A Re-examination of Some Popular Security Return Anomalies

Michael J. Brennan
Phone: (310) 825-3587, email: michael.brennan@anderson.ucla.edu
The Anderson School, UCLA; and London Business School

Tarun Chordia
Phone: (615) 322-3644; email: chordiat@ctrvax.vanderbilt.edu
Owen Graduate School of Management, Vanderbilt University

Avanidhar Subrahmanyam
Phone: (310) 825-5355; email: subra@anderson.ucla.edu
The Anderson School, UCLA

September 22, 1997

We are especially grateful to an anonymous referee for insightful and constructive suggestions. We also thank Eugene Fama, Ken French, Will Goetzmann, Craig Holden, Bob Jennings, Bruce Lehmann, Richard Roll, and seminar participants at Yale University and Indiana University for useful comments; and Eugene Fama, Ken French, Marc Reinganum, and Hans Stoll for providing data used in this study. We are responsible for remaining errors.

Abstract

A Re-examination of Some Popular Security Return Anomalies

We re-examine the relation between stock returns, measures of risk, and a set of non-risk security characteristics, including the book-to-market ratio, firm size, the bid-ask spread, the stock price, the dividend yield, and lagged returns. Our primary objective is to determine whether these non-risk characteristics have marginal explanatory power relative to the loadings on the Connor and Korajczyk (1988) risk factors. Fama-MacBeth type regressions using risk adjusted returns on individual securities shed light on earlier anomalous findings, and reveal new relations. The widely cited book-to-market and momentum effects are somewhat attenuated once account is taken of the Connor and Korajczyk factors. Firm size is shown to have no incremental explanatory power for returns in the presence of trading volume, suggesting that firm size is a proxy for liquidity. Finally, dividend yield effects consistent with a changing tax code are detected.

Introduction

The early empirical research on the determinants of expected returns on common stocks was concerned with detecting an association between average returns on beta-sorted portfolios and their betas, as predicted by the CAPM.[1] This work was refined by the introduction of statistical tests of the null hypothesis that expected returns are determined *solely* by betas.[2] Following the development of the APT, a similar series of tests was conducted, in which proxies for the APT factors and factor loadings replaced the market portfolio and betas of the CAPM.[3] Starting with the work of Black and Scholes (1972), Basu (1977), and Banz (1981), researchers began to test these asset pricing models against specific alternatives; these alternative hypotheses posited that expected returns on securities, instead of being determined solely by the risk characteristics of the securities, as measured by betas or factor loadings, were also affected by non-risk security characteristics such as size, market-to-book ratios, dividend yields, and earnings-price ratios. While the role of some of these non-risk characteristics, such as dividends, could be rationalized in an equilibrium model, or could possibly be accounted for by their statistical properties as proxies for expected returns, the roles of other characteristics such as firm size, have remained more elusive, so that their apparent importance for expected returns leaves the empirical validity of the rational asset pricing paradigm open to question.

In an important series of papers, Fama and French (1992, 1993, 1996) (FF) have argued for the continuing validity of the rational pricing paradigm by showing that, with the exception of the

[1] For example, Black, Jensen and Scholes (1972).

[2] Gibbons (1982), Stambaugh(1982).

[3] Roll and Ross (1980), Brown and Weinstein (1983), Lehman and Modest (1988).

momentum strategy of Jegadeesh and Titman (1993), the cross-sectional variation in expected returns associated with these non-risk characteristics can be captured by only two characteristics, namely the firm's size and its book-to-market ratio;[4] and that, moreover, (FF, 1993) these firm characteristics proxy for the security's loadings on priced factors. They show that the firm size and book-to-market effects can be accounted for within a three factor model in which the factors are the returns on the market portfolio, and on two zero net investment portfolios, one of which is long in high market to book and short in low book to market securities (HML), and the other of which is long in small firms and short in large firms (SMB).

Daniel and Titman (DT) (1996) address the issue of whether the return patterns of characteristic-sorted portfolios are consistent with a factor model in which factor loadings correspond to firm characteristics such as size and book to market ratio. DT consider a generalized class of factor models in which the loading on a distress factor is allowed to evolve stochastically and to be proxied by the book-to-market ratio; they also consider the possibility that factor risk premia vary over time in such a way that distressed firms have higher risk premia. Despite these generalizations, they conclude that portfolios of firms that have similar characteristics (size and book-to-market), but different loadings on the Fama French factors, have similar returns, so that it is the security characteristics and not their loadings on the Fama-French factors that determine expected returns. Thus DT's results are in contrast to the notion that the FF results provide a risk-based explanation of expected returns.

An important aspect of much of this research is that the returns that are analyzed are the

[4]Fama and French (1992) show that firm size and the ratio of book to market equity capture the cross-sectional relation between average returns and earnings yield and leverage.

returns on portfolios that are constructed by sorting securities on some criterion of interest: the portfolios are formed either to mitigate problems caused by using *estimated* betas as independent variables in a two-step estimation procedure or, when a one-step estimation procedure is used, to allow estimation of the covariance matrix of residual returns. This causes two quite different types of problem. First, as Roll (1977) pointed out, the portfolio formation process, by concealing possibly return relevant security characteristics within portfolio averages, may make it difficult to reject the null hypothesis of no effect on security returns. Lo and MacKinlay (1990) make an almost precisely opposite point, that if the researcher forms portfolios on the basis of characteristics which prior research has shown to be related to average returns, he will be inclined to reject the null hypothesis too often due to a "data-snooping" bias.[5] The resulting problem of inference is illustrated in Fama and French (1996), and Brennan, Chordia and Subrahmanyam (1996), where results are presented for six and seven sets of portfolios respectively, and quite different results are obtained according to the different criteria used in portfolio formation. Further, the portfolio approach breaks down completely if the researcher wishes to examine simultaneously the effects of several characteristics on expected returns, because the portfolio formation procedure will tend to induce extreme multi-collinearity between portfolio characteristics.

In this paper we provide new evidence on the extent to which expected returns can be

---

[5] Table 5 of Lo and Mackinlay (1990) shows that if the R^2 between the sorting characteristic used to form portfolios and the estimated "’s is 0.005, then the probability that a standard F test will reject at the 5% level is 11.8% if 1000 securities are sorted into 10 portfolios of 100 securities, *even though the underlying data satisfy the null hypothesis*. If the R^2 is 0.01 the size of a 5% test rises to 36.7% for 1000 securities sorted into 10 portfolios of 100 securities, *even though the underlying data satisfy the null hypothesis*. If no portfolio aggregation had been performed the size of these tests would be 5%!

explained by risk factors rather than by non-risk characteristics. Our approach differs from that of Fama and French and Daniel and Titman in three principal ways. First, rather than specifying the risk factors *a priori*, we follow the intuition of the APT, that the risk factors should be those which capture the variation of returns in large well-diversified portfolios,[6] and use the principal components approach of Connor and Korajczyk (1988) to estimate risk factors. Thus, our null hypothesis is that expected returns depend only on loadings on the Connor-Korajczyk factors. Secondly, rather than limiting ourselves to the small number of firm characteristics that Fama and French have found to be associated with average returns, notably size and book-to-market ratio, we estimate simultaneously the marginal effects of nine firm characteristics, including dividend yield, and measures of market liquidity such as share price, trading volume and the bid-ask spread, for which there is theoretical warrant, as well as measures of momentum for which there is currently no theoretical rationale but strong empirical evidence. We are able to consider these several characteristics simultaneously because instead of examining the returns on portfolios, we examine the risk-adjusted returns on individual securities.[7] Under the null hypothesis, these risk-adjusted returns should be independent of other (non-risk) security characteristics. Not only does this approach allow us to consider the effects of a large number of firm characteristics simultaneously,

[6] Campbell (1996), using the intuition of Merton's intertemporal CAPM, argues that "priced factors should be found not by running a factor analysis on the covariance matrix of returns ...Instead, innovations in variables that have been found to forecast stock returns and labor income should be used." It seems likely to us that variables that have a significant effect on the future investment opportunity set are also likely to have a significant effect on contemporaneous returns, so that their traces will be evident in the covariance matrix of returns.

[7] Papers that use risk-unadjusted returns for cross-sectional analyses on individual securities include Fama and French (1992), Litzenberger and Ramaswamy (1979), Miller and Scholes (1982), and Lehmann (1990).

but it also avoids the data-snooping biases that are inherent in the portfolio based approaches as discussed above.[8] Our approach also avoids the errors-in-variables bias created by errors in estimating factor loadings, since errors in the factor loadings are impounded in the dependent variable.

We find that after risk-adjustment using the Connor-Korajczyk (1988) factors, mean returns remain significantly related to only two of the several firm characteristics we consider, namely, trading volume and dividend yield, as well as lagged returns. When the analysis is repeated using the Fama-French portfolios as factors, the stock price becomes significant and the dividend yield becomes insignificant (although the point estimate of the latter coefficient increase by over 60%), and the size and significance of the coefficients on the lagged return variables increases. We find that firm size continues to be negatively associated with the risk adjusted returns on individual securities, whether the risk-adjustment is made with the C-K factors or with the Fama-French portfolios, unless the dollar volume of trading is included as an independent variable. However, the negative size effect disappears when the dollar volume of trading is included in the regression, suggesting that the oft-cited size effect is really a trading volume effect; this is consistent with trading volume acting as a proxy for the liquidity of the market in the firm's shares.[9]

We also find that the C-K factors have squared Sharpe ratios that fall within the range suggested as reasonable by the analysis of MacKinlay (1995), while the FF factors appear to offer a much higher reward for risk ratio. These findings are confirmed when we regress the FF factors

[8] Of course, we are guilty of data-snooping in a different sense: The security characteristics we have chosen to consider are motivated by previous results. But we do avoid the aggravation of the problem caused by sorting to form portfolios.

[9] Glosten and Harris (1988) and Brennan and Subrahmanyam (1995) have shown that trading volume is a major determinant of market liquidity.

on the C-K factors; we are able to reject the null hypothesis that the intercepts from these regressions are jointly zero. This is consistent with the observation that the C-K factors were selected on the basis of their ability to describe the variance-covariance matrix, while the FF factors were selected based on their ability to explain the cross-sectional structure of expected equity returns.

The remainder of the paper is organized as follows. In Section I we describe the empirical hypotheses we test. In Section II the data are described and in Section III the empirical results are presented. Section IV compares the FF and C-K factors, and Section V concludes.

I

Hypotheses

Our null hypothesis is a 5-factor version of the APT which implies that the expected excess return on security j is determined solely by the loadings of the security's return on the five factors, β_{jk} (k = 1,..,5). Consider the following equation:

$$E[\tilde{R}_j] - R_F = c_0 + \sum_{k=1}^{5} \lambda_k \beta_{jk} + \sum_{m=1}^{M} c_m Z_{mj} \qquad \textbf{(1)}$$

where R_j is the return on security j, R_F is the risk free interest rate, β_{jk} is the loading of security j on factor k, λ_k is the risk premium associated with factor k, Z_{mj} (m = 1,..,M) is the value of (non-risk) characteristic m for security j, and c_m is the premium per unit of characteristic m. Our null hypothesis is that $c_m = 0$ (m = 0, 1,..,M). We include 9 security characteristics (including three momentum-based lagged return variables) as possible determinants of expected returns.

The five risk factors are taken to be the first five (asymptotic) principal components of excess stock returns estimated over the sample period. In deciding which firm characteristics to include as possible determinants of expected returns, attention was given to those variables that had been found to be important in prior studies, as well as those for which there exists a theoretical rationale. Thus firm size is included because of the widespread evidence of a "small firm effect."[10] While Fama and French (1993) implicitly treat firm size as a loading on a risk factor, Daniel and Titman (1996) suggest that the firm size itself is a stronger determinant of expected returns than is the loading on the FF "size factor." Ball (1978), suggests that firm size may appear as a determinant of expected returns because of imperfect risk adjustment in the empirical analysis (this point is formalized in Berk (1995)). It is therefore important to assess whether size has any residual explanatory power for expected returns once account is taken of the five risk factors[11] and other firm characteristics. We also include the ratio of book-to-market equity because this has been found to be strongly associated with average returns.[12] It has been hypothesized that the low price effect documented by Miller and Scholes (1982) reflects the fact that firms with low prices are often in financial distress, and that financial institutions may be reluctant to invest in them on account of the prudent man rule.[13] Therefore we include the reciprocal of share price as a possible determinant of expected returns.

The bid-ask spread is included because this variable is associated with liquidity, and the work

---

[10] Banz (1981) and Fama and French (1992).

[11] Lehmann and Modest (1988) found that their implementation of a five-factor APT was unable to account for the size anomaly.

[12] See Fama and French (1992) and Lakonishok, Shleifer and Vishny (1994).

[13] Falkenstein (1996) shows that mutual funds "show an aversion to low-price stocks."

of Amihud and Mendelson (1986), and Brennan and Subrahmanyam (1996), suggests that expected returns are affected by liquidity. Over our sample period bid-ask spread data are available only annually.[14] However, a number of authors (e.g., Stoll (1978)) find trading volume to be the most important determinant of the bid-ask spread, and since this variable is available on a monthly basis, we also include dollar trading volume as a separate variable in our analysis.[15] We include dividend yield because Brennan (1970) suggests that differential taxation of dividends and capital gains could make this variable relevant, and the resulting empirical work of Litzenberger and Ramaswamy (1979) and Miller and Scholes (1978, 1982) has been inconclusive. Finally, we include lagged return variables because Jegadeesh and Titman (1993) have shown these to be relevant, and by including them we should improve the efficiency of the estimates of the coefficients of the other variables.

II

Data

The basic data consist of monthly returns and other characteristics for a sample of NYSE securities for the period January 1966 to December 1989. The sample was restricted to NYSE securities in order to ensure availability of data on the bid-ask spread. This requirement also determined the sample period. To be included in the sample for a given month a security had to satisfy the following criteria: (1) Its return in the current month and in 24 of the previous 60 months be available from CRSP, and sufficient data be available to calculate the size, price, dollar volume,

[14] We thank Hans Stoll and Marc Reinganum for providing us with the bid-ask spread data.

[15] Petersen and Fialkowski (1994) find that the quoted spread is only loosely associated with the effective spread; therefore it is possible that trading volume provides a better measure of liquidity than does the quoted spread.

and dividend yield as of the previous month; (2) Sufficient data be available on the COMPUSTAT tapes to calculate the book to market ratio as of December of the previous year; (3) The average bid-ask spread be available for the previous year. This screening process yielded an average of 911 stocks per month. In comparison, Fama and French (1992) had an average annual sample size of 2267 stocks. The Fama and French sample, however, also included AMEX and NASDAQ stocks.

For each security the following variables were calculated each month as follows:

SIZE - the natural logarithm of the market value of the equity of the firm as of the end of the previous month.

BM - the natural logarithm of the ratio of the book value of equity plus deferred taxes to the market value of equity, using the the end of the previous calendar year market value and the most recent book value available at the end of the previous calendar year.[16] Following Fama and French (1992), book-to-market ratio values greater than the 0.995 fractile or less than the 0.005 fractile are set to equal the 0.995 and 0.005 fractile values, respectively.

SPREAD - the natural logarithm of the average bid-ask spread as a proportion of the closing stock price for the previous year (calculated as the average of beginning and the end of year closing bid-ask spread relative to the mean quote).

DVOL - the natural logarithm of the dollar volume of trading in the security in the previous month.

PRICE - the natural logarithm of the reciprocal of the share price as reported at the end of the previous month.

YLD - the dividend yield as measured by the sum of all dividends paid over the previous 12 months,

[16] It was found that the results were unchanged when this variable was lagged by one year.

divided by the share price at the end of the previous month.

RET2-3 - the natural logarithm of the cumulative return over the two months ending at the beginning of the previous month.

RET4-6 - the natural logarithm of the cumulative return over the three months ending three months previously.

RET7-12 - the natural logarithm of the cumulative return over the 6 months ending 6 months previously.

The lagged return variables were constructed to exclude the return during the immediate prior month in order to avoid any spurious association between the prior month return and the current month return caused by thin trading or bid-ask spread effects.

Table 1 reports the grand time-series and cross-sectional means, medians, and standard deviations of the raw (i.e., unlogged) security characteristics, and displays the summary statistics associated with both trimmed and untrimmed values of the book to market ratio. Note that the variables display considerable skewness. It is for this reason that in our empirical analysis we employ logarithmic transforms of all these variables except the dividend yield (which may be zero). Finally, for all of the regressions reported below, the transformed firm characteristics variables for a given month were expressed as deviations from their cross-sectional means for that month; this implies that the expected return for a security with average values of these characteristics will be determined solely by its factor loadings and the factor risk premia. Table 2 reports the averages of the month by month cross-sectional correlations of the variables that we use in our analysis. The largest correlations with SIZE are DVOL (positive), SPREAD (negative), and PRICE (negative); with DVOL, they are SPREAD (negative), and PRICE (negative); with SPREAD, PRICE (negative). The other correlations are smaller than 0.4 in absolute value.

Five factors were estimated by the asymptotic principal components technique developed by Connor and Korajczyk (1988)[17] (henceforth C-K factors) applied to returns in excess of the risk-free rate on all securities listed continuously over the estimation period. In order to keep the estimation process computationally manageable, the factors were estimated separately over each of two subperiods: January 1963 to December 1976 and January 1977 to December 1989. The risk free interest rate was taken as the 1 month risk free rate from the CRSP bond files.

III

Statistical Model

As we have argued above, empirical findings based on the returns on portfolios are hard to interpret. Therefore, we report the results from analyzing the returns on *individual securities.*The null hypothesis against which we evaluate the influence of the non-risk security characteristics is the five-factor APT. Thus, assume that returns are generated by a five-factor approximate factor model:[18]

$$\tilde{R}_{jt} = E[\tilde{R}_{jt}] + \sum_{k=1}^{k=5} \beta_{jk} \tilde{f}_{kt} + \tilde{e}_{jt} \tag{2}$$

where $\tilde{f}_{kt}$ are mean zero and $E[\tilde{e}_{jt} \mid \tilde{f}_{kt}] = 0$. Then the exact or equilibrium version of the APT

[17] Connor and Korajczyk (1993) "find evidence for one to six pervasive factors generating returns on the NYSE and AMEX over the period 1967 to 1991."

[18] See Connor and Korajczyk (1988) for example, for the definition of an approximate factor model.

implies that expected returns may be written as :

$$E\,[\tilde{R}_{jt}] - R_{Ft} = \sum_{k=1}^{k=5} \lambda_{kt}\, \beta_{jk} \tag{3}$$

where R_{Ft} is the return on the riskless asset, and 8_{kt} is the risk premium for factor k. Substituting from (3) in (2), the APT implies that realized returns are given by:

$$\tilde{R}_{jt} - R_{Ft} = \sum_{k=1}^{5} \beta_{jk}\, \tilde{F}_{kt} + \tilde{e}_{jt} \tag{4}$$

where $\tilde{F}_{kt} \equiv \lambda_{kt} + \tilde{f}_{kt}$ is the factor plus its associated risk premium.

Our goal is to test whether security characteristics have incremental explanatory power for returns relative to the Connor-Korajczyk (C-K) factors.

A standard application of the Fama-MacBeth (1973) procedure would involve estimation of the following equation:

$$\tilde{R}_{jt} - R_{Ft} = c_0 + \sum_{k=1}^{5} \beta_{jk}\, \tilde{F}_{kt} + \sum_{m=1}^{M} c_m Z_{mjt} + \tilde{e}_{jt} \tag{5}$$

where Z_{mjt} is the value of characteristic m for security j in month t. Under the null hypothesis that expected returns depend only on the risk characteristics of the returns, as represented by $\$_{jk,}$ the loadings on the C-K factors, the coefficients c_m (m = 1,..,M) will be equal to zero. This hypothesis can be tested in principle by estimating the factor loadings for each month using prior data, estimating a cross-section regression for each month in which the independent variables are the factor loadings and non-risk characterstics, and then averaging the monthly coefficient estimates over time and calculating their time-series standard errors. This simple Fama-MacBeth approach, however, presents problems because the factor loadings are measured with error. One approach to dealing with this problem is to use the information from the first stage regressions (in which the factor loadings are estimated) to correct the coefficient estimates in the second stage regressions.[19] We employ an approach which does not rely on information taken from the first stage regressions to correct the bias.

First, each year, from 1966 to 1989, factor loadings, β_{jk}, were estimated for all securities that had at least 24 return observations over the prior 60 months using the Dimson (1979) procedure with one lag to allow for thin trading. The estimated risk-adjusted return on each of the securities, R^*_{jt}, for each month t of the following year was then calculated as:

$$\tilde{R}^*_{jt} \equiv \tilde{R}_{jt} - R_{Ft} - \sum_{k=1}^{5} \hat{\beta}_{jk} \tilde{F}_{kt} \qquad \textbf{(6)}$$

These risk-adjusted returns constitute the raw material for the estimates that we present below of the equation:

---

[19] This is the approach followed by Litzenberger and Ramaswamy (1979) and Lehmann (1990).

$$\tilde{R}^*_{jt} = c_0 + \sum_{m=1}^{M} c_m Z_{mjt} + \tilde{e}_{jt} \tag{7}$$

Each month an estimate of the vector of characteristic rewards $\underline{\hat{c}}_t$, is calculated from:

$$\underline{\hat{c}}_t = (\underline{Z}'_t \underline{\Omega}_t^{-1} \underline{Z}_t)^{-1} \underline{Z}'_t \underline{\Omega}_t^{-1} \underline{R}^*_t \tag{8}$$

where $\underline{Z}_t$ is the vector of firm characteristics in month t, $\underline{R}^*_t$ is the vector of estimated risk-adjusted returns, and $\underline{\Omega}_t$ is a weighting matrix. Note that although the factor loadings, β_{jk}, are estimated with error, this error affects only the dependent variable, $\underline{R}^*_t$, in the generalized least squares regression estimator (8), and while the factor loadings will be correlated with the security characteristics, $\underline{Z}_t$, there is no *a priori* reason to believe that *errors* in the estimated loadings will be correlated with the security characteristics, so the estimated coefficient vector, $\underline{\hat{c}}_t$, should be unbiased.

For each characteristic, m (m = 0,1,..,M) (including the constant term) the coefficient estimates, for each month from January 1966 to December 1989, are aggregated into an overall estimate in one of two ways. The first, which we call the *raw estimate*, $\hat{c}_{mr}$, is given by:

$$\hat{c}_{mr} = (\underline{i}' \underline{V}_m^{-1} \underline{i})^{-1} \underline{i}' \underline{V}_m^{-1} \underline{\hat{c}}_m \tag{9}$$

where $\underline{i}$ is the unit vector, $\underline{\hat{c}}_m$ is the vector of monthly estimates of c_m, and $\underline{V}_m$ is a weighting matrix. While there is no *a priori* reason to believe that the errors in the estimated factor loadings will be correlated with the security characteristics, $\underline{Z}_t$, to the extent that they are correlated, the estimated coefficients of the firm characteristics, $\hat{c}_{mt}$, will be correlated with the factor realizations. Therefore, as a check on the robustness of our results, the monthly regression coefficients on each of the characteristics were regressed in time series on the C-K factor realizations to yield what we call the *purged estimate*, $\hat{c}_{mp}$, which allows for the possibility that the estimation errors in the monthly estimates depend on the factor realizations. This estimator, which purges the monthly estimates of the factor dependent component, is given by:

$$\hat{c}_{mp} = \underline{e}'(\underline{F}^{*\prime}\underline{V}_m^{-1}\underline{F}^{*})^{-1}\underline{F}^{*\prime}\underline{V}_m^{-1}\underline{\hat{c}}_m \tag{10}$$

where $\underline{e}$ is a 6-element vector (1 0 0 0 0 0)' which serves to pick out the constant of the regression, and $\underline{F}^*$ is the matrix of factor portfolio returns augmented by a vector of units.

Initially, we shall follow standard practice and assume that the matrices $\underline{\Omega}_t$ and $\underline{V}_m$ are proportional to the identity matrix. This yields the standard Fama-MacBeth estimator for the raw estimator, and the constant term from the OLS regression of the month-by-month Fama-MacBeth estimates on the factor portfolio returns for the purged estimator.[20] The standard error of the estimate is taken from the time series of monthly estimates in the case of the raw estimator, and from the standard error of the constant from the OLS regression in the case of the purged estimator.

To begin our analysis we present the results of simple Fama-MacBeth regressions of the returns on the security characteristics SIZE, BM, PRICE, SPREAD, YLD, and the lagged returns

[20] Separate estimates are calculated corresponding to the two subperiods for which the principal components were estimated; these were then aggregated using precision weights.

without risk adjustment of the returns. These results are reported in the first column of Table 3. As can be seen, the coefficients of SIZE and BM are respectively negative and positive, and both are statistically significant, which is consistent with earlier studies such as Fama and French (1992). In addition, the coefficient of the inverse price level variable is positive and significant, while that of the the spread variable is negative and significant. There are also strong positive lagged return effects.

Having shown that average security returns are significantly associated with firm characteristics, we now consider whether this relation is maintained for *risk-adjusted* returns.The raw and purged estimates of the characteristic rewards, $\hat{\mathbf{c}}_{mr}$ and $\hat{\mathbf{c}}_{mp}$, for risk adjusted returns are reported in the second and third columns of Table 3. SIZE remains significantly negatively associated with returns while BM becomes insignificant. The lagged return variables and PRICE and SPREAD remain significant, while the dividend yield variable, YLD, now has a strongly significant positive effect. One effect of using risk-adjusted returns rather than raw excess returns in the cross-section regressions (7) is to reduce the correlation between the errors in the monthly cross-section regressions. This increases the efficiency of the monthly estimates of the coefficients. As a result, the standard error of the coefficient of YLD, for example, is halved when the risk-adjusted returns are used in the cross-section regressions. The factor model performs well in pricing securities with *average* characteristics in that, as predicted by the APT, the intercept of the regression is small[21] and insignificant. The purged estimates shown in the third column of Table 3 are close to the raw estimates, as anticipated.

[21] *Minus* 1.2 basis points and 0.2 basis points per month for the raw and purged estimates respectively .

The significant coefficient on SPREAD in Table 3 is puzzling at first sight because its negative sign is inconsistent with a liquidity effect. Note however, that since SPREAD is measured as the logarithm of the average of the bid-ask spreads at the beginning and end of the previous year, this variable is approximately proportional to the logarithm of the reciprocal of the average price level during the previous year. Furthermore, since PRICE is the logarithm of the reciprocal of the price at the beginning of the current month, the difference between these variables is a noisy estimator of the return over the months up to the beginning of the *current* month. But the lagged return variables measure returns up to the beginning of the *previous* month. Therefore, a linear combination of SPREAD, PRICE, and the lagged return variables provides a noisy estimate of the return in the *previous* month. We conjecture therefore that SPREAD acquires its significance because of its association with the the stock return over the previous month, and the well-known tendency towards negative first order correlation in stock returns created by bid-ask effects and thin trading.[22]

To eliminate this possible source of bias the analysis was repeated lagging all variables which involve the beginning of month price by one month.[23] The raw and purged results are reported in the first two columns of Table 4. As conjectured, SPREAD and PRICE become insignificant, and the coefficients of the lagged return variables become smaller although they remain significant; the other coefficients are not qualitatively affected. Therefore in the remainder of this study SIZE, PRICE and YLD are measured with a one month lag.

The first two columns of Table 4 show that the only firm characteristics to have a reliable

[22] See Jegadeesh (1990). It is easy to show that thin trading will cause *risk-adjusted* returns to exhibit first order negative serial correlation.

[23] That is, SIZE, PRICE, and YLD.

association with risk-adjusted returns are YLD, SIZE, RET2-3, and RET4-6. It is possible that SIZE may acquire its significance as a proxy for liquidity, since our SPREAD variable is at best a noisy measure of liquidity, because it is available only on an annual basis, and in any case has only a low correlation with other measures of liquidity.[24] Therefore the logarithm of average daily trading volume (DVOL) was included as an additional measure of liquidity. The results are presented in the second two columns of Table 4. Consistent with its role as a measure of liquidity, the coefficient of DVOL is negative and highly significant; moreover, the coefficient of SIZE is now *positive* and significant. This suggests that the negative relation between firm size and returns that was first noted by Banz (1981) is in fact due to the association between size and trading volume, and that the latter is important because of its role as a determinant of liquidity.[25] We shall explore below whether the *positive* coefficient on SIZE when DVOL is introduced can be attributed to an incorrect specification of the return volume relation.

To this point we have followed Fama and MacBeth in using weighting matrices, $\underline{\Omega}_t$ and $\underline{V}_m$, that are proportional to the identity matrix. In the last four columns of the table, weighted least squares estimates are presented. The weighting matrix for the month t estimator, $\underline{\Omega}_t$, is taken as a diagonal matrix[26] whose typical element is $\sigma^2\{\tilde{e}_{jt}\}$, the residual variance from the multiple

---

[24] Brennan and Subrahmanyam (1996) report that the proportional spread has a correlation of only 0.38 with their (Kyle-lambda based) estimate of the average cost of transacting; while Petersen and Fialkowski (1994) report a correlation of only 0.10 between the quoted spread and the effective spread that investors pay.

[25] Stoll (1978) and Brennan and Subrahmanyam (1995) show that trading volume is the single most important determinant of the bid-ask spread and the price-level-adjusted Kyle lambda, respectively.

[26] Recall that the dependent variable is the risk-adjusted return which eliminates covariance arising from the dependence of the raw returns on the common risk factors.

regression used to estimate the factor loadings on which the estimated risk-adjusted return for security j in month t is based. The matrix $\underline{\mathbf{V}}_m$ used to aggregate the monthly estimators is taken as a diagonal matrix whose typical element is $\sigma^2\{\hat{c}_{mt}\}$, the estimated variance of the monthly estimator from (8). The standard errors of the aggregated estimates are calculated as the standard error from the WLS regression of the monthly estimates on a vector of units (the raw estimator), and the factor realizations (the purged estimator).[27]

Columns (5) and (6) of Table 4 present the weighted least squares parameter estimates. Besides reducing the estimated standard errors of the coefficients, the main effect of the WLS procedure is to reduce the magnitude of the coefficient of the dividend yield variable by about 75%; however, it remains statistically significant. Columns (7) and (8) of the table report the result of including a squared term in DVOL in the regression. Consistent with the notion that successive increases in volume have a diminishing effect on required returns, the coefficient on this variable is positive, although it is only marginally significant (p-value of 0.07); however, inclusion of $DVOL^2$ reduces the t-statistic on SIZE to 1.82 with a p-value of 0.07. The fact that the sign of coefficient of the SIZE variable is changed by introducing DVOL, and that its significance becomes only marginal when the squared DVOL term is introduced, while DVOL remains highly significant, strongly suggests that it is trading volume and not firm size *per se* which affects expected returns, and that the remaining role of SIZE is probably due to mis-specification of the functional relation between returns and trading volume. The coefficient of the dividend yield variable remains positive and significant, which is consistent with tax effects. The coefficient implies that a 1% dividend yield

[27] This is more conservative than assuming that the only source of variation in the monthly estimates is the measurement error which is taken into account by the weighting matrix.

would be associated with an increase in expected return of 1.06 basis points per month, or approximately 12.7 basis points per year. Although the lagged return variables remain statistically significant, their economic significance is marginal. They imply that a 10% return in one month will be followed by a further 0.5% abnormal return over the following 12 months. In sum, we find, after accounting for the Connor-Korajczyk factors, that apart from this modest evidence of slow adjustment, the only non-risk firm characteristics that have a significant effect on expected returns are trading volume and dividend yield. These findings are consistent with a liquidity effect of the type described by Amihud and Mendelson (1986), and a tax effect of the type described by Brennan (1970).[28]

Figure 1 plots risk-adjusted cumulative excess returns to portfolios that have a unit exposure to one of the characteristics we have considered and zero exposure to the others, using the Connor-Korajczyk factors. For characteristic m (m = 1,..,9) the figures plot C_{mt} where:

$$C_{mt} = t\,\gamma_{m0} + \sum_{\tau=1}^{t} u_{m\tau} \tag{11}$$

where $(_{m0}$ is the intercept from the WLS regression of the monthly estimates on the factors and u_{mJ} is the error term. The plots offer a visual representation of the behavior of the monthly estimates of the rewards to the non-risk characteristics. The cumulative reward for firm size is fairly flat except for a period in the early 1970's where it rises steeply. The cumulative reward for book to market ratio fluctuates without any apparent trend. The reward for PRICE is generally positive in the first half

[28]Excluding data for the months of January from the sample did not materially affect the results in Table 4. These results are not reported for brevity, but are available upon request.

of the sample period and negative in the second. The reward to dollar volume is calculated for a security whose deviation from the mean (log) dollar volume is one standard deviation:[29] the steady downward trend in the cumulative reward reflects the significant negative reward to volume. In contrast, the cumulative reward to SPREAD is highly irregular. Interestingly, the cumulative reward for dividend yield, increasing in the early years, appears to level off in the early 80s. The levelling off coincides closely with the tax measures which generally reduced the tax penalty on dividends.[30] This supports the notion that the dividend yield effect is a tax effect along the lines suggested by Brennan (1970). The graphs for the lagged returns appear more like wandering series.

For comparison purposes, we repeat the analysis using the factors proposed by Fama and French to risk adjust the returns; the results appear in Table 5.[31] The results are qualitatively similar, with three significant exceptions: first, the lagged return effects are about 50% larger and more significant,[32] PRICE is significant, and although the point estimate of the YLD coefficient about 50% higher than when the C-K factors are used, this coefficient is not statistically significant. However, the parameter estimates appear to be less efficient when the Fama-French factors are

[29] It is necessary to specify the magnitude of the dollar volume whose reward is calculated because of the non-linear reward to DVOL.

[30] The maximum personal tax rate was reduced from 70% to 50% in 1983, to 33% in 1987/8. The capital gains tax exclusion changed from 60% to zero in 1987.

[31] The Fama and French factors are available only from July 1963, so the estimation period for the factor loading in the first month of the regression period (January 1966) consisted of 30 observations, the next month, 31, and so on till the 60 month level was reached from which point it was kept constant at 60 months.

[32] Whereas the C-K factor results implied that a 10% return in one month would be followed by an abnormal return of 0.5% over the next 12 months, the FF results imply that the same 10% return would be followed by an abnormal return of 1.44%. When there is no risk adjustment, the corresponding figure is 1.28%.

used.[33]

Table 4 suggests that the C-K factors appear to capture part of the momentum effect that has been previously documented by Jegadeesh and Titman (1993) and Fama and French (1996). Jegadeesh and Titman argue, within the context of *single* factor model of security returns, that the momentum effect cannot be due to time-varying risk premia on the factor. To investigate this phenomenon empirically within the context of our multi-factor model, we regress the five CK factors on their lagged 3, 6, and 12 month wealth relatives,[34] separately for each of the two estimation periods. The results are presented in Table 6. We find evidence of both positive and negative autocorrelation, particularly in the first subperiod. Focusing on the significant results in the first subperiod (Panel A), Factor 3 is negatively correlated with its 3-month laggged return, while Factor 4 is positively correlated with its 12 month lagged return, and negatively correlated with its 6 month lagged return. The last row of Panel A reports the F-statistic for the null hypothesis that cross-lags of other factors have no joint explanatory power in the relevant autoregression. Adding cross-lags appears to enhance the predictive power of only the Factor 1 autoregression. The full results using cross-lags for this factor are not reported in the interest of brevity. However, we found that the 6th lag of Factor 2 and the 12th lag of Factor 3 were significant in the regression of Factor 1 on its own and cross-laggged returns. Panel B of Table 8 reveals that there is only weak evidence of autocorrelation in the second subperiod. Overall, our analysis suggests that a significant part of the previously documented momentum effect is due, not to delayed price reactions to firm specific

---

[33] This reflects the fact that they represent the covariance matrix less well than do the C-K factors.See Daniel and Titman (1996).

[34] The lagged wealth relatives are defined in a manner consistent with RET2-3 etc. to end *one month prior* to the current month.

information, but to autocorrelation in the factor excess returns which seem to have been especially important in the first half of the sample period.

IV

Factor Risk Premia

Lakonishok et al. (1994) argue that the relative performances over a sample period that is similar to ours[35] of "value" and "glamour" stocks, as classified for example by their book-to-market ratios, "pose a stiff challenge to any risk-based explanation", and claim that if there is to be a risk based explanation of their findings then the risk premium on the factor that covaries with the difference in their returns "should also be quite high". Mackinlay (1995) makes a similar point, more generally; that risk-based explanations of asset pricing anomalies are bounded by the plausibility of the (squared) Sharpe ratio of the tangency portfolio that they imply. Mackinlay argues that a plausible upper bound on the squared Sharpe ratio for the tangency portfolio using monthly returns is 0.031, assuming no imperfections.

Table 7 reports the mean excess returns on the five C-K factor portfolios for each estimation subperiod. Since we have no assurance that the factor rotations are the same for the two subperiods, they should be treated separately. However, we include a P^2 test of the joint hypothesis that the mean excess returns of each factor are zero for both subperiods. Only for factor 1 can we reject the null. Using standard t-tests for the subperiods, the mean excess return on factor 3 is significant and negative in the first subperiod but insignificantly different from zero in the second. We also report the mean excess return for each factor portfolio, obtained simply by combining the series for the two subperiods. Only factor 1 is significant. The table also reports the squared Sharpe ratio(SSR) for each

[35] April 1963 to April 1990.

factor for each subperiod, as well as an aggregate SSR for the factor portfolios, which is the SSR of the tangency portfolio formed from the factor portfolios.[36] The largest SSR's are 0.0108 and 0.0186 for factor 1 in subperiods 1 and 2 respectively, and 0.0254 for factor 3 in subperiod 1. These all fall comfortably within Mackinlay's acceptable range. However, the estimated SSR for the tangency portfolio is approximately 0.04 for *both* subperiods, well above the plausible range estimated by Mackinlay. On the other hand, the p-values for these values of the SSR under the null hypothesis that the factor risk premia are jointly *zero* are 0.31 and 0.24.[37] Thus, in accounting for the return anomalies with the C-K factors, we have not had to rely on factors with implausibly high risk premia - indeed we cannot even reject the null hypothesis that the five portfolio factor risk premia are jointly zero.

As shown in Table 8, the aggregate SSR for the FF factors[38] is 0.056 with a p-value of 0.00 under the null hypothesis. The main contributor to the aggregate SSR of the FF portfolios is the HML portfolio; this portfolio alone has a SSR of 0.0381, which is four times that of the market portfolio. This is not surprising because the size and market to book ratio factors were chosen on the basis of these characteristics' role in explaining the cross-sectional structure of equity returns.

Table 9 reports the results of regressing the three FF factors on the C-K factors. The FF market factor is strongly associated with the first C-K factor in both subperiods, and overall the C-K

---

[36] Since the portfolio returns are orthogonal the SSR of the tangency portfolio is simply the sum of the individual portfolio SSR's. See Mackinlay (1995) and references therein.

[37] Under the null hypothesis [(T-N)/N]SSR is distributed central F(N, T-N) where N is the number of portfolios and T is the number of time-series observations. See Mackinlay (1995).

[38] This is given by $\underline{\mu}' \underline{S}^{-1} \underline{\mu}$ where $\underline{\mu}$ is the vector of mean excess returns on the three factor portfolios and $\underline{S}$ is the variance covariance matrix.

factors explain 96-97% of the variance. However, the market factor alpha is negative and strongly significant. This is in contrast to the results reported in Tables 3 and 4 which imply that the C-K factors do a good job of capturing the returns on securities with average characteristics. The FF market portfolio has characteristics which differ from those of the average security - in particular, since it is a value weighted index, it is heavily weighted towards securities of large firms or firms with a heavy trading volume and, as we have seen in Tables 3 and 4, these securities, which are more liquid, have expected returns that are significantly less than can be accounted for by the C-K factors alone. Since we have reliable evidence of additional size or liquidity effects that are not accounted for by the C-K factors, it is a little puzzling that the C-K factors price the SMB factor returns so well - the estimated alpha for this portfolio is statistically insignificant. However, since the SMB factor is the difference between the returns on two value weighted portfolios constructed according to firm size relative to that of the median NYSE firm (see Fama and French (1993)), it gives little weight to the most illiquid (i.e., the smallest) firms. As expected, however, the C-K factors fail to explain the HML factor returns.

V

Conclusion

In this paper we make several contributions. First, we test a risk-based asset pricing model against non-risk alternatives using data on individual securities. This is important since, as Lo and MacKinlay (1990) have shown, the use of portfolios is very likely to give rise to a data snooping bias. On a substantive basis we have shown that when individual security returns are risk-adjusted using a 5 factor APT, the significance of the book-to-market ratio disappears, and the firm size effect is subsumed under a trading volume effect which we associate with liquidity. There is also evidence of a dividend yield effect, although this disappears in the early 1980's as the bias against dividends is eliminated from the tax code. The lagged return effects discovered by Jegadeesh and Titman (1993) are found to be partly associated with autocorrelation in the factor returns and the residual component, although statistically significant, appears to be of only marginal economic significance. Finally, we find that there is no evidence of a return effect associated with the stock price or the bid-ask spread.

We have limited our analysis to NYSE securities over the period 1966 to 1989 because of data requirements. It would be interesting to extend the study to a broader cross-section of securities and a longer sample period. This extension is left for future research

References

Amihud, Y., and H. Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223-249.

Ball, R., 1978, Anomalies in relationships between securities' yields and yield surrogates, *Journal of Financial Economics* 6, 102-126.

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

Basu, S., 1977, The investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient markets hypothesis, *Journal of Finance* 32, 663-682.

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

Black, F., M. Jensen, and M. Scholes, 1972, The capital asset pricing model: some empirical tests, in Michael Jensen (ed.), *Studies in the Theory of Capital Markets,* Praeger Publishers, New York.

Black, F., and M. Scholes, 1972, The effects of dividend yield and dividend policy on common stock prices and returns, *Journal of Financial Economics* 1, 1-22.

Brennan, M.J., 1970, Taxes, market valuation, and corporate financial policy, *National Tax Journal* 4, 417-427.

Brennan, M.J., and A. Subrahmanyam, 1995, Investment analysis and price formation in securities markets, *Journal of Financial Economics* 38, 361-382.

Brennan, M.J., and A. Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 341-364.

Brennan, M.J., T. Chordia, and A. Subrahmanyam, 1996, Cross-sectional determinants of expected returns, forthcoming, *On Finance: In Honor of Fischer Black* (David Modest, ed.), Oxford University Press.

Brown, S., and M. Weinstein, 1983, A new approach to testing asset pricing models: the bilinear paradigm, *Journal of Finance* 38, 711-743.

Campbell, J., 1996, Understanding risk and return, *Journal of Political Economy* 104, 298-345.

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

Connor, G., and R. A, Korajczyk, 1993, A test for the number of factors in an approximate factor model, *Journal of Finance* 48, 1263-1291.

Daniel, K., and S. Titman, 1996, Evidence on the characteristics of cross sectional variation in stock returns, forthcoming, *Journal of Finance.*

Dimson, E., 1979, Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* 7, 197-226.

Falkenstein, E.G., 1996, Preferences for stock characteristics as revealed by mutual fund holdings, *Journal of Finance* 51 111-135.

Fama, E.F., and J. MacBeth, 1973, Risk and return: some empirical tests, *Journal of Political Economy* 81, 607-636.

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

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

Fama, E.F., and K.R. French, 1996, Multifactor explanations for asset pricing anomalies, *Journal of Finance* 51, 55-84.

Gibbons, M.R., 1982, Multivariate tests of financial models: a new approach, *Journal of Financial Economics* 10, 3-27.

Jegadeesh, N., 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881-898.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-92.

Jegadeesh, N., and S. Titman, 1995, Overreaction, delayed reaction, and contrarian profits, *Review of Financial Studies* 8, 973-993.

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

Lehman, B., 1990, Residual risk revisited, *Journal of Econometrics* 45, 71-97.

Lehman, B., and D.A. Modest, 1988, The empirical foundations of the arbitrage pricing theory, *Journal of Financial Economics* 21, 213-54

Litzenberger, R.H, and K. Ramaswamy, 1979, Dividends, short-selling restrictions, tax-induced investor clienteles and market equilibrium, *Journal of Financial Economics* 7, 163-196.

Lo, A.W., and A.C. MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431-468.

MacKinlay, A.C., 1995, Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3-28.

Miller, M.M., and M.S. Scholes, 1978, Dividends and taxes, *Journal of Financial Economics* 6, 333-364.

Miller, M.M., and M.S. Scholes, 1982, Dividends and taxes: some empirical evidence, *Journal of Political Economy* 90, 1118-1141.

Petersen, M., and D. Fialkowski, 1994, Posted versus effective spreads: good prices or bad quotes, *Journal of Financial Economics* 35, 269-292.

Roll, R., 1977, A critique of the asset pricing theory's tests: on past and potential testability of theory, *Journal of Financial Economics* 4, 129-176.

Roll, R., and S.A. Ross, 1980, An empirical investigation of the arbitrage pricing theory, *Journal of Finance* 35, 1073-1103.

Sharpe, W.F., 1982, Factors in NYSE security returns 1931-79, *Journal of Portfolio Management* 8, 5-19.

Stoll, H., 1978, The pricing of dealer services: an empirical study of NASDAQ stocks, *Journal of Finance* 33, 1152-1173.

Table 1
Summary statistics

The summary statistics represent the grand cross-section, time-series average for an average of 911 NYSE stocks over 288 months from Jan. 1966 through Dec. 1989. Each stock had to satisfy the following criteria (1) Its return in the current month and in 24 of the previous 60 months be available from CRSP, and sufficient data be available to calculate the size, price, dollar volume, and dividend yield as of the previous month; (2) Sufficient data be available on the COMPUSTAT tapes to calculate the book to market ratio as of December of the previous year; (3) The average bid-ask spread be available for the previous year. The row titled Book to Market Ratio (trimmed) provides summary statistics for this variable after book-to-market ratio values greater than the 0.995 fractile or less than the 0.005 fractile are set to equal the 0.995 and 0.005 fractile values, respectively.

Variable	Mean	Median	Std. Dev.
Firm Size ($bill)	0.921	0.210	3.028
Book to Market Ratio	1.624	0.920	9.683
Book to Market Ratio (trimmed)	1.298	0.920	3.286
Dollar Trading Volume ($mill per month)	36.434	5.152	143.299
Proportional Bid-ask Spread (%)	1.462	1.159	1.248
Share Price ($)	28.957	23.750	51.809
Dividend Yield (%)	3.794	3.373	3.160

Table 2
Correlation Matrix of Transformed Firm Characteristics

This table presents time series averages of monthly cross-sectional correlations between transformed firm characteristics used in pricing regressions. The variables relate to an average of 911 stocks over 288 months from Jan 1966 through Dec 1989. RETURN denotes the excess monthly return, i.e., the raw return less the risk-free return. SIZE represents logarithm of the market capitalization of firms in billions of dollars. BM is the logarithm of the ratio of book value of equity plus deferred taxes to market capitalization, with the exception that book-to-market ratio values greater than the 0.995 fractile or less than the 0.005 fractile are set to equal the 0.995 and 0.005 fractile values, respectively. DVOL is the logarithm of the dollar trading volume. SPREAD is the logarithm of the relative bid-ask spread. PRICE is the logarithm of the share price reciprocal; YLD is the logarithm of the dividend yield; RET2-3 is the natural logarithm of the cumulative return over the two months ending at the beginning of the previous month; RET4-6 and RET7-12 are defined similarly.

	RETURN	SIZE	BM	DVOL	SPREAD	PRICE	YLD	RET2-3	RET4-6	RET7-12
RETURN	1.00	-0.016	0.025	-0.022	0.017	0.005	0.015	0.012	0.020	0.026
SIZE	-0.016	1.00	-0.344	0.864	-0.719	-0.692	0.124	0.049	0.058	0.085
BM	0.026	-0.335	1.00	-0.308	0.256	0.320	0.220	0.018	-0.011	-0.103
DVOL	-0.022	0.864	-0.310	1.00	-0.658	-0.619	-0.038	0.118	0.097	0.123
SPREAD	0.017	-0.719	0.250	-0.658	1.00	0.716	0.190	0.010	-0.001	-0.041
PRICE	0.005	-0.692	0.288	-0.619	0.716	1.00	-0.148	-0.146	-0.167	-0.220
YLD	0.015	0.124	0.190	-0.038	0.190	-0.148	1.00	-0.074	-0.075	-0.058
RET2-3	0.012	0.049	0.018	0.118	0.010	-0.146	-0.074	1.00	0.015	0.050
RET4-6	0.020	0.058	-0.011	0.097	-0.001	-0.167	-0.075	0.015	1.00	0.064
RET7-12	0.026	0.085	-0.103	0.123	-0.041	-0.220	-0.068	0.050	0.064	1.00

Table 3
Fama-MacBeth Regression Estimates of equation (7) using individual security data.
Coefficient estimates are time-series averages of cross-sectional OLS regressions. The dependent variable in the first column is simply the excess return, while in the second and third columns it is the excess returns risk-adjusted using the C-K factors (Dimson betas with one lag are used). The independent variables are the firm characteristics, measured as the deviation from the cross-sectional mean in each period. The estimates in the column labeled "Raw" are the coefficients estimated using equations (8) and (9), while those in the column labeled "Purged" are from equations (8) and (10). The sample and the variables are defined in Tables 1 and 2. All coefficients are multiplied by 100. t-statistics are in parentheses.

	Excess returns	Risk-adjusted returns using the Connor-Korajczyk factors	
		Raw	Purged
Intercept	0.682 (2.08)	-0.012 (0.40)	0.002 (0.05)
SIZE	-0.094 (2.30)	-0.102 (3.00)	-0.124 (3.05)
BM	0.138 (2.05)	0.031 (0.48)	0.030 (0.36)
PRICE	0.307 (2.57)	0.390 (3.88)	0.286 (2.34)
SPREAD	-0.157 (2.42)	-0.227 (3.59)	-0.215 (2.79)
YLD	3.201 (1.17)	8.399 (6.34)	5.929 (5.35)
RET2-3	0.980 (2.47)	1.578 (4.39)	1.323 (3.29)
RET4-6	1.169 (3.30)	1.170 (3.58)	0.877 (2.89)
RET7-12	1.213 (5.10)	0.416 (1.63)	0.467 (2.22)

Table 4
Fama-MacBeth Regression Estimates of equation (7) using individual security data, using one additional lag for variables which involve monthly price levels.
Coefficient estimates are time-series averages of cross-sectional regressions. The dependent variable is the excess return adjusted for the **Connor- Korajczyk factors** (Dimson betas with one lag are used). The independent variables are the firm characteristics, measured as the deviation from the cross-sectional mean in each period. All of the independent variables which involve the use of monthly prices are lagged two periods. The estimates in the column labeled "Raw" are time series averages of the montl\hly estimates, while those in the columns labeled "Purged" are the intercept terms from regressions of the monthly estimates on the factors. The sample and the variables are defined in Tables 1 and 2. All coefficients are multiplied by 100. t-statistics are in parentheses.

	Without dollar volume		With log dollar volume		Weighted least squares			
	(1) Raw	(2) Purged	(3) Raw	(4) Purged	(5) Raw	(6) Purged	(7) Raw	(8) Purged
Intercept	-0.004 (0.15)	0.005 (0.13)	-0.005 (0.15)	0.005 (0.17)	-0.033 (1.24)	-0.049 (2.00)	-0.032 (1.19)	-0.047 (1.95)
SIZE	-0.094 (2.71)	-0.116 (2.81)	0.168 (3.05)	0.164 (3.01)	0.104 (2.63)	0.092 (2.25)	0.085 (2.17)	0.073 (1.82)
BM	0.043 (0.70)	0.040 (0.48)	0.057 (0.91)	0.054 (0.94)	0.022 (0.49)	0.022 (0.50)	0.016 (0.36)	0.016 (0.36)
PRICE	0.030 (0.31)	-0.070 (0.59)	0.069 (0.72)	-0.031 (0.34)	-0.055 (0.80)	-0.117 (1.73)	-0.060 (0.86)	-0.123 (1.83)
DVOL			-0.290 (5.82)	-0.289 (5.84)	0.198 (5.29)	-0.195 (5.12)	0.351 (2.96)	-0.378 (3.24)
$DVOL^2$							0.009 (1.41)	0.011 (1.76))
SPREAD	0.080 (1.27)	0.089 (1.12)	-0.019 (0.30)	-0.008 (0.14)	-0.001 (0.16)	0.020 (0.41)	0.005 (0.10)	0.027 (0.57)
YLD	9.003 (6.84)	6.374 (5.72)	5.877 (4.65)	4.270 (4.40)	0.696 (1.97)	1.103 (2.96)	0.699 (2.00)	1.068 (2.93)
RET2-3	1.246 (3.40)	1.020 (2.45)	1.543 (4.13)	1.313 (3.74)	1.112 (3.43)	1.080 (3.36)	1.122 (3.43)	1.092 (3.38)
RET4-6	0.787 (2.41)	0.506 (1.64)	0.892 (2.72)	0.619 (2.32)	0.778 (2.97)	0.534 (2.20)	0.795 (3.04)	0.557 (2.29)
RET7-12	0.134 (0.518)	0.200 (0.96)	0.225 (0.86)	0.277 (1.44)	0.332 (1.73)	0.292 (1.78)	0.338 (1.75)	0.305 (1.86)

Table 5
Fama-MacBeth Regression Estimates of equation (7) using individual security data, using one additional lag for variables which involve monthly price levels.
Coefficient estimates are time-series averages of cross-sectional OLS regressions. The dependent variable is the excess return adjusted for the **Fama-French factors** (Dimson betas with one lag are used). The independent variables are the firm characteristics, measured as the deviation from the cross-sectional mean in each period. All of the independent variables wich involve the use of monthly prices are lagged two periods. The estimates in the column labeled "Raw" are time series averages of the monthly estimates, while those in the columns labeled "Purged" are the intercept terms from regressions of the monthly estimates on the factors. The sample and the variables are defined in Tables 1 and 2. All coefficients are multiplied by 100. t-statistics are in parentheses.

	Without dollar volume		With dollar volume variables	
	Raw	Purged	Raw	Purged
Intercept	0.077 (1.40)	0.064 (1.13)	0.077 (1.40)	0.064 (1.13)
SIZE	-0.050 (1.59)	-0.077 (2.44)	0.146 (2.48)	0.105 (1.74)
BM	0.058 (0.98)	0.064 (1.06)	0.063 (1.06)	0.068 (1.12)
PRICE	-0.085 (0.86)	-0.242 (2.52)	-0.057 (0.58)	-0.213 (2.25)
DVOL			-0.236 (3.94)	-0.228 (3.69)
$DVOL^2$			-0.002 (0.20)	0.001 (0.12)
SPREAD	0.102 (1.56)	0.091 (1.35)	0.021 (0.34)	0.019 (0.30)
YLD	4.619 (2.69)	4.241 (2.42)	1.893 (1.18)	1.746 (1.06)
RET2-3	0.907 (2.25)	1.197 (2.92)	1.189 (2.96)	1.476 (3.61)
RET4-6	0.889 (2.51)	1.211 (3.39)	0.993 (2.81)	1.320 (3.71)
RET7-12	0.741 (2.64)	1.161 (4.35)	0.838 (3.03)	1.246 (4.75)

Table 6
CK factor autoregressions

This table provides the results of autoregressing the Connor-Korajczyk factors on the three, six, and twelve-lagged nonoverlapping wealth relatives (respectively defined as LAG2-3, LAG4-6, and LAG7-12) obtained by investing in the factor portfolio corresponding to each factor. The two panels correspond to the two subperiods over which the factors are estimated. The F-test reported in the last row corresponds to the null hypothesis that cross-lags of wealth relatives have no additional explanatory power.

A. 1st subperiod (1963-1976)

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Intercept	-0.024 (0.31)	0.014 (0.18)	-0.052 (0.46)	-0.006 (0.09)	-0.017 (0.21)
LAG2-3	0.013 (0.24)	-0.029 (0.54)	-0.164 (2.32)	-0.054 (1.05)	-0.040 (0.66)
LAG4-6	0.009 (0.20)	0.011 (0.24)	0.153 (2.58)	-0.039 (0.97)	0.034 (0.66)
LAG7-12	0.009 (0.31)	-0.002 (0.06)	0.058 (1.09)	0.100 (3.55)	0.026 (0.78)
Adjusted R^2	-0.019	-0.018	0.102	0.070	-0.010
F-test (p-value)	2.180 (0.02)	1.168 (0.31)	0.739 (0.71)	1.728 (0.07)	0.810 (0.64

B. 2nd subperiod (1977-1989)

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Intercept	0.190 (1.97)	-0.043 (0.61)	-0.108 (1.34)	-0.028 (0.38)	0.006 (0.07)
LAG2-3	-0.125 (2.16)	-0.003 (0.06)	-0.005 (0.07)	-0.031 (0.53)	-0.048 (0.78)
LAG4-6	-0.004 (0.07)	0.034 (0.77)	0.062 (1.24)	0.010 (0.21)	-0.037 (0.72)
LAG7-12	-0.044 (1.29)	0.017 (0.54)	0.054 (1.43)	0.050 (1.62)	0.079 (2.01)
Adjusted R^2	0.021	-0.014	0.031	-0.002	0.011
F-test (p-value)	1.061 (0.40)	0.980 (0.47)	1.223 (0.27)	1.584 (0.10)	0.887 (0.56)

Table 7: Excess Returns and Sharpe Ratios for the Connor-Korajcyk Factors.
Mean monthly excess returns and Sharpe ratios for the 5 Connor Korajcyk factors for the two sub-period estimations. The Sharpe Ratio corresponding to a factor is the ratio of the mean excess return of a factor to its standard deviation. The aggregate squared Sharpe Ratio is the sum of the individual factor portfolio squared Sharpe Ratios, and is the estimated squared Sharpe Ratio of the tangency portfolio formed from the 5 (orthogonal) factor portfolios. The P^2 statistics are for the null hypothesis that the mean excess returns for the portfolios in each of the two subperiods are both equal to zero. The figures for the combined periods simply treat the returns on the factor portfolios in the two subperiods as a sngle series, despite the fact that they come from two different principal components estimations.

Factor	1	2	3	4	5	Aggregate
Period 1: January 1963 to December 1976						
Mean Excess Return (x 100) (t-ratio)	0.801 (1.35)	-0.448 (0.75)	-1.225 (2.07)	0.043 (0.07)	0.125 (0.21)	
Squared Sharpe Ratio (p-value)	0.0108	0.0033	0.0254	0.0000	0.0003	0.0399 (0.31)
Period 2: January 1977 to December 1989						
Mean Excess Return (x 100) (t-ratio)	1.115 (2.26)	0.457 (0.83)	0.208 (0.56)	0.149 (0.27)	0.125 (0.44)	
Squared Sharpe Ratio (p-value)	0.0344	0.0047	0.00209	0.0005	0.0013	0.0429 (0.24)
Combined Periods Squared Sharpe Ratio Mean Excess Return (x100) (t-ratio))	 0.0186 0.948 (2.42)	 0.0000 -0.024 (0.06)	 0.0074 -0.554 (1.53)	 0.0002 0.093 (0.23)	 0.0000 -0.052 (0.12)	 0.0262 (0.15)
Overall Period: H_0: Mean Excess Return = 0 P^2 (p-value)	6.94 (0.03)	1.27 (0.53)	4.61 (0.10)	0.08 (0.96)	0.24 (0.89)	

Table 8:Excess Returns and Sharpe Ratios for the Fama-French Factors.
Mean monthly excess returns for the Fama-French factors for the period July 1963 to December 1989. SMB is the difference between the returns of a small and a large firm portfolio and HML is the difference between the returns on a high book to market ratio portfolio and a low book to market ratio portfolio.The Sharpe Ratio corresponding to a factor is the ratio of the mean excess return of a factor to its standard deviation. The aggregate squared Sharpe Ratio is the estimated squared Sharpe Ratio of the tangency portfolio formed from the 3 portfolios.

Factor	Market	SMB	HML	Aggregate
Mean Excess Return (x 100) (t-ratio)	0.423 (1.67)	0.307 (1.88)	0.500 (3.48)	
Squared Sharpe Ratio (p-value)	0.0088	0.0111	0.0381	0.0745 (0.00)

Table 9: Regressions of the Fama-French factors on the estimated Connor-Korajczyk portfolios.
The marker factor is the excess return on the FF market portfolio. SMB is the difference between the returns of a small and a large firm portfolio; HML is the difference between the returns on a high book to market ratio portfolio and a low book to market ratio portfolio. C-K_k (k = 1,..,5) denotes the Connor-Korajczyk factor portfolio returns.

January 1963 to December 1976	Constant	C-K_1	C-K_2	C-K_3	C-K_4	C-K_5	R^2 Nobs
Market Factor	-0.0041 (5.32)	0.518 (56.67)	0.051 (5.64)	-0.165 (18.17)	-0.033 (3.73)	0.043 (4.66)	0.96 132
SMB	0.0015 (1.12)	0.273 (18.45)	-0.221 (15.01)	0.132 (9.04)	0.030 (2.12)	0.043 (2.94)	0.83 132
HML	0.0086 (6.02)	-0.0089 (0.50)	0.060 (3.55)	0.168 (10.11)	0.062 (3.78)	-0.200 (11.93)	0.68 132
January 1977 to December 1989							
Market Factor	-0.0014 (2.25)	0.715 (65.12)	-0.0725 (7.34)	0.188 (12.75)	-0.003 (0.36)	-0.022 (2.51)	0.97 156
SMB	0.0015 (1.77)	0.246 (16.37)	-0.081 (6.07)	-0.431 (21.34)	-0.063 (4.76)	-0.062 (5.21)	0.83 156
HML	0.0049 (3.16)	-0.134 (4.90)	0.183 (7.55)	-0.136 (3.69)	-0.087 (3.62)	0.078 (3.56)	0.46 156
January 1963 to December 1989							
Market Factor	-0.0018 (1.87)	0.613 (45.66)	0.007 (0.53)	-0.076 (5.32)	-0.042 (3.36)	0.021 (1.66)	0.88 288
SMB	0.0012 (0.95)	0.223 (12.87)	-0.182 (11.010	0.010 (0.56)	0.022 (1.334)	-0.011 (0.65)	0.51 288
HML	0.0058 (4.12)	-0.073 (3.74)	0.095 (5.12)	0.100 (4.80)	0.026 (1.41)	-0.084 (4.68)	0.25 288

Legend for Figure 1

(Figure follows over the next five pages)

Figure 1 plots risk-adjusted cumulative excess returns to portfolios that have a unit exposure to one of the characteristics we have considered and zero exposure to the others. For characteristic m (m = 0, 1,..,14) the figures plot C_{mt} where:

$$C_{mt} = t\,\gamma_{m0} + \sum_{\tau=1}^{t} u_{m\tau}$$

where ($_{m0}$ and u_{mJ} are the intercept and residuals from the WLS regression of the monthly characteristic coefficients on the factors.

Intercepts from the regressions of the estimated Connor-Korajczyk factors on the estimated Fama-French factors

Intercept estimates are multiplied by 100. t-statistics are in parentheses

Factor	1st sub-period (Jan. 63-Dec. 76)	2nd subperiod (Jan. 77-Dec.89)
1	-0.005 (0.06)	0.041 (0.51)
2	-0.492 (0.90)	0.278 (0.64)
3	-2.097 (4.25)	0.307 (1.67)
4	-0.065 (0.09)	0.371 (0.70)
5	1.260 (2.12)	-0.664 (1.22)