THE JOURNAL OF FINANCE • VOL. LI, NO. 1 • MARCH 1996

The Conditional CAPM and the Cross-Section of Expected Returns

RAVI JAGANNATHAN and ZHENYU WANG*

ABSTRACT

Most empirical studies of the static CAPM assume that betas remain constant over time and that the return on the value-weighted portfolio of all stocks is a proxy for the return on aggregate wealth. The general consensus is that the static CAPM is unable to explain satisfactorily the cross-section of average returns on stocks. We assume that the CAPM holds in a conditional sense, i.e., betas and the market risk premium vary over time. We include the return on human capital when measuring the return on aggregate wealth. Our specification performs well in explaining the cross-section of average returns.

A SUBSTANTIAL PART OF the research effort in finance is directed toward improving our understanding of how investors value risky cash flows. It is generally agreed that investors demand a higher expected return for investment in riskier projects, or securities. However, we still do not fully understand how investors assess the risk of the cash flow on a project and how they determine what risk premium to demand. Several capital asset-pricing models have been suggested in the literature that describe how investors assess risk and value risky cash flows. Among them, the Sharpe-Lintner-Black Capital Asset Pricing Model (CAPM)[1] is the one that financial managers use most often for assessing the risk of the cash flow from a project and for arriving at the appropriate

* Jagannathan is from the Carlson School of Management, University of Minnesota; Hong Kong University of Science and Technology; and the Federal Reserve Bank of Minneapolis. Wang is from Columbia University. The authors benefited from discussions with S. Rao Aiyagari, Gordon Alexander, Michael Brennan, Charlie Calomiris, Mark Carhart, Eugene Fama, Wayne Ferson, Murray Frank, John Geweke, Lars Peter Hansen, Campbell Harvey, Pat Hess, Allan Kleidon, Peter Knez, Robert Korajczyk, Bruce Lehmann, Steve LeRoy, David Modest, Edward Prescott, Judy Rayburn, Richard Roll, Jay Shanken, Michael Sher, Stephen D. Smith, and Robert Stambaugh, as well as with participants at numerous finance workshops in the United States, Canada, and East Asia. Special thanks go to the anonymous referee and the managing editor of the journal for valuable comments. We are grateful to Eugene Fama for providing us with the Fama-French factors and Raymond A. Dragan for editorial assistance. All errors in this paper are the authors' responsibility. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. Ravi Jagannathan gratefully acknowledges financial support from the National Science Foundation (grant SBR-9409824). Zhenyu Wang gratefully acknowledges financial support from the Alfred P. Sloan Foundation (doctoral dissertation fellowship, grant DD-518). An earlier version of the paper appeared under the title, "The CAPM Is Alive and Well." The compressed archive of the data and the FORTRAN programs used for this paper can be obtained via anonymous FTP at ftp.socsci.umn.edu. The path is outgoing/wang/capm.tar.Z.

[1] See Sharpe (1964), Lintner (1965), and Black (1972).

discount rate to use in valuing the project. According to the CAPM, (a) the risk of a project is measured by the beta of the cash flow with respect to the return on the market portfolio of all assets in the economy, and (b) the relation between required expected return and beta is linear.

Over the past two decades a number of studies have empirically examined the performance of the static version of the CAPM in explaining the cross-section of realized average returns. The results reported in these studies support the view that it is possible to construct a set of portfolios such that the static CAPM is unable to explain the cross-sectional variation in average returns among them.[2] In particular, portfolios containing stocks with relatively small capitalization appear to earn higher returns on average than those predicted by the CAPM.[3]

In spite of the lack of empirical support, the CAPM is still the preferred model for classroom use in MBA and other managerial finance courses. In a way it reminds us of cartoon characters like Wile E. Coyote who have the ability to come back to original shape after being blown to pieces or hammered out of shape. Maybe the CAPM survives because (a) the empirical support for other asset-pricing models is no better,[4] (b) the theory behind the CAPM has an intuitive appeal that other models lack, and (c) the economic importance of the empirical evidence against the CAPM reported in empirical studies is ambiguous.

In their widely cited study, Fama and French (1992) present evidence suggesting that the inability of the static CAPM to explain the cross-section of average returns that has been reported in the literature may be economically important. Using return data on a large collection of assets, they examine the static version of the CAPM and find that the "relation between market beta and average return is flat."[5] The CAPM is widely viewed as one of the two or three major contributions of academic research to financial managers during the postwar era. As Fama and French point out, the robustness of the size effect and the absence of a relation between beta and average return are so contrary to the CAPM that they shake the foundations on which MBA and other managerial course materials in finance are built.

The CAPM was derived by examining the behavior of investors in a hypothetical model-economy in which they live for only one period. In the real world investors live for many periods. Therefore, in the empirical examination of the CAPM, using data from the real world, it is necessary to make certain assumptions. One of the commonly made assumptions is that the betas of the assets remain constant over time. In our view, this is not a particularly reasonable assumption since the relative risk of a firm's cash flow is likely to vary over the

[2] See Banz (1981), Reinganum (1981), Gibbons (1982), Basu (1983), Chan, Chen, and Hsieh (1985), Shanken (1985), and Bhandari (1988).

[3] Hansen and Jagannathan (1994) find that this is true even after controlling for systematic risk using a variety of other measures.

[4] See Hansen and Singleton (1982), Connor and Korajczyk (1988a and 1988b), Lehmann and Modest (1988), and Hansen and Jagannathan (1991 and 1994).

[5] Also see Jegadeesh (1992), who obtains results similar to Fama and French.

business cycle. During a recession, for example, financial leverage of firms in relatively poor shape may increase sharply relative to other firms, causing their stock betas to rise. Also, to the extent that the business cycle is induced by technology or taste shocks, the relative share of different sectors in the economy fluctuates, inducing fluctuations in the betas of firms in these sectors. Hence, betas and expected returns will in general depend on the nature of the information available at any given point in time and vary over time. In this study, therefore, we assume that the conditional version of the CAPM holds, i.e., the expected return on an asset based on the information available at any given point in time is linear in its conditional beta.

Though several researchers have empirically examined the conditional version of the CAPM, no one to our knowledge has directly studied the ability of the conditional CAPM to explain the cross-sectional variation in average returns on a large collection of stock portfolios. The focus of our paper is to fill this gap in the literature. For this purpose, we first derive the unconditional model implied by the conditional CAPM. We show that when the conditional version of the CAPM holds (i.e., when betas and expected returns are allowed to vary over the business cycle), a two-factor model obtains unconditionally. Average returns are jointly linear in the average beta and in a measure of "beta instability," which we show how to calculate. The fact that the implied unconditional model nests the static CAPM facilitates direct comparison of their relative performance.

Using the value-weighted index from CRSP as the market portfolio, we find that the unconditional model implied by the conditional CAPM explains nearly 30 percent of the cross-sectional variation in average returns of 100 stock portfolios similar to those used in Fama and French (1992). This is a substantial improvement when compared to the 1 percent explained by the static CAPM. The rejection by the data and the size effect are much weaker than those for the static CAPM.

In order to implement the CAPM, for practical purposes, it is commonly assumed that the return on the value-weighted portfolio of all stocks listed on the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX) (as well as those traded on Nasdaq) is a reasonable proxy for the return on the market portfolio of all assets. In view of this, another possible interpretation of the evidence is that the particular proxy Fama and French (1992) use for the return on the market portfolio of all assets is a major cause for the unsatisfactory performance of the CAPM. Hence, in measuring the return on aggregate wealth, we follow Mayers (1972) and include a measure of return on human capital. We find that when human capital is also included in measuring wealth, the unconditional model implied by the conditional CAPM is able to explain over 50 percent of the cross-sectional variation in average returns, and the data fail to reject the model. More importantly, size and book-to-market variables have little ability to explain what is left unexplained.

The rest of the paper is organized as follows: In Section I, we show that when the CAPM holds in a conditional sense (i.e., expected returns and betas vary over time in a systematic stochastic manner), unconditional expected returns on assets will be linear in (a) the average beta and (b) a measure of beta instability over time. When betas remain constant over time, this model collapses to the familiar static CAPM. In Section II, we show how to examine this model empirically. Section III describes the data and presents the empirical results. We draw our conclusions in Section IV.

I. Models for the Expected Stock Returns

A. The Sharpe-Lintner-Black (Static) CAPM

Let R_i denote the return on any asset i and R_m be the return on the market portfolio of all assets in the economy. The Black (1972) version of the CAPM is

$$E[R_i] = \gamma_0 + \gamma_1\beta_i, \quad (1)$$

where β_i is defined as

$$\beta_i = \text{Cov}(R_i, R_m)/\text{Var}[R_m],$$

and $E[\cdot]$ denotes the expectation, $\text{Cov}(\cdot)$ denotes the covariance, and $\text{Var}[\cdot]$ denotes the variance.

In their widely cited study, Fama and French (1992) empirically examine the CAPM given above and find that the estimated value of γ_1 is close to zero. They interpret the "flat" relation between average return and beta as strong evidence against the CAPM.

While a "flat" relation between average return (the sample analog of the unconditional expected return) and beta may be evidence against the static CAPM, it is not necessarily evidence against the conditional CAPM. The CAPM was developed within the framework of a hypothetical single-period model economy. The real world, however, is dynamic and hence, as pointed out earlier, expected returns and betas are likely to vary over time. Even when expected returns are linear in betas for every time period, based on the information available at the time, the relation between the unconditional expected return and the unconditional beta could be "flat."[6] The following example illustrates this point.

Consider a hypothetical economy in which the CAPM holds period by period. Suppose that the econometrician considers only two stocks and that there are only two possible types of dates in the world. The betas of the first stock in the two date-types are, respectively, 0.5 and 1.25 (corresponding to an average beta of 0.875). The corresponding betas of the second stock are 1.5 and 0.75 (corresponding to an average beta of 1.125). Suppose that the expected risk premium on the market is 10 percent on the first date and 20 percent on the second date. Then, if the CAPM holds in each period, the expected risk premium on the first stock will be 5 percent on the first date and 25 percent on the second date. The expected risk premium on the second stock will be 15

[6] This is because an asset that is on the conditional mean-variance frontier need not be on the unconditional frontier, as Dybvig and Ross (1985) and Hansen and Richard (1987) point out.

percent on both dates. Hence, an econometrician who ignores the fact that betas and risk premiums vary over time will mistakenly conclude that the CAPM does not hold, since the two stocks earn an average risk premium of 15 percent, but their average betas differ. While the numbers we use in this example are rather extreme and unrealistic, they do illustrate the pitfalls involved in any empirical study of the CAPM that ignores time variation in betas.

The need to take time variation in betas into account is also demonstrated by the commercial success of firms like BARRA, which provide beta estimates for risk management and valuation purposes, using elaborate time-series models. Several empirical studies of beta-pricing models reported in the literature find that betas exhibit statistically significant variability over time.[7] Moreover, in empirical studies that examine the reaction of stock prices to certain events (referred to as "event studies" in the financial economics literature), it has become common practice to allow for time variations in betas, following Mandelker (1974). Hence, the inconclusive nature of the empirical evidence for the static CAPM may well be due to systematic stochastic changes affecting the environment that generates returns, as pointed out by Black (1993) and Chan and Lakonishok (1993).

In the next section, we will therefore assume that the CAPM holds in a conditional sense, i.e., it holds at every point in time, based on whatever information is available at that instant. We will then derive an unconditional asset-pricing model starting from the conditional version of the CAPM.

B. The Conditional CAPM

We use the subscript t to indicate the relevant time period. For example, R_{it} denotes the gross (one plus the rate of) return on asset i in period t, and R_{mt} the gross return on the aggregate wealth portfolio of all assets in the economy in period t. We refer to R_{mt} as the market return. Let I_{t-1} denote the common information set of the investors at the end of period $t - 1$. We assume that all the time series in this paper are covariance stationary and that all the conditional and unconditional moments that we use in the paper exist.

Risk-averse rational investors living in a dynamic economy will typically anticipate and hedge against the possibility that investment opportunities in the future may change adversely. Because of this hedging need that arises in a dynamic economy, the conditionally expected return on an asset will typically be jointly linear in the conditional market beta and "hedge portfolio betas."[8] However, following Merton (1980), we will assume that the hedging motives are not sufficiently important, and hence the CAPM will hold in a conditional sense as given below.

[7] See Bollerslev, Engle, and Wooldridge (1988), Harvey (1989), Ferson and Harvey (1991, 1993), and Ferson and Korajczyk (1993).

[8] See Merton (1973) and Long (1974).

THE CONDITIONAL CAPM. *For each asset i and in each period t,*

$$E[R_{it}|I_{t-1}] = \gamma_{0t-1} + \gamma_{1t-1}\beta_{it-1}, \tag{2}$$

where β_{it-1} *is the conditional beta of asset i defined as*

$$\beta_{it-1} = \mathrm{Cov}(R_{it}, R_{mt}|I_{t-1})/\mathrm{Var}(R_{mt}|I_{t-1}), \tag{3}$$

γ_{0t-1} *is the conditional expected return on a "zero-beta" portfolio, and* γ_{1t-1} *is the conditional market risk premium.*

Since our aim is to explain the cross-sectional variations in the unconditional expected return on different assets, we take the unconditional expectation of both sides of equation (2) to get

$$E[R_{it}] = \gamma_0 + \gamma_1\bar{\beta}_i + \mathrm{Cov}(\gamma_{1t-1}, \beta_{it-1}), \tag{4}$$

where

$$\gamma_0 = E[\gamma_{0t-1}] \qquad \gamma_1 = E[\gamma_{1t-1}] \qquad \bar{\beta}_i = E[\beta_{it-1}].$$

Here, γ_1 is the expected market risk premium, and $\bar{\beta}_i$ is the expected beta.[9] If the covariance between the conditional beta of asset i and the conditional market risk premium is zero (or a linear function of the expected beta) for every arbitrarily chosen asset i, then equation (4) resembles the static CAPM, i.e., the expected return is a linear function of the expected beta. However, in general, the conditional risk premium on the market and conditional betas are correlated. During bad economic times when the expected market risk premium is relatively high, firms on the "fringe" and more leveraged firms are more likely to face financial difficulties and thus have higher conditional betas. If the uncertainty associated with future growth opportunities is the cause for the higher beta of firms on the "fringe," then their conditional betas will be relatively low during bad economic times, resulting in natural perverse market timing. This is because during bad times the uncertainty as well as the value of future growth opportunities is reduced, and this effect may more than offset the effect of increased leverage.

In fact, we know from earlier studies that the expected risk premium on the market as well as conditional betas are not constant (Keim and Stambaugh (1986), Breen, Glosten, and Jagannathan (1989)), and vary over the business cycle (Fama and French (1989), Chen (1991), and Ferson and Harvey (1991)). Therefore, in general the last term in equation (4) is not zero, and the unconditional expected return is not a linear function of the expected beta alone.

Notice that the last term in equation (4) depends only on the part of the conditional beta that is in the linear span of the market risk premium. This motivates us to decompose the conditional beta of any asset i into two orthogonal components by projecting the conditional beta on the market risk pre-

[9] Note that expected betas are not the same as unconditional betas, but we will relate the two in the next subsection.

mium. For each asset i, we define the *beta-prem sensitivity* (denoted by ϑ_i) and *residual beta* (denoted by η_{it-1}) as follows:

$$\vartheta_i = \mathrm{Cov}(\beta_{it-1}, \gamma_{1t-1})/\mathrm{Var}(\gamma_{1t-1}) \tag{5}$$

$$\eta_{it-1} = \beta_{it-1} - \bar{\beta}_i - \vartheta_i(\gamma_{1t-1} - \gamma_1). \tag{6}$$

In the above expression, beta-prem sensitivity, ϑ_i, measures the sensitivity of conditional beta to the market risk premium. It can be verified that, for each asset i, we have

$$\beta_{it-1} = \bar{\beta}_i + \vartheta_i(\gamma_{1t-1} - \gamma_1) + \eta_{it-1}, \tag{7}$$

$$\mathrm{E}[\eta_{it-1}] = 0, \tag{8}$$

$$\mathrm{E}[\eta_{it-1}\gamma_{it-1}] = 0. \tag{9}$$

Equation (7) decomposes each conditional beta (which is a random variable) into three orthogonal parts. The first part is the expected beta, which is a constant. The second part is a random variable that is perfectly correlated with the market risk premium. The last part is on average zero and uncorrelated with the market risk premium.

C. Implications for Unconditional Expected Returns

Substituting (7) into (4) gives

$$\mathrm{E}[R_{it}] = \gamma_0 + \gamma_1\bar{\beta}_i + \mathrm{Var}(\gamma_{1t-1})\vartheta_i. \tag{10}$$

Hence, cross-sectionally, the unconditional expected return on any asset i is a linear function of its expected beta and its beta-prem sensitivity. The larger this sensitivity, the larger is the variability of the above second part of the conditional beta. In this sense, the beta-prem sensitivity of an asset measures the instability of the asset's beta over the business cycle. Stocks with higher expected betas have higher unconditional expected returns. Likewise, stocks with betas that are prone to vary with the market risk premium and hence are less stable over the business cycle also have higher unconditional expected returns. Hence, the one-factor conditional CAPM leads to a two-factor model for unconditional expected returns.

A complete test of the conditional CAPM specification given in (2) requires estimation of expected beta, $\bar{\beta}_i$, and beta-prem sensitivity, ϑ_i, given in (10) as well as other parameters. This requires additional restrictive assumptions regarding the nature of the stochastic process governing the joint temporal evolution of conditional market betas and the conditional market risk premium.[10] However, our objective is to examine whether the unconditional expected

[10] See Bodurtha and Mark (1991) for an empirical examination of the conditional CAPM specification under more restrictive assumptions.

returns are consistent with the conditional CAPM. Because of the limited scope of our study, we can get by with somewhat less restrictive assumptions.

It can be seen from equation (10) that the residual betas do not affect the unconditional expected return. So, when considering unconditional returns, we can ignore η_{it-1} and concentrate on the first two parts of each conditional beta. Since we cannot estimate ϑ_i and $\bar{\beta}_i$, we look directly at how the stock returns react to the market return on average and how they respond to the changes of the market risk premium. This leads us to define the following two types of unconditional betas:

$$\beta_i \equiv \mathrm{Cov}(R_{it}, R_{mt})/\mathrm{Var}(R_{mt}), \tag{11}$$

$$\beta_i^\gamma \equiv \mathrm{Cov}(R_{it}, \gamma_{1t-1})/\mathrm{Var}(\gamma_{1t-1}). \tag{12}$$

We refer to the first unconditional beta as the *market beta* and the second as the *premium beta*. They measure the *average market risk* and *beta-instability risk*, respectively.

In Appendix A, we show that under rather mild assumptions, the unconditional expected return is a linear function of the above two unconditional betas. This is summarized as the following theorem:

THEOREM 1. *If β_i^γ is not a linear function of β_i, then there are some constants a_0, a_1, and a_2 such that the equation*

$$\mathrm{E}[R_{it}] = a_0 + a_1\beta_i + a_2\beta_i^\gamma \tag{13}$$

holds for every asset i.

The two-beta model presented here is not a special case of the multi-beta capital asset-pricing models commonly seen in finance literature. For example, according to the general equilibrium multi-beta model of Merton (1973), the conditionally expected return is linear in several conditional betas, one of which is the market beta. In contrast, we assume that the conditionally expected return is linear in the conditional market beta alone. From this, we show that the unconditional expected return is linear in the market beta and the premium beta. Also, there are several important differences between the two-beta model given above and the two-beta version of the linear factor models that owe their origins to the model first proposed by Ross (1976). First, we do not assume that returns have a linear factor structure as is commonly assumed in linear factor models. Second, γ_{1t-1} is a predetermined variable and is not a factor in the sense commonly understood.

II. Econometric Specifications and Tests

A. Empirical Specifications

The model given in equation (13) forms the basis for our empirical work. In order to empirically examine whether equation (13) can explain the cross-section of expected returns on stocks, we need some further assumptions to estimate the model using time-series data. First, we need observations on the conditional market risk premium γ_{1t-1} for computing β_i^γ. Since the conditional market risk premium depends on the nature of the information available to the investors and how they make use of it, we have to take a stand on the information set for investors. Second, the return on the aggregate wealth portfolio of all assets in the economy is not observable. Hence, we need to use a proxy for R_{mt} as well. We discuss these issues in the rest of this subsection.

A.1. The Proxy for the Conditional Market Risk Premium, γ_{1t-1}

There is a general agreement in the literature that stock prices vary over the business cycle. Hence, one may suspect that the market risk premium will also vary over the business cycle.[11] This observation suggests making use of the same variables that help predict the business cycle for forecasting the market risk premium as well.

While a number of variables may help predict future economic conditions, we need to restrict attention to a small number of such variables in order to ensure that we are able to estimate the parameters of interest with some degree of precision. For convenience, we have decided to restrict our attention in this study to only one forecasting variable. To determine which variable we should pick, we examined the literature on business-cycle forecasting. Our reading of this literature suggests that, in general, interest-rate variables are likely to be most helpful in predicting future business conditions. Stock and Watson (1989) examine several variables and find that the spread between six-month commercial paper and six-month Treasury bill rates and the spread between ten- and one-year Treasury bond rates both outperform nearly every other variable as a forecaster of the business cycle. Bernanke (1990), who runs "a horse race" between a number of interest-rate variables, finds that the best single variable is the spread between the commercial paper rate and Treasury bill rate first used by Stock and Watson.

Based on these findings, we choose the yield spread between BAA- and AAA-rated bonds, denoted by R_{t-1}^{prem}, as a proxy for the market risk premium. The variable R_{t-1}^{prem} is similar to the spread between commercial paper and the Treasury bill rates, but it has been used extensively in finance. In addition, we also assume that the market risk premium is a linear function of R_{t-1}^{prem}, i.e.,

ASSUMPTION 1. *There are some constants κ_0, κ_1 such that*

$$\gamma_{1t-1} = \kappa_0 + \kappa_1 R_{t-1}^{\text{prem}}. \tag{14}$$

For each asset i, we define *prem-beta* as

$$\beta_i^{\text{prem}} = \text{Cov}(R_{it}, R_{t-1}^{\text{prem}})/\text{Var}(R_{t-1}^{\text{prem}}). \tag{15}$$

[11] Keim and Stambaugh (1986), Fama and French (1989), and Chen (1991) provide empirical evidence that supports this view.

Under Assumption 1, the expected return is linear in its prem-beta and its market beta. To see this, we can substitute (14) into (12) and make use of (15) and Theorem 1 to obtain the following corollary:

COROLLARY 1. *Suppose that β_i^γ is not a linear function of β_i and that Assumption 1 holds, then there are some constants c_0, c_m, and c_{prem} such that the equation*

$$E[R_{it}] = c_0 + c_m\beta_i + c_{\text{prem}}\beta_i^{\text{prem}} \tag{16}$$

holds for every asset i.

A.2. The Proxy for the Return on the Wealth Portfolio, R_{mt}

In empirical studies of the CAPM it is commonly assumed that the return on the value-weighted portfolio of all stocks traded in the United States is a good proxy for the return on the portfolio of the aggregate wealth. Let R_t^{vw} denote the return on the value-weighted *stock index* portfolio. The implicit assumption is that the market return is a linear function of the stock index, i.e., there are some constants ϕ_0 and ϕ_{vw} such that

$$R_{mt} = \phi_0 + \phi_{\text{vw}}R_t^{\text{vw}}. \tag{17}$$

Let us define the *vw-beta* as

$$\beta_i^{\text{vw}} = \text{Cov}(R_{it}, R_t^{\text{vw}})/\text{Var}(R_t^{\text{vw}}). \tag{18}$$

Suppose that the static CAPM in equation (1) holds unconditionally as well. In this case, we can substitute (17) into equation (11) and use equation (18) and the static CAPM to obtain the following linear relation between the unconditional expected return and the vw-beta:

$$E[R_{it}] = c_0 + c_{\text{vw}}\beta_i^{\text{vw}}, \tag{19}$$

where c_0 and c_{vw} are some constants.

This is the specification that is commonly used in empirical studies of the static CAPM. Hence, tests of the CAPM based on this specification can be interpreted as a joint test of two hypotheses: (i) the static CAPM holds, and (ii) the market return is a linear function of the stock index return. Consequently, the results of these investigations are open to various interpretations. In particular, the reason for the empirical rejections of equation (19) may be that the static CAPM does not hold. Alternatively, it may also be the case that the static CAPM holds, but the return on the stock index portfolio is a poor proxy for the return on the aggregate wealth. Roll (1977) makes a related observation that the market portfolio is not observable. It is possible that the value-weighted index of stocks is a poor proxy for the portfolio of the aggregate wealth; hence, this might be the reason for the poor performance of the CAPM under empirical examination. In fact, Mayers (1972) points out that human capital forms a substantial part of the total capital in the economy. Following

Mayers' suggestion, we therefore consider extending the proxy for the market return to include a measure of return on human capital.

To appreciate the need for examining other proxies for systematic risk, note that stocks form only a small part of the aggregate wealth. The monthly per capita income in the United States from dividends during the period 1959:1–1992:12 was less than 3 percent of the monthly personal income from all sources, whereas income from salaries and wages was about 63 percent during the same period. While these income flows ignore capital gains, these proportions remained relatively steady during this period.[12] This suggests that common stocks of all corporations constitute about a thirtieth of national income and probably national wealth as well. Another way to see this is as follows: As Diaz-Gimenez *et al.* (1992) point out, almost two-thirds of nongovernment tangible assets are owned by the household sector, and only one-third is owned by the corporate sector (p. 536, op. cit.). Approximately a third of the corporate assets are financed by equity (see Table 2, op. cit.). Hence, it appears that the return on stocks alone is unlikely to measure the return on aggregate wealth sufficiently accurately.

Apparently, the observation that stocks form only a small part of the total wealth is what motivated Stambaugh (1981 and 1982) to examine the sensitivity of the CAPM to different proxies for the market portfolio. In his seminal comparative study of the various market proxies, he finds that "even when stocks represent only 10 percent of the portfolio value, inferences about the CAPM are virtually identical to those obtained with a stock-only portfolio." However, he does not consider the return on human capital in his otherwise extensive study.

The commonly held view appears to be that human capital is not tradable and hence should be treated differently from other capital (see Mayers (1972)). This view is not entirely justified. First, note that mortgage loans, which are in most cases borrowing against future income, constitute about a third of all outstanding loans. At the end of 1986 the total market value of equities held by the households category was 0.80 GNP, whereas the outstanding stock of mortgages (0.60 GNP), consumer credit (0.16 GNP), and bank loans to the household sector (0.04 GNP) also amounted to 0.80 GNP (Table 4, Diaz-Gimenez *et al.* (1992)). Second, active insurance markets exist for hedging the risk in human capital. Examples include life insurance, unemployment insurance, and medical insurance. Hence, it does not appear inappropriate, as a first approximation, to take the view that human capital is just like any other form of physical capital, cash flows from which are traded through issuance of financial assets.

There is, however, an important difference between human capital and other physical assets owned by corporations. Typically, the entire cash flow that arises from the use of the physical assets employed by firms is promised away by issuing financial securities. This is not the case with human capital, where

[12] See Table 2.2 in *National Income and Product Account of the U.S.* published by the Bureau of Economic Analysis, the U.S. Department of Commerce.

only a portion of the labor income is secured by issuing mortgages. Also, in contrast to stocks, which are the residual claimants in the firm, the cash flow from labor income promised to mortgages comes from the top. Hence, factors that affect the return on human capital cannot be identified precisely enough by examining returns on financial assets like mortgages. We therefore follow a different strategy in measuring the return on human capital.

We assume that the return on human capital is an exact linear function of the growth rate in per capita labor income. While the use of the growth rate in per capita labor income is rather ad hoc, we can provide some rationale for using it. For example, suppose that, to a first order of approximation, the expected rate of return on human capital is a constant r, and that date-t per capita labor income L_t follows an autoregressive process of the form

$$L_t = (1 + g)L_{t-1} + \varepsilon_t.$$

In such a case, the realized capital-gain part of the rate of return on human capital (not corrected for additional investment in human capital made during the period) will be the realized growth rate in per capita labor income. To see this, note that under these assumptions, wealth due to human capital is

$$W_t = \frac{L_t}{r - g}.$$

The rate of change in wealth is then given by

$$R_t^{\text{labor}} = \frac{L_t - L_{t-1}}{L_{t-1}}.$$

Fama and Schwert (1977) arrive at a similar measure based on different lines of reasoning. Following the inter-temporal asset-pricing model, Campbell derives a measure for the return on human capital, which is the above current growth rate of labor income, plus a term that depends on expected future growth rates of labor income and expected future asset returns (see equation (3.3) in Campbell (1993b)). If both the forecastable part of the growth rates of labor income and the forecastable part of the returns on assets are not important, the term added to the above current growth rate of labor income will be very small. In this case, Campbell's measure and Fama-Schwert's measure for the return on human capital are approximately the same (see Campbell (1993b) for details). Motivated by these observations, we make the simple ad hoc assumption that the return on human capital is a linear function of the growth rate in per capita labor income.

It is possible that even when stocks constitute only a small fraction of total wealth, the stock-index portfolio return could well be an excellent proxy for the return on the portfolio of the aggregate wealth.[13] However, to allow for the possibility that this may not be the case, in our empirical work we consider the following proxy that incorporates a measure of the return on human capital. Let R_t^{labor} denote the growth rate in per capita labor income, which proxies for the return on human capital. We assume that the market return is a linear function of R_t^{vw} and R_t^{labor}, i.e.,

ASSUMPTION 2. *There are some constants ϕ_0, ϕ_{vw}, and ϕ_{labor} such that*

$$R_{mt} = \phi_0 + \phi_{\text{vw}} R_t^{\text{vw}} + \phi_{\text{labor}} R_t^{\text{labor}}. \tag{20}$$

Let us define the *labor-beta* as

$$\beta_i^{\text{labor}} = \text{Cov}(R_{it}, R_t^{\text{labor}})/\text{Var}(R_t^{\text{labor}}). \tag{21}$$

Then, by substituting (20) into (11), it follows from (18) and (21) that

$$\beta_i = b_{\text{vw}} \beta_i^{\text{vw}} + b_{\text{labor}} \beta_i^{\text{labor}}. \tag{22}$$

Under Assumptions 1 and 2, the unconditional expected return on any asset is a linear function of its vw-beta, prem-beta, and labor-beta. This can be seen by substituting equation (22) into the equation in Corollary 1 to get

COROLLARY 2. *Suppose that β_i^γ is not a linear function of β_i and that Assumptions 1 and 2 hold, then there are some constants c_0, c_{vw}, c_{prem}, and c_{labor} such that the equation*

$$\text{E}[R_{it}] = c_0 + c_{\text{vw}} \beta_i^{\text{vw}} + c_{\text{prem}} \beta_i^{\text{prem}} + c_{\text{labor}} \beta_i^{\text{labor}} \tag{23}$$

holds for every asset i.

We consider this to be the Premium-Labor model, and it forms the basis for the empirical study that follows. In the rest of this paper, it will be referred to as the "*PL-model.*"

B. Econometric Tests

There are several ways to examine whether data are consistent with the PL-model. According to this model, the unconditional expected return on any asset is a linear function of its three betas only. A natural specification test is to examine whether any other variable has the ability to explain the cross-section of average returns not explained by the three-beta model. In particular, we investigate whether there are residual size effects in the PL-model. The rationale for testing a model against size effects has been discussed by Berk (1995). The *size* of a stock is defined as the logarithm of the market value of the stock. Let $\log(\text{ME}_i)$ denote the time-series average of size for asset i. We

[13] See Shanken (1987) and Kandel and Stambaugh (1987 and 1990), who show how the correlation between the market index proxy return and the unobserved wealth return is related to the mean-variance efficiency of the market-index proxy portfolio.

examine whether any residual size effects exist by including $\log(\mathrm{ME}_i)$ into the PL-model to get

$$E[R_{it}] = c_0 + c_{\text{size}} \log(\mathrm{ME}_i) + c_{\text{vw}}\beta_i^{\text{vw}} + c_{\text{prem}}\beta_i^{\text{prem}} + c_{\text{labor}}\beta_i^{\text{labor}}. \quad (24)$$

If the PL-model holds, then the coefficient c_{size} should be zero, i.e., there should be no residual size effects.

The unconditional models in equations (23) and (24) can be consistently estimated by the cross-sectional regression (CSR) method proposed by Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973). Notice that the PL-model nests the static CAPM as a special case. It facilitates direct comparison of the two models. For comparing the relative performance of the different empirical specifications, we use the R^2 in the cross-sectional regression as an informal and intuitive measure, which shows the fraction of the cross-sectional variation of average returns that can be explained by the model. We are also interested in examining whether c_{vw}, c_{prem}, c_{labor}, and c_{size} are different from zero after allowing for estimation errors. For this purpose, we need to estimate the sampling errors associated with the estimators for these parameters. In Appendix B, we show that the standard errors computed in the Fama-MacBeth procedure are biased, since it does not take into account the sampling errors in the estimated betas. Following the approach suggested by Shanken (1992), we derive a formula for correcting the bias (see Appendix B for details). In deriving the formula for the bias-correction, we made rather strong assumptions (see Assumptions 4 and 5 in Appendix B). Since these assumptions may not be satisfied in practice, we also evaluate the various CAPM specifications using the Generalized Method of Moments, which requires much weaker statistical assumptions.

For this purpose, consider the testable implications that arise from the following moment restriction imposed by the PL-model. Following Dybvig and Ingersoll (1982) we substitute the definition of β_i^{vw}, β_i^{labor}, and β_i^{prem} into the PL-model and rearrange the terms to get

$$E[R_{it}(\delta_0 + \delta_{\text{vw}}R_t^{\text{vw}} + \delta_{\text{prem}}R_{t-1}^{\text{prem}} + \delta_{\text{labor}}R_t^{\text{labor}})] = 1, \quad (25)$$

where δ_0, δ_{vw}, δ_{prem}, and δ_{labor} are the constants defined as follows:

$$\delta_0 = \frac{1}{c_0} + \frac{1}{c_0}\left[\frac{c_{\text{vw}}E[R_t^{\text{vw}}]}{\mathrm{Var}(R_t^{\text{vw}})} + \frac{c_{\text{prem}}E[R_t^{\text{prem}}]}{\mathrm{Var}(R_{t-1}^{\text{prem}})} + \frac{c_{\text{labor}}E[R_t^{\text{labor}}]}{\mathrm{Var}(R_t^{\text{labor}})}\right]$$

$$\delta_{\text{vw}} = -\frac{c_{\text{vw}}}{c_0\mathrm{Var}(R_t^{\text{vw}})} \qquad \delta_{\text{prem}} = -\frac{c_{\text{prem}}}{c_0\mathrm{Var}(R_{t-1}^{\text{prem}})} \qquad \delta_{\text{labor}} = -\frac{c_{\text{labor}}}{c_0\mathrm{Var}(R_t^{\text{labor}})}.$$

It is well known to financial economists (see Ross (1976)) that, so long as the financial market satisfies the law of one price, there will be at least some random variable d_t such that

$$E[R_{it}d_t] = 1, \quad (26)$$

where d_t is generally referred to as a *stochastic discount factor*. Hansen and Richard (1987) point out that an asset-pricing model (like the CAPM) specifies the nature of this stochastic discount factor in terms of potentially observable variables. In our case the stochastic discount factor is given by

$$d_t(\delta) = \delta_0 + \delta_{vw} R_t^{vw} + \delta_{prem} R_{t-1}^{prem} + \delta_{labor} R_t^{labor}, \qquad (27)$$

which depends on the parameters $\delta \equiv (\delta_0, \delta_{vw}, \delta_{prem}, \delta_{labor})'$.[14]

Suppose that there are N assets used in our econometric tests. Let 1_N be the N-dimensional vector of 1s, and

$$\mathrm{R}_t \equiv (R_{1t}, \ldots, R_{Nt})'$$

$$\mathrm{Y}_t \equiv (1, R_t^{vw}, R_t^{prem}, R_t^{labor})'.$$

Then $d_t(\delta) = \mathrm{Y}_t'\delta$. If we let $\mathrm{w}_t(\delta) = R_t d_t(\delta) - 1_N$, then $\mathrm{E}[\mathrm{w}_t(\delta)]$ is the vector of pricing errors of the model. Equation (25) implies that, when the PL-model is correctly specified, the N-dimensional pricing errors, $\mathrm{E}[\mathrm{w}_t(\delta)]$, should be zero. We can evaluate the relative performance of several competing model specifications by comparing the size of pricing errors. For this purpose, we therefore study the quadratic form $\mathrm{E}[\mathrm{w}_t(\delta)]'\mathrm{A}\mathrm{E}[\mathrm{w}_t(\delta)]$, where A is a positive definite matrix (called weighting matrix). We should choose δ to minimize the pricing error by minimizing the value of the quadratic form, which leads to estimation of the parameters δ by the Generalized Method of Moments.

[14] Notice that we can rewrite the conditional CAPM given in equation (2) to get the following conditional stochastic discount factor representation: $\mathrm{E}[R_{it}d_t|I_{t-1}] = 1$, where

$$d_t = \kappa_{0t-1} + \kappa_{1t-1}R_{mt}, \qquad \kappa_{0t-1} = \frac{1}{\gamma_{0t-1}} + \left[\frac{\gamma_{1t-1}}{\gamma_{0t-1}\mathrm{Var}(R_{mt}|I_{t-1})}\right]\mathrm{E}[R_{mt}|I_{t-1}],$$

and

$$\kappa_{1t-1} = -\frac{\gamma_{1t-1}}{\gamma_{0t-1}\mathrm{Var}(R_{mt}|I_{t-1})}.$$

Cochrane (1992) suggests examining $\mathrm{E}[R_{it}d_t|I_{t-1}] = 1$ empirically by assuming that κ_{0t-1} and κ_{1t-1} are linear functions of variables in the date $t-1$ information set I_{t-1}. If one assumes, as in Carhart *et al.* (1995), that (i) $\kappa_{1t-1} = \kappa_1$ (a constant) and (ii) $\kappa_{0t-1} = \kappa_{01} + \kappa_{02}R_{t-1}^{prem}$, the stochastic discount factor then becomes $d_t = \kappa_{01} + \kappa_{02}R_{t-1}^{prem} + \kappa_1 R_t^{vw}$, which resembles the one given in equation (25) with $\delta_{labor} = 0$. However, these assumptions are rather unreasonable. First, since κ_{1t-1} is a function of the conditional market risk premium, conditional zero-beta rate, and conditional variance of the market portfolio, it should be time-varying in nature. It is not reasonable to assume that κ_{1t-1} is a constant when the purpose is to evaluate the conditional CAPM with time-varying expected returns, variances, and covariances. Second, with assumption (i), it follows that $\kappa_{0t-1} = (1/\gamma_{0t-1}) - \kappa_1\mathrm{E}[R_{mt}I_{t-1}]$. Thus, assumption (ii) implies that the conditionally expected market return is a linear function of R_{t-1}^{prem} and the *inverse* of the zero-beta rate. This is very hard to justify because the conditionally expected market return should, according to the conditional CAPM, be the sum of the market risk premium and the zero-beta rate (see equation (A1) in Appendix A). In contrast, to derive model (25), we only assume that the conditionally expected market risk premium is a linear function of the variables in the information set. This assumption can be justified under joint normality.

Now the issue is how to choose the weighting matrix A. The most famous choice is the "optimal" weighting matrix suggested by Hansen and Singleton (1982). For expositional simplicity, let us now assume that w_t is i.i.d. over time. In this case, the weighting matrix suggested by Hansen and Singleton (1982) is $A = [\mathrm{Var}(w_t(\delta))]^{-1}$. With this choice for the weighting matrix, they show that the minimized value of the sample analog of the quadratic form asymptotically has a χ^2 distribution with $N - K$ degrees of freedom, where K is the number of unknown parameters in the model. This asymptotic distribution can be used to test whether the pricing errors are zero. However, this weighting matrix will be different for different model specifications, and thus, we cannot use the value of the quadratic form to compare the relative size of the pricing errors associated with different models. Especially, if a model contains "more noise," i.e., the variance of $w_t(\delta)$ is larger, then the value of the quadratic form will be smaller. In this case, it would be misleading to conclude that the "more noisy" the model, the better it performs.

We therefore choose the weighting matrix suggested by Hansen and Jagannathan (1994), which is $A = (E[R_t R_t'])^{-1}$. Since this weighting matrix remains the same across various competing model specifications, it allows us to compare the performance of those models by the value of the quadratic form. Hansen and Jagannathan (1994) show that the value of the quadratic form is the squared distance from the candidate stochastic discount factor of a given model to the set of all the discount factors that price the N assets correctly. Thus, we refer to the square root of the quadratic form with this weighting matrix as the *Hansen-Jagannathan distance*, or simply, *HJ-distance*. Hansen and Jagannathan (1994) also show that the HJ-distance is the pricing error for the portfolio that is most mispriced by the model (see Appendix C for details). Since the weighting matrix suggested by Hansen and Jagannathan (1994) is generally not "optimal" in the sense of Hansen (1982), the minimized value of the sample analog of the quadratic form does not have a χ^2 distribution, and we thus cannot directly use Hansen's (1982) J-test for the overidentifying restrictions. In Appendix C, we therefore extend Hansen's (1982) results and show how to calculate the asymptotic distributions of the minimized quadratic form in the Generalized Method of Moments when the weighting matrix is chosen arbitrarily.

III. Empirical Results

A. Description of the Data

Though Fama and French (1992) use returns to common stocks of nonfinancial corporations listed in NYSE, AMEX (1962–90), and Nasdaq (1973–90) that are covered by CRSP as well as COMPUSTAT in their study, we study the returns to stocks of nonfinancial firms listed in NYSE and AMEX (1962–90) covered by CRSP alone. Nasdaq stocks are not included because we do not have monthly data for Nasdaq stocks available to us. This should not be an issue since Fama and French (1992) report that their results do not depend on the inclusion of Nasdaq stocks.

It is well known that firms in the COMPUSTAT database may have some survivorship bias,[15] since stocks move in and out of the COMPUSTAT list depending on their performance. Kothari, Shanken, and Sloan (1995) provide indirect evidence for the existence of such bias—they point out that the annual returns are about 10 percentage points more for small firms in COMPUSTAT when compared to small firms that are only in CRSP. Breen and Korajczyk (1994) provide some direct evidence that supports the view that selection bias may be an important issue for tests that use standard sources of accounting data like COMPUSTAT. Fama and French (1993) attempt to address this problem by omitting the first two years of data, in response to COMPUSTAT's claim that no more than two years of data are included at the time a firm is added to the COMPUSTAT list. However, it is not clear whether this completely eliminates the bias in the COMPUSTAT tapes. With this in mind, we do not examine the relation between book-to-market equity and the cross-section of returns.[16] Hence, we are not constrained to limit our attention to stocks that are in CRSP as well as COMPUSTAT.[17]

We create 100 portfolios of NYSE and AMEX stocks as in Fama and French (1992). For every calendar year, starting in 1963, we first sort firms into size deciles based on their market value at the end of June. For each size decile, we estimate the beta of each firm, using 24 to 60 months of past-return data and the CRSP value-weighted index as the market index proxy. Following Fama and French (1992), we denote this beta as the "pre-ranking" beta estimate, or "pre-beta" for short. We then sort firms within each size decile into beta deciles based on their pre-betas. This gives us 100 portfolios, and we compute the return on each of these portfolios for the next 12 calendar months by equally weighting the returns on stocks in the portfolio. We repeat this procedure for each calendar year. This gives a time series of monthly returns (July 1963–December 1990, i.e., 330 observations) for each of the 100 portfolios.

The Fama and French (1992) sorting procedure produces an impressive dispersion in the characteristics of interest. Time-series averages of portfolio returns are given in Panel A of Table I. The rates of return range from a low of 0.51 percent to a high of 1.71 percent per month. The β_i^{vw}s of the portfolios are presented in Panel B of Table I. They range from a low of 0.57 to a high of 1.70. We calculate the size of a portfolio as the equally-weighted average of the

[15] See Chari, Jagannathan, and Ofer (1986).

[16] Chan, Jegadeesh, and Lakonishok (1995) report that the sample selection bias, if any, does not explain the superior performance of value stocks for the top quintile of the NYSE-AMEX stocks.

[17] Davis (1994), using a database that is free of survivorship bias, finds that book-to-market equity has significant explanatory power with respect to the cross-section of realized stock returns during the period of July 1940 through June 1963. It is worthwhile pointing out that we do not claim that the "book-to-market" variable does not help predict future returns on a stock. We are only pointing out that the COMPUSTAT data set has some selection bias and hence is unsuitable for econometric evaluation of asset-pricing models until we have a clearer understanding of the bias.

Table I

Basic Characteristics of the 100 Portfolios

Using stocks of nonfinancial firms listed in the NYSE and AMEX covered by CRSP, the 100 portfolios are formed in the same way as in Fama and French (1992). For every calendar year, starting from 1963, we first sort firms into size deciles based on their market value at the end of June. For each size category, we estimate the pre-beta of each firm by the slope coefficient in the regression of the 24 to 60 months of past-return data on a constant and the CRSP value-weighted index of the corresponding months. We then sort firms within each size decile into beta deciles based on their pre-betas. This gives 100 portfolios, and we compute the return on each of these portfolios for the next 12 calendar months by equally weighting the returns on stocks in the portfolio. We repeat this procedure for each calendar year. This gives a time series of monthly returns (July 1963–December 1990, i.e., 330 observations) for each of the 100 size-beta portfolios. β_i^{vw} is the slope in the regression of portfolio i's return on the CRSP value-weighted stock index return and a constant for the entire 330-month period. A portfolio size is calculated as the equally-weighted average of the logarithm of the market value (in million dollars) of the stocks in the portfolio. β_i^{prem} and β_i^{labor} are calculated in a similar way. The numbers given in Panel D are the part of β_i^{prem} orthogonal to a constant and β_i^{vw}, and the numbers in Panel E are the part of β_i^{labor} orthogonal to a constant, β_i^{vw} and β_i^{prem}.

	β-L	β-2	β-3	β-4	β-5	β-6	β-7	β-8	β-9	β-H
Panel A: Time-Series Averages of Returns										
Size-S	1.44	1.53	1.56	1.71	1.36	1.44	1.37	1.33	1.46	1.34
Size-2	1.13	1.22	1.09	1.19	1.38	1.37	1.37	1.30	1.15	0.95
Size-3	1.26	1.27	1.22	1.26	1.16	1.29	1.34	1.19	1.12	0.89
Size-4	1.37	1.47	1.40	1.28	1.01	1.39	1.11	1.33	1.07	0.95
Size-5	0.97	1.53	1.10	1.28	1.18	1.04	1.35	1.07	1.23	0.82
Size-6	1.07	1.36	1.34	1.12	1.25	1.27	0.84	0.94	0.92	0.77
Size-7	0.99	1.18	1.13	1.19	0.96	0.99	1.11	0.91	0.90	0.83
Size-8	0.95	1.19	1.02	1.39	1.18	1.24	0.94	1.02	0.88	1.08
Size-9	0.94	0.92	1.05	1.17	1.15	1.03	1.02	0.84	0.80	0.51
Size-B	1.06	0.97	1.02	0.94	0.83	0.93	0.82	0.83	0.61	0.72
Panel B: The Estimated β_i^{vw}s										
Size-S	0.90	0.99	1.01	1.13	1.17	1.21	1.20	1.31	1.44	1.54
Size-2	0.83	1.00	1.09	1.12	1.18	1.29	1.33	1.39	1.48	1.63
Size-3	0.78	0.93	1.09	1.11	1.18	1.27	1.29	1.40	1.42	1.70
Size-4	0.75	0.91	1.05	1.13	1.19	1.32	1.25	1.32	1.56	1.61
Size-5	0.57	0.78	1.10	1.10	1.12	1.20	1.25	1.43	1.45	1.54
Size-6	0.62	0.77	0.88	1.01	1.08	1.25	1.22	1.34	1.32	1.59
Size-7	0.64	0.84	1.01	1.07	1.16	1.21	1.26	1.26	1.31	1.54
Size-8	0.64	0.73	0.91	1.04	1.07	1.17	1.22	1.19	1.23	1.50
Size-9	0.62	0.78	0.88	0.96	1.04	1.05	1.13	1.17	1.22	1.34
Size-B	0.68	0.76	0.80	1.00	0.97	1.00	1.04	1.09	1.10	1.28
Panel C: The Time-Series Averages of Size (log $million)										
Size-S	2.48	2.50	2.49	2.48	2.48	2.50	2.46	2.46	2.46	2.34
Size-2	3.71	3.72	3.73	3.73	3.71	3.71	3.72	3.72	3.72	3.72
Size-3	4.21	4.21	4.21	4.21	4.21	4.23	4.21	4.22	4.21	4.20
Size-4	4.67	4.65	4.64	4.65	4.65	4.65	4.65	4.64	4.64	4.64
Size-5	5.07	5.09	5.07	5.08	5.08	5.07	5.07	5.07	5.07	5.05
Size-6	5.47	5.48	5.47	5.48	5.48	5.48	5.48	5.48	5.47	5.48
Size-7	5.91	5.92	5.93	5.92	5.92	5.89	5.91	5.90	5.92	5.90
Size-8	6.44	6.42	6.43	6.39	6.43	6.41	6.43	6.42	6.40	6.40
Size-9	6.98	6.98	7.00	6.98	6.96	6.97	6.95	6.96	6.95	6.97
Size-B	8.11	8.26	8.22	8.19	8.16	8.18	8.06	8.03	7.92	7.81

Table I—*Continued*

	β-L	β-2	β-3	β-4	β-5	β-6	β-7	β-8	β-9	β-H
Panel D: The Estimated β_i^{prem} that is Orthogonal to β_i^{vw}										
Size-S	−0.03	0.30	0.20	0.08	0.43	0.25	0.16	0.25	0.46	0.24
Size-2	−0.47	0.41	0.47	−0.23	0.02	0.32	0.40	0.66	0.34	−0.03
Size-3	−0.04	0.10	0.41	0.27	−0.01	0.06	−0.50	−0.04	0.18	−0.38
Size-4	−0.04	0.37	−0.07	−0.28	0.24	0.13	0.71	−0.28	−0.26	−0.18
Size-5	−0.06	−0.13	0.18	0.08	0.06	0.17	0.11	0.14	−0.19	−0.79
Size-6	0.21	0.38	0.38	−0.01	0.12	0.62	−0.30	0.02	−0.45	−0.49
Size-7	−0.06	0.44	0.44	0.26	0.19	0.15	−0.07	0.06	0.31	0.05
Size-8	−0.44	−0.33	−0.02	0.16	0.41	0.11	0.05	−0.34	−0.69	−0.37
Size-9	−0.26	−0.34	0.14	0.09	−0.14	−0.13	0.15	−0.53	−0.03	−0.40
Size-B	−0.45	−0.07	−0.02	−0.26	−0.09	−0.64	−0.53	−0.34	−0.89	−0.25
Panel E: The Estimated β_i^{labor} that is Orthogonal to β_i^{prem} and β_i^{vw}										
Size-S	1.23	1.15	0.71	0.48	0.65	0.19	1.38	0.73	1.06	0.73
Size-2	0.60	0.55	0.20	−0.10	0.51	0.53	−0.30	−0.84	0.35	−0.30
Size-3	0.32	0.19	−0.44	0.79	−0.29	0.07	0.94	0.13	0.13	0.52
Size-4	−0.04	0.48	0.46	−0.17	−0.05	−0.18	−1.40	0.15	0.30	0.42
Size-5	0.19	0.27	−0.15	0.30	0.16	−0.58	0.02	0.12	1.10	0.16
Size-6	−0.09	−0.05	−0.05	−0.08	0.02	−0.66	0.01	−0.52	−0.01	−1.04
Size-7	−0.66	−0.22	−0.38	−0.26	0.07	−0.07	−0.58	−0.21	−0.16	−0.45
Size-8	−0.25	0.16	−0.40	0.20	−0.53	−0.35	−0.76	−0.50	0.73	−0.78
Size-9	−0.38	−0.16	−0.27	−0.64	−0.07	−0.21	−0.25	−0.29	−0.39	−0.67
Size-B	0.32	−0.04	−0.04	−0.37	−0.24	0.01	−0.23	−0.43	0.02	−1.24

logarithm of market value of stocks (in million dollars). The time-series averages of portfolio size are presented in Panel C of Table I. They range from a low of 2.34 to a high of 8.26. Properties of these three characteristics of the portfolios are very similar to those of the portfolios formed by Fama and French (1992). The numbers given in Panel D are the part of β_i^{prem} orthogonal to a constant and β_i^{vw}, and the numbers in Panel E are the part of β_i^{labor} orthogonal to a constant, β_i^{vw} and β_i^{prem}.

The BAA- and AAA-bond yields are taken from Table 1.35 in the *Federal Reserve Bulletin* published by the Board of Governors of the Federal Reserve System. The data on personal income and population are taken from Table 2.2 in the *National Income and Product Account of the U.S.A.* published by the Bureau of Economic Analysis, U.S. Department of Commerce. The labor income used in this study is the difference between the total personal income and the dividend income. We construct the growth rate in per capita monthly labor income series using the formula,

$$R_t^{\text{labor}} = [L_{t-1} + L_{t-2}]/[L_{t-2} + L_{t-3}],$$

where R_t^{labor} denotes the growth rate in labor income that becomes known at the end of month t and L_{t-1} denotes the per capita labor income for month $t-1$, which becomes known at the end of month t. This dating convention is

Table II

Evaluation of Various CAPM Specifications

This table gives the estimates for the cross-sectional regression model

$$E[R_{it}] = c_0 + c_{size}\log(ME_i) + c_{vw}\beta_i^{vw} + c_{prem}\beta_i^{prem} + c_{labor}\beta_i^{labor}$$

and the model for the moments

$$E[R_{it}(\delta_0 + \delta_{vw}R_t^{vw} + \delta_{prem}R_t^{prem} + \delta_{labor}R_t^{labor})] = 1,$$

with either a subset or all of the variables. Here, R_{it} is the return on portfolio i ($i = 1, 2, \ldots, 100$) in month t (July 1963–December 1990), R_t^{vw} is the return on the value-weighted index of stocks, R_{t-1}^{prem} is the yield spread between low- and high-grade corporate bonds, and R_t^{labor} is the growth rate in per capita labor income. The β_i^{vw} is the slope coefficient in the OLS regression of R_{it} on a constant and R_t^{vw}. The other betas are estimated in a similar way. The portfolio size, log(ME_i), is calculated as the equally-weighted average of the logarithm of the market value (in million dollars) of the stocks in portfolio i. The regression models are estimated by using the Fama-MacBeth procedure. The "corrected t- and p-values" take sampling errors in the estimated betas into account. The models for the moments are estimated by using the Generalized Method of Moments with the Hansen-Jagannathan weighting matrix. The minimized value of the GMM criterion function is the first item under the "HJ-dist," with the associated p-value immediately below it. All the R-squares and p-values are reported as percentages.

Panel A: The Static CAPM without Human Capital

Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{size}	R-square
Estimate:	1.24	−0.10				1.35
t-value:	5.17	−0.28				
p-value:	0.00	78.00				
Corrected-t:	5.16	−0.28				
Corrected-p:	0.00	78.01				
Estimate:	2.08	−0.32			−0.11	57.56
t-value:	5.79	−0.94			−2.30	
p-value:	0.00	34.54			2.14	
Corrected-t:	5.77	−0.94			−2.30	
Corrected-p:	0.00	34.60			2.17	
Coefficient:	δ_0	δ_{vw}	δ_{prem}	δ_{labor}		HJ-dist
Estimate:	0.97	1.55				0.6548
t-value:	89.01	1.09				
p-value:	0.00	27.59				0.22

Panel B: The Conditional CAPM without Human Capital

Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{size}	R-square
Estimate:	0.81	−0.31	0.36			29.32
t-value:	2.72	−0.87	3.28			
p-value:	0.66	38.45	0.10			
Corrected-t:	2.19	−0.70	2.67			
Corrected-p:	2.87	48.43	0.75			
Estimate:	1.77	−0.38	0.16		−0.10	61.66
t-value:	4.75	−1.10	2.50		−1.93	
p-value:	0.00	27.17	1.26		5.35	
Corrected-t:	4.53	−1.05	2.40		−1.84	
Corrected-p:	0.00	29.53	1.66		6.59	

Table II—*Continued*

Panel B:—*Continued*						
Coefficient:	δ_0	δ_{vw}	δ_{prem}	δ_{labor}		HJ-dist
Estimate:	1.48	2.05	−45.94			0.6425
t-value:	6.71	1.47	−2.36			
p-value:	0.00	14.14	1.83			0.98
Panel C: The Conditional CAPM with Human Capital						
Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{size}	R-square
Estimate:	1.24	−0.40	0.34	0.22		55.21
t-value:	5.51	−1.18	3.31	2.31		
p-value:	0.00	23.76	0.09	2.07		
Corrected-t:	4.10	−0.88	2.48	1.73		
Corrected-p:	0.00	37.99	1.31	8.44		
Estimate:	1.70	−0.40	0.20	0.10	−0.07	64.73
t-value:	4.61	−1.18	3.00	2.09	−1.45	
p-value:	0.00	23.98	0.27	3.62	14.74	
Corrected-t:	4.14	−1.06	2.72	1.89	−1.30	
Corrected-p:	0.00	29.07	0.66	5.87	19.29	
Coefficient:	δ_0	δ_{vw}	δ_{prem}	δ_{labor}		HJ-dist
Estimate:	2.26	1.81	−65.72	−97.72		0.6184
t-value:	6.39	1.26	−3.10	−2.94		
p-value:	0.00	20.65	0.20	0.33		19.38
Panel D: The Static CAPM with Human Capital						
Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{size}	R-square
Estimate:	1.67	−0.22		0.23		30.46
t-value:	6.91	−0.63		2.37		
p-value:	0.00	53.19		1.77		
Corrected-t:	5.71	−0.52		1.97		
Corrected-p:	0.00	60.49		4.87		
Estimate:	2.09	−0.32		0.05	−0.10	58.55
t-value:	5.80	−0.96		1.22	−2.15	
p-value:	0.00	33.78		22.29	3.19	
Corrected-t:	5.70	−0.95		1.20	−2.11	
Corrected-p:	0.00	34.46		22.93	3.48	
Coefficient:	δ_0	δ_{vw}	δ_{prem}	δ_{labor}		HJ-dist
Estimate:	1.37	1.22		−68.68		0.6422
t-value:	7.73	0.85		−2.32		
p-value:	0.00	39.65		2.01		1.94

consistent with the fact that the monthly labor income data are typically published with a one-month delay. We use a two-month moving average in per capita labor income to minimize the influence of measurement errors.

B. The Main Results

Using return data on the 100 portfolios described earlier, we first examine the traditional empirical CAPM specification,

$$\mathrm{E}[R_{it}] = c_0 + c_{\mathrm{vw}}\beta_i^{\mathrm{vw}}. \tag{28}$$

The results are presented in Panel A of Table II. The t-value for c_{vw} is -0.28, and the corresponding p-value is 78 percent. The R^2 of the regression is only 1.35 percent, i.e., only 1.35 percent of the cross-sectional variation in average returns can be explained by this specification. The correction to the standard errors for estimation errors in the betas does not appear to be important. After the correction, the t-value remains as -0.28. Hence, we can conclude that c_{vw} is not significantly different from zero after allowing for sampling errors. When size is added to the model, the t-value for size is -2.30 and the corresponding p-value is only 2.14 percent. The R^2 goes up to 57.56 percent. The corrected t-value is not very different. The strong size effect suggests that the conventional specification of the CAPM is inconsistent with the data. In the GMM test that uses the Hansen-Jagannathan weighting matrix, the estimated HJ-distance is 0.6548 and the corresponding p-value is 0.22 percent, indicating that the pricing error is significantly different from zero. The p-value for the coefficient δ_{vw} in the moment restriction of the model is 27.59 percent, suggesting that R_t^{vw} does not play a significant role in constructing a stochastic discount factor that helps to explain the cross-sectional dispersion in expected returns on the 100 portfolios in our study. These results are consistent with what has been reported in the literature.

We next allow betas to vary over time, i.e., assume that the conditional CAPM holds, but still use the stock index as a proxy for the market return. This gives the following specification:

$$\mathrm{E}[R_{it}] = c_0 + c_{\mathrm{vw}}\beta_i^{\mathrm{vw}} + c_{\mathrm{prem}}\beta_i^{\mathrm{prem}}. \tag{29}$$

The results are presented in Panel B of Table II. The estimated value of c_{prem}, using the Fama-MacBeth regression, is significantly different from zero. The t-value for c_{prem} is 3.28 with a p-value of 0.10 percent. The R^2 is 29.32 percent, which is a substantial improvement compared with 1.35 percent for the model in (28). The t-value for c_{prem} is 2.67 when the standard errors are corrected and the associated p-value is 0.75 percent. When size is added to equation (29), the t-value for c_{size} is -1.93 (p-value = 5.35 percent). When the standard errors are corrected, the t-value drops to -1.84 (p-value = 6.59 percent). Although there are still some size effects in model (29), they are much weaker than those in model (28). The GMM test with the HJ weighting matrix gives an estimated value of 0.6425 for HJ-distance with p-value of 0.98 percent. Hence, this specification reduces the pricing errors, but they are still significantly different from zero. The p-value for δ_{prem} in the moment restriction is 1.83 percent, which indicates that R_t^{prem} is a significant and important component of the stochastic discount factor.

We now consider the main model developed in this paper:

$$E[R_{it}] = c_0 + c_{vw}\beta_i^{vw} + c_{prem}\beta_i^{prem} + c_{labor}\beta_i^{labor}, \qquad (30)$$

where the return on the market portfolio of all assets is assumed to be a linear function of the stock index and the growth rate of per capita labor income. Equation (30) is the same PL-model in equation (23). The estimation results are presented in Panel C of Table II. The estimated value of c_{labor}, using the Fama-MacBeth regression, is significantly different from zero (t-value = 2.31, p-value = 2.07 percent). The R^2 increases to 55.21 percent. However, when the standard errors are corrected, the t-value for c_{labor} drops to 1.73 (p-value = 8.44 percent). The coefficient c_{prem} remains significant. When size is added to the model, the t-value for the size coefficient is -1.45 and the associated p-value is 14.74 percent, which shows that size does not explain what is left unexplained in this model after controlling for sampling errors. When the standard errors are corrected, the p-value for size becomes even larger, reinforcing our conclusions. In the GMM test with the HJ weighting matrix, the estimated HJ-distance drops sharply to 0.6184 and the p-value jumps to 19.38 percent. Hence, the pricing errors of the PL-model are much smaller and not significantly different from zero. Notice that both R_t^{prem} and R_t^{labor} are significant in the GMM test, which is consistent with the results obtained from the Fama-MacBeth regression. While the point estimate of the slope coefficient c_{vw} is negative, it is not significantly different from zero, after allowing for sampling errors. Also, the estimated value of the average zero-beta rate is rather high when compared to the average T-bill rate and the average risk premium of stocks. Hence there is cause for concern even though our CAPM specification does substantially better than the static CAPM in explaining the cross-section of average returns on stocks. It appears that we are still missing some important aspect of reality in our modeling exercise.

In order to visually compare the performance of the different specifications, we plot the fitted expected return, computed by using the estimated parameter values in a model specification, against the realized average return. If the fitted expected returns and the realized average returns are the same, then all the points should lie on the 45-degree line through the origin. When β_i^{vw} alone is used, the fitted expected returns are all about the same, whereas the realized average returns vary substantially across the 100 portfolios (Figure 1). The performance substantially improves when β_i^{prem} and β_i^{labor} are also used (Figure 3). The fit is about as good as the model with size and β_i^{vw} (Figure 2). The distribution of the points around the 45-degree line in Figure 3 suggests that the improved performance of the CAPM using the specifications we suggest in this paper is not due to a few outliers. The distribution of the points around the 45-degree line does not significantly change when we add log(ME) as an additional explanatory variable (Figure 4).

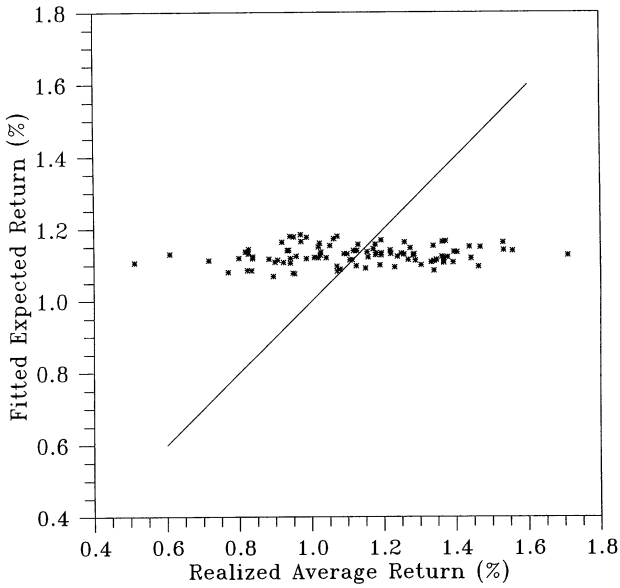

Figure 1. Fitted expected returns versus realized average returns. Each scatter point in the graph represents a portfolio, with the *realized average return* as the horizontal axis and the *fitted expected return* as the vertical axis. For each portfolio i, the realized average return is the time-series average of the portfolio return, and the fitted expected return is the fitted value for the expected return, $E[R_i]$, in the following regression model:

$$E[R_i] = c_0 + c_{vw}\beta_i^{vw},$$

where β_i^{vw} is the slope coefficient in the OLS regression of the portfolio return on a constant and the return on the value-weighted index portfolio of stocks. The straight line in the graph is the 45° line from the origin.

We may suspect that R_t^{labor} is the driving force behind the results for our main model. To determine if this is the case, we examine the following model:

$$E[R_{it}] = c_0 + c_{vw}\beta_i^{vw} + c_{labor}\beta_i^{labor}, \tag{31}$$

which can be obtained from the static CAPM by including the growth rate of labor income into the proxy for the market return. The estimated results for this specification are presented in Panel D of Table II. The coefficient corresponding to the growth rate of labor income is significant, both in the Fama-MacBeth regression and the GMM test using the HJ weighting matrix.[18] However, there is a strong residual size effect in the Fama-MacBeth regression. The HJ-distance is just slightly lower than that of model (28), and the

[18] Our empirical specification with labor income is similar to that used by Fama and Schwert (1977) when betas do not vary over time. The difference is that we use lagged labor income since labor income is published with a one-month lag. For a more detailed discussion of this issue, see Jagannathan and Wang (1993).

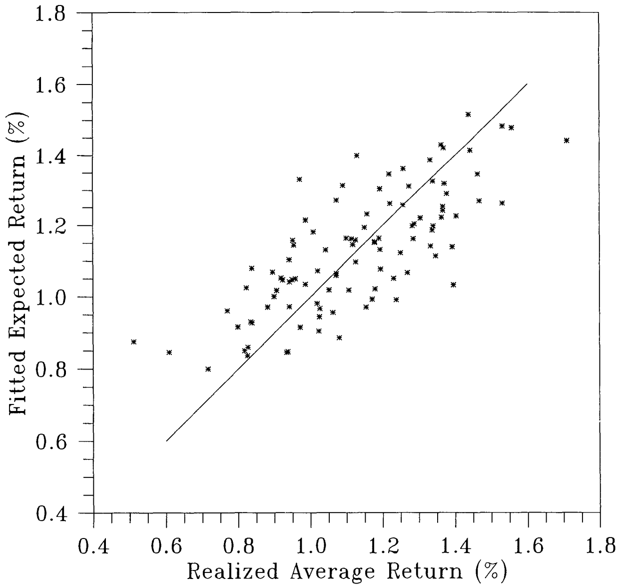

Figure 2. Fitted expected returns versus realized average returns. Each scatter point in the graph represents a portfolio, with the *realized average return* as the horizontal axis and the *fitted expected return* as the vertical axis. For each portfolio i, the realized average return is the time-series average of the portfolio return, and the fitted expected return is the fitted value for the expected return, $\mathrm{E}[R_i]$, in the following regression model:

$$\mathrm{E}[R_i] = c_0 + c_{\mathrm{size}} \log(\mathrm{ME}_i) + c_{\mathrm{vw}} \beta_i^{\mathrm{vw}},$$

where β_i^{vw} is the slope coefficient in the OLS regression of the portfolio return on a constant and the return on the value-weighted index portfolio of stocks, and the portfolio size, $\log(\mathrm{ME}_i)$, is calculated as the equally-weighted average of the logarithm of the market value (in million dollars) of the stocks in portfolio i. The straight line in the graph is the 45° line from the origin.

p-value is only 1.94 percent. Thus, the pricing error of this model is still substantial. This suggests that it is necessary to allow for time variations in betas as well in order to explain the cross-section of expected returns on stocks.

C. Additional Investigations

The unconditional model we develop in this paper to some extent resembles the multi-factor model specified by Chen, Roll, and Ross (1986). A natural question that arises is whether the "lagged-prem factor" and the "labor-income-growth-rate factor" that we use in our specifications are just proxies for the macroeconomic factors that are identified by Chen, Roll, and Ross in their earlier work. Following them, we consider, besides the value-weighted stock index, four additional factors: (a) UTS_t is the monthly return spread between the long-term government bond and Treasury bill, (b) UPR_t is the return

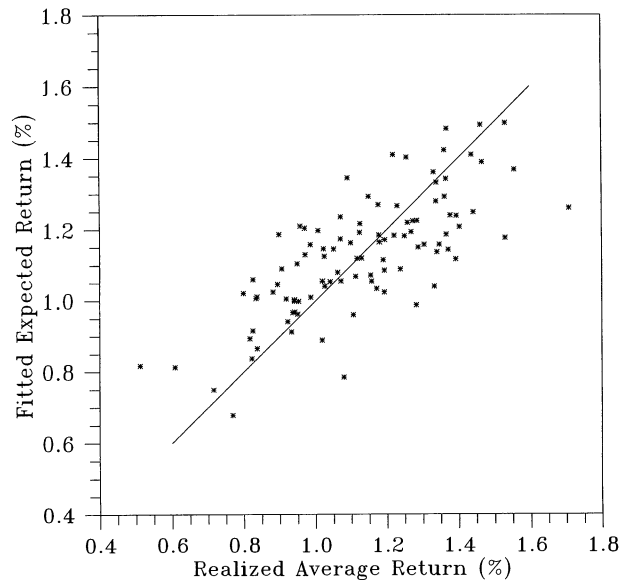

Figure 3. Fitted expected returns versus realized average returns. Each scatter point in the graph represents a portfolio, with the *realized average return* as the horizontal axis and the *fitted expected return* as the vertical axis. For each portfolio i, the realized average return is the time-series average of the portfolio return, and the fitted expected return is the fitted value for the expected return, $E[R_i]$, in the following regression model:

$$E[R_i] = c_0 + c_{vw}\beta_i^{vw} + c_{prem}\beta_i^{prem} + c_{labor}\beta_i^{labor},$$

where β_i^{vw} is the slope coefficient in the OLS regression of the portfolio return on a constant and the return on the value-weighted index portfolio of stocks, β_i^{prem} is the slope coefficient in the OLS regression of the portfolio return on a constant and the yield spread between low- and high-grade corporate bonds, and β_i^{labor} is the slope coefficient in the OLS regression of the portfolio return on a constant and the growth rate in per capita labor income. The straight line in the graph is the 45° line from the origin.

differential between a long-term corporate bond and long-term government bond, (c) MP_t is the growth rate in monthly industrial production in the United States, and (d) UI_t is the change of inflation rate. The betas are estimated using contemporaneous values of these variables. The UTS_t and MP_t are the same variable used by Chen, Roll, and Ross. While the UPR_t and UI_t used in our test should be similar to those corresponding factors used by Chen, Roll, and Ross, they may not be exactly the same since we do not have access to their data. The data series on inflation, corporate-bond return, and long-term government bond return are from Ibbotson Associates.[19] Monthly industrial production data are obtained from Table 2.10 in the *Federal Reserve Bulletin* published by the Board of Governors of the Federal Reserve System. We

[19] See *Stocks, Bonds, Bills and Inflation*, 1991 Year Book by Ibbotson Associates Inc.

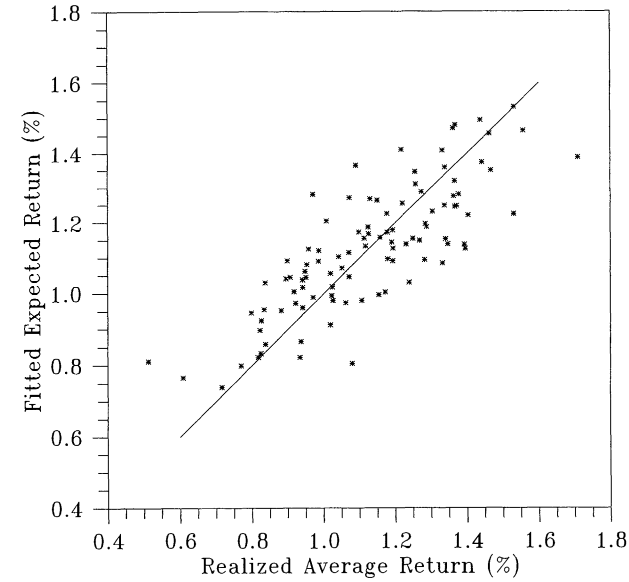

Figure 4. Fitted expected returns versus realized average returns. Each scatter point in the graph represents a portfolio, with the *realized average return* as the horizontal axis and the *fitted expected return* as the vertical axis. For each portfolio i, the realized average return is the time-series average of the portfolio return, and the fitted expected return is the fitted value for the expected return, $E[R_i]$, in the following regression model:

$$E[R_i] = c_0 + c_{size} \log(ME_i) + c_{vw}\beta_i^{vw} + c_{prem}\beta_i^{prem} + c_{labor}\beta_i^{labor},$$

where β_i^{vw} is the slope coefficient in the OLS regression of the portfolio return on a constant and the return on the value-weighted index portfolio of stocks, β_i^{prem} is the slope coefficient in the OLS regression of the portfolio return on a constant and the yield spread between low- and high-grade corporate bonds, β_i^{labor} is the slope coefficient in the OLS regression of the portfolio return on a constant and the growth rate in per capita labor income, and the portfolio size, $\log(ME_i)$, is calculated as the equally-weighted average of the logarithm of the market value (in million dollars) of the stocks in portfolio i. The straight line in the graph is the 45° line from the origin.

consider the following models:

$$E[R_{it}] = c_0 + c_{vw}\beta_i^{vw} + c_{UTS}\beta_i^{UTS} + c_{UPR}\beta_i^{UPR} + c_{MP}\beta_i^{MP} + c_{UI}\beta_i^{UI} \tag{32}$$

$$E[R_{it}] = c_0 + c_{vw}\beta_i^{vw} + c_{prem}\beta_i^{prem} + c_{labor}\beta_i^{labor} + c_{UTS}\beta_i^{UTS} + c_{UPR}\beta_i^{UPR} + c_{MP}\beta_i^{MP} + c_{UI}\beta_i^{UI}, \tag{33}$$

where all the betas are calculated in the same way as β_i^{vw}.

The results are given in Table III. The top half of the table gives the estimates for model (32). The R^2 for the model is 38.96 percent, which is substantially less than the R^2 for the PL-model (55.21 percent). The HJ-distance for equation (32) is 0.6529, which is larger than the HJ-distance for the PL-model (0.6184). So, the R^2 and the HJ-distance consistently indicate that the PL-model performs better than model (32). The p-value shows that the

Table III

Comparison with the Factors Used by Chen, Roll, and Ross (1986)

This table gives the estimates for the cross-sectional regression model

$$E[R_{it}] = c_0 + c_{vw}\beta_i^{vw} + c_{prem}\beta_i^{prem} + c_{labor}\beta_i^{labor} + c_{UTS}\beta_i^{UTS} + c_{UPR}\beta_i^{UPR} + c_{MP}\beta_i^{MP} + c_{UI}\beta_i^{UI}$$

and the model for the moments

$$E[R_{it}(\delta_0 + \delta_{vw}R_t^{vw} + \delta_{prem}R_t^{prem} + \delta_{labor}R_t^{labor} + \delta_{UTS}UTS_t + \delta_{UPR}UPR_t + \delta_{MP}MP_t + \delta_{UI}UI_t)] = 1,$$

with either a subset or all of the variables. Here, R_{it} is the return on portfolio i ($i = 1, 2, \ldots, 100$) in month t (July 1963–December 1990), R_t^{vw} is the return on the value-weighted index of stocks, R_{t-1}^{prem} is the yield spread between low- and high-grade corporate bonds, R_t^{labor} is the growth rate in per capita labor income, UTS_t is the return spread between long-term government bonds and Treasury bills, UPR_t is the return differential between long-term corporate and long-term government bonds, MP_t is the growth rate in monthly industrial production in the United States, and UI_t is the change of inflation rate. The β_i^{vw} is the slope coefficient in the OLS regression of R_{it} on a constant and R_t^{vw}. The other betas are estimated in a similar way. The regression models are estimated by using the Fama-MacBeth procedure. The "corrected t- and p-values" take sampling errors in the estimated betas into account. The models for the moments are estimated by using the Generalized Method of Moments with the Hansen-Jagannathan weighting matrix. The minimized value of the GMM criterion function is the first item under the "HJ-dist," with the associated p-value immediately below it. All the R-squares and p-values are reported as percentages.

Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{UTS}	c_{UPR}	c_{MP}	c_{UI}	R-square
Estimate:	1.80	−0.44			−1.07	0.39	−0.02	−0.07	38.96
t-value:	7.18	−1.28			−2.44	1.63	−0.17	−1.95	
p-value:	0.00	20.14			1.46	10.33	86.27	5.13	
Corrected-t:	6.17	−1.10			−2.12	1.41	−0.15	−1.68	
Corrected-p:	0.00	26.99			3.38	15.93	88.19	9.34	
Coefficient:	δ_0	δ_{vw}	δ_{prem}	δ_{labor}	δ_{UTS}	δ_{UPR}	δ_{MP}	δ_{UI}	HJ-dist
Estimate:	0.97	1.15			2.81	3.40	2.30	8.08	0.6529
t-value:	28.47	0.76			0.88	0.45	0.19	0.23	
p-value:	0.00	44.77			37.96	65.19	85.12	81.70	0.16
Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{UTS}	c_{UPR}	c_{MP}	c_{UI}	R-square
Estimate:	1.37	−0.51	0.29	0.18	−0.17	0.19	0.07	−0.03	57.87
t-value:	6.33	−1.46	3.54	2.44	−0.46	0.92	0.61	−0.99	
p-value:	0.00	14.50	0.04	1.47	64.75	35.72	54.26	32.11	
Corrected-t:	4.97	−1.15	2.81	1.93	−0.36	0.72	0.48	−0.78	
Corrected-p:	0.00	25.17	0.50	5.39	71.91	46.89	63.24	43.53	
Coefficient:	δ_0	δ_{vw}	δ_{prem}	δ_{labor}	δ_{UTS}	δ_{UPR}	δ_{MP}	δ_{UI}	HJ-dist
Estimate:	2.38	1.73	−72.05	−104.33	−0.27	7.31	−5.91	−9.81	0.6152
t-value:	6.14	1.13	−3.19	−2.84	−0.08	0.86	−0.43	−0.27	
p-value:	0.00	25.86	0.14	0.46	94.01	38.95	66.68	78.80	22.06

HJ-distance for model (32) is significantly different from zero. The lower half of the table gives the results for model (33). Both the R^2 and the HJ-distance indicate that inclusion of the four additional factors in Chen, Roll, and Ross (1986) does not substantially improve the performance of the PL-model. More importantly, none of the coefficients corresponding to the factors in Chen, Roll, and Ross (1986) is significantly different from zero after taking sampling errors into account.

Earlier, we examined the possibility of model misspecification by checking whether firm size can explain the cross-sectional variation of expected returns that cannot be explained by our conditional CAPM specification. An alternative is to examine whether the betas with respect to the size and book-to-market factors, SMB_t and HML_t, introduced in Fama and French (1993), can explain the cross-sectional variation of expected returns not explained by our model. Although Berk (1995) shows that the log of size and the log of book-to-market equity should be correlated with expected returns in the cross-section, his observation does not imply that this correlation can be captured by the two factors. Hence, we are interested in examining whether the two factors that Fama and French (1993) identify from the data are proxying for the risk associated with the return on human capital and beta instability that we model.[20] For this purpose, we consider the following models:

$$\mathrm{E}[R_{it}] = c_0 + c_{\mathrm{vw}}\beta_i^{\mathrm{vw}} + c_{\mathrm{SMB}}\beta_i^{\mathrm{SMB}} + c_{\mathrm{HML}}\beta_i^{\mathrm{HML}} \tag{34}$$

$$\begin{aligned}\mathrm{E}[R_{it}] = c_0 &+ c_{\mathrm{vw}}\beta_i^{\mathrm{vw}} + c_{\mathrm{prem}}\beta_i^{\mathrm{prem}} + c_{\mathrm{labor}}\beta_i^{\mathrm{labor}} \\ &+ c_{\mathrm{SMB}}\beta_i^{\mathrm{SMB}} + c_{\mathrm{HML}}\beta_i^{\mathrm{HML}},\end{aligned} \tag{35}$$

where all the betas are calculated in the same way as β_i^{vw}.

The empirical results are given in Table IV. The top half of the table gives the estimates for model (34). The R^2 is 55.12 percent, which is not much different from 55.21 percent, the R^2 for the PL-model. This means that the PL-model fits the data at least as well as model (34) does. Also, the estimated value of the zero-beta rate is not very different from the one obtained using the PL-model. However, the HJ-distance for model (34) is 0.6432, which is clearly larger than that for the PL-model (0.6184). In other words, although the two models do equally well on average, the pricing error for the portfolio that is most mispriced by model (34) is larger than the pricing error for the portfolio that is most mispriced by the PL-model. The p-value also shows that the HJ-distance for model (34) is significantly different from zero. The lower half of the table gives the estimates for model (35). The R^2 goes up from 55.21 to 64.04 percent when the two factors in Fama and French (1993) are included. This is about the same increase that is obtained when size is included. None of

[20] Daniel and Titman (1995) find that only part of the return premia on small capitalization and high book-to-market stocks can be explained by the betas with respect to the two factors introduced by Fama and French (1993). Hansen and Jagannathan (1994) point out that any given misspecified model can be "fixed" by adding a particular "modifying portfolio payoff" to the stochastic discount factor associated with the model. Equivalently, any given misspecified linear beta-pricing model can be "fixed" by adding one more beta, where the additional beta is computed with respect to the return on the "modifying portfolio." The results in Fama and French (1993) suggest that the modifying portfolio associated with the static CAPM is a portfolio of only two factors—the size and book-to-market factors. However, there is no theoretical explanation for this empirical regularity.

Table IV

Comparison with the Factors Used by Fama and French (1993)

This table gives the estimates for the cross-sectional regression model

$$E[R_{it}] = c_0 + c_{vw}\beta_i^{vw} + c_{prem}\beta_i^{prem} + c_{labor}\beta_i^{labor} + c_{SMB}\beta_i^{SMB} + c_{HML}\beta_i^{HML}$$

and the model for the moments

$$E[R_{it}(\delta_0 + \delta_{vw}R_t^{vw} + \delta_{prem}R_t^{prem} + \delta_{labor}R_t^{labor} + \delta_{SMB}SMB_t + \delta_{HML}HML_t)] = 1,$$

with either a subset or all of the variables. Here, R_{it} is the return on portfolio i ($i = 1, 2, \ldots, 100$) in month t (July 1963–December 1990), R_t^{vw} is the return on the value-weighted index of stocks, R_{t-1}^{prem} is the yield spread between low- and high-grade corporate bonds, R_t^{labor} is the growth rate in per capita labor income, and SMB_t and HML_t denote the respective Fama and French (1993) factors that are designed to capture the risks related to firm size and book-to-market equity. The β_i^{vw} is the slope coefficient in the OLS regression of R_{it} on a constant and R_t^{vw}. The other betas are estimated in a similar way. The regression models are estimated by using the Fama-MacBeth procedure. The "corrected t- and p-values" take sampling errors in the estimated betas into account. The models for the moments are estimated by using the Generalized Method of Moments with the Hansen-Jagannathan weighting matrix. The minimized value of the GMM criterion function is the first item under the "HJ-dist," with the associated p-value immediately below it. All the R-squares and p-values are reported as percentages.

Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{SMB}	c_{HML}	R-square
Estimate:	1.39	−0.45			0.33	0.25	55.12
t-value:	6.07	−0.95			1.53	0.96	
p-value:	0.00	34.34			12.60	33.59	
Corrected-t:	5.99	−0.94			1.51	0.95	
Corrected-p:	0.00	34.97			13.12	34.19	
Coefficient:	δ_0	δ_{vw}	δ_{prem}	δ_{labor}	δ_{SMB}	δ_{HML}	HJ-dist
Estimate:	0.98	2.62			−4.56	−0.94	0.6432
t-value:	35.00	1.35			−2.10	−0.28	
p-value:	0.00	17.78			3.60	77.91	0.65
Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{SMB}	c_{HML}	R-square
Estimate:	1.20	−0.38	0.22	0.11	0.16	0.22	64.04
t-value:	5.24	−0.80	3.32	2.25	0.78	0.84	
p-value:	0.00	42.41	0.09	2.44	43.79	40.24	
Corrected-t:	4.60	−0.70	2.95	1.99	0.68	0.74	
Corrected-p:	0.00	48.22	0.32	4.69	49.49	46.11	
Coefficient:	δ_0	δ_{vw}	δ_{prem}	δ_{labor}	δ_{SMB}	δ_{HML}	HJ-dist
Estimate:	2.17	2.62	−62.00	−89.33	−3.30	−0.59	0.6123
t-value:	6.09	1.26	−2.94	−2.67	−1.42	−0.18	
p-value:	0.00	20.90	0.32	0.77	15.52	85.98	18.58

the coefficients corresponding to the two factors in Fama and French (1993) is statistically significantly different from zero after taking sampling errors into account. These results suggest that the two Fama and French (1993) factors may proxy for the risk associated with the return on human capital and beta instability.

In our study, we focus on the Black version of the conditional CAPM, which assumes that the borrowing and the lending rates are different. The zero-beta rate in such an environment should lie between the riskless borrowing and riskless lending rates. To examine whether this is true, we proceed as follows. We first assume that the riskless lending and borrowing rates are the same as the interest rate on one-month Treasury bills (T-bills). In that case the model should assign the right average return to T-bills as well. Let $R_{\mathrm{TBill}t}$ denote the monthly return on the T-bills. Applying the PL-model to the T-bill, we have

$$\mathrm{E}[R_{\mathrm{TBill}t}] = c_0 + c_{\mathrm{vw}}\beta_{\mathrm{TBill}}^{\mathrm{vw}} + c_{\mathrm{prem}}\beta_{\mathrm{TBill}}^{\mathrm{prem}} + c_{\mathrm{labor}}\beta_{\mathrm{TBill}}^{\mathrm{labor}}. \quad (36)$$

Subtracting the above equation from the PL-model gives the relation between expected excess returns and betas for the 100 portfolios:

$$\mathrm{E}[\tilde{R}_{it}] = c_{\mathrm{vw}}\tilde{\beta}_i^{\mathrm{vw}} + c_{\mathrm{prem}}\tilde{\beta}_i^{\mathrm{prem}} + c_{\mathrm{labor}}\tilde{\beta}_i^{\mathrm{labor}} \qquad i = 1, \ldots, 100, \quad (37)$$

where $\tilde{R}_{it} = R_{it} - R_{\mathrm{TBill}t}$ and $\tilde{\beta}_i^{\mathrm{vw}} = \mathrm{Cov}(\tilde{R}_{it}, R_t^{\mathrm{vw}})/\mathrm{Var}(R_t^{\mathrm{vw}})$ for $i = 1, \ldots, 100$, and other $\tilde{\beta}$s are defined in a similar way. If the borrowing and lending rates are different, the relation given in (37) should be modified to include a positive intercept term, which should equal the difference between the average zero-beta rate and the average T-bill rate. One way to examine model misspecification is to estimate the above relation with an intercept term and test if it is positive and reasonable given our priors regarding what the difference between the zero-beta rate and T-bill rate should be. The moment restrictions implied by this model are

$$\mathrm{E}[\tilde{R}_{it}(1 + \tilde{\delta}_{\mathrm{vw}}R_t^{\mathrm{vw}} + \tilde{\delta}_{\mathrm{prem}}R_{t-1}^{\mathrm{prem}} + \tilde{\delta}_{\mathrm{labor}}R_t^{\mathrm{labor}})] = 0,$$

where $\tilde{\delta}_{\mathrm{vw}}$, $\tilde{\delta}_{\mathrm{prem}}$, and $\tilde{\delta}_{\mathrm{labor}}$ are the constants defined as follows:

$$\tilde{\delta}_{\mathrm{vw}} = -\frac{c_{\mathrm{vw}}}{\tilde{\delta}_0\mathrm{Var}(R_t^{\mathrm{vw}})} \qquad \tilde{\delta}_{\mathrm{prem}} = -\frac{c_{\mathrm{prem}}}{\tilde{\delta}_0\mathrm{Var}(R_{t-1}^{\mathrm{prem}})} \qquad \tilde{\delta}_{\mathrm{labor}} = -\frac{c_{\mathrm{labor}}}{\tilde{\delta}_0\mathrm{Var}(R_t^{\mathrm{labor}})}.$$

$$\tilde{\delta}_0 = 1 + \frac{c_{\mathrm{vw}}\mathrm{E}[R_t^{\mathrm{vw}}]}{\mathrm{Var}(R_t^{\mathrm{vw}})} + \frac{c_{\mathrm{prem}}\mathrm{E}[R_t^{\mathrm{prem}}]}{\mathrm{Var}(R_{t-1}^{\mathrm{prem}})} + \frac{c_{\mathrm{labor}}\mathrm{E}[R_t^{\mathrm{labor}}]}{\mathrm{Var}(R_t^{\mathrm{labor}})}.$$

These moment restrictions can be tested using the Generalized Method of Moments as described in Section II-B. Notice that one should not subtract the T-bill return from any of the factors when calculating betas or constructing the stochastic discount factor since the zero-beta return may be different from the T-bill return. In contrast, the three-factor model for the excess returns specified in Fama and French (1993) is

$$\mathrm{E}[\tilde{R}_{it}] = c_{\mathrm{vw}}\ddot{\beta}_i^{\mathrm{vw}} + c_{\mathrm{SMB}}\tilde{\beta}_i^{\mathrm{SMB}} + c_{\mathrm{HML}}\tilde{\beta}_i^{\mathrm{HML}}, \quad (38)$$

where $\ddot{\beta}_i^{\text{vw}} = \text{Cov}(\tilde{R}_{it}, R_t^{\text{vw}} - R_{\text{TBill}t})/\text{Var}(R_t^{\text{vw}} - R_{\text{TBill}t})$, which is the slope of regressing the excess asset return on the *excess* return of the CRSP value-weighted portfolio.

The results for excess returns are presented in Table V. The intercepts in the regressions for both the models are significantly different from zero, which suggests that the zero-beta rate is different from the T-bill rate. Neither our conditional CAPM specification nor the Fama-French three-factor specification assigns the right value to T-bills. The fact that the intercept term exceeds one percent per month suggests that both specifications are missing some important aspect of reality.

Ferson and Foerster (1994) point out that the GMM has rather poor finite sample properties. It is for this reason that we chose not to test the conditional CAPM directly, but rather to test its unconditional implications. This reduces the dimensionality of the problem and hence is likely to result in better finite sample statistical properties. This is also the reason we chose to use the weighting matrix suggested by Hansen and Jagannathan (1994) instead of the optimal GMM weighting matrix. Since the HJ weighting matrix does not depend on the unknown parameters that are being estimated, it is likely to improve the statistical properties of the GMM tests in finite samples.[21] Zhou (1994) provides some evidence that supports this view. In addition, we also formed portfolios of stocks by first sorting them into size quintiles and then pre-beta quintiles. We then repeated all the tests, using the time series of monthly returns on these 25 portfolios. The results for the 25 portfolios are qualitatively similar to those for the 100 portfolios reported in this paper.

IV. Conclusion

There are two major difficulties in examining the empirical support for the static CAPM. First, the real world is inherently dynamic and not static. Second, the return on the portfolio of aggregate wealth is not observable. These issues are typically ignored in empirical studies of the CAPM. It is commonly assumed that betas of assets remain constant over time, and the return on stocks measures the return on the aggregate wealth portfolio. Under these assumptions, Fama and French (1992) find that the relation between average return and beta is flat and that there is a strong size effect.

We argue that those two assumptions are not reasonable. Relaxing the first assumption naturally leads us to examine the conditional CAPM. We demonstrate that the empirical support for our conditional CAPM specification is rather strong. When betas and expected returns are allowed to vary over time by assuming that the CAPM holds period by period, the size effects and the statistical rejections of the model specifications become much weaker. When a proxy for the return on human capital is also included in measuring the return

[21] However, this issue is not explored in this paper. In a separate paper, we will compare the sampling properties of the optimal and HJ weighting matrices.

Table V

Tests Using the Time Series of Monthly Excess Returns on the 100 Size-Beta Sorted Portfolios

This table gives the estimates for the following two regression models:

$$\mathrm{E}[\tilde{R}_{it}] = c_{\mathrm{vw}}\tilde{\beta}_i^{\mathrm{vw}} + c_{\mathrm{prem}}\tilde{\beta}_i^{\mathrm{prem}} + c_{\mathrm{labor}}\tilde{\beta}_i^{\mathrm{labor}} \qquad \mathrm{E}[\tilde{R}_{it}] = c_{\mathrm{vw}}\ddot{\beta}_i^{\mathrm{vw}} + c_{\mathrm{SMB}}\tilde{\beta}_i^{\mathrm{SMB}} + c_{\mathrm{HML}}\tilde{\beta}_i^{\mathrm{HML}}$$

and the two models for the moments

$$\mathrm{E}[\tilde{R}_{it}(1 + \tilde{\delta}_{vw}R_t^{\mathrm{vw}} + \tilde{\delta}_{\mathrm{prem}}R_{t-1}^{\mathrm{prem}} + \tilde{\delta}_{\mathrm{labor}}R_t^{\mathrm{labor}})] = 0$$

$$\mathrm{E}[\tilde{R}_{it}(1 + \tilde{\delta}_{vw}\tilde{R}_t^{\mathrm{vw}} + \tilde{\delta}_{\mathrm{SMB}}R_t^{\mathrm{SMB}} + \tilde{\delta}_{\mathrm{HML}}R_t^{\mathrm{HML}})] = 0.$$

Here, $\tilde{R}_{it} = R_{it} - R_t^{\mathrm{TBill}}$, where R_{it} is the return on portfolio i ($i = 1, 2, \ldots, 100$) in month t (July 1963–December 1990) and R_t^{TBill} is the return on the T-bill. R_t^{vw} is the return on the value-weighted index of stocks and $\tilde{R}_t^{\mathrm{vw}} = R_t^{\mathrm{vw}} - R_t^{\mathrm{TBill}}$. R_{t-1}^{prem} is the yield spread between low- and high-grade corporate bonds, R_t^{labor} is the growth rate in per capita labor income, and SMB_t and HML_t denote the respective Fama and French (1993) factors that are designed to capture the risks related to firm size and book-to-market equity. The $\tilde{\beta}_i^{\mathrm{vw}}$ is the slope coefficient in the OLS regression of $\tilde{R}_{it}$ on a constant and R_t^{vw}. The other $\tilde{\beta}$s are estimated in a similar way. The $\ddot{\beta}_i^{\mathrm{vw}}$ is the slope coefficient in the OLS regression of $\tilde{R}_{it}$ on a constant and $\tilde{R}_t^{\mathrm{vw}}$. The regression models are estimated by using the Fama-MacBeth procedure. The "corrected t- and p-values" take sampling errors in the estimated betas into account. The models for the moments are estimated by using the Generalized Method of Moments with the Hansen-Jagannathan weighting matrix. The minimized value of the GMM criterion function is the first item under the "HJ-dist," with the associated p-value immediately below it. All the R-squares and p-values are reported as percentages.

Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{SMB}	c_{HML}	R-square
Estimate:	0.79	−0.40	0.34	0.22			55.21
t-value:	3.58	−1.18	3.31	2.31			
p-value:	0.03	23.76	0.09	2.07			
Corrected-t:	2.66	−0.88	2.48	1.73			
Corrected-p:	0.78	37.99	1.31	8.44			
Coefficient:		$\tilde{\delta}_{\mathrm{vw}}$	$\tilde{\delta}_{\mathrm{prem}}$	$\tilde{\delta}_{\mathrm{labor}}$	$\tilde{\delta}_{\mathrm{SMB}}$	$\tilde{\delta}_{\mathrm{HML}}$	HJ-dist
Estimate:		−0.10	−48.21	−59.92			0.1443
t-value:		−0.25	−13.13	−9.25			
p-value:		80.10	0.00	0.00			96.49
Coefficient:	c_0	c_{vw}	c_{prem}	c_{labor}	c_{SMB}	c_{HML}	R-square
Estimate:	0.86	−0.47			0.33	0.24	55.20
t-value:	3.76	−0.99			1.56	0.92	
p-value:	0.02	32.20			11.91	35.97	
Corrected-t:	3.71	−0.98			1.54	0.90	
Corrected-p:	0.02	32.84			12.42	36.57	
Coefficient:		$\tilde{\delta}_{\mathrm{vw}}$	$\tilde{\delta}_{\mathrm{prem}}$	$\tilde{\delta}_{\mathrm{labor}}$	$\tilde{\delta}_{\mathrm{SMB}}$	$\tilde{\delta}_{\mathrm{HML}}$	HJ-dist
Estimate:		−4.58			−0.45	−9.94	0.5348
t-value:		−3.34			−0.23	−3.80	
p-value:		0.08			81.79	0.01	26.48

on aggregate wealth, the pricing errors of the model are not significant at conventional levels. More importantly, firm size does not have any additional explanatory power.

Although the conditional model performs substantially better than the static model, we still advocate caution in interpreting these results as strong support for the conditional CAPM for the following reasons:

First, our modeling of the time variations in betas is rather simple. If one were to take seriously the criticism that the real world is inherently dynamic, then it may be necessary to model explicitly what is missing in a static model. In particular, in a dynamic world, investors may care about hedging against a variety of risks that do not arise in a static economy. One possibility is to extend Merton's inter-temporal CAPM for empirical analysis, along the lines suggested by Campbell (1993a and 1993b). However, the dynamic conditional CAPM has an undesirable feature. The econometrician has to take a stand on the nature of the information available to the investors. For example, while deriving the unconditional multi-factor model implied by the conditional CAPM, we assumed that the conditional market risk premium is a linear function of the yield spread between low- and high-grade bonds. An alternative is to follow Bansal, Hsieh, and Viswanathan (1993) and Bansal and Viswanathan (1993) and consider unconditional nonlinear factor models which may be relatively more robust to information-set misspecification.

Second, a number of events occur at deterministic monthly and yearly frequencies. It may be reasonable to expect that such events may influence the behavior of asset prices at these frequencies. Since such events are outside the scope of asset-pricing models like the CAPM, one strategy would be to study the performance of models by using annual data over a sufficiently long period of time, as in Amihud, Christensen, and Mendelson (1992), Jagannathan and Wang (1992), and Kothari, Shanken, and Sloan (1995). Such an approach has its own shortcomings, the most important of which is that the economy may not really be stationary. There is some need for developing statistical sampling theories for making inferences that are robust to the presence of such features, possibly along the lines of Bossaerts (1994).

Finally, we have to keep in mind that the CAPM, like any other model, is only an approximation of reality. Hence, it would be rather surprising if it turns out to be "100 percent accurate." The interesting question is not whether a particular asset-pricing model can be rejected by the data. The question is: "How inaccurate is the model?" Fama and French (1992) show that the static version of the CAPM is very inaccurate. We find that the conditional version of the CAPM explains the cross-section of stock returns rather well. In doing so, we implicitly assume that the portfolio of stocks used in our study is economically important. As we point out in Appendix B, it is possible to mask or highlight the model specification error through appropriate choice of the portfolios. We will not be surprised if subsequent studies form a set of portfolios for which the model we examine in this study performs rather differently. In order to reconcile these differing views, we need to devise methods for evaluating the economic importance of the data sets used in empirical studies of asset-pricing models. We intend to focus on this issue in our future research.

The conditional CAPM we study in this article is very different from what is commonly understood as the CAPM, and resembles the multi-factor model of Ross (1976). The model we evaluate has three betas, whereas the standard CAPM has only one beta. We chose this model because (i) the use of a better proxy for the return on the market portfolio results in a two-beta model in place of the classical one-beta model, and (ii) when the CAPM holds in a conditional sense, unconditional expected returns will be linear in the unconditional beta as well as a measure of beta-instability over time. When the CAPM holds conditionally, we need more than the unconditional beta calculated by using the value-weighted stock index to explain the cross-section of unconditional expected returns.

Appendix A: Modeling the Unconditional Expected Returns

We first show that when betas vary over time, $(\beta_i, \beta_i^\gamma)$ is a linear function of $(\bar{\beta}_i, \vartheta_i)$. We then show that if ϑ_i is a linear function of $\bar{\beta}_i$, the static CAPM will obtain even in the unconditional sense—i.e., unconditional expected returns will be linear in the market beta β_i. In this case, β_i^γ is also a linear function of β_i. Finally, we show that when β_i^γ is not a linear function of β_i, which should be the usual case, $(\beta_i, \beta_i^\gamma)$ will contain all the necessary information contained in $(\bar{\beta}_i, \vartheta_i)$. Hence, expected returns will be linear in $(\bar{\beta}_i, \vartheta_i)$ as well as $(\beta_i, \beta_i^\gamma)$.

To show that $(\beta_i, \beta_i^\gamma)$ is a linear function of $(\bar{\beta}_i, \vartheta_i)$, note that the market return R_{mt} also satisfies the conditional CAPM. This gives the following equations:

$$\mathrm{E}[R_{mt}|I_{t-1}] = \gamma_{0t-1} + \gamma_{1t-1} \tag{A1}$$

$$\gamma_{1t-1} = \mathrm{E}[R_{mt} - \gamma_{0t-1}|I_{t-1}]. \tag{A2}$$

We then define ε_{it} as

$$\varepsilon_{it} = R_{it} - \gamma_{0t-1} - (R_{mt} - \gamma_{0t-1})\beta_{it-1}. \tag{A3}$$

It follows from equations (2), (3), and (A3) that

$$\mathrm{E}[\varepsilon_{it}|I_{t-1}] = 0 \tag{A4}$$

$$\mathrm{E}[\varepsilon_{it}R_{mt}|I_{t-1}] = 0. \tag{A5}$$

These two equations together imply the following orthogonality conditions:

$$\mathrm{E}[\varepsilon_{it}] = 0 \tag{A6}$$

$$\mathrm{E}[\varepsilon_{it}R_{mt}] = 0 \tag{A7}$$

$$\mathrm{E}[\varepsilon_{it}\gamma_{1t-1}] = 0. \tag{A8}$$

We can substitute equation (7) into (A3) to obtain

$$R_{it} = \gamma_{0t-1} + (R_{mt} - \gamma_{0t-1})\bar{\beta}_i + (R_{mt} - \gamma_{0t-1})(\gamma_{1t-1} - \gamma_1)\vartheta_i + (R_{mt} - \gamma_{0t-1})\eta_{it-1} + \varepsilon_{it}. \tag{A9}$$

From the definition of covariance, the expression given above for R_{it} in (A9), and the orthogonality conditions in (A6), (A7), and (A8), we obtain

$$\begin{aligned} \mathrm{Cov}(R_{it}, R_{mt}) &= \mathrm{Var}(R_{mt})\beta_i \\ &= \mathrm{Cov}(\gamma_{0t-1}, R_{mt}) + \mathrm{Cov}(R_{mt} - \gamma_{0t-1}, R_{mt})\bar{\beta}_i \\ &\quad + \mathrm{Cov}((R_{mt} - \gamma_{0t-1})(\gamma_{1t-1} - \gamma_1), R_{mt})\vartheta_i \\ &\quad + \mathrm{Cov}((R_{mt} - \gamma_{0t-1})\eta_{it-1}, R_{mt}), \end{aligned} \tag{A10}$$

$$\begin{aligned} \mathrm{Cov}(R_{it}, \gamma_{1t-1}) &= \mathrm{Var}(\gamma_{1t-1})\beta_i^{\gamma} \\ &= \mathrm{Cov}(\gamma_{0t-1}, \gamma_{1t-1}) + \mathrm{Cov}(R_{mt} - \gamma_{0t-1}, \gamma_{1t-1})\bar{\beta}_i \\ &\quad + \mathrm{Cov}((R_{mt} - \gamma_{0t-1})(\gamma_{1t-1} - \gamma_1), \gamma_{1t-1})\vartheta_i \\ &\quad + \mathrm{Cov}((R_{mt} - \gamma_{0t-1})\eta_{it-1}, \gamma_{1t-1}). \end{aligned} \tag{A11}$$

Let us denote the conditional variance of the market return by $v_{t-1} = \mathrm{E}[R_{mt}|I_{t-1}]$. Using equations (8), (9), (A1), and (A2), one can show that the last term in equation (A10) is

$$\mathrm{Cov}((R_{mt} - \gamma_{0t-1})\eta_{it-1}, R_{mt}) = \mathrm{E}[(v_{t-1} + \gamma_{1t-1}^2 + \gamma_{1t-1}\gamma_{0t-1})\eta_{it-1}], \tag{A12}$$

and the last term in equation (A11) is

$$\mathrm{Cov}((R_{mt} - \gamma_{0t-1})\eta_{it-1}, \gamma_{1t-1}) = \mathrm{E}[\gamma_{1t-1}^2\eta_{it-1}]. \tag{A13}$$

Then, equations (A10) and (A11) imply that there will be a linear relation between $(\beta_i, \beta_i^{\gamma})$ and $(\bar{\beta}_i, \vartheta_i)$, if the expressions in (A12) and (A13) are zero. Hence, we make the following additional assumption throughout the paper unless mentioned otherwise:

ASSUMPTION 3. *For each asset i, the residual beta η_{it-1} satisfies*

$$\mathrm{E}[\eta_{it-1}v_{t-1}] = 0 \tag{A14}$$

$$\mathrm{E}[\eta_{it-1}\gamma_{1t-1}^2] = 0 \tag{A15}$$

$$\mathrm{E}[\eta_{it-1}\gamma_{1t-1}\gamma_{0t-1}] = 0. \tag{A16}$$

According to the first equation in Assumption 3, the residual betas are uncorrelated with the conditional volatility of the market return. If the market return is conditionally homoskedastic, which is an assumption sometimes made by researchers, then the first equation in Assumption 3 is a consequence of equation (8). Since the residual betas do not affect the unconditional expected return, as was shown in Section I.*C*, we can ignore η_{it-1} by assuming that they are random noises that are uncorrelated with the market conditions, then all the equations in Assumption 3 will hold. Under Assumption 3, the following linear function for the betas follows from equations (A10) and (A11):

LEMMA 1. *There are constants* $\{b_{kl} : k = 1, 2; l = 0, 1, 2\}$ *such that*

$$\begin{pmatrix} \beta_i \\ \beta_i^{\gamma} \end{pmatrix} = \begin{pmatrix} b_{10} \\ b_{20} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} \bar{\beta}_i \\ \vartheta_i \end{pmatrix}. \tag{A17}$$

If the beta-prem sensitivity is a linear function of the expected beta, then the unconditional expected return will be linear in the market beta, i.e., the CAPM will hold unconditionally as well. This observation leads to the next lemma:

LEMMA 2. *If* ϑ_i *is a linear function of* $\bar{\beta}_i$, *then there are some constants* a_0 *and* a_1 *such that the equation*

$$\mathrm{E}[R_{it}] = a_0 + a_1\beta_i \tag{A18}$$

holds for every asset i, i.e., the static CAPM will hold for unconditional expected returns.

Proof. Let $\vartheta_i = d_0 + d_1\bar{\beta}_i$ and substitute it into equation (10) to get

$$\mathrm{E}[R_{it}] = [\gamma_0 + d_0\mathrm{Var}(\gamma_{1t-1})] + [\gamma_1 + d_1\mathrm{Var}(\gamma_{1t-1})]\bar{\beta}_i. \tag{A19}$$

Substitute $\vartheta_i = d_0 + d_1\bar{\beta}_i$ into the first equation in (A17) to obtain

$$\beta_i = (b_{10} + d_0 b_{12}) + (b_{11} + d_1 b_{12})\bar{\beta}_i, \tag{A20}$$

which implies that β_i is a linear function of $\bar{\beta}_i$. Since β_i is not a constant across assets, $b_{11} + d_1 b_{12}$ must be nonzero. So, we can substitute (A20) into equation (A19) to obtain equation (A18), which completes the proof.

One important special case arises when the conditional betas are uncorrelated with the market risk premium. In this case, we have $\vartheta_i = 0$, and thus, the single beta model (A18) in Theorem 2 holds. Chan and Chen (1988) derived equation (A18) by assuming

$$\beta_{it-1} = \bar{\beta}_i + \lambda_{t-1}(\bar{\beta}_i - \bar{\beta}^*) + \eta^*_{it-1}, \tag{A21}$$

where $\bar{\beta}^*$ is the cross-sectional average of $\bar{\beta}_i$, λ_{t-1} has zero mean, and η^*_{it-1} is the random noise. With this specification for the conditional betas, we have

$$\vartheta_i = \frac{\mathrm{Cov}(\lambda_{t-1}, \gamma_{1t-1})}{\mathrm{Var}(\gamma_{1t-1})}(\bar{\beta}_i - \bar{\beta}^*). \tag{A22}$$

Therefore, the assumption made by Chan and Chen (1988) implies that the beta-prem sensitivity ϑ_i is a linear function of $\bar{\beta}_i$, and hence, equation (A18) holds under their assumption.

The restriction that the beta-prem sensitivity is a linear function of the expected beta implies the following restriction on the premium betas:

LEMMA 3. *If ϑ_i is a linear function of $\bar{\beta}_i$, then β_i^γ is a linear function of β_i.*

Proof. We substitute $\vartheta_i = d_0 + d_1\bar{\beta}_i$ into equation (A17) to obtain

$$\beta_i = [b_{10} + d_0 b_{12}] + [b_{11} + d_1 b_{12}]\bar{\beta}_i \tag{A23}$$

$$\beta_i^\gamma = [b_{20} + d_0 b_{22}] + [b_{21} + d_1 b_{22}]\bar{\beta}_i. \tag{A24}$$

Since β_i is not constant across assets, we must have $b_{11} + d_1 b_{12} \neq 0$. We can then substitute equation (A23) into equation (A24) to express β_i^γ as a linear function of β_i, which completes the proof.

If the premium beta is not a linear function of the market beta, then, by Lemma 3, beta-prem sensitivity cannot be a linear function of the expected beta. In this case, the single-beta model in Lemma 2 will not hold, i.e., the CAPM will not hold unconditionally, even though it holds in a conditional sense. Instead, the unconditional expected return will be a linear function of *two* variables—the market beta and the premium beta. This fact is stated as Theorem 1 in Section I.C.

Now, let us prove Theorem 1. We first prove that the 2 by 2 matrix in equation (A17) is invertible. Suppose it is singular; then there is a nonzero vector (x, y) such that

$$(x, y)\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = 0, \tag{A25}$$

which implies that $x\beta_i + y\beta_i^\gamma$ is a constant across assets. Since β_i is not a constant across assets, we must have $y \neq 0$. But this means that β_i^γ is a linear function of β_i, which contradicts the assumption in Theorem 1. Now we can invert equation (A17) such that $(\bar{\beta}_i, \vartheta_i)$ are linear functions of $(\beta_i, \beta_i^\gamma)$ and then substitute them into equation (10) to obtain equation (13).

Appendix B: The Cross-Sectional Regressions

For the purpose of developing the sampling theory, it is more convenient to write all the unconditional models that we have discussed into the following form:

$$\mathrm{E}[R_{it}] = \sum_{k=1}^{K_1} c_{1k} z_{ik} + \sum_{k=1}^{K_2} c_{2k} \beta_{ik}, \tag{B1}$$

where $\{z_{ik}\}_{k=1,\ldots,K_1}$ are K_1 observable characteristics of asset i, $\beta_{ik} = \mathrm{Cov}(R_{it}, y_{kt})/\mathrm{Var}(y_{kt})$, with $\{y_{kt}\}_{k=1,\ldots,K_2}$ being K_2 economic variables, and $\{c_{jk}\}$ are some coefficients. As an example, for equation (24), we can let $K_1 = 2$, $z_{i1} \equiv 1$, $z_{i2} = \log(\mathrm{ME}_i)$, $K_2 = 3$, $y_{1t} = R_t^{\mathrm{vw}}$, $y_{2t} = R_{t-1}^{\mathrm{prem}}$, $y_{3t} = R_t^{\mathrm{labor}}$, $\beta_{i1} = \beta_i^{\mathrm{vw}}$, $\beta_{i2} = \beta_i^{\mathrm{prem}}$, and $\beta_{i3} = \beta_i^{\mathrm{labor}}$.

Equation (B1) can be written in a more concise form as

$$\mathrm{E}[\mathrm{R}_t] = \mathrm{Zc}_1 + \mathrm{Bc}_2 = \mathrm{Xc}, \tag{B2}$$

where $\mathrm{R}_t = (R_{1t}, \ldots, R_{Nt})'$, $\mathrm{c}_1 = (c_{11}, \ldots, c_{1K1})'$, $\mathrm{c}_2 = (c_{21}, \ldots, c_{2K2})'$, $\mathrm{c} = (\mathrm{c}_1' : \mathrm{c}_2')'$, and $\mathrm{X} = (\mathrm{Z} : \mathrm{B})$, where

$$\mathrm{Z} = \begin{pmatrix} z_{11} & \cdots & z_{1K_1} \\ \vdots & \ddots & \vdots \\ z_{N1} & \cdots & z_{NK_1} \end{pmatrix} \qquad \mathrm{B} = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1K_2} \\ \vdots & \ddots & \vdots \\ \beta_{N1} & \cdots & \beta_{NK_2} \end{pmatrix}.$$

In the cross-sectional regression method, we first estimate β_{ik} by the slope coefficient in the univariate regression of R_{it} on y_{kt} and a constant over time. Let $\hat{\beta}_{ik}$ be the estimated slope coefficient in this regression. Replacing all the betas in B by their estimates, we obtain an estimate of B which we denote by $\hat{\mathrm{B}}$. Let $\hat{\mathrm{X}} = (\mathrm{Z} : \hat{\mathrm{B}})$ and $\bar{\mathrm{R}}$ be the time-series average of R_t, i.e., $\bar{\mathrm{R}} = (1/\mathrm{T}) \sum_{t=1}^{T} \mathrm{R}_t$. The estimator of the parameters, denoted by $\hat{\mathrm{c}}$, in the cross-sectional regression method is obtained by regressing $\bar{\mathrm{R}}$ on $\hat{\mathrm{X}}$, that is, $\hat{\mathrm{c}} = (\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}\, \hat{\mathrm{X}}'\bar{\mathrm{R}}$. Here, we assume that both X and $\hat{\mathrm{X}}$ have the rank $K_1 + K_2$. If $\mathrm{plim}_{T\to\infty} \hat{\mathrm{B}} = \mathrm{B}$ and $\mathrm{plim}_{T\to\infty} \bar{\mathrm{R}} = \mathrm{E}[\mathrm{R}_t]$, then $\mathrm{plim}_{T\to\infty} \hat{\mathrm{c}} = \mathrm{c}$, i.e., $\hat{\mathrm{c}}$ is a consistent estimator of c.

Although the cross-sectional regression method does not provide a test for the linearity imposed by the model, it is still a very natural and intuitive tool for checking the ability of an unconditional model to explain the cross-sectional variation of average returns. The R^2 of the cross-sectional regression associated with a particular empirical specification provides a natural measure of how well that particular model does in explaining the cross-section of average returns. However, it is necessary to use caution in interpreting a low R^2 as indicating that a particular specification is bad in any absolute sense.

To see why, consider a hypothetical economy where the econometrician has observations on four assets. The betas with respect to a proxy market portfolio for the four assets are 0.5, 0.5, 2, and 2. The corresponding expected rates of returns are 12, 8, 24, and 20 percent. There are no measurement errors involved here. It can be verified that, in this case, the estimated regression equation will be

$$R_i = 6 + 8\beta_i + \hat{\varepsilon}_i \qquad i = 1, 2, 3, 4$$

and that the R^2 of the regression is 95 percent. Now consider forming four other portfolios (by an invertible linear transformation) from the four given assets as follows. Let $z = R_3 - R_4$ denote the payoff on the zero investment portfolio constructed by going long one dollar on the third asset and going short one dollar on the fourth asset. The beta of the payoff, z, is 0 by construction. Define the return on the four new portfolios by $R_1^* = R_1 + 3z$; $R_2^* = R_2 + 3z$; $R_3^* = R_3$; and $R_4^* = R_4$. Notice that the original four assets can be constructed as portfolios of these four new portfolios. The betas of the four portfolios defined this way are 0.5, 0.5, 2, and 2, respectively. The expected returns on these portfolios are 24, 20, 24, and 20 percent, respectively. Clearly, when these four portfolios are used, the relation between expected return and beta is flat (i.e., the R^2 is 0 percent). This shortcoming is not an issue for the way we use R^2 to compare the performance of different competing specifications of the CAPM, since we use the same set of portfolios across all model specifications.[22]

To assess the sampling errors associated with the estimated parameters, Fama and MacBeth (1973) suggest regressing R_t, instead of $\bar{R}$, on $\hat{X}$ for each period t to obtain

$$\hat{\mathrm{c}}_t = (\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}\hat{\mathrm{X}}'\mathrm{R}_t \tag{B3}$$

and then estimate the covariance matrix of $\sqrt{T}(\hat{\mathrm{c}} - \mathrm{c})$ by

$$\hat{\mathrm{V}} = \frac{1}{T}\sum_{t=1}^{T}(\hat{\mathrm{c}}_t - \bar{\mathrm{c}})(\hat{\mathrm{c}}_t - \bar{\mathrm{c}})', \tag{B4}$$

where $\bar{\mathrm{c}} = (1/T)\sum_{t=1}^{T}\hat{\mathrm{c}}_t$. It is easy to see that $\bar{\mathrm{c}} = \hat{\mathrm{c}}$. Substituting $\hat{\mathrm{c}}_t = (\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}\hat{\mathrm{X}}'\mathrm{R}_t$ into equation (B4), we have

$$\hat{\mathrm{V}} = (\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}\hat{\mathrm{X}}'\left[\frac{1}{T}\sum_{t=1}^{T}(\mathrm{R}_t - \bar{\mathrm{R}})(\mathrm{R}_t - \bar{\mathrm{R}})'\right]\hat{\mathrm{X}}(\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}. \tag{B5}$$

In order to understand the properties of the estimated covariance matrix $\hat{\mathrm{V}}$ provided by the Fama-MacBeth procedure, it is convenient to define $\mu = \mathrm{E}[\mathrm{R}_t]$ and use equation (B2) to write the average return as

$$\bar{\mathrm{R}} = \hat{\mathrm{X}}\mathrm{c} + (\bar{\mathrm{R}} - \mu) - (\hat{\mathrm{X}} - \mathrm{X})\mathrm{c}. \tag{B6}$$

Substituting it into the definition for $\hat{\mathrm{c}}$, we can obtain

$$\hat{\mathrm{c}} - \mathrm{c} = (\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}\hat{\mathrm{X}}'(\bar{\mathrm{R}} - \mu) - (\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}\hat{\mathrm{X}}'(\hat{\mathrm{B}} - \mathrm{B})\mathrm{c}_2. \tag{B7}$$

[22] Kandel and Stambaugh (1995) suggest using an alternative measure of goodness of fit that is invariant to portfolio formation for examining the performance of a given model. However, for comparing the relative performance of different models using the same set of assets, the OLS R^2 measure is quite appropriate.

Suppose that $\sqrt{T}(\bar{\mathrm{R}} - \mu)$ converges to a random variable $\tilde{\mathrm{u}}$ in distribution and $\sqrt{T}(\hat{\mathrm{B}} - \mathrm{B})$ converges to a random variable $\tilde{\mathrm{H}}$ in distribution. If $\mathrm{plim}_{T\to\infty} \hat{\mathrm{B}} = \mathrm{B}$, then

$$\sqrt{T}(\hat{\mathrm{c}} - \mathrm{c}) \overset{d}{\rightarrow} (\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'(\tilde{\mathrm{u}} - \tilde{\mathrm{H}}\mathrm{c}_2). \tag{B8}$$

Here, $(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\tilde{\mathrm{u}}$ is the sampling error of $\hat{\mathrm{c}}$ from replacing expected returns by average returns, and $(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\tilde{\mathrm{H}}\mathrm{c}_2$ is the sampling error from replacing true beta by their estimates. The conventional consistent estimate for the variance of $\tilde{\mathrm{u}}$ is

$$\frac{1}{T}\sum_{t=1}^{T} (\mathrm{R}_t - \bar{\mathrm{R}})(\mathrm{R}_t - \bar{\mathrm{R}})'.$$

Hence, a consistent estimate for the variance of $(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\tilde{\mathrm{u}}$ is given by

$$(\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}\hat{\mathrm{X}}'\left[\frac{1}{T}\sum_{T=1}^{T} (\mathrm{R}_t - \bar{\mathrm{R}})(\mathrm{R}_t - \bar{\mathrm{R}})'\right]\hat{\mathrm{X}}(\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1},$$

which is exactly $\hat{\mathrm{V}}$ in view of equation (B5). If we can ignore the sampling error $\tilde{\mathrm{H}}$ that is due to the errors associated with the estimated betas, then the consistent estimate for the variance of $\hat{\mathrm{c}}$ is given by $\hat{\mathrm{V}}$ obtained from the standard Fama-MacBeth procedure. If $\tilde{\mathrm{H}}$ is not negligible, the standard error of $\hat{\mathrm{c}}$ provided by the Fama-MacBeth procedure will generally be biased.

In general, it is difficult to assess the magnitude of the bias of the Fama-MacBeth procedure. However, under some additional assumptions, Shanken (1992) derives an expression for the bias when the betas are estimated using multiple regression. Since we use betas estimated from univariate OLS regressions, Shanken's formula is not directly applicable in our case.

In what follows, using methods similar to those used by Shanken (1992), we derive an expression for the sampling errors associated with parameters estimated using the cross-sectional regression method. We need to introduce two additional assumptions. For $i = 1, \ldots, N$, $k = 1, \ldots, K_2$, and $t = 1, \ldots, T$, define

$$\alpha_{ik} = \mathrm{E}[R_{it}] - \beta_{ik}\mathrm{E}[y_{kt}] \tag{B9}$$

$$e_{ikt} = R_{it} - \alpha_{ik} - \beta_{ik}y_{kt}. \tag{B10}$$

We then have

$$R_{it} = \alpha_{ik} + \beta_{ik}y_{kt} + e_{ikt} \tag{B11}$$

$$\mathrm{E}[e_{ikt}] = 0 \tag{B12}$$

$$\mathrm{E}[e_{ikt}y_{kt}] = 0. \tag{B13}$$

The two additional assumptions are as follows:

ASSUMPTION 4. *We assume that, for* $i, j = 1, \ldots, N$ *and* $k, l = 1, \ldots, K_2$, $\mathrm{E}[e_{ikt}|\{y_{ns}\}_{n=1,\ldots,K_2;\, s=1,\ldots,T}] = 0$ *and* $\mathrm{E}[e_{ikt}e_{jlt}|\{y_{ns}\}_{n=1,\ldots,K_2;\, s=1,\ldots,T}] = \sigma_{ijkl}$. *We then define*

$$\Sigma_{kl}\begin{pmatrix} \sigma_{11kl} & \cdots & \sigma_{1Nkl} \\ \vdots & \ddots & \vdots \\ \sigma_{N1kl} & \cdots & \sigma_{NNkl} \end{pmatrix}. \tag{B14}$$

ASSUMPTION 5. *Let* $\bar{y}_k = (1/T)\sum_{t=1}^{T} y_{kt}$. *We assume that the probability limit*

$$\operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} (y_{kt} - \bar{y}_k)(y_{lt} - \bar{y}_l) = \omega_{kl}$$

exists for $k, l = 1, \ldots, K_2$, *and* $\omega_{kk} > 0$ *for* $k = 1, \ldots, K_2$.

Under Assumptions 4 and 5, the limiting distribution of the estimated parameter vector is given by the following:

THEOREM 2. *Suppose that when* $T \to \infty$, $\sqrt{T}(\bar{\mathrm{R}}' - \mu', \mathrm{c}_2'(\hat{\mathrm{B}} - \mathrm{B})')'$ *converges to a joint normal distribution* $(\tilde{\mathrm{u}}', \mathrm{c}_2'\tilde{\mathrm{H}}')'$, *with zero mean. Suppose that*

$$\operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} (\mathrm{R}_t - \bar{\mathrm{R}})(\mathrm{R}_t - \bar{\mathrm{R}})' = \mathrm{Var}(\tilde{u}) \tag{B15}$$

$$\operatorname*{plim}_{T\to\infty} \hat{\mathrm{B}} = \mathrm{B}. \tag{B16}$$

Then, under Assumptions 4 *and* 5, $\sqrt{T}(\hat{\mathrm{c}} - \mathrm{c})$ *converges to a normal distribution with zero mean and variance* V + W, *where*

$$\mathrm{V} = \operatorname*{plim}_{T\to\infty} \hat{\mathrm{V}} \tag{B17}$$

$$\mathrm{W} = \sum_{k,l=1}^{K_2} c_{2k}c_{2l}(\omega_{kk}^{-1}\omega_{kl}\omega_{ll}^{-1})(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\Sigma_{kl}\mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}. \tag{B18}$$

Hence, under Assumptions 4 and 5, the bias of the Fama-MacBeth procedure is W. To obtain a consistent estimate of the sampling errors, we first use the Fama-MacBeth procedure to obtain $\hat{\mathrm{V}}$, and then apply Theorem 2 to obtain a consistent estimate of W as

$$\hat{\mathrm{W}} = \sum_{k,l=1}^{K_2} \hat{c}_{2k}\hat{c}_{2l}(\hat{\omega}_{kk}^{-1}\hat{\omega}_{kl}\hat{\omega}_{ll}^{-1})(\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}\hat{\mathrm{X}}'\hat{\Sigma}_{kl}\hat{\mathrm{X}}(\hat{\mathrm{X}}'\hat{\mathrm{X}})^{-1}, \tag{B19}$$

where $\hat{\omega}_{kl}$ and $\hat{\Sigma}_{kl}$ are the sample analogs of ω_{kl} and Σ_{kl}.

Theorem 2 can be proved as follows. First, we introduce some additional notations. By I_N (I_T) we denote the $N(T)$-dimensional identity matrix. By 1_T we denote a T-dimensional vector, with each of its elements equal to 1. Let $\bar{y}_k$ be the time-series sample average of y_{kt} and define

$$\mathrm{y}_k \equiv (y_{k1} - \bar{y}_k, \ldots, y_{kT} - \bar{y}_k)'$$

$$\mathrm{b}_k \equiv (\beta_{1k}, \ldots, \beta_{Nk})'$$

$$\hat{\mathrm{b}}_k \equiv (\hat{\beta}_{1k}, \ldots, \hat{\beta}_{Nk})'$$

$$\mathrm{e}_k \equiv (e_{1k1}, \ldots, e_{1kT}, \ldots, e_{Nk1}, \ldots, e_{NkT})'$$

$$\mathrm{Y} \equiv (\{y_{kt}\}_{k=1,\ldots,K_2,\, t=1,\ldots,T}).$$

It follows from (B11) that

$$\bar{\mathrm{R}} - \mu = \frac{1}{T}(1_N \otimes 1_T')\mathrm{e}_l, \qquad l = 1, \ldots, K_2, \tag{B20}$$

where $\otimes$ denotes the Kroneker product. By the definition of $\bar{\beta}_{ik}$, we have

$$\hat{\mathrm{b}}_k - \mathrm{b}_k = [\mathrm{I}_N \otimes ((\mathrm{y}_k'\mathrm{y}_k)^{-1}\mathrm{y}_k')]\mathrm{e}_k. \tag{B21}$$

In view of Assumption 4, equations (B20) and (B21) together imply

$$\mathrm{E}[(\hat{\mathrm{b}}_k - \mathrm{b}_k)(\bar{\mathrm{R}} - \mu)'|\mathrm{Y}] \tag{B22}$$

$$= \frac{1}{T}[\mathrm{I}_N \otimes ((\mathrm{y}_k'\mathrm{y}_k)^{-1}\mathrm{y}_k')]\mathrm{E}[\mathrm{e}_k\mathrm{e}_l'|\mathrm{Y}](\mathrm{I}_N \otimes 1_T) \tag{B23}$$

$$= \frac{1}{T}[\mathrm{I}_N \otimes ((\mathrm{y}_k'\mathrm{y}_k)^{-1}\mathrm{y}_k')](\Sigma_{kl} \otimes \mathrm{I}_T)(\mathrm{I}_N \otimes 1_T) \tag{B24}$$

$$= \frac{1}{T}\Sigma_{kl} \otimes [(\mathrm{y}_k'\mathrm{y}_k)^{-1}\mathrm{y}_k'1_T] \tag{B25}$$

$$= 0. \tag{B26}$$

This follows from the fact that $\mathrm{y}_k'1_T = 0$. Hence, $\hat{\mathrm{b}}_k - \mathrm{b}_k$ is uncorrelated with $\bar{\mathrm{R}} - \mu$. Assumption 4 and equation (B20) also imply that Y is uncorrelated with $\bar{\mathrm{R}} - \mu$. Therefore, $\tilde{\mathrm{u}}$ and $\tilde{\mathrm{H}}\mathrm{c}_2$ should be uncorrelated with each other, and the asymptotic variance of $\sqrt{T}(\hat{\mathrm{c}} - \mathrm{c})$ is given by

$$(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'[\mathrm{Var}(\tilde{\mathrm{u}}) + \mathrm{Var}(\tilde{\mathrm{H}}\mathrm{c}_2)]\mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}. \tag{B27}$$

By Assumptions 4 and 5, we have

$$
\begin{aligned}
&TE[(\hat{b}_k - b_k)(\hat{b}_l - b_l)'|y] \\
&= T[I_N \otimes ((y_k'y_k)^{-1}y_k')](\Sigma_{kl} \otimes I_T)[I_N \otimes ((y_l'y_l)^{-1}y_l')] \\
&= \Sigma_{kl} \otimes \left(\left(\frac{1}{T}\,y_k'y_k\right)^{-1}\left(\frac{1}{T}\,y_k'y_l\right)\left(\frac{1}{T}\,y_l'y_l\right)^{-1}\right) \\
&\to \omega_{kk}^{-1}\omega_{kl}\omega_{ll}^{-1}\Sigma_{kl} \qquad (\text{as } T \to +\infty).
\end{aligned}
$$

Thus,

$$
\mathrm{Var}(\tilde{H}c_2) = \sum_{k,l=1}^{K_2} c_{2k}c_{2l}\omega_{kk}^{-1}\omega_{kl}w_{ll}^{-1}\Sigma_{kl},
$$

and

$$
\begin{aligned}
W &= (X'X)^{-1}X'\mathrm{Var}(\tilde{H}c_2)X(X'X)^{-1} \\
&= \sum_{k,l=1}^{K_2} c_{2k}c_{2l}\omega_{kk}^{-1}\omega_{kl}\omega_{ll}^{-1}(X'X)^{-1}X'\Sigma_{kl}X(X'X)^{-1}.
\end{aligned}
$$

From equations (B5) and (B15) and expression (B27), it follows that

$$
V = (X'X)^{-1}X'\mathrm{Var}(\tilde{u})X(X'X)^{-1} = \mathrm{plim}\,\hat{V}.
$$

This completes the proof.

Appendix C: The Hansen-Jagannathan Distance

If there is only one asset, then it is relatively straightforward to compare the performance of the different versions of the unconditional model implied by the conditional CAPM. All we have to do is to compare the pricing error—i.e., the difference between the market price of an asset and the hypothetical price assigned to it by the stochastic discount factor implied by a particular empirical specification. When there are many assets (100 in our study), it is rather difficult to compare the pricing errors across the different candidate stochastic discount factors for the model.

In view of this, we follow Hansen and Jagannathan (1994), who suggest examining the pricing error on the portfolio that is most mispriced by a given model. There is a practical problem in implementing this simple idea. Suppose there are at least two assets which do not have the same pricing error for a given candidate stochastic discount factor. Let R_{1t} and R_{2t} denote the corresponding gross returns. The date $t - 1$ prices of these payoffs are both 1, i.e., by investing one dollar at date $t - 1$ in asset i, the investor gets the payoff R_{it} at date t. A given asset-pricing model may not assign a price of 1 at date $t - 1$ to the payoff R_{it}. Suppose the pricing error is ψ_i, i.e., the model assigns a price of $1 + \psi_i$. Consider forming a zero-investment portfolio by going long one dollar in security 1 and short one dollar in security 2. The pricing error on this zero-investment portfolio is $\psi_1 - \psi_2$. So long as this is not zero, the pricing error on any portfolio of the two assets with a price of one dollar can be made arbitrarily large by adding a scale multiple of this zero-investment portfolio. The same problem arises if instead of examining the pricing error we examine the difference between the expected return on a portfolio and the expected return assigned by a particular asset-pricing model to that portfolio. To overcome this problem, it is necessary to examine the pricing error on portfolios that have the same "size." Hansen and Jagannathan (1994) suggest using the second moment of the payoff as a measure of "size," i.e., examine the portfolio which has the maximum pricing error among all portfolio payoffs that have the unit second moment.

Consider a portfolio of the N primitive assets defined by the vector of portfolio weights x. The date t payoff on this portfolio is given by $x'R_t$. It has a price of $x'1_N$ at the beginning of each date. The pricing error on this portfolio is $x'E[w_t(\delta)]$. The second moment of this portfolio payoff is $E[x'R_t]^2$, i.e., the norm of this portfolio is $\sqrt{E[x'R_t]^2}$. For a given vector of parameters δ, Hansen and Jagannathan (1994) show that the maximum pricing error per unit norm on any portfolio of these N assets is given by

$$\text{Dist}(\delta) \equiv \sqrt{E[w_t(\delta)]' G^{-1} E[w_t(\delta)]}, \qquad \text{(C1)}$$

where $G = E[R_t R_t']$ and is assumed to be nonsingular. We refer to Dist as the *HJ-distance*. It is also the least-square distance between the given candidate stochastic discount factor and the nearest point to it in the set of all discount factors that price assets correctly. (See Hansen and Jagannathan (1994) for details.)

Since the vector, δ, of parameters describing a particular asset-pricing model is unknown, a natural way to estimate them is to choose those values for δ that minimize Dist given in (C1). We can then assess the specification error of a given stochastic discount factor by examining the maximum pricing error Dist associated with it, as suggested by Hansen and Jagannathan (1994).

Let

$$D_T = \frac{1}{T} \sum_{t=1}^{T} R_t Y_t' \qquad \text{(C2)}$$

$$w_T(\delta) = \frac{1}{T} \sum_{t=1}^{T} w_t(\delta) = D_T \delta - 1_N \qquad \text{(C3)}$$

$$G_T = \frac{1}{T} \sum_{t=1}^{T} R_t R_t'. \qquad \text{(C4)}$$

The sample analog of the HJ-distance defined in (C1) is thus

$$\mathrm{Dist}_T(\delta) = \sqrt{\min_{\delta} \mathrm{w}_T'(\delta)\mathrm{G}_T^{-1}\mathrm{w}_T(\delta)}. \tag{C5}$$

We will therefore estimate δ_T by minimizing the sample analog of (C1), i.e., choose δ_T as the solution to

$$\min_{\delta} \mathrm{w}_T(\delta)'\mathrm{G}_T^{-1}\mathrm{w}_T(\delta). \tag{C6}$$

The first order condition of the minimization problem is

$$\mathrm{D}_T'\mathrm{G}_T^{-1}\mathrm{w}_T(\delta_T) = 0, \tag{C7}$$

which gives

$$\delta_T = (\mathrm{D}_T'\mathrm{G}_T^{-1}\mathrm{D}_T)^{-1}\mathrm{D}_T'\mathrm{G}_T^{-1}1_N. \tag{C8}$$

The estimator δ_T is equivalent to a GMM estimator defined by Hansen (1982) with the moment restriction $\mathrm{E}[\mathrm{w}(\delta)] = 0$ and the weighting matrix G^{-1}. It is also an extremum estimator described in Amemiya (1985). Therefore, under some regularity conditions, δ_T is consistent and has an asymptotic normal distribution. For details, we refer readers to Hansen (1982) or Chapter 4 of Amemiya (1985). We refer to G^{-1} as the HJ weighting matrix.

If the weighting matrix is optimal in the sense of Hansen (1982), then $T[\mathrm{Dist}_T(\delta_T)]^2$ is asymptotically a random variable of χ^2 distribution with $N - K$ degrees of freedom, where K is the dimension of the vector δ of unknown parameters. However, G^{-1} is generally not optimal, and thus the distribution of $T[\mathrm{Dist}_T(\delta_T)]^2$ is not $\chi^2(N - K)$. The following theorem shows that the asymptotic distribution of $T[\mathrm{Dist}_T(\delta_T)]^2$ is a weighted sum of χ^2 distributed random variables, each of which has 1 degree of freedom.

THEOREM 3. *Suppose that for some δ_0 we have $\sqrt{T}\mathrm{w}_T(\delta_0) \xrightarrow{d} N(0_N, \mathrm{S})$, where* S *is a positive definite matrix. Assume $\mathrm{D}_T \xrightarrow{p} \mathrm{D}$, where* D *is an $N \times K$ matrix of rank K, and assume $\mathrm{G}_T \xrightarrow{p} \mathrm{G}$, where* G *is nonsingular. Let*

$$\mathrm{A} = \mathrm{S}^{1/2}\mathrm{G}^{-1/2}(\mathrm{I}_N - (\mathrm{G}^{-1/2})'\mathrm{D}[\mathrm{D}'\mathrm{G}^{-1}\mathrm{D}]^{-1}\mathrm{D}'\mathrm{G}^{-1/2})(\mathrm{G}^{-1/2})'(\mathrm{S}^{1/2})', \tag{C9}$$

where $\mathrm{S}^{1/2}$ and $\mathrm{G}^{1/2}$ are the upper-triangle matrices from the Cholesky decomposition of S *and* G, *and I_N is the N-dimensional identity matrix. Then* A *has exactly* $\mathrm{N} - \mathrm{K}$ *nonzero eigenvalues, which are positive and denoted by $\lambda_1, \ldots, \lambda_{N-K}$, and the asymptotic sampling distribution of the HJ-distance is*

$$T[\mathrm{Dist}_T(\delta_T)]^2 \xrightarrow{d} \sum_{j=1}^{N-K} \lambda_j v_j \quad \text{as} \quad T \to \infty, \tag{C10}$$

where $v_1, \ldots, v_{N-K}$ are independent $\chi^2(1)$ random variables.

Notice that, when all the eigenvalues are unity, $T[\text{Dist}_T(\delta_T)]^2$ has an asymptotic chi-square distribution with $N - K$ degrees of freedom. In this case, G^{-1} is optimal.

As long as we have a consistent estimate S_T of the matrix S, we can estimate the matrix A defined in Theorem 3 by

$$A_T = S_T^{1/2} G_T^{-1/2} (I_N - (G_T^{-1/2})' D_T [D_T' G_T^{-1} D_T]^{-1} D_T' G_T^{-1/2})(G_T^{-1/2})'(S_T^{1/2})'. \quad \text{(C11)}$$

Then we can estimate the λ_js by the positive eigenvalues of A_T.

Let u be the asymptotic distribution of $T[\text{Dist}_T(\delta_T)]^2$, i.e.,

$$u \equiv \sum_{j=1}^{N-K} \lambda_j v_j,$$

and let $\psi(u)$ be the probability distribution function of u. Although $\psi(u)$ is not a known distribution function, we can still conveniently compute the p-value to test the null hypothesis that the discount factors are specified correctly. Let

$$\{v_{ij}\}_{i=1,\ldots,T^*;\, j=1,\ldots,N-K}$$

denote $T^*(N - K)$ independent random draws from a $\chi^2(1)$ distribution. These random draws can be easily obtained on computer. Then, we can obtain a set of independent samples, $\{u_i\}_{i=1}^{T^*}$, by letting

$$u_i = \sum_{j=1}^{N-K} \lambda_j v_{ij}.$$

By the Law of Large Numbers, for each nonnegative number a, we have, as $T^* \rightarrow \infty$,

$$\frac{1}{T^*} \sum_{i=1}^{T^*} I(u_i \le a) \xrightarrow{p} \int_0^a d\psi(u) = \text{Prob}\{u \le a\},$$

where $I(u \le a)$ is an index function defined as

$$I(u \le a) = \begin{cases} 1 & \text{if } u \le a \\ 0 & \text{if } u > a. \end{cases} \quad \text{(C12)}$$

Here is the proof of Theorem 3. It follows from equation (C3) that

$$w_T(\delta_T) = w_T(\delta_0) + D_T(\delta_T - \delta_0). \quad \text{(C13)}$$

Multiplying both sides of equation (C13) by $D_T' G_T^{-1}$ and applying the first order condition (C7), we obtain

$$\delta_T - \delta_0 = -(D_T' G_T^{-1} D_T)^{-1} D_T' G_T^{-1} w_T(\delta_0). \quad \text{(C14)}$$

Substituting (C14) into (C13) gives

$$w_T(\delta_T) = (I_N - D_T(D_T'G_T^{-1}D_T)^{-1}D_T'G_T^{-1})w_T(\delta_0). \tag{C15}$$

After substituting (C15) into (C5) and some algebraic simplifications, we obtain

$$[\mathrm{Dist}_T(\delta_T)]^2 = w_T(\delta_0)'(G_T^{-1} - G_T^{-1}D_T(D_T'G_T^{-1}D_T)^{-1}D_T'G_T^{-1})w_T(\delta_0), \tag{C16}$$

which gives

$$T[\mathrm{Dist}_T(\delta_T)]^2 \xrightarrow{d} Z'(G^{-1} - G^{-1}D(D'G^{-1}D)^{-1}DG^{-1})Z, \tag{C17}$$

where Z is the N-dimensional random vector of normal distribution with zero mean and covariance matrix S.

Let z be the N-dimensional random vector of normal distribution with zero mean and covariance matrix I_N. Then $Z = (S^{1/2})'z$. Substituting this into equation (C17), we have

$$T[\mathrm{Dist}_T(\delta_T)]^2 \xrightarrow{d} z'Az, \tag{C18}$$

where A is defined in (C9) and is obviously symmetric and positive semi-definite.

It is easy to check that

$$I_N - (G^{-1/2})'D(D'G^{-1}D)^{-1}D'G^{-1/2}$$

is symmetric and idempotent, and that its trace is $N - K$. Thus, we know that its rank is $N - K$, which implies that the rank of A is also $N - K$. It follows that A has exactly $N - K$ positive eigenvalues, denoted by $\lambda_1, \ldots, \lambda_{N-K}$. Then there is an orthogonal matrix H and a diagonal matrix Λ such that

$$\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_{N-K}, 0_K'\} \tag{C19}$$

and $A = H'\Lambda H$. Let $x = Hz$, then $x \sim N(0_N, I_N)$. Then we have

$$T[\mathrm{Dist}_T(\delta_T)]^2 \xrightarrow{d} x\Lambda x = \sum_{j=1}^{N-K} \lambda_j x_j^2. \tag{C20}$$

Letting $v_j = x_j^2$ completes the proof.

REFERENCES

Amemiya, Takeshi, 1985, *Advanced Econometrics* (Harvard University Press, Cambridge, Mass.).

Amihud, Yakov, Bent Jesper Christensen, and Haim Mendelson, 1992, Further evidence in the risk-return relationship, Working paper No. S-93-11, New York University.

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

Bansal, Ravi, and S. Viswanathan, 1993, No arbitrage and arbitrage pricing: A new approach, *Journal of Finance* 48, 1231–1262.

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

Basu, Sanjoy, 1983, The relationship between earnings yield, market value, and return for NYSE common stocks: Further evidence, *Journal of Financial Economics* 12, 51–74.

Berk, Jonathan B., 1995, A critique of size-related anomalies, *Review of Financial Studies* 8, 275–286.

Bernanke, Ben S., 1990, On the predictive power of interest rates and interest rate spreads, *New England Economic Review* (Federal Reserve Bank of Boston, Nov./Dec.), 51–68.

Bhandari, Laxmi Chand, 1988, Debt/equity ratio and expected common stock returns: Empirical evidence, *Journal of Finance* 43, 507–528.

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

Black, Fischer, 1993, Beta and return, *Journal of Portfolio Management* 20, 8–18.

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

Bodurtha, James N., Jr., and Nelson C. Mark, 1991, Testing the CAPM with time-varying risks and returns, *Journal of Finance* 46, 1485–1505.

Bollerslev, Tim, Robert F. Engle, and Jeffrey M. Wooldridge, 1988, A capital asset pricing model with time varying covariances, *Journal of Political Economy* 96, 116–131.

Bossaerts, Peter, 1994, Time series analysis of inefficient financial markets, Manuscript, California Institute of Technology.

Breen, William J., Larry R. Glosten, and Ravi Jagannathan, 1989, Economic significance of predictable variations in stock index returns, *Journal of Finance* 44, 1177–1190.

Breen, William J., and Robert A. Korajczyk, 1994, On selection biases in book-to-market based tests of asset pricing models, Working paper No. 167, Northwestern University.

Campbell, John Y., 1993a, Intertemporal asset pricing without consumption data, *American Economic Review* 83, 487–512.

Campbell, John Y., 1993b, Understanding risk and returns, Working paper No. 4554, NBER.

Carhart, Mark M., Robert J. Krail, Ross L. Stevens, and Kelly D. Welch, 1995, Testing the conditional CAPM, Manuscript, University of Chicago.

Chan, K. C., and Nai-fu Chen, 1988, An unconditional asset-pricing test and the role of firm size as an instrumental variable for risk, *Journal of Finance* 43, 309–325.

Chan, K. C., Nai-fu Chen, and David A. Hsieh, 1985, An exploratory investigation of the firm size effect, *Journal of Financial Economics* 14, 451–471.

Chan, Louis K. C., Narasimhan Jegadeesh, and Josef Lakonishok, 1995, Evaluating the performance of value versus glamour stocks: The impact of selection bias, *Journal of Financial Economics* 38, 269–296.

Chan, Louis K. C., and Josef Lakonishok, 1993, Are the reports of beta's death premature? *Journal of Portfolio Management* 19, 51–62.

Chari, V. V., Ravi Jagannathan, and Aharon R. Ofer, 1986, Seasonalities in security returns: The case of earnings announcements, *Journal Financial Economics* 21, 101–122.

Chen, Nai-Fu, 1991, Financial investment opportunities and the macroeconomy, *Journal of Finance* 46, 529–554.

Chen, Nai-Fu, Richard Roll, and Stephen A. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383–404.

Cochrane, John H., 1992, A cross-sectional test of a production based asset pricing model, Working paper No. 4025, NBER.

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

Connor, Gregory, and Robert A. Korajczyk, 1988b, The arbitrage pricing theory and multi-factor models of asset returns, Working paper, Kellogg Graduate School of Management, Northwestern University.

Daniel, Kent, and Sheridan Titman, 1995, Evidence on the characteristics of cross sectional variation in stock returns, Manuscript, University of Chicago.

Davis, James L., 1994, The cross-section of realized stock returns: The pre-COMPUSTAT evidence, *Journal of Finance* 49, 1579–1593.

Diaz-Gimenez, Javier, Edward C. Prescott, Terry Fitzgerald, and Fernando Alvarez, 1992, Banking in computable general equilibrium economics, *Journal of Economic Dynamics and Control* 16, 533–559.

Dybvig, P. H., and J. E. Ingersoll, Jr., 1982, Mean-variance theory in capital markets, *Journal of Business* 55, 233–251.

Dybvig, P. H., and Stephen A. Ross, 1985, Differential information and performance measurement using a security market line, *Journal of Finance* 40, 383–400.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and the expected returns on bonds and stocks, *Journal of Financial Economics* 25, 23–50.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–466.

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

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

Fama, Eugene F., and G. William Schwert, 1977, Human capital and capital market equilibrium, *Journal of Financial Economics* 4, 115–146.

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

Ferson, Wayne E., and Campbell R. Harvey, 1993, The risk and predictability of international equity returns, *Review of Financial Studies* 6(3), 527–566.

Ferson, Wayne E., and S. R. Foerster, 1994, Finite sample properties of the generalized method of moments in tests of conditional asset pricing models, *Journal of Financial Economics* 36, 29–35.

Ferson, Wayne E., and Robert A. Korajczyk, 1993, Do arbitrage pricing models explain the predictability of stock returns? Working paper, Kellogg Graduate School of Management, Northwestern University.

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

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

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

Hansen, Lars Peter, and Ravi Jagannathan, 1994, Assessing specification errors in stochastic discount factor models, Technical working paper No. 153, NBER.

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

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

Harvey, Campbell R., 1989, Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24, 289–318.

Jagannathan, Ravi, and Zhenyu Wang, 1992, The cross-section of expected stock returns: Do size and book to market equity measure systematic risk better than beta? Notes for The Berkeley Program in Finance on "*Are Betas Irrelevant? Evidence and Implications for Asset Management*," September 13–15, 1992, Santa Barbara, California.

Jagannathan, Ravi, and Zhenyu Wang, 1993, The CAPM is alive and well, Staff report 165, Federal Reserve Bank of Minneapolis.

Jegadeesh, Narasimhan, 1992, Does market risk really explain the size effects? *Journal of Financial and Quantitative Analysis* 27, 337–351.

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

Kandel, Shmuel, and Robert F. Stambaugh, 1990, A mean-variance framework for tests of asset pricing models, *Review of Financial Studies* 2, 125–156.

Kandel, Shmuel, and Robert F. Stambaugh, 1995, Portfolio inefficiency and the cross-section of expected returns, *Journal of Finance* 50, 157–184.

Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting returns in the stock and bond markets, *Journal of Financial Economics* 17, 357–390.

Kothari, S. P., Jay Shanken, and Richard G. Sloan, 1995, Another look at the cross-section of expected stock returns, *Journal of Finance* 50, 185–224.

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

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

Long, J. B., 1974, Stock prices, inflation, and the term structure of interest rates, *Journal of Financial Economics* 1, 131–170.

Mandelker, G., 1974, Risk and return: The case of merging firms, *Journal of Financial Economics* 4, 303–335.

Mayers, David, 1972, Nonmarketable assets and capital market equilibrium under uncertainty, in Michael C. Jensen, Ed.: *Studies in the Theory of Capital Markets* (Praeger, New York), pp. 223–248

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

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

Reinganum, Mark R., 1981, Misspecification of capital asset pricing: Empirical anomalies based on earnings yields and market values, *Journal of Financial Economics* 9, 19–46.

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

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

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

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

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

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

Stambaugh, Robert F., 1981, Missing assets, measuring the market, and testing the capital asset pricing model, Doctoral dissertation, University of Chicago.

Stambaugh, Robert F., 1982, On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis, *Journal of Financial Economics* 10, 237–268.

Stock, James, and Mark Watson, 1989, New indexes of coincident and leading economic indicators, in Olivier J. Blanchard and Stanley Fischer, Eds.: *NBER Macroeconomics Annual 1989*.

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