normal, maximum likelihood setting. In this case, the log of the likelihood function to be maximized is:

$$\ln L = \frac{NT}{2} \ln(2\pi) - \frac{T}{2} \ln |\Omega| - \frac{1}{2} \sum_t (r_t - \alpha - \beta r_{pt})' \Omega^{-1} (r_t - \alpha - \beta r_{pt}). \quad (50)$$

Standard tests for the hypothesis that $\alpha = 0$ are compared by Buse (1982) and Gibbons, Ross and Shanken (1989), and most of the standard tests have been used to test the efficiency of stock market indexes, as in the CAPM. Examples include the likelihood ratio test [Gibbons (1982)], the Lagrange multiplier test [Stambaugh (1982)] and the Wald test [Gibbons, Ross and Shanken (1989)]. The Wald test is of particular interest. This is not because of its sampling performance, which is typically the worst of the three, but because it leads to a graphical interpretation that provides some economic intuition for the tests. Since the likelihood ratio and Lagrange multiplier tests are simple transformations of the Wald statistic, as shown by Buse (1982), a similar intuition would apply.

We first need some facts about squared Sharpe ratios. The Sharpe ratio of r_p is $E(r_p)/\sigma(r_p)$, the ratio of the expected excess return to the standard deviation. Let $S^2(r)$ be the maximum squared Sharpe ratio that can be obtained using fixed-weight portfolios of the N assets:

$$S^2(r) \equiv \text{Max}_x \frac{(x'E(r))^2}{x'\Sigma x} = E(r)'\Sigma^{-1}E(r), \quad (51)$$

where the second equality follows from solving the calculus problem. The maximum squared Sharpe ratio in a sample of assets is related to the squared Sharpe ratio of a tested portfolio, r_p , included among the test assets, through a quadratic form in the alphas. I call this result the:

Law of Conservation of Squared Sharpe Ratios.

$$S^2(r) = \alpha'\Sigma^{-1}\alpha + S^2(r_p). \quad (52)$$

²² The notation is as follows: $\sqrt{T}(\hat{C} - C)$ converges in distribution to a vector with mean zero and variance, $ACov(\hat{C})$. Thus, the asymptotic approximation to the finite sample variance of \hat{C} is $(1/T)ACov(\hat{C})$.

A proof uses the fact that, in the stacked regression model, $\alpha' \Sigma^{-1} \beta = 0$ ²³.

Proof:

$$\begin{aligned}
 S^2(r) &= E(r)' \Sigma^{-1} E(r), \\
 &= [\alpha + \beta E(r_p)]' \Sigma^{-1} [\alpha + \beta E(r_p)], \\
 &= \alpha' \Sigma^{-1} \alpha + 2E(r_p)' \alpha' \Sigma^{-1} \beta + E(r_p)' \beta' \Sigma^{-1} \beta E(r_p), \\
 &= \alpha' \Sigma^{-1} \alpha + E(r_p)' \beta' \Sigma^{-1} \beta E(r_p), \\
 &= \alpha' \Sigma^{-1} \alpha + E(r_p)' \text{Var}(r_p)^{-1} E(r_p), \\
 &= \alpha' \Sigma^{-1} \alpha + S^2(r_p).
 \end{aligned}$$

□

The law states that the highest squared Sharpe ratio obtainable in the sample is equal to the squared Sharpe ratio of the tested portfolio, plus a sort of squared Sharpe ratio, based on the alphas. If the alphas are zero, the two Sharpe ratios are the same and the tested portfolio is efficient. When the tested portfolio is not efficient, the quadratic form in the alphas tells how far it is from efficient. This is similar to our previous discussion of how a quadratic form in the APT pricing errors bounds the correlation between a combination of the APT factor portfolios and a minimum variance efficient portfolio. MacKinlay (1995) develops the interpretation of portfolios whose weights are proportional to $\Sigma^{-1} \alpha$, which have many interesting properties.

The law of conservation of squared Sharpe ratios provides a graphical interpretation of the Wald test statistic. Using the law, and the fact that in a multivariate regression model the covariance matrix of the intercept estimator is proportional to the covariance

²³ This occurs when the right-hand side variable(s) are simple combinations of the test assets. In a stacked regression model: $r = \alpha + r_p \beta' + u$, where $r_p = rW$ is a combination of the test assets with weight given by the $n \times k$ matrix, W . Using the definition $\beta' = (W' \Sigma W)^{-1} W' \Sigma$, where Σ is the covariance matrix of r , then

$$\begin{aligned}
 \alpha' \Sigma^{-1} \beta &= E(r - r_p \beta') \Sigma^{-1} [(W' \Sigma W)^{-1} W' \Sigma]', \\
 &= E(r - r_p \beta') \Sigma^{-1} \Sigma W (W' \Sigma W)^{-1}, \\
 &= E(r - rW (W' \Sigma W)^{-1} W' \Sigma) W (W' \Sigma W)^{-1}, \\
 &= E\{rW (W' \Sigma W)^{-1} - rW (W' \Sigma W)^{-1}\}, \\
 &= 0.
 \end{aligned}$$

Note also that $\text{Var}(r_p) = (W' \Sigma W)$, and $\beta' \Sigma^{-1} \beta = (W' \Sigma W)^{-1}$.

matrix of the left-hand side asset returns, or $\text{Cov}(\alpha) = (1 + S^2(r_p))\Sigma$, we may write the Wald test as:

$$\begin{aligned} \text{Wald} &= T\alpha' \text{Cov}(\alpha)^{-1}\alpha, \\ &= T(1 + S^2(r_p))^{-1}\alpha'\Sigma^{-1}\alpha, \\ &= T(1 + S^2(r_p))^{-1}[S^2(r) - S^2(r_p)]. \end{aligned} \quad (53)$$

Thus, the test may be interpreted as a normalized difference between $S^2(r)$, the maximum squared Sharpe ratio in the sample of tested assets, and $S^2(r_p)$, the squared ratio for the tested portfolio. If the tested portfolio presents a Sharpe ratio that is “close” to the sample efficient portfolio, the value of the test statistic is small and we should not reject efficiency. If the tested portfolio lies far inside the sample mean variance frontier, we are likely to reject its efficiency.

4.3.1. Comparing the SDF and beta-pricing approaches

Before the mid 1980s, most of the empirical asset pricing literature used the beta-pricing representation (9) and regression-based approaches or MLE. Then, the SDF representation combined with the GMM began to take hold. The latter combination is appealing, since $E_t\{mR - 1\} = 0$ leads naturally to moment conditions for the GMM, and it is easy to multiply by lagged instruments, in order to use conditioning information. Recent studies have started to explore the tradeoffs between these approaches; see Kan and Zhou (1999), Jagannathan and Wang (2002) and Cochrane (2001).

We have seen that both cross-sectional and time-series regressions are special cases of the GMM. So is maximum likelihood. If we use the GMM on the first order conditions, or “scores” of the likelihood function (50), we get quasi-maximum likelihood estimators. If we further impose normality, then the information matrix identity leads to the MLE standard errors for the parameters, and therefore to the Cramer–Rao lower bound [see Hamilton (1994, Chapter 14)]. The implication is that the tradeoff between the approaches has little to do with GMM versus MLE or regression, but has everything to do with the set of moments that are examined. The SDF representation and the beta pricing formulation can lead us to examine different moments. When they do, their empirical results can differ.

This can be illustrated in the context of a recent debate. Kan and Zhou (1999) consider returns in excess of a risk-free rate, r , comparing beta pricing with the SDF approach. Ignoring conditioning information, beta pricing says $r_t = \beta(f_t + \lambda) + u_t$, where $f_t \equiv F_t - E(F_t)$ is the mean-centered factor, and the moments are $E(u_t) = E(u_t f_t) = 0$. The stochastic discount factor is $m_t = 1 - b'f_t$, and the SDF moment conditions for the excess returns are $E\{r_t(1 - b'f_t)\} = 0$. Kan and Zhou find that the SDF approach is much less efficient than beta pricing. However, the moments being used are not the same. Kan and Zhou implicitly assume that $E(F_t)$ is known,

ignoring the moment condition for estimating $E(F_t)$, and the sampling variation that this moment condition generates. Jagannathan and Wang (2002) and Cochrane (2001) show that when the two methods correctly exploit the same moments, they deliver nearly identical results.

5. Conditional performance evaluation

Classical measures of investment performance compare the return of a managed portfolio to that of a benchmark. For example, an alpha for a fund may be calculated as the average return in excess of a risk-free rate, minus a fixed beta times the average excess return of a benchmark portfolio. Using the market portfolio of the CAPM as the benchmark, Jensen (1968) advocated such a risk-adjusted measure of performance. These classical measures are “unconditional”, in the sense that the expected returns and betas in the model are unconditional moments, estimated by past averages.

If expected returns and risks vary over time, the classical approach is likely to be unreliable. For example, if the risk exposure of a managed portfolio varies predictably with the business cycle, but the manager has no superior forecasting ability, a traditional approach to performance measurement will confuse the common variation in fund risk and expected market returns with abnormal performance. In conditional performance evaluation, we model the expected returns and risk measures, attempting to account for their changes with the state of the economy, and thus controlling for common variation.

The problem of confounding variation in mutual fund risks and market returns has long been recognized [e.g., Jensen (1972), Grant (1977)], but previous studies interpreted it as reflecting superior information or market timing ability. Conditional performance evaluation takes the view that a managed portfolio strategy which can be replicated using readily available public information should not be judged as having superior performance. For example: in a conditional approach, a mechanical market timing rule using lagged interest rate data is not a value-adding strategy. Only managers who correctly use more information than is generally publicly available, are considered to have potentially superior ability. Conditional performance evaluation is therefore consistent with a version of market efficiency, in the semi-strong form sense of Fama (1970).

The beauty of a conditional approach to performance evaluation is that it can accommodate whatever standard of superior information is held to be appropriate, by the choice of the lagged instruments that are used to represent the public information. Incorporating a given set of lagged instruments, managers who trade mechanically in response to these variables get no credit. In practice, the trading behavior of managers may overlay complex portfolio dynamics on the dynamics of the underlying assets they trade. The desire to handle such dynamic strategies further motivates a conditional approach.

5.1. A numerical example

The appeal of a conditional model for performance evaluation can be illustrated with the following highly stylized numerical example. Assume that there are two equally-likely states of the market as reflected in investors' expectations; say, a "Bull" state and a "Bear" state. In a Bull market, assume that the expected return of the S&P500 is 20%, and in a Bear market, it is 0%. The risk-free return to cash is 5%. Assume that all investors share these views – the current state of expected market returns is common knowledge. In this case, an investment strategy using as its only information the current state of the market, will not yield abnormal returns.

Now, imagine a mutual fund which holds the S&P500 in a Bull market and holds cash in a Bear market. Conditional on a Bull market, the beta of the fund is 1.0, the fund's expected return is 20%, equal to the S&P500, and the fund's alpha is zero. Conditional on a Bear market, the fund's beta is 0.0, the expected return of the fund is the risk-free return, 5%, and the alpha is, again, zero. A conditional approach to performance evaluation correctly reports an alpha of zero in each state.

By contrast, an unconditional approach to performance evaluation incorrectly reports an alpha greater than zero for our hypothetical mutual fund. The unconditional beta of the fund²⁴ is 0.75. The unconditional expected return of the fund is $.5(.20) + .5(.05) = 0.125$. The unconditional expected return of the S&P500 is $.5(.20) + .5(.0) = .10$, and the unconditional alpha of the fund is therefore: $(.125 - .05) - 0.75(.10 - .05) = 0.0375$. The unconditional approach leads to the mistaken conclusion that the manager has positive abnormal performance. But the manager's performance does not reflect superior skill or ability, it just reflects the fund's decision to take on more market risk in times when the risk is more highly rewarded in the market. Investors who have access to the same information about the

²⁴ The calculation is as follows. The unconditional beta is $\text{Cov}(F, M) / \text{Var}(M)$, where F is the fund return and M is the market return. The numerator is:

$$\begin{aligned} \text{Cov}(F, M) &= E \{ (F - E(F))(M - E(M)) \mid \text{Bull} \} \times \text{Prob}(\text{Bull}) \\ &\quad + E \{ (F - E(F))(M - E(M)) \mid \text{Bear} \} \times \text{Prob}(\text{Bear}) \\ &= \{ (0.20 - 0.125)(0.20 - 0.10) \} \times 0.5 + \{ (0.05 - 0.125)(0 - 0.10) \} \times 0.5 \\ &= 0.0075. \end{aligned}$$

The denominator is:

$$\begin{aligned} \text{Var}(M) &= E \{ (M - E(M))^2 \mid \text{Bull} \} \times \text{Prob}(\text{Bull}) \\ &\quad + E \{ (M - E(M))^2 \mid \text{Bear} \} \times \text{Prob}(\text{Bear}) \\ &= \{ (0.20 - 0.10)^2 \} \times 0.5 + \{ (0.0 - 0.10)^2 \} \times 0.5 \\ &= 0.01. \end{aligned}$$

The beta is therefore $0.0075/0.01 = 0.75$.

economic state, would not be willing to pay the fund management fees to use this common knowledge.

5.1. Stochastic discount factor formulation

For a given SDF we may define a fund's conditional SDF alpha following Chen and Knez (1996) and Farnsworth et al. (2002) as:

$$\alpha_{pt} \equiv E(m_{t+1}R_{p,t+1} \mid Z_t) - 1, \quad (54)$$

where one dollar invested with the fund at time t returns $R_{p,t+1}$ dollars at time $t + 1$. If the SDF prices a set of "primitive" assets, R_{t+1} , then α_{pt} will be zero when the fund (costlessly) forms a portfolio of the primitive assets, if the portfolio strategy uses only the public information at time t . In that case $R_{p,t+1} = x(Z_t)' R_{t+1}$, where $x(Z_t)$ is the portfolio weight vector. Then Equation (2) implies that $\alpha_{pt} = [E(m_{t+1}x(Z_t)' R_{t+1} \mid Z_t)] - 1 = x(Z_t)' [E(m_{t+1} R_{t+1} \mid Z_t)] - 1 = x(Z_t)' \underline{1} - 1 = 0$.

Consider an example where m_{t+1} is the intertemporal marginal rate of substitution for a representative investor, and Equation (2) is the Euler equation which must be satisfied in equilibrium. If the consumer has access to a fund for which the conditional alpha is not zero he or she will wish to adjust the portfolio, purchasing more of the fund if alpha is positive and less if alpha is negative.

The SDF alpha depends on the model for the SDF, and the SDF is not unique unless markets are complete. Thus, different SDFs can produce different measured performance. This mirrors the classical approaches to performance evaluation, where performance is sensitive to the benchmark²⁵.

While α_{pt} is in general a function of Z_t , it is simpler to discuss the estimation of $\alpha_p = E(\alpha_{pt})$. The parameter α_p is the expectation of the conditional alpha, defined by Equation (54). Thus, we examine the average abnormal performance of a fund²⁶. A useful approach for estimating SDF alphas in this case is to form a system of equations as follows:

$$\begin{aligned} u1_t &= [m_{t+1}R_{t+1} - \underline{1}] \otimes Z_t, \\ u2_t &= \alpha_p - m_{t+1}R_{p,t+1} + 1. \end{aligned} \quad (55)$$

The sample moment condition is $g = T^{-1} \sum_t (u1_t', u2_t)'$. We can use the GMM to simultaneously estimate the parameters of the SDF model and the fund's SDF alpha.

²⁵ Roll (1978), Dybvig and Ross (1985), Brown and Brown (1987), Chen, Copeland and Mayers (1987), Lehmann and Modest (1988) and Grinblatt and Titman (1989b) address this issue in the beta-pricing context. Farnsworth et al. (2002) provide empirical evidence for the SDF setting.

²⁶ For a discussion of time-varying conditional alphas, see Christopherson, Ferson and Glassman (1998a).

5.1.1. Invariance to the number of funds

The system (55) may be estimated using a two-step approach, where the parameters of the model for m_{t+1} are estimated in the first step and the fitted SDF is used to estimate alphas in the second step. Farnsworth et al. (2002) find that simultaneous estimation is dramatically more efficient. However, a potential problem with the simultaneous approach is that the number of moment conditions grows substantially if many funds are to be evaluated, and there are more funds than months in most of the available data sets.

Fortunately, Farnsworth et al. (2002) show that we can estimate the joint system separately for each fund without loss of generality. Estimating a version of system (55) for one fund at a time is equivalent to estimating a system with many funds simultaneously. The estimates of α_p and the standard errors for any subset of funds is invariant to the presence of another subset of funds in the system.

5.1.2. Additional issues

Farnsworth et al. (2002) consider two sets of linear factor models for m_{t+1} . One is based on nontraded factors (e.g., industrial production) and another is based on traded factors (e.g. the S&P500 index). For the traded factor models, they find that it is important to impose the restriction that the model price the traded factors. For example, in the unconditional CAPM, $m_{t+1} = a + bR_{m,t+1}$, where $R_{m,t+1}$ is the gross market return. Requiring the model to price the market return and also a zero beta return we have:

$$E \{ [a + bR_{m,t+1}] R_{m,t+1} \} = 1 \quad \text{and} \quad E \{ [a + bR_{m,t+1}] R_{0,t+1} \} = 1. \quad (56)$$

These two conditions identify the parameters $a(\cdot)$ and $b(\cdot)$ as functions of the first and second moments of the market index and the zero-beta return, as shown previously in Lemma 1.

Farnsworth et al. also find that it is important to impose the restriction that the model prices the risk-free asset. This identifies the conditional mean of the SDF: $E(m_{t+1}|Z_t) = R_{ft}^{-1}$, when R_{ft} is included in Z_t . Non-traded factor models, in particular, are much less accurate when they aren't forced to price the risk-free asset.

5.2. Beta-pricing formulation

Ferson and Schadt (1996) modify Jensen's alpha and two simple market-timing models to incorporate conditioning information. They start with a conditional CAPM, which implies that Equation (57) is satisfied for the assets available to portfolio managers.

They show it is easy to extend the analysis beyond the CAPM, for a conditional multiple-beta model.

$$\begin{aligned} r_{it+1} &= \beta_{im}(Z_t) r_{mt+1} + u_{i,t+1}, \quad i = 0, \dots, N, \quad t = 0, \dots, T-1, \\ E(u_{i,t+1} | Z_t) &= 0, \\ E(u_{i,t+1} r_{mt+1} | Z_t) &= 0, \end{aligned} \quad (57)$$

The $\beta_{im}(Z_t)$ are the time t conditional market betas of the excess return of asset i . The second Equation of (57) follows from the conditional CAPM assumption and the third equation says that the $\beta_{im}(Z_t)$ are conditional regression coefficients. Equation (57) implies that a portfolio strategy which depends only on the public information Z_t will satisfy a similar regression. The intercept, or “alpha” of the regression should be zero, and the error term should not be related to the public information variables²⁷.

Under the hypothesis that the manager uses no more information than Z_t , then the portfolio beta, $\beta_{pm}(Z_t)$, is a function only of Z_t . Using a Taylor series we can approximate this function linearly:

$$\beta_{pm}(Z_t) = b_{0p} + B'_p z_t, \quad (58)$$

where $z_t = Z_t - E(Z)$ is a vector of the deviations of Z_t from the unconditional means, and B_p is a vector with dimension equal to the dimension of Z_t . The coefficient b_{0p} may be interpreted as an “average beta”, i.e., the unconditional mean of the conditional beta: $E(\beta_{pm}(Z_t))$. The elements of B_p are the response coefficients of the conditional beta with respect to the information variables Z_t .

Equations (57) and (58) imply a regression of a managed portfolio excess return on the market factor excess return and its product with the lagged information:

$$r_{pt+1} = \alpha_p + \delta_{1p} r_{mt+1} + \delta'_{2p} (z_t r_{mt+1}) + \varepsilon_{pt+1}, \quad (59)$$

where the model implies $\alpha_p = 0$, $\delta_{1p} = b_{0p}$, and $\delta_{2p} = B_p$.

The regression (59) may be interpreted as a multi-factor model, where the excess market return is the first factor and the products of the market and the lagged information variables are additional factors. The additional factors may be interpreted as the returns to dynamic strategies, which hold z_t units of the market index, financed by borrowing or selling z_t in Treasury bills. The coefficient α_p is the average difference between the managed portfolio excess return and the excess return to the dynamic strategies, which replicate its time-varying risk exposure. A manager with a positive

²⁷ That is, if $R_{p,t+1} = x(Z_t)' R_{t+1}$, where $x(\cdot)$ is an N -vector of weights and R_{t+1} is the N -vector of the available risky security returns, then the portfolio excess return will satisfy the conditional CAPM, with $\beta_{pm}(Z_t) = x(Z_t)' \beta_m(Z_t)$, where $\beta_m(Z_t)$ is the vector of the securities' conditional betas. The error term in the regression for the portfolio strategy is $u_{p,t+1} = x(Z_t)' u_{t+1}$, where u_{t+1} is the vector of the $u_{i,t+1}$'s, and therefore $E(u_{p,t+1} | Z_t) = E(x(Z_t)' u_{t+1} | Z_t) = x(Z_t)' E(u_{t+1} | Z_t) = 0$.

alpha in this setting is one whose average return is higher than the average return of the conditional-beta-replicating strategies.

5.3. Using portfolio weights

The previously discussed performance measurement techniques are all returns-based. The strength of returns-based methodologies is their minimal information requirements. One needs only returns on the managed portfolio and data for the model of m_{t+1} . However, this ignores potentially useful information that is often available in practice: the composition of the managed portfolio. Grinblatt and Titman (1993) propose a weight-based measure of mutual fund performance. Their measure combines portfolio weights with unconditional moments to measure performance. Ferson and Khang (2002) argue that the use of portfolio weights may be especially important in a conditional setting. When expected returns are time-varying and managers trade between return observation dates, returns-based approaches are likely to be biased. This “interim trading bias” can be avoided by using portfolio weights in a conditional setting.

The intuition behind weight-based performance measures can be motivated with a single-period model where an investor maximizes the expected utility of terminal wealth.

$$\text{Max}_x E \{ U (W_0 [R_f + x'r]) \mid Z, S \}, \quad (60)$$

where R_f is the risk-free rate, r is the vector of risky asset returns in excess of the risk-free rate, W_0 is the initial wealth, x is the vector of portfolio weights on the risky assets, Z is public information available at time 0, and S is private information available at time 0. Private information, by definition, is correlated with r , conditional on Z . If returns are conditionally normal, the first and second order conditions for the maximization when the investor has nonincreasing absolute risk aversion imply [see Khang (1997)] that

$$E \{ x(Z, S)' [r - E(r \mid Z)] \mid Z \} > 0, \quad (61)$$

where $x(Z, S)$ is the optimal weight vector and $r - E(r \mid Z)$ are the unexpected, or abnormal returns, from the perspective of an observer with the public information. Conditional on the public information, the sum of the conditional covariances between the weights of a manager with private information, S , and the abnormal returns for the securities in a portfolio is positive. If the manager has no private information, S , then the covariance is zero.

Ferson and Khang (2002) study a Conditional Weight Measure (CWM) that follows from Equation (61). They introduce a “benchmark” weight, x_b , that is in the public information set Z , so Equation (61) implies

$$E \{ [x(Z, S) - x_b]' [r - E(r \mid Z)] \mid Z \} > 0, \quad (62)$$

if the manager has superior information, S . Because x_b is a constant given Z , it will not affect the conditional covariance. Weight changes are advantageous on statistical

grounds, as the levels of the weights may be nonstationary. Other benchmark weights could be used, when a particular benchmark may be suggested by the application.

5.3.1. Conditional performance attribution

Traditional regression-based analysis sometimes interprets the regression as providing a decomposition of the sources of a fund's returns. For example, fund beta times the market excess return is the component of the fund's excess return due to overall market exposure. Conditional performance measures allow refinements of such decompositions. For example, in the Ferson and Schadt regression (59), we have a component due to average market exposure and one due to the mechanical use of the instruments, Z_t , to track the time-varying exposure. The average conditional alpha is the difference between the fund return and the dynamic beta-matched strategy. Weight-based measures allow a similar decomposition.

Consider the following identity for the unconditional covariance:

$$\sum_j \text{Cov}(\Delta x_j, r_j) = \sum_j E\{\text{Cov}(\Delta x_j, r_j|Z)\} + \sum_j \text{Cov}\{E(\Delta x_j|Z), E(r_j|Z)\}, \quad (63)$$

where $\Delta x_j \equiv x_{jt} - x_{bjt}$. The left-hand side is the unconditional weight measure (UWM) as in Grinblatt and Titman (1993). The second term is the "average" conditional weight measure, equal to the unconditional mean of Equation (62). The third term captures the variation in the weight changes associated with changes in the expected returns, conditioned on public information. By comparing the conditional and unconditional measures, the third term may be calculated as a residual.

Equation (63) decomposes the manager's total return from active trading into a component attributable to private information (the first term on the right) and a component attributable to the public information. For example, the second component may be compared with the investor's cost of monitoring the public information. The first component is the performance the investor could not obtain without the manager, even if he chose to monitor the public information. Isolating this component enables an investor to compensate a manager for his use of private information.

5.3.2. Interim trading bias

The conditional weight-based approach can control an interim trading bias, which arises when we depart from the assumption that returns are independently and identically distributed over time (iid), and is therefore especially relevant to a conditional setting. Consider an example where returns are measured over two "periods", but a manager trades each period. The manager has neutral performance, but the portfolio weights for the second period can be a function of public information at the intervening date. If returns are iid, this creates no bias, as there is no information at the intervening date that is correlated with the second period return. However, if expected returns vary with public information, then a manager who observes and trades

on public information at the intervening date generates a return for the second period from the conditional distribution. His two-period portfolio strategy will contain more than the public information at the beginning of the first period, and a returns-based measure over the two periods will detect this as “superior” information.

Goetzmann, Ingersoll and Ivkovic (2000) address interim trading bias by simulating the multiperiod returns generated by the option to trade between return observation dates. A conditional weight-based measure avoids the problem by examining the conditional covariance between the manager’s weights at the beginning of the first period and the subsequent two-period returns. The ability of the manager to trade at the intervening period thus creates no interim trading bias.

Of course, managers may engage in interim trading based on superior information to enhance performance, and a weight-based measure will not record these interim trading effects. Interim trading thus presents a bias under the null hypothesis that managers possess only public information. Under the alternative hypothesis of superior ability, a weight-based measure may have limited power to detect the ability. Thus, the cost of using a weight-based measure to avoid bias is a potential loss of power. Ferson and Khang (2002) evaluate these tradeoffs, and conclude that the conditional weight-based measure is attractive.

5.4. Conditional market-timing models

In a market-timing context, the goal of conditional performance evaluation is to distinguish timing ability that merely reflects publicly available information, as captured by the set of lagged instrumental variables, from timing based on better information. We may call such informed timing ability *conditional market timing*.

A classic market-timing regression, when there is no conditioning information, is the quadratic regression of Treynor and Mazuy (1966):

$$r_{pt+1} = a_p + b_p r_{mt+1} + \gamma_{tmu} [r_{m,t+1}]^2 + v_{pt+1}, \quad (64)$$

where the coefficient γ_{tmu} measures market timing ability. Admati et al. (1986) describe a model in which a manager with constant absolute risk aversion in a normally distributed world, observes at time t a private signal equal to the future market return plus noise, $r_{m,t+1} + \eta$. The manager’s response is change the portfolio beta as a linear function of the signal. They show that the γ_{tmu} coefficient in regression (64) is positive if the manager increases market exposure when the signal about the future market return is positive.

In a conditional model, the part of the correlation of fund betas with the future market return that can be attributed to the public information, is not considered to reflect market timing ability. Ferson and Schadt (1996) develop a conditional version of the Treynor–Mazuy regression:

$$r_{pt+1} = a_p + b_p r_{mt+1} + C_p'(z_t r_{mt+1}) + \gamma_{tmc} [r_{m,t+1}]^2 + v_{pt+1}, \quad (65)$$

where the coefficient vector C_p captures the linear response of the manager’s beta to the public information, Z_t . The term $C_p'(z_t r_{mt+1})$ controls for the public information

effect, which would bias the coefficients in the original Treynor–Mazuy model. The coefficient γ_{mc} measures the sensitivity of the manager's beta to the private market timing signal.

Merton and Henriksson (1981) and Henriksson (1984) describe an alternative model of market timing in which the quadratic term in Equation (64) is replaced by an option payoff, $\text{Max}(0, r_{m,t+1})$. This reflects the idea that market timers may be thought of as delivering (hopefully, attractively priced) put options on the market index. Ferson and Schadt (1996) develop a conditional version of this model as well.

Becker et al. (1999) develop conditional market-timing models with explicit performance benchmarks. In this case, managers maximize the utility of their portfolio returns in excess of a benchmark portfolio return. In practice, performance benchmarks often represent an important component of managers' incentive systems. Such benchmarks have been controversial in the academic literature. Starks (1987), Grinblatt and Titman (1989c) and Admati and Pfleiderer (1997) argue that benchmarks don't properly align managers' incentives. Carpenter, Dybvig and Farnsworth (2000) provide a theoretical justification of benchmarks, used in combination with investment restrictions. Becker et al. simultaneously estimate the fund managers' risk aversion for tracking error and the precision of the market-timing signal, in a sample of more than 400 U.S. mutual funds for 1976–94, including a subsample with explicit asset allocation objectives. The estimates suggest that U.S. equity mutual funds behave as risk averse, benchmark investors, but little evidence of timing ability is found.

5.5. Empirical evidence on conditional performance

Traditional measures of the average abnormal performance of mutual funds, like Jensen's alpha, are observed to be negative more often than positive across the many studies. For example, Jensen (1968) used the CAPM to conclude that a typical fund has neutral performance, only after adding back expenses. Traditional measures of market timing often find that any significant market timing ability is perversely "negative", suggesting that investors could time the market by doing the opposite of a typical fund. Such results make little economic sense, which suggests that they may be spurious.

Conditional performance evaluation takes the view that a mechanically managed portfolio strategy using only public information does not have abnormal performance. A manager's return is therefore compared with such a benchmark, mechanically constructed using public information to match the time-varying risk of the fund. The empirical evidence suggests that conditional performance measures can produce results different from the classical methods.

Ferson and Schadt (1996) find evidence that funds' risk exposures change in response to public information on the economy, such as level of interest rates and dividend yields. Using conditional models Ferson and Schadt (1996), Kryzanowski, Lalancette and To (1997) and Zheng (1999) find that the distribution of mutual fund alphas shifts to the right and is centered near zero. Ferson and Warther (1996) attribute differences between unconditional and conditional alphas to predictable flows of public

money into funds. Inflows are correlated with reduced market exposure, at times when the public expects high returns, due to larger cash holdings at such times. In pension funds, which are not subject to high frequency flows of public money, no overall shift in the distribution of fund alphas is found when moving to conditional models [Christopherson et al. 1998b]).

Once we control for public information variables, there seems to be little evidence that mutual funds have conditional timing ability for the level of the market return. Busse (1999) asks whether fund returns contain information about market volatility. He finds evidence using daily data that funds may shift their market exposures in response to changes in second moments. Further research in this direction is clearly warranted.

Farnsworth et al. (2002) use a variety of SDF models to evaluate performance in a monthly sample of US equity mutual funds. They find that many of the SDF models are biased. The average bias is about -0.19% per month for unconditional models, -0.12% for conditional models. This is less than two standard errors, as a typical standard error is 0.1% per month. They find that the average mutual fund alpha is no worse than a hypothetical stock-picking fund with neutral performance. Adding back average expenses of about 0.17% per month to the mutual fund alphas (since the actual funds pay expenses, while the hypothetical funds do not), the average fund's performance is slightly higher than hypothetical funds with no ability.

Ferson and Khang (2002) develop the conditional, weight-based approach to measuring performance. Using a sample of equity pension fund managers, 1985–1994, they find that the traditional, returns-based alphas of the funds are positive, consistent with previous studies of pension fund performance. However, these alphas are smaller than the potential effects of interim trading bias. By using instruments for public information combined with portfolio weights, their conditional weight-based measures find that the pension funds also have neutral performance. Thus, the empirical evidence based on conditional performance measures suggests that abnormal fund performance, controlling for public information, is rare.

6. Conclusions

This chapter has reviewed tests of multifactor asset pricing models, volatility bounds and portfolio performance. We developed three essentially equivalent paradigms: beta pricing, stochastic discount factors and minimum variance efficiency, and we discussed each approach in the context of *conditional asset-pricing* models. These models are stated in terms of expected returns and risk measures, conditioned on available information about the state of the economy. Conditional models are most interesting when there are observable instruments that can track time-varying expected returns and security risks. The evidence for such predictability in returns is both extensive and controversial. Conditional asset pricing models should provide a useful framework for many continuing investigations.

The three paradigms of empirical asset pricing have traditionally been linked with particular empirical methods. The stochastic discount factor paradigm seems to fit naturally with the generalized method of moments. Mean variance efficiency tests have (since the early 1980s) most commonly employed multivariate regression, and multibeta models seem to cry out for cross-sectional regressions. But this pairing of the models and methods is not sacrosanct. In fact, any of these empirical methods can be paired with any of the paradigms, and recent studies are beginning to explore the possibilities and tradeoffs. I am coming to the view that the set of moments the investigator chooses to examine is the key issue. Different approaches can lead one to examine different moments, and when they do, the results will differ.

Conditional performance evaluation provides an example, where the models and methods meet the data. Empirical work in this area essentially applies conditional asset pricing models to the returns of managed portfolios. The evidence to date shows that conditional models make a difference. I expect these approaches to yield more interesting insights in the future, about the behavior and performance of mutual funds, pension funds, hedge funds and other professionally managed portfolios.

Conditional asset pricing and conditional performance evaluation are still relatively young. They are roughly “teenagers” at the time of this writing. I expect them to contribute a lot more to our state of knowledge as they mature with future research. I hope that this chapter helps to facilitate some of that research.

References

- Admati, A., and P. Pfleiderer (1997), “Performance benchmarks: does it all add up?” *Journal of Business* 70(3):323–350.
- Admati, A., S. Bhattacharya, S. Ross and P. Pfleiderer (1986), “On timing and selectivity”, *Journal of Finance* 41:715–730.
- Ang, A., and G. Bekaert (2001), “Stock return predictability: is it there?” Working Paper (Columbia University, New York).
- Balduzzi, P., and H. Kallal (1997), “Risk premia and variance bounds”, *The Journal of Finance* 52(5): 1913–1949.
- Balduzzi, P., and T. Yao (2001), “Does heterogeneity matter for asset pricing?” Working Paper (Boston College).
- Bansal, R., and B.N. Lehmann (1997), “Growth-optimal restrictions on asset pricing models”, *Macroeconomic Dynamics* 1:1–22.
- Basu, D., and A. Stremme (2003), “Portfolio efficiency and stochastic discount factor bounds with conditioning information: an empirical study”, Working Paper (University of Warwick, Coventry, UK).
- Becker, C., W. Ferson, D. Myers and M. Schill (1999), “Conditional market timing with benchmark investors”, *Journal of Financial Economics* 52:119–148.
- Beja, A. (1971), “The structure of the cost of capital under uncertainty”, *Review of Economic Studies* 4:359–369.
- Bekaert, G., and J. Liu (2002), “Conditioning information and variance bounds on pricing kernels”, *Review of Financial Studies*, forthcoming.
- Berk, J. (1995), “A critique of size-related anomalies”, *Review of Financial Studies* 8:275–286.
- Berk, J. (2000), “Sorting out sorts”, *Journal of Finance* 55:407–427.

- Bhattacharya, S. (1981), "Notes on multiperiod valuation and the pricing of options", *Journal of Finance* 36:163–180.
- Black, F. (1972), "Capital market equilibrium with restricted borrowing", *Journal of Business* 45: 444–454.
- Black, F., and M. Scholes (1974), "The effects of dividend yield and dividend policy on common stock prices and returns", *Journal of Financial Economics* 1:1–22.
- Black, F., M. Jensen and M. Scholes (1972), "The capital asset pricing model: some empirical tests", in: M. Jensen, ed., *Studies in the Theory of Capital Markets* (Praeger, New York).
- Brav, A., G. Constantinides and C. Geczy (2002), "Asset pricing with heterogeneous consumers and limited participation: empirical evidence", *Journal of Political Economy* 110(4).
- Breeden, D. (1979), "An intertemporal asset pricing model with stochastic consumption and investment opportunities", *Journal of Financial Economics* 7:265–296.
- Brown, K., and G. Brown (1987), "Does the market portfolio's composition matter?" *Journal of Portfolio Management* 13:26–32.
- Bruner, R.F., K. Eades, R. Harris and R. Higgins (1998), "Best practices in estimating the cost of capital: survey and synthesis", *Journal of Financial Practice and Education* (Spring/Summer), pp. 13–27.
- Burmeister, E., and M.B. McElroy (1988), "Joint estimation of factor sensitivities and risk premia for the arbitrage pricing theory", *Journal of Finance* 43:721–733.
- Buse, A. (1982), "The likelihood ratio, Wald and Lagrange multiplier tests: an expository note", *American Statistician* 36:153–157.
- Busse, J. (1999), "Volatility timing in mutual timing: evidence from daily returns", *Review of Financial Studies* 12:1009–1041.
- Campbell, J.Y. (1987), "Stock returns and the term structure", *Journal of Financial Economics* 18: 373–399.
- Campbell, J.Y. (1993), "Intertemporal asset pricing without consumption data", *American Economic Review* 83:487–512.
- Campbell, J.Y., and L. Viceira (1999), "Consumption and portfolio decisions when expected returns are time-varying", *Quarterly Journal of Economics* 114:433–495.
- Campbell, J.Y., A. Lo and A.C. MacKinlay (1997), *The Econometrics of Financial Markets* (Princeton University Press, Princeton, NJ).
- Carpenter, J., P.H. Dybvig and H. Farnsworth (2000), "Portfolio performance and agency", Working Paper (Washington University, St. Louis).
- Carroll, R.J. (1989), "Redescending M -estimators", in: Samuel Kotz, ed., *Encyclopedia of Statistical Sciences, Supplement Volume* (Wiley, New York) pp. 134–137.
- Chamberlain, G. (1983), "Funds, factors and diversification in arbitrage pricing models", *Econometrica* 51:1305–1324.
- Chamberlain, G., and M. Rothschild (1983), "Arbitrage, factor structure and mean variance analysis on large asset markets", *Econometrica* 51:1281–1304.
- Chen, N., R. Roll and S. Ross (1986), "Economic forces and the stock market", *Journal of Business* 59:383–403.
- Chen, N.-f., T. Copeland and D. Mayers (1987), "A comparison of single and multifactor performance methodologies", *Journal of Financial and Quantitative Analysis* 22:401–417.
- Chen, Z., and P.J. Knez (1996), "Portfolio performance measurement: theory and applications", *Review of Financial Studies* 9:511–556.
- Christopherson, J.A., W. Ferson and D.A. Glassman (1998a), "Conditioning manager alphas on economic information: another look at the persistence of performance", *Review of Financial Studies* 11:111–142.
- Christopherson, J.A., W. Ferson and D.A. Glassman (1998b), "Conditional measures of performance and persistence for pension funds", in: A. Chen, ed., *Research in Finance, Vol. 16* (JAI Press, Stamford, CT) pp. 1–46.

- Cochrane, J., and L. Hansen (1992), "Asset pricing lessons for macroeconomics", in: O. Blanchard and S. Fisher, eds., *The Macroeconomics Annual* (MIT Press, Cambridge, MA) pp. 115–165.
- Cochrane, J.H. (1996), "A cross-sectional test of a production based asset pricing model", *Journal of Political Economy* 104:572–621.
- Cochrane, J.H. (2001), *Asset Pricing* (Princeton University Press, New Jersey).
- Cochrane, J.H., and J. Saa'-Requejo (2000), "Beyond arbitrage: good-deal pricing of derivatives in incomplete markets", *Journal of Political Economy* 108:79–119.
- Connor, G. (1984), "A unified beta pricing theory", *Journal of Economic Theory* 34:13–31.
- Connor, G., and R. Korajczyk (1988), "Risk and return in an equilibrium APT: application of a new test methodology", *Journal of Financial Economics* 21:255–290.
- Connor, G., and R. Korajczyk (1995), "Arbitrage pricing theory and multifactor models of asset returns", in: R. Jarrow, V. Maksimovic and W.T. Ziemba, eds., *Handbooks in Operations Research and Management Science: Finance* (Elsevier, Amsterdam) pp. 87–137.
- Connor, G., and R.A. Korajczyk (1986), "Performance measurement with the arbitrage pricing theory: a new framework for analysis", *Journal of Financial Economics* 15:373–394.
- Conrad, J., and G. Kaul (1989), "Mean reversion in short-horizon expected returns", *Review of Financial Studies* 2:225–240.
- Constantinides, G.M. (1982), "Intertemporal asset pricing with heterogeneous consumers and without demand aggregation", *Journal of Business* 55:253–267.
- Cox, J.C., J.E. Ingersoll and S.A. Ross (1985), "A theory of the term structure of interest rates", *Econometrica* 53:385–408.
- Cragg, J.G., and B.G. Malkiel (1982), *Expectations and the Structure of Share Prices* (University of Chicago Press, Chicago).
- Daniel, K., and S. Titman (1997), "Evidence on the characteristics of cross sectional variation in stock returns", *Journal of Finance* 52:1–33.
- DeBondt, W., and R. Thaler (1985), "Does the stock market overreact?" *Journal of Finance* 40:793–805.
- Douglas, G. (1969), "Risk in equity markets: an empirical appraisal of market efficiency", *Yale Economics Essays* 9:3–45.
- Dybvig, P., and J. Ingersoll (1982), "Mean variance theory in complete markets", *Journal of Business* 55:233–252.
- Dybvig, P.H. (1983), "An explicit bound on individual assets' deviations from APT pricing in a finite economy", *Journal of Financial Economics* 12:483–496.
- Dybvig, P.H., and S.A. Ross (1985), "Performance measurement using differential information and a security market line", *Journal of Finance* 40:383–399.
- Fama, E., and K. French (1988), "Permanent and temporary components of stock prices", *Journal of Political Economy* 96:246–273.
- Fama, E., and K. French (1989), "Business conditions and expected returns on stocks and bonds", *Journal of Financial Economics* 25:23–49.
- Fama, E., and K. French (1993), "Common risk factors in the returns of stocks and bonds", *Journal of Financial Economics* 33:3–56.
- Fama, E.F. (1970), "Efficient capital markets: a review of theory and empirical work", *Journal of Finance* 25:383–417.
- Fama, E.F. (1976), *Foundations of Finance* (Basic Books, New York).
- Fama, E.F. (1991), "Efficient capital markets II", *Journal of Finance* 46:1575–1617.
- Fama, E.F., and K.R. French (1996), "Multifactor explanations of asset pricing anomalies", *Journal of Finance* 51:55–87.
- Fama, E.F., and K.R. French (1997), "Industry costs of equity", *Journal of Financial Economics* 43: 153–194.
- Fama, E.F., and J.D. MacBeth (1973), "Risk, return and equilibrium: empirical tests", *Journal of Political Economy* 81:607–636.

- Fama, E.F., and G.W. Schwert (1977), "Asset returns and inflation", *Journal of Financial Economics* 5:115–146.
- Farnsworth, H., W. Ferson, D. Jackson and S. Todd (2002), "Performance evaluation with stochastic discount factors", *Journal of Business* 75:473–504.
- Ferson, W.E. (1983), "Expectations of real interest rates and aggregate consumption: empirical tests", *Journal of Financial and Quantitative Analysis* 18:477–497.
- Ferson, W.E. (1989), "Changes in expected security returns, risk and the level of interest rates", *Journal of Finance* 44:1191–1217.
- Ferson, W.E. (1995), "Theory and empirical testing of asset pricing models", in: R. Jarrow, V. Maksimovic and W.T. Ziemba, eds., *Handbooks in Operations Research and Management Science: Finance* (Elsevier, Amsterdam) pp. 145–200.
- Ferson, W.E., and C.R. Harvey (1991), "The variation of economic risk premiums", *Journal of Political Economy* 99:385–415.
- Ferson, W.E., and C.R. Harvey (1993), "The risk and predictability of international equity returns", *Review of Financial Studies* 6:527–566.
- Ferson, W.E., and C.R. Harvey (1999), "Economic, financial and fundamental global risk in and out of EMU", *Swedish Economic Policy Review* 6:123–184.
- Ferson, W.E., and R. Jagannathan (1996), "Econometric evaluation of asset pricing models", in: G.S. Maddala and C.R. Rao, eds., *Handbook of Statistics, Vol. 14: Statistical Methods in Finance* (Elsevier, Amsterdam) pp. 1–30.
- Ferson, W.E., and K. Khang (2002), "Conditional performance measurement using portfolio weights: evidence for pension funds", *Journal of Financial Economics* 65:249–282.
- Ferson, W.E., and R.A. Korajczyk (1995), "Do arbitrage pricing models explain the predictability of stock return?", *Journal of Business* 68(3):309–349.
- Ferson, W.E., and R.W. Schadt (1996), "Measuring fund strategy and performance in changing economic conditions", *Journal of Finance* 51:425–462.
- Ferson, W.E., and A.F. Siegel (2001), "The efficient use of conditioning information in portfolios", *Journal of Finance* 56(3):967–982.
- Ferson, W.E., and A.F. Siegel (2002a), "Stochastic discount factor bounds with conditioning information", *Review of Financial Studies*, forthcoming.
- Ferson, W.E., and A.F. Siegel (2002b), "Testing portfolio efficiency with conditioning information", Working Paper (Boston College).
- Ferson, W.E., and V.A. Warther (1996), "Evaluating fund performance in a dynamic market", *Financial Analysts Journal* 52(6):20–28.
- Ferson, W.E., S. Sarkissian and T. Simin (1999), "The alpha factor asset pricing model: a parable", *Journal of Financial Markets* 2:49–68.
- Ferson, W.E., S. Sarkissian and T. Simin (2002), "Spurious regressions in financial economics?" *Journal of Finance*, forthcoming.
- Ferson, W.E., A. Siegel and P. Xu (2002), "Mimicking portfolios with conditioning information", Working Paper (Boston College).
- Foster, D., T. Smith and R. Whaley (1997), "Assessing goodness-of-fit of asset pricing models: the distribution of the maximal R-squared", *Journal of Finance* 52:591–607.
- Gallant, R.A., L.P. Hansen and G. Tauchen (1990), "Using the conditional moments of asset payoffs to infer the volatility of intertemporal marginal rates of substitution", *Journal of Econometrics* 45: 141–179.
- Gibbons, M.R. (1982), "Multivariate tests of financial models", *Journal of Financial Economics* 10:3–27.
- Gibbons, M.R., S.A. Ross and J. Shanken (1989), "A test of the efficiency of a given portfolio", *Econometrica* 57:1121–1152.
- Goetzmann, W.N., J. Ingersoll and Z. Ivkovic (2000), "Monthly measurement of daily timers", *Journal of Financial and Quantitative Analysis* 35:257–290.

- Goodall, C. (1983), “*M*-estimators of location: an outline of the theory”, in: D.C. Hoaglin, F. Mosteller and J.W. Tukey, eds., *Understanding Robust and Exploratory Data Analysis* (Wiley, New York) pp. 339–403.
- Goyal, A., and I. Welch (1999), “Predicting the equity premium”, Working Paper (University of California, Los Angeles).
- Graham, J.R., and C.R. Harvey (2001), “The theory and practice of corporate finance: evidence from the field”, *Journal of Financial Economics* 60:187–243.
- Grant, D. (1977), “Portfolio performance and the ‘cost’ of timing decisions”, *Journal of Finance* 32:837–846.
- Grinblatt, M., and S. Titman (1983), “Factor pricing in a finite economy”, *Journal of Financial Economics* 12:497–508.
- Grinblatt, M., and S. Titman (1987), “The relation between mean-variance efficiency and arbitrage pricing”, *Journal of Business* 60:97–112.
- Grinblatt, M., and S. Titman (1989a), “Mutual fund performance: an analysis of quarterly portfolio holdings”, *Journal of Business* 62:393–416.
- Grinblatt, M., and S. Titman (1989b), “Portfolio performance evaluation: old issues and new insights”, *Review of Financial Studies* 2:393–421.
- Grinblatt, M., and S. Titman (1989c), “Adverse risk incentives in the design of performance-based contracts”, *Management Science* 35:807–822.
- Grinblatt, M., and S. Titman (1993), “Performance measurement without benchmarks: an examination of mutual fund returns”, *Journal of Business* 60:97–112.
- Grossman, S., and R.J. Shiller (1982), “Consumption correlatedness and risk measurement in economies with nontraded assets and heterogeneous information”, *Journal of Financial Economics* 10:195–210.
- Hamilton, J.D. (1994), *Time-Series Analysis* (Princeton University Press, Princeton, NJ).
- Hampel, F.R. (1974), “The influence curve and its role in robust estimation”, *Journal of the American Statistical Association* 69:383–393.
- Hansen, L.P. (1982), “Large sample properties of the generalized method of moments estimators”, *Econometrica* 50:1029–1054.
- Hansen, L.P., and R. Hodrick (1980), “Forward exchange rates as optimal predictors of future spot rates: an econometric analysis”, *Journal of Political Economy* 88:829–853.
- Hansen, L.P., and R. Jagannathan (1991), “Implications of security market data for models of dynamic economies”, *Journal of Political Economy* 99:225–262.
- Hansen, L.P., and R. Jagannathan (1997), “Assessing specification errors in stochastic discount factor models”, *Journal of Finance* 52:557–590.
- Hansen, L.P., and S.F. Richard (1987), “The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models”, *Econometrica* 55:587–613.
- Hansen, L.P., J. Heaton and E. Luttmer (1995), “Econometric evaluation of asset pricing models”, *Review of Financial Studies* 8:237–274.
- Harrison, M., and D. Kreps (1979), “Martingales and arbitrage in multi-period securities markets”, *Journal of Economic Theory* 20:381–408.
- Harvey, C.R. (1991), “The world price of covariance risk”, *Journal of Finance* 46:111–157.
- Harvey, C.R., and C.M. Kirby (1996), “Instrumental variables estimation of conditional beta pricing models”, in: G.S. Maddala and C.R. Rao, eds., *Handbook of Statistics*, Vol 14 (Elsevier, New York) Ch. 2.
- Henriksson, R.D. (1984), “Market timing and mutual fund performance: an empirical investigation”, *Journal of Business* 57:73–96.
- Huberman, G., S.A. Kandel and R.F. Stambaugh (1987), “Mimicking portfolios and exact arbitrage pricing”, *Journal of Finance* 42:1–10.
- Ingersoll, J.E. (1987), *Theory of Financial Decision Making* (Rowman and Littlefield, Savage, MD).
- Jagannathan, R., and Z. Wang (2002), “Empirical evaluation of asset pricing models: a comparison of SDF and beta methods”, *Journal of Finance* 57:2337–2367.

- Jegadeesh, N., and S. Titman (1993), "Returns to buying winners and selling losers: implications for stock market efficiency", *Journal of Finance* 48:65–91.
- Jensen, M. (1968), "The performance of mutual funds in the period 1945–1964", *Journal of Finance* 23:389–46.
- Jensen, M.C. (1972), "Optimal utilization of market forecasts and the evaluation of investment performance", in: G.P. Szego and K. Shell, eds., *Mathematical Methods in Investment and Finance* (Elsevier, Amsterdam).
- Jobson, J.D., and R. Korkie (1982), "Potential performance and tests of portfolio efficiency", *Journal of Financial Economics* 10:433–466.
- Jobson, J.D., and R. Korkie (1985), "Some tests of asset pricing with multivariate normality", *Canadian Journal of Administrative Sciences* 2:114–138.
- Kan, R., and G. Zhou (1999), "A critique of the stochastic discount factor methodology", *Journal of Finance* 54:1221–1248.
- Kandel, S., and R.F. Stambaugh (1989), "A mean variance framework for tests of asset pricing models", *Review of Financial Studies* 2:125–156.
- Kaul, G. (1996), "Predictable components in stock returns", in: G.S. Maddala and C.R. Rao, eds., *Handbook of Statistics*, Vol. 14 (Elsevier, Amsterdam) pp. 269–296.
- Keim, D.B. (1983), "Size-related anomalies and stock return seasonality: further empirical evidence", *Journal of Financial Economics* 12:13–32.
- Khang, K. (1997), "Performance measurement using portfolio weights and conditioning information: an examination of pension fund equity manager performance", Ph.D dissertation (University of Washington, Seattle, WA) unpublished.
- Kim, M., C.R. Nelson and R. Startz (1991), "Mean reversion in stock returns? A reappraisal of the statistical evidence", *Review of Economic Studies* 58:515–528.
- Kim, T.S., and E. Omberg (1996), "Dynamic nonmyopic portfolio behavior", *Review of Financial Studies* 9:141–161.
- Kirby, C. (1998), "The restrictions on predictability implied by asset pricing models", *Review of Financial Studies* 11:343–382.
- Kryzanowski, L., S. Lalancette and M.C. To (1997), "Performance attribution using an APT with prespecified macrofactors and time-varying risk premia and betas", *Journal of Financial and Quantitative Analysis* 32:205–224.
- Lehmann, B.N., and D.M. Modest (1987), "Mutual fund performance evaluation: a comparison of benchmarks and benchmark comparisons", *Journal of Finance* 42:233–265.
- Lehmann, B.N., and D.M. Modest (1988), "The empirical foundations of the arbitrage pricing theory", *Journal of Financial Economics* 21:213–254.
- Lettau, M., and S. Ludvigson (2001), "Consumption, aggregate wealth and expected stock returns", *Journal of Finance* 56:815–849.
- Lintner, J. (1965), "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets", *Review of Economics and Statistics* 47:13–37.
- Litzenberger, R., and K. Ramaswamy (1979), "The effect of personal taxes and dividends on capital asset prices: theory and evidence", *Journal of Financial Economics* 7:163–196.
- Litzenberger, R., and K. Ramaswamy (1982), "The effects of dividends on common stock prices: tax effects or information effects?" *Journal of Finance* 37:429–433.
- Lo, A.W., and A.C. MacKinlay (1990), "Data snooping in tests of financial asset pricing models", *Review of Financial Studies* 1:41–66.
- Long, J. (1974), "Stock prices, inflation, and the term structure of interest rates", *Journal of Financial Economics* 1:131–170.
- Lucas Jr, R.E. (1978), "Asset prices in an exchange economy", *Econometrica* 46:1429–1445.
- MacKinlay, A.C. (1995), "Multifactor models do not explain deviations from the CAPM", *Journal of Financial Economics* 38:3–28.

- MacKinlay, A.C., and M.P. Richardson (1991), "Using the generalized method of moments to test mean-variance efficiency", *Journal of Finance* 46:511–528.
- Mehra, R., and E. Prescott (1985), "The equity premium: a puzzle", *Journal of Monetary Economics* 15:145–162.
- Merton, R.C. (1973), "An intertemporal capital asset pricing model", *Econometrica* 41:867–887.
- Merton, R.C., and R. Henriksson (1981), "On market timing and investment performance II: Statistical procedures for evaluating forecasting skills", *Journal of Business* 54:513–533.
- Newey, W., and K.D. West (1987), "A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica* 55:703–708.
- Pesaran, M.H., and A. Timmermann (1995), "Predictability of stock returns: robustness and economic significance", *Journal of Finance* 50:1201–1228.
- Pontiff, J., and L. Schall (1998), "Book-to-market as a predictor of market returns", *Journal of Financial Economics* 49:141–160.
- Reinganum, M.R. (1981), "Misspecification of capital asset pricing: empirical anomalies based on earnings yields and market values", *Journal of Financial Economics* 9:19–46.
- Roll, R. (1977), "A critique of the asset pricing theory's tests – part 1: On past and potential testability of the theory", *Journal of Financial Economics* 4:129–176.
- Roll, R. (1978), "Ambiguity when performance is measured by the security market line", *Journal of Finance* 33:1051–1069.
- Roll, R., and S.A. Ross (1980), "An empirical investigation of the arbitrage pricing theory", *Journal of Finance* 35:1073–1103.
- Roll, R.R. (1984), "A simple implicit measure of the effective bid-ask spread in an efficient market", *Journal of Finance* 39:1127–1140.
- Roll, R.R. (1988), "R²", *Journal of Finance* 43:541–566.
- Ross, S. (1977), "Risk, return and arbitrage", in: I. Friend and J. Bicksler, eds., *Risk and Return in Finance* (Ballinger, Cambridge, MA).
- Ross, S.A. (1976), "The arbitrage pricing theory of capital asset pricing", *Journal of Economic Theory* 13:341–360.
- Rubinstein, M. (1974), "An aggregation theorem for securities markets", *Journal of Financial Economics* 1:225–244.
- Rubinstein, M. (1976), "The valuation of uncertain income streams and the pricing of options", *Bell Journal of Economics and Management Science* 7:407–425.
- Sarkissian, S. (2002), "Incomplete consumption risk sharing and currency risk premiums", *Review of Financial Studies*, forthcoming.
- Scholes, M., and J. Williams (1977), "Estimating beta from nonsynchronous data", *Journal of Financial Economics* 5:309–327.
- Shanken, J. (1987), "Multivariate proxies and asset pricing relations: living with the Roll critique", *Journal of Financial Economics* 18:91–110.
- Shanken, J. (1992), "On the estimation of beta pricing models", *Review of Financial Studies* 5:1–34.
- Sharpe, W.F. (1964), "Capital asset prices: a theory of market equilibrium under conditions of risk", *Journal of Finance* 19:425–442.
- Sharpe, W.F. (1977), "The capital asset pricing model: a multi-beta interpretation", in: H. Levy and M. Sarnat, eds., *Financial Decision Making under Uncertainty* (Academic Press, New York).
- Simin, T. (2002), "The (poor) predictive performance of asset pricing models", Working Paper (Pennsylvania State University).
- Singleton, K. (1990), "Specification and estimation of intertemporal asset pricing models", in: B. Freidman and F. Hahn, eds., *Handbook of Monetary Economics* (Elsevier, Amsterdam).
- Snow, K.N. (1991), "Diagnosing asset pricing models using the distribution of asset returns", *Journal of Finance* 46:955–983.
- Solnik, B. (1993), "The unconditional performance of international asset allocation strategies using conditioning information", *Journal of Empirical Finance* 1:33–55.

- Stambaugh, R.F. (1982), "On the exclusion of assets from tests of the two-parameter model", *Journal of Financial Economics* 10:235–268.
- Stambaugh, R.F. (1983), "Testing the CAPM with broader market indexes: a problem of mean deficiency", *Journal of Banking and Finance* 7:5–16.
- Stambaugh, R.F. (1999), "Predictive regressions", *Journal of Financial Economics* 54:315–421.
- Starks, L. (1987), "Performance incentive fees: an agency theoretic approach", *Journal of Financial and Quantitative Analysis* 22:17–32.
- Theil, H. (1971), *Principles of Econometrics* (Wiley, New York).
- Treynor, J., and K. Mazuy (1966), "Can mutual funds outguess the market?" *Harvard Business Review* 44:131–136.
- Wheatley, S. (1989), "A critique of latent variable tests of asset pricing models", *Journal of Financial Economics* 23:325–338.
- Wilson, R.B. (1968), "The theory of syndicates", *Econometrica* 36:119–131.
- Zeldes, S. (1989), "Consumption and liquidity constraints: an empirical investigation", *Journal of Political Economy* 97:305–346.
- Zheng, L. (1999), "Is money smart? A study of mutual fund investors' fund selection ability", *Journal of Finance* 54:901–933.