Chapter 8

ASSET PRICING MODELS

WAYNE E. FERSON, *Boston College, USA*

Abstract

The asset pricing models of financial economics describe the prices and expected rates of return of securities based on arbitrage or equilibrium theories. These models are reviewed from an empirical perspective, emphasizing the relationships among the various models.

Keywords: financial assets; arbitrage; portfolio optimization; stochastic discount factor; beta pricing model; intertemporal marginal rate of substitution; systematic risk; Capital Asset Pricing Model; consumption; risk aversion; habit persistence; durable goods; mean variance efficiency; factor models; arbitrage pricing model

Asset pricing models 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, and futures contracts. The asset pricing models of financial economics are based on two central concepts. The first is the "no arbitrage principle," which states that market forces tend to align the prices of financial assets so as to eliminate arbitrage opportunities. An arbitrage opportunity arises if assets can be combined in a portfolio with zero cost, no chance of a loss, and a positive probability of gain. Arbitrage opportunities tend to be eliminated in financial markets because prices adjust as investors attempt to trade to exploit the arbitrage opportunity. For example,

if there is an arbitrage opportunity because the price of security A is too low, then traders' efforts to purchase security A will tend to drive up its price, which will tend to eliminate the arbitrage opportunity. The arbitrage pricing model (APT), (Ross, 1976) is a well-known asset pricing model based on arbitrage principles.

The second central concept in asset pricing is "financial market equilibrium." Investors' desired holdings of financial assets are derived from an optimization problem. A necessary condition for financial market equilibrium in a market with no frictions is that the first-order conditions of the investor's optimization problem are satisfied. This requires that investors are indifferent at the margin to small changes in their asset holdings. Equilibrium asset pricing models follow from the first-order conditions for the investors' portfolio choice problem, and a market-clearing condition. The market-clearing condition states that the aggregate of investors' desired asset holdings must equal the aggregate "market portfolio" of securities in supply.

Differences among the various asset pricing models arise from differences in their assumptions about investors' preferences, endowments, production and information sets, the process governing the arrival of news in the financial markets, and the types of frictions in the markets. Recently, models have been developed that emphasize the role of human imperfections in this process. For a review of this "behavioral finance" perspective, see Barberis and Shleifer (2003).

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

$$P_t = E_t\{m_{t+1}(P_{t+1} + D_{t+1})\} \qquad (8.1)$$

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). By recursive substitution in Equation (8.1), the future price may be eliminated to express the current price as a function of the future cash flows and SDFs only: $P_t = E_t\{\Sigma_{j>0}(\Pi_{k=1,\ldots,j} m_{t+k})D_{t+j}\}$ Prices are obtained by "discounting" the payoffs, or multiplying by SDFs, so that the expected "present value" of the payoff is equal to the price.

We say that a SDF "prices" the assets if Equation (8.1) is satisfied. Any particular asset pricing model may be viewed simply as a specification for the stochastic discount factor. The random variable m_{t+1} is also known as the benchmark pricing variable, equivalent martingale measure, Radon–Nicodym derivative, or intertemporal marginal rate of substitution, depending on the context. The representation in Equation (8.1) 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 (8.1) is equivalent to:

$$E_t\ (m_{t+1}\ \mathbf{R}_{t+1} - \mathbf{1}) = 0, \qquad (8.2)$$

where $\mathbf{R}_{t+1}$ is the vector of primitive asset gross returns and $\mathbf{1}$ is an $\mathbf{N}$-vector of ones. The gross return $R_{i,t+1}$ is defined as $(P_{i,t+1} + D_{i,t+1})/P_{i,t}$, where $P_{i,t}$ is the price of the asset i at time t and $D_{i,t+1}$ is the payment received at time $t + 1$. Empirical tests of asset pricing models often work directly with asset returns in Equation (8.2) and the relevant definition of m_{t+1}.

Without more structure the Equations (8.1,8.2) have no content, because it is 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 returns. For example, take a sample of asset gross returns with a nonsingular covariance matrix and let m_{t+1} be $.[\mathbf{1}'(E_t\{\mathbf{R}_{t+1}\mathbf{R}_{t+1'}\} - \mathbf{1}]\mathbf{R}_{t+1}$ Substitution in to Equation (8.2) shows that this SDF will always "work" in any sample of returns. The ability to construct an SDF as a function of the returns that prices all of the included assets, is essentially equivalent to the ability to construct a minimum-variance efficient portfolio and use in as the "factor" in a beta pricing model, as described below.

With the restriction that m_{t+1} is a strictly positive random variable, Equation (8.1) becomes equivalent to the no arbitrage principle, which says that all portfolios of assets with payoffs that can never be negative but are positive with positive probability, must have positive prices (Beja, 1971; Rubinstein, 1976; Ross, 1977; Harrison and Kreps, 1979; Hansen and Richard, 1987.)

While the no arbitrage principle places restrictions on m_{t+1}, empirical work more typically explores the implications of equilibrium models for the SDF based on investor optimization. A representative consumer–investor's optimization implies the Bellman equation:

$$J(W_t,s_t) \equiv \max E_t\{U(C_t,.) + J(W_{t+1},s_{t+1})\} \qquad (8.3)$$

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

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 product of the marginal utility of consumption expenditures, $Uc(C_t,.) > 0$ (where a subscript denotes partial derivative), multiplied by the price of the asset, and which is measured in the same units as the consumption expenditures. 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: $P_{i,t} E_t\{U_c (C_t,.)\} = E_t\{(P_{i,t+1} + D_{i,t+1}) J_w(W_{t+1}, s_{t+1})\}$ which is equivalent to Equation (8.1), with

$$m_{t+1} = \frac{J_w(W_{t+1}, s_{t+1})}{E_t\{U_c (C_t,.)\}}. \qquad (8.4)$$

The m_{t+1} in Equation (8.4) is the "intertemporal marginal rate of substitution" (IMRS) of the consumer–investor.

Asset pricing models typically focus on the relation of security returns to aggregate quantities. It is therefore necessary to aggregate the first-order conditions of individuals to obtain equilibrium expressions in terms of aggregate quantities. Then, Equation (8.4) may be considered to hold for a representative investor who holds all the securities and consumes the aggregate quantities. Theoretical conditions that justify the use of aggregate quantities are discussed by Gorman (1953), Wilson (1968), Rubinstein (1974), and Constantinides (1982), among others. When these conditions fail, investors' heterogeneity will affect the form of the asset pricing relation. The effects of heterogeneity are examined by Lintner (1965), Brennan and Kraus (1978), Lee et al. (1990), Constantinides and Duffie (1996), and Sarkissian (2003), among others.

Typically, empirical work in asset pricing focuses on expressions for expected returns and excess rates of return. The expected excess returns are modeled in relation 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 (8.2) holds for both assets it implies:

$$E_t\{m_{t+1} r_{i,t+1}\} = 0 \text{ for all } i. \qquad (8.5)$$

Use the definition of covariance to expand Equation (8.5) into the product of expectations plus the covariance, obtaining:

$$E_t\{r_{i,t+1}\} = \frac{\text{Cov}_t(r_{i,t+1}; -m_{t+1})}{E_t\{m_{t+1}\}}, \text{ for all } i, \qquad (8.6)$$

where $\text{Cov}_t(.;.)$ is the conditional covariance. Equation (8.6) is a general expression for the expected excess return from which most of the expressions in the literature can be derived.

Equation (8.6) implies that the covariance of return with m_{t+1}, is a general measure of "systematic risk." This risk is systematic in the sense that any fluctuations in the asset return that are uncorrelated with fluctuations in the SDF are not "priced," meaning that these fluctuations do not command a risk premium. For example, in the conditional regression $r_{it+1} = a_{it} + b_{it} m_{t+1} + u_{it+1}$, then $\text{Cov}_t(u_{it+1}, m_{t+1}) = 0$. Only the part of the variance in a risky asset return that is correlated with the SDF is priced as risk.

Equation (8.6) displays that a security will earn a positive risk premium if its return is negatively correlated with the SDF. When the SDF is an aggregate IMRS, negative correlation means that the asset is likely to return more than expected when the marginal utility in the future period is low, and less than expected when the marginal utility and the value of the payoffs, is high. For a given expected payoff, the more negative the covariance of the asset's payoffs with the IMRS, the less desirable the distribution of the random return, the lower the value of the asset and the larger the expected compensation for holding the asset given the lower price.

8.1. The Capital Asset Pricing Model

One of the first equilibrium asset pricing models was the Capital Asset Pricing Model (CAPM),

developed by Sharpe (1964), Lintner (1965), and Mossin (1966). The CAPM remains one of the foundations of financial economics, and a huge number of theoretical papers refine the assumptions and provide derivations of the CAPM. The CAPM states that expected asset returns are given by a linear function of the assets' "betas," which are their regression coefficients against the market portfolio. Let R_{mt} denote the gross return for the market portfolio of all assets in the economy. Then, according to the CAPM,

$$E(R_{it+1}) = \delta_0 + \delta_1 \, \beta_i, \tag{8.7}$$

where $\beta_i = \text{Cov}(R_i, R_m)/\text{Var}(R_m)$.

In Equation (8.7), $\delta_0 = E(R_{0t+1})$, where the return R_{0t+1} is referred to as a "zero-beta asset" to R_{mt+1} because the condition $\text{Cov}(R_{0t+1}, R_{mt+1}) = 0$.

To derive the CAPM, it is simplest to assume that the investor's objective function in Equation (8.3) is quadratic, so that $J(W_{t+1}, S_{t+1}) = V\{E_t(R_{pt+1}), \text{Var}_t(R_{pt+1})\}$ where R_{pt+1} is the investor's optimal portfolio. The function $V(.,.)$ is increasing in its first argument and decreasing in the second if investors are risk averse. In this case, the SDF of Equation (8.4) specializes as: $m_{t+1} = a_t + b_t R_{pt+1}$. In equilibrium, the representative agent must hold the market portfolio, so $R_{pt+1} = R_{mt+1}$. Equation (8.7) then follows from Equation (8.6), with this substitution.

8.2. Consumption-based Asset Pricing Models

Consumption models may be derived from Equation (8.4) by exploiting the envelope condition, $U_c(.) = J_w(.)$, which states that the marginal utility of current consumption must be equal to the marginal utility of current 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 $\Sigma_t \beta^t U(C_t)$, where β is a time preference parameter

and $U(.)$ is increasing and concave in current consumption, C_t. Breeden's model is a linearization of Equation (8.1), which follows from the assumption that asset values and consumption follow diffusion processes (Bhattacharya, 1981; Grossman and Shiller, 1982). A discrete-time version follows Lucas (1978), assuming a power utility function:

$$U(C) = [C^{1-\alpha} - 1]/(1 - \alpha), \tag{8.8}$$

where $\alpha > 0$ is the concavity parameter of the period utility function. This function displays constant relative risk aversion equal to α. "Relative risk aversion" in consumption is defined as: $Cu''(C)/u'(C)$. Absolute risk aversion is defined as: $u''(C)/u'(C)$. Ferson (1983) studied a consumption-based asset pricing model with constant absolute risk aversion.

Using Equation (8.8) and the envelope condition, the IMRS in Equation (8.4) becomes:

$$m_{t+1} = \beta(C_{t+1}/C_t)^{-\alpha}. \tag{8.9}$$

A large body of literature in the 1980s tested the pricing Equation (8.1) with the SDF given by the consumption model (Equation (8.9)). See, for example, Hansen and Singleton (1982, 1983), Ferson (1983), and Ferson and Merrick (1987).

More recent work generalizes the consumption-based model to allow for "nonseparabilities" in the $U_c(C_t,.)$ function in Equation (8.4), as may be implied by the durability of consumer goods, habit persistence in the preferences for consumption, nonseparability of preferences across states of nature, and other refinements. Singleton (1990), Ferson (1995), and Cochrane (2001) review this literature; Sarkissian (2003) provides a recent empirical example with references. The rest of this section provides a brief historical overview of empirical work on nonseparable-consumption models.

Dunn and Singleton (1986) and Eichenbaum et al. (1988) developed consumption models with durable goods. Durability introduces nonseparability over time, since the actual consumption at a given date depends on the consumer's previous expenditures. The consumer optimizes over the

current expenditures C_t, accounting for the fact that durable goods purchased today increase consumption at future dates, and thereby lower future marginal utilities. Thus, $U_c(C_t, .)$ in Equation (8.4) depends on expenditures prior to date t.

Another form of time nonseparability arises if the utility function exhibits "habit persistence." Habit persistence means that consumption at two points in time are complements. For example, the utility of current consumption may be evaluated relative to what was consumed in the past, so the previous standard of living influences the utility derived from current consumption. Such models are derived by Ryder and Heal (1973), Becker and Murphy (1988), Sundaresan (1989), Constantinides (1990), and Campbell and Cochrane (1999), among others.

Ferson and Constantinides (1991) model both durability and habit persistence in consumption expenditures. They show that the two combine as opposing effects. In an example based on the utility function of Equation (8.8), and where the "memory" is truncated at a single-lag, the derived utility of expenditures is:

$$U(C_t, .) = (1 - \alpha)^{-1} \Sigma_t \, \beta_t \, (C_t + bC_{t-1})^{1-\alpha}, \quad (8.10)$$

where the coefficient b is positive and measures the rate of depreciation if the good is durable and there is no habit persistence. Habit persistence implies that the lagged expenditures enter with a negative effect ($b < 0$). Empirical evidence on similar habit models is provided by Heaton (1993) and Braun et al. (1993), who find evidence for habit in international consumption and returns data.

Consumption expenditure data are highly seasonal, and Ferson and Harvey (1992) argue that the Commerce Department's X11 seasonal adjustment program may induce spurious time series behavior in the seasonally adjusted consumption data that most empirical studies have used. Using data that are not adjusted, they find strong evidence for a seasonal habit model.

Abel (1990) studied a form of habit persistence in which the consumer evaluates current consump-

tion relative to the aggregate consumption in the previous period, and which the consumer takes as exogenous. The idea is that people care about "keeping up with the Joneses." Campbell and Cochrane (1999) developed another model in which the habit stock is taken as exogenous (or "external") by the consumer. The habit stock in this case is modeled as a highly persistent weighted average of past aggregate consumptions. This approach results in a simpler and more tractable model, since the consumer's optimization does not have to take account of the effects of current decisions on the future habit stock In addition, by modeling the habit stock as an exogenous time series process, Campbell and Cochranes' model provides more degrees of freedom to match asset market data.

Epstein and Zin (1989, 1991) consider a class of recursive preferences that can be written as: $J_t = F(C_t, CEQ_t(J_{t+1}))$. $CEQ_t(.)$ is a time t "certainty equivalent" for the future lifetime utility J_{t+1}. The function $F(., CEQ_t(.))$ generalizes the usual expected utility function and may be nontime-separable. They derive a special case of the recursive preference model in which the preferences are:

$$J_t = \left[(1 - \beta)C_t^p + \beta \, E_t \, (J_{t+1}^{1-\alpha})^{p/(1-\alpha)} \right]^{1/p}. \quad (8.11)$$

They show that the IMRS for a representative agent becomes (when $p \neq 0$, $1 - \alpha \neq 0$):

$$m_{t+1} = [\beta(C_{t+1}/C_t)^{p-1}]^{(1-\alpha)/p} \, \{R_{m,t+1}\}^{((1-\alpha-p)/p)}. \quad (8.12)$$

The coefficient of relative risk aversion for timeless consumption gambles is α and the elasticity of substitution for deterministic consumption is $(1 - p)^{-1}$. If $\alpha = 1 - p$, the model reduces to the time-separable power utility model. If $\alpha = 1$, the log utility model of Rubinstein (1976) is obtained. Campbell (1993) shows that the Epstein–Zin model can be transformed to an empirically tractable model without consumption data. He used a linearization of the budget constraint that makes it

possible to substitute for consumption in terms of the factors that drive the optimal consumption function. Expected asset returns are then determined by their covariances with the underlying factors.

8.3. Multi-Beta Asset Pricing Models

Beta pricing models are a class of asset pricing models that imply the expected returns of securities are related to their sensitivity to changes in the underlying factors that measure the state of the economy. Sensitivity is measured by the securities' "beta" coefficients. For each of the relevant state variables, there is a market-wide price of beta measured in the form of an increment to the expected return (a "risk premium") per unit of beta.

The CAPM represented in Equation (8.7) is the premier example of a single-beta pricing model. Multiple-beta models were developed in continuous time by Merton (1973), Breeden (1979), and Cox et al. (1985). Long (1974), Sharpe (1977), Cragg and Malkiel (1982) and Connor (1984). Dybvig (1983), Grinblatt and Titman (1983), and Shanken (1987) provide multi-beta interpretations of equilibrium models in discrete time. Multiple-beta models follow when m_{t+1} can be written as a function of several factors. Equation (8.3) suggests that likely candidates for the factors are variables that proxy for consumer wealth, consumption expenditures, or the state variables – the sufficient statistics for the marginal utility of future wealth in an optimal consumption–investment plan. A multi-beta model asserts that the expected return is a linear function of several betas, i.e.

$$E(R_{it+1}) = \delta_0 + \Sigma_{j=1,\dots,K} \, \beta_{ij} \, \delta_j, \qquad (8.13)$$

where the β_{ij}, $j = 1, \dots, K$, are the multiple regression coefficients of the return of asset i on K economy-wide risk factors, f_j, $j = 1, \dots, K$. The coefficient δ_0 is the expected return on an asset that has $\beta_{0j} = 0$, for $j = 1, \dots, K$, i.e. it is the expected return on a zero-(multiple) beta asset. If

there is a risk-free asset, then δ_0 is the return for this asset. The coefficient δ_k, corresponding to the k'th factor has the following interpretation: it is the expected return differential, or premium, for a portfolio that has $\beta_{ik} = 1$ and $\beta_{ij} = 0$ for all $j \neq k$, measured in excess of the zero-beta asset's expected return. In other words, it is the expected return premium per unit of beta risk for the risk factor, k.

A multi-beta model, under certain assumptions, is equivalent to the SDF representation of Equation (8.2). This equivalence was first discussed, for the case of the CAPM, by Dybvig and Ingersoll (1982). The general multifactor case is derived by Ferson (1995) and Ferson and Jagannathan (1996), who show that the multi-beta expected return model of Equation (8.13) is equivalent to Equation (8.2), when the SDF is linear in the factors: $m_{t+1} = a_t + \Sigma_j b_{jt} \, f_{jt+1}$.

The logic of the equivalence between multi-beta pricing and the SDF representation of asset pricing models is easily seen using a regression example. Consider a regression of asset returns onto the factors, f_j of the multi-beta model. The regression model is $R_{it+1} = a_i + \Sigma_j \beta_{ijt} \, f_{jt} + u_{it+1}$. Substitute the regression equation into the right hand side of Equation (8.6) and assume that $\text{Cov}_t(u_{i,t+1}, m_{t+1}) = 0$. The result is:

$$
\begin{aligned}
E_t(R_{it+1}) = \delta_{0t} + \Sigma_{j=1,\dots K} \, \beta_{ijt} \\
[\text{Cov}_t\{f_{jt+1}, -m_{t+1}\}/E_t(m_{t+1})]
\end{aligned}
\qquad (8.14)
$$

which is a version of the multi-beta Equation (8.13). The market-wide risk premium for factor j is $\delta_{jt} = [\text{Cov}_t\{f_{jt+1}, -m_{t+1}\}/E_t(m_{t+1})]$ In the special case where the factor f_{jt+1} is a traded asset return, Equation (8.14) implies that $\delta_{jt} = E_t(f_{j,t+1}) - \delta_{0t}$; the expected risk premium equals the factor portfolio's expected excess return.

Equation (8.14) is useful because it provides intuition about the signs and magnitudes of expected risk premiums for particular factors. The intuition is essentially the same as in Equation (8.6). If a risk factor f_{jt+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 negatively related to marginal utility should carry a positive premium, because the big payoffs disappointingly come when the value of payoffs is low. This implies a low present value, and thus a high expected return. With a positive covariance the opposite occurs. If the factor is high when payoffs are highly valued, assets with a positive beta on the factor have a payoff distribution that is "better" than risk free. Thus, the expected return premium is negative, and such assets can have expected returns below that of a risk-free asset.

8.4. Relation to Mean–Variance Efficiency

The concept of a "minimum-variance portfolio" is central in the asset pricing literature. A portfolio R_{pt+1} is minimum variance if and only if no portfolio with the same expected return has a smaller variance. Roll (1977) and others have shown that a portfolio is minimum variance if and only if a single-beta pricing model holds, using the portfolio as the risk factor.[1] According to the CAPM, the market portfolio with return R_{mt+1} is minimum variance. If investors are risk averse, the CAPM also implies that R_{mt+1} is on the positively sloped portion of the minimum-variance frontier, or "mean–variance efficient." This implies that the coefficient δ_1 in Equation (8.7) is positive, which says that there is a positive tradeoff between market risk and expected return when investors are risk averse.

Multiple-beta asset pricing models imply that combinations of particular portfolios are minimum-variance efficient. Equation (8.13) is equivalent to the statement that a combination of K factor-portfolios is minimum-variance efficient, when the factors are traded assets. This result is proved by Grinblatt and Titman (1987), Shanken (1987), and Huberman et al. (1987). The correspondence between multi-beta pricing and mean variance efficiency is exploited by Jobson and Korkie (1982), Gibbons et al. (1989), Kandel and Stambaugh (1989), and Ferson and Siegel (2005), among others, to develop tests of multi-beta models based on mean variance efficiency.

8.5. Factor Models

A beta pricing model has no empirical content until the factors are specified, since there will almost always be a minimum-variance portfolio which satisfies Equation (8.13), with $K = 1$. Therefore, the empirical content of the model is the discipline imposed in selecting the factors. There have been four main approaches to finding empirical factors. The first approach is to specify empirical proxies for factors specified by the theory. For example, the CAPM says that the "market portfolio" of all capital assets is the factor, and early studies concentrated on finding good measures for the market portfolio. A second approach is to use factor analytic or principal components methods. This approach is motivated by the APT, as described below. A third approach chooses the risk factors as economic variables or portfolios, based on intuition such as that provided by Equations (8.3) and (8.4). With this approach, likely candidates for the factors are proxies for consumer wealth, consumer expenditures, and variables that may be sufficient statistics for the marginal utility of future wealth in an optimal consumption–investment plan. For examples of this approach, see Chen et al. (1986), Ferson and Harvey (1991), Campbell (1993), and Cochrane (1996). A fourth approach to factor selection forms portfolios by ranking stocks on firm characteristics that are correlated with the cross-section of average returns. For example, Fama and French (1993, 1996) use the ratio of book value to market price, and the relative market value (size) of the firm to form their "factors."

Lo and MacKinlay (1990), MacKinlay (1995), and Ferson et al. (1999) provide critiques of the approach of sorting stocks on empirically motivated characteristics in order to form asset pricing factors. Lo and MacKinlay examine the approach as a version of data mining. MacKinlay argues that the factors generated in this fashion by Fama and

French (1993, 1996) are statistically unlikely to reflect market risk premiums. Ferson, Sarkissian, and Simin show that a hypothetical characteristic, bearing an anomalous relation to returns, but completely unrelated to risk, can be repackaged as a spurious "risk factor" with this approach. Berk (1995) emphasizes that the price of a 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 ratios of cash flow to price. Some of the most empirically powerful characteristics for the cross-sectional prediction of stock returns are ratios, with market price per share in the denominator. However, patterns that are related to the cross-section of asset risks are also likely to be captured by sorting stocks on such ratios. Thus, the approach of sorting stocks on patterns in average returns to form factors is potentially dangerous, because it is likely to "work" when it "should" work, and it is also likely to work when it should not. At the time this chapter was written the controversy over such empirically motivated factors was unresolved.

8.6. Factor Models and the Arbitrage Pricing Model

The Arbitrage Pricing Model based on the APT of Ross (1976) is an example of a multiple-beta asset pricing model, although in the APT Equation (8.13) is an approximation. The expected returns are approximately a linear function of the relevant betas as the number of securities in the market grows without bound. Connor (1984) provided sufficient conditions for Equation (8.13) to hold exactly in an economy with an infinite number of assets, in general equilibrium. This version of the multiple-beta model, the exact APT, has received wide attention in the finance literature. See Connor and Korajczyk (1988), Lehmann and Modest (1988), Chen, (1983) and Burmeister, and McElroy (1988) for discussions on estimating and testing the model when the factor realizations are not observable, under auxiliary assumptions.

This section describes the Arbitrage Pricing Theory (APT) of Ross (1976), and how it is related to factor models and to the general SDF representation for asset pricing models, as in Equation (8.2). For this purpose, we suppress the time subscripts and related notation. Assume that the following data-generating model describes equity returns in excess of a risk-free asset:

$$r_i = E(r_i) + \beta'_{if} f + e_i, \qquad (8.15)$$

where $E(f) = 0 = E(e_{if})$, all i, and $f_t = F_t - E(F_t)$ 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 returns can then be expressed as:

$$\text{Cov}(R) \equiv \Sigma = BB' + V, \qquad (8.16)$$

where V is the covariance matrix of the residual vector, $\mathbf{e}$, B is the $N \times K$ matrix of the vectors, $\beta_{\mathbf{i}}$, and Σ is assumed to be nonsingular for all N. An "exact" factor structure assumes that V is diagonal. An approximate factor model, as described by Chamberlain (1983) and Chamberlain and Rothschild (1983), assumes that the eigenvalues of V are bounded as $N \to \infty$, while the K non-zero-eigenvalues of BB' become infinite as $N \to \infty$. Thus, the covariance matrix Σ has K unbounded and N–K bounded eigenvalues, as N becomes large.

The factor model represented in Equation (8.16) decomposes the variances of returns into "pervasive" and "nonsystematic" risks. If $\mathbf{x}$ is an N-vector of portfolio weights, the portfolio variance is $\mathbf{x}'\Sigma\mathbf{x}$, where $\lambda_{\max}(\Sigma)\mathbf{x}'\mathbf{x} \geq \mathbf{x}'\Sigma\mathbf{x} \geq \lambda_{\min}(\Sigma)\mathbf{x}'\mathbf{x}$, $\lambda_{\min}(\Sigma)$ being the smallest eigenvalue of Σ and $\lambda_{\max}(\Sigma)$ being the largest. Following Chamberlain (1983), a portfolio is "well diversified" if $\mathbf{x}'\mathbf{x} \to 0$ as N grows without bound. For example, an equally weighted portfolio is well diversified; in this case $\mathbf{x}'\mathbf{x} = (1/N) \to 0$. The bounded eigenvalues 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 (8.13), 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, also 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 0 as N grows.

The pricing errors in a beta pricing model are related to those of a SDF representation. If we define $\alpha_m = E(m\mathbf{R} - 1)$, where m is linear in the APT factors, then it follows 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 percent, 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 SDF and beta pricing paradigms.

The "exact" version of the APT derived by Connor (1984) asserts that $\alpha'\alpha \to 0$ as N grows without bound, and thus the pricing errors of all assets go to zero as the market gets large. 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 have $E(m\mathbf{R}-1) = 0$ when m is linear in the APT factors, and a combination of the factors is a minimum-variance efficient portfolio in the large market.

8.7. Summary

The asset pricing models of financial economics are based on an assumption that rules out arbitrage opportunities, or they rely on explicit equilibrium conditions. Empirically, there are three central representations. The first is the minimum-variance efficiency of a portfolio. The second is the beta pricing model stated in terms of risk factors, and the third is the SDF representation. These three representations are closely related, and become equivalent under ancilliary assumptions. Together they provide a rich and flexible framework for empirical analysis.

NOTE

1. It is assumed that the portfolio R_{pt+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.

REFERENCES

Abel, A.B. (1990). "Asset prices under habit formation and catching up with the Joneses." *American Economic Review,* 80: 38–42.

Barberis, N. and Shleifer, A. (2003). "Behavioral Finance," in G.M. Constantinides, M. Harris and R.M. Stulz (eds.) *Handbook of the Economics of Finance*. North Holland: Elsevier.

Becker, G. and Murphy, K.M. (1988). "A theory of rational addiction." *Journal of Political Economy*, 96: 675–700.

Beja, A. (1971). "The structure of the cost of capital under uncertainty." *Review of Economic Studies,* 4: 359–369.

Berk, J. (1995). "A critique of size-related anomalies." *Review of Financial Studies*, 8: 275–286.

Bhattacharya, S. (1981). "Notes on multiperiod valuation and the pricing of options." *Journal of Finance*, 36: 163–180.

Burmeister, E. and McElroy, M.B. (1988). "Joint estimation of factor sensitivities and risk premia for the Arbitrage Pricing Theory." *Journal of Finance*, 43(3): 721–733.

Braun, P., Ferson, W.E., and Constantinides, G.M. (1993). "Time nonseparability in aggregate consumption: international evidence." *European Economic Review*, 37: 897–920.

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

Brennan, M.J. and Kraus, A. (1978). "Necessary Conditions for Aggregation in Securities Markets." *Journal of Financial and Quantitative Analysis*, 407–418.

Campbell, J.Y. (1993). "Intertemporal Asset Pricing without consumption data." *American Economic Review*, 83: 487–512.

Campbell, J.Y. and Cochrane J.H. (1999). "By force of habit: a consumption-based explanation of aggregate stock market behavior." *Journal of Political Economy*, 107: 205–251.

Chamberlain, G. (1983). "Funds, factors and diversification in arbitrage pricing models." *Econometrica*, 51: 1305–1324.

Chamberlain, G. and Rothschild, M. (1983). "Arbitrage, factor structure and mean variance analysis on large asset markets." *Econometrica*, 51: 1281–1304.

Chen, N.F. (1983). "Some empirical tests of the theory of arbitrage pricing." *Journal of Finance*, 38: 1393–1414.

Chen, N., Roll, R., and Ross, S. (1986). "Economic forces and the stock market." *Journal of Business*, 59: 383–403.

Cochrane, J.H. (1996). "A cross-sectional test of a production based asset pricing model, working paper." *Journal of Political Economy*, 104: 572–621.

Cochrane, J.H. (2001). *Asset Pricing*. Princeton, NJ: Princeton University.

Connor, G. (1984). "A unified beta pricing Theory." *Journal of Economic Theory*, 34: 13–31.

Connor, G. and Korajczyk, R. (1988). "Risk and return in an equilibrium APT: application of a new test methodology." *Journal of Financial Economics*, 21: 255–290.

Constantinides, G.M. (1982). "Intertemporal Asset Pricing with heterogeneous consumers and without demand Aggregation." *Journal of Business*, 55: 253–267.

Constantinides, G.M. (1990). "Habit formation: a resolution of the equity premium puzzle." *Journal of Political Economy*, 98: 519–543.

Constantinides, G.M. and Duffie, D. (1996). "Asset pricing with heterogenous consummers." *Journal of Political Economy*, 104: 219–240

Cox, J.C., Ingersoll, J.E., and Ross, S.A. (1985). "A theory of the term structure of interest rates." *Econometrica*, 53: 385–408.

Cragg, J.G. and Malkiel, B.G. (1982). *Expectations and the Structure of Share Prices*. Chicago: University of Chicago Press.

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

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. and Ingersoll, J. (1982). "Mean variance theory in complete markets." *Journal of Business*, 55: 233–252.

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

Epstein, L.G. and Zin, S.E. (1989). "Substitution, risk aversion and the temporal behavior of asset returns: a theoretical approach." *Econometrica*, 57: 937–970.

Epstein, L.G. and Zin, S.E. (1991). "Substitution, risk aversion and the temporal behavior of asset returns." *Journal of Political Economy*, 99: 263–286.

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

Fama, E.F. and French, K.R. (1996). "Multifactor explanations of asset pricing anomalies." *Journal of Finance*, 51: 55–87.

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. (1995). "Theory and empirical testing of asset pricing models," in J. Maksimovic, and Ziemba (eds.) *Handbooks in Operations Research and Management Science*. Amsterdam: Elsevier.

Ferson, W.E. and Constantinides, G.M. (1991). "Habit persistence and durability in aggregate consumption: empirical tests." *Journal of Financial Economics*, 29: 199–240.

Ferson, W.E. and Harvey, C.R. (1991). "The Variation of Economic Risk Premiums." *Journal of Political Economy*, 99: 385–415.

Ferson, W.E. and Harvey, C.R. (1992). "Seasonality and consumption based Asset Pricing Models." *Journal of Finance*, 47: 511–552.

Ferson, W.E. and Jagannathan, R. (1996). "Econometric evaluation of Asset Pricing Models," in G.S. Maddala and C.R. Rao (eds.) *Handbook of Statistics: volume 14: Statistical Methods in Finance*. Amsterdam, the Netherlands.

Ferson, W.E. and Merrick, J.J. (1987) "Non-stationarity and stage of the business cycle effects in consumption-based asset pricing relations." *Journal of Financial Economics*, 18: 127–146.

Ferson, W.E., Sarkissian, S., and Simin, T. (1999). "The alpha factor Asset Pricing Model: a parable." *Journal of Financial Markets*, 2: 49–68.

Ferson, W.E. and Siegel, A.F. (2005). "Testing portfolio efficiency with conditioning information," Working paper, Boston College.

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

Gorman, W.M. (1953). "Community Preference Fields." *Econometrica*, 21: 63–80.

Grinblatt, M. and Titman, S. (1983). "Factor pricing in a finite economy." *Journal of Financial Economics*, 12: 497–508.

Grinblatt, M. and Titman, S. (1987). "The relation between mean-variance efficiency and arbitrage pricing." *Journal of Business*, 60: 97–112.

Grossman, S. and Shiller, R.J. (1982). "Consumption correlatedness and risk measurement in economies with nontraded assets and heterogeneous information." *Journal of Financial Economics*, 10: 195–210.

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

Hansen, L.P. and Singleton, K.J. (1982). "Generalized instrumental variables estimation of nonlinear rational expectations models." *Econometrica*, 50: 1269–1285.

Hansen, L.P. and Singleton, K.J. (1983). "Stochastic consumption, risk aversion and the temporal behavior of asset returns." *Journal of Political Economy*, 91: 249–266.

Harrison, M. and Kreps, D. (1979). "Martingales and arbitrage in multi-period securities markets." *Journal of Economic Theory*, 20: 381–408.

Heaton, J. (1993). "The interaction between time-non-separable preferences and time aggregation." *Econometrica*, 61(2): 353–385.

Huberman, G., Kandel, S.A., and Stambaugh, R.F. (1987). "Mimicking portfolios and exact arbitrage pricing." *Journal of Finance*, 42: 1–10.

Jobson, J.D. and Korkie, R. (1982). "Potential performance and tests of portfolio efficiency." *Journal of Financial Economics*, 10: 433–466.

Kandel, S. and Stambaugh, R.F. (1989). "A mean variance framework for tests of asset pricing models." *Review of Financial Studies*, 2: 125–156.

Lee, C.F., Wu, C., and Wei, K.C.J. (1990). "The heterogeneous investment horizon and the capital asset pricing model: theory and implications." *Journal of Financial and Quantitative Analysis,* 25: 361–376.

Lehmann, B.N. and Modest, D.M. (1988). "The empirical foundations of the arbitrage pricing theory." *Journal of Financial Economics*, 21: 213–254.

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.

Lo, A.W. and MacKinlay, A.C. (1990). "Data snooping in tests of financial asset pricing models." *Review of Financial Studies*, 3: 431–467.

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.

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

Mossin, J. (1966). "Equilibrium in a capital asset market." *Econometrica*, 34(4): 768–783.

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.

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

Ross, S. (1977). "Risk, return and arbitrage," in I. Friend and J. Bicksler (eds.) *Risk and Return in Finance Ballinger*. Cambridge: Mass.

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.

Ryder, Jr. H.E. and Heal, G.M. (1973). "Optimal growth with intertemporally dependent preferences." *Review of Economic Studies*, 40: 1–31.

Sarkissian, S. (2003). "Incomplete consumption risk sharing and currency premiums." *Review of Financial Studies*, 16: 983–1005.

Shanken, J. (1987). "Multivariate Proxies and Asset Pricing Relations: Living with the Roll Critique." *Journal of Financial Economics*, 18: 91–110.

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*. New York: Academic Press.

Singleton, K.J. (1990). "Specification and estimation of intertemporal Asset Pricing Modelas," in E. Friedman and F. Hahn (eds.) *Handbook of Monetary Economics*. North Holland: Elsevier.

Sundaresan, S. M. (1989). "Intertemporally dependent preferences and the volatility of consumption and wealth." *Review of Financial Studies*, 2: 73–89.

Wilson, R. B. (1968). "The theory of syndicates." *Econometrica*, 36: 119–131.