February 22, 1999 Comments Welcome

Sorting out "Sorting out Sorts"

Kent D. Daniel and Sheridan Titman†

†Kellogg School of Management, Northwestern University and NBER; and University of Texas at Austin and NBER. We thank Ravi Jagannathan, Mitchell Petersen and Canice Prendergast for helpful comments and suggestions.

Sorting out "Sorting out Sorts"

Berk (1998) describes an econometric bias that can arise when one is trying to distinguish between two highly correlated potential determinants of expected return. The paper contributes to the literature by illustrating how the errors-in-variables problem pointed out by Miller and Scholes (1972) can potentially be exacerbated when tests are run on portfolios sorted on variables that may proxy for the expected rate of return.[1] Although Berk's general point is useful, the application of his critique to our paper Daniel and Titman (1997), is misleading. The tests we employ in our paper are not the tests Berk describes, and are subject to none of the biases he describes. This brief note clarifies Berk's point, and describes the steps we took to avoid not only the biases, but power problems associated with the issues Berk raises.

The econometric tests in question examine whether factor loadings have a discernible effect on average returns after controlling for stock characteristics like size and book-to-market ratios. Our analysis focuses on the Fama and French (1993) three-factor model that popularized the idea that factor portfolios could be formed based on characteristics. The premise of Fama and French's model is very reasonable: if there are priced factors that are responsible for the size and value premium, then sorting stocks into portfolios based on their size and book-to-market ratios are likely to result in diversified portfolios that span the factor space. Their method is supported by the arguments of Ball (1978) or Berk (1995) that price-scaled variables such as size and book-to-market should be a good proxy for expected returns. Indeed, the Fama and French results are consistent with this interpretation: First, there are strong covariances between the returns of the stocks within their size and book-to-market factor-mimicking portfolios suggesting that these stocks load on a common, distress-related risk factor.[2] Second, the returns of their portfolios appear to do a good job of explaining the returns of other size and book-to-market sorted portfolios.

Our paper, Daniel and Titman (1997), suggests that while the Fama and French results are consistent with their interpretation, they are also consistent with alternative interpretations. We propose as an *ad-hoc* alternative hypothesis a *characteristics model*, which specifies that a stock's average return is determined directly by its characteristics, and is independent of that stock's factor sensitivities. We argue that Fama and French's asset pricing tests

[1] Miller and Scholes (1972) enumerate a set of econometric issues related to tests by Douglas (1969) and Lintner (1965), among which is the "errors of measurement in b_1 [market beta] and the attenuation bias." Among others, Fama and MacBeth (1973) and Black, Jensen, and Scholes (1972) develop procedures to reduce or eliminate this bias.

[2] In Section 3 of Daniel and Titman (1997) we run an alternative test, and argue that the results of this test show that the high covariance is not due to the presence of a separate, distress related factor. We do not concentrate on this result here, as Berk's comments do not concern this part of our paper.

would be unlikely to reject the factor model even if the characteristics model were true and the factor model false. The reason is that the characteristics and factor loadings of Fama and French's test portfolios are highly correlated: their test portfolios consist of either high book-to-market stocks with high HML loadings or low book-to-market stocks with low HML loadings. Thus, the expected returns for these portfolios are the same under both the factor model and the characteristics model, and tests run on these portfolios will have no power to distinguish between the two models. To discriminate between these models thus requires that we construct portfolios of high book-to-market stocks that behave like (*i.e.*, covary with) growth stocks, or low book-to-market stocks that behave like value stocks. To construct these portfolios we first sort stocks into portfolios according to their characteristics, and then do a second sort according to factor loadings.

Berk (1998) explains how our method of sorting stocks into test portfolios, in combination with poorly designed tests, can lead to incorrect inferences. His argument is that if stocks are sorted into groups with very similar book-to-market ratios, there may be very little variation in the true factor loadings or in the expected returns. Hence a second sort on estimated loadings might divide stocks into groups according to measurement error, and might thus produce no variation in average returns or in true factor loadings. Observing this, a researcher might incorrectly conclude that expected returns were unrelated to true loadings. Additionally, Berk shows that a regression of average returns on the measured loadings would suffer from an errors-in-variables problem that would bias the coefficients on the measured loadings downward, which could, in turn, lead to a false rejection of the factor model (*i.e.,* it would have incorrect size).

It was precisely because of this errors-in-variables problem that, in Daniel and Titman (1997) we *did not use tests like those that Berk describes.* We did not regress returns on factor loadings. Rather, to test the null of the Fama and French three-factor model, we run regression tests of our double-sorted portfolios on the factor-mimicking portfolio returns:[3]

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,Mkt}(R_{Mkt,t} - R_{f,t}) + \beta_{i,HML}R_{HML,t} + \beta_{i,SMB}R_{SMB,t} + \epsilon_{i,t}$$

and test whether α_i is different from zero. This test does not regress returns on factor loadings, but rather returns on returns, and hence is not subject to the errors-in-variables bias.

Of course, if the dispersion of factor loadings in the second sort were small, then our factor-model tests would have little power and we would be incapable of rejecting the three factor model even if it were wrong (*i.e.*, we would not have sufficient power against the

[3]Black, Jensen, and Scholes (1972) propose this sort of regression test as a way of eliminating the errors-in-variables bias (see their equation (6) and footnote 3).

characteristics alternative). The fact that we do indeed reject the factor model null is *prima facie* evidence that our portfolio sorting methods give us adequate power against the factor model.[4]

In addition to our test of the factor model we construct a very similar test with the characteristic-model as the null hypothesis. We find that we cannot reject the characteristics model. However, this failure to reject could result from a lack of power: if our second sort were to produce almost no dispersion in factor loadings, then we might be unable to find any return variation associated with factor loading variation because the variation in true factor loadings is small, rather than because it is not there.

Because we were concerned about this possibility, we took steps to verify that the power problem was not extreme. We constructed a set of characteristic-balanced portfolios: zero-investment portfolios which buy and sell stocks with equal characteristics. If Berk's conjecture were correct, and if characteristics and true factor loadings were almost perfectly correlated, the factor loadings of these characteristic-balanced portfolios would be close to zero. We found instead that the characteristic-balanced portfolio (which we use in our joint test) has an estimated loading on the HML factor of -0.724, with a T-statistic of -12.3. Thus, the test indicates that the characteristics and factor loadings are not so highly correlated as to make discriminating between the two models impossible.

Our characteristic-model and factor-model tests are really very similar: To test the factor model, we examine whether the factor risk-premia are reliably different than we would expect from the Fama and French three-factor model. We find that the return premia are reliably different from what the factor model predicts and thus we can statistically reject the model. To test the characteristics model we examine whether the factor return premia are reliably different than zero. We find that the return premia are not reliably different than zero and, therefore, we cannot reject the characteristics model. Of course we can never say that the factor premia are exactly zero and that we *accept* the characteristics model, only that we fail to reject it.

In summary, Berk illustrates the potential pitfalls that can arise from poorly designed tests of factor models. These insights may be useful for researchers studying the determinants of expected returns. However, not all tests are subject to the errors-in-variables bias Berk discusses. The tests in Daniel and Titman (1997) do not suffer from this bias, and hence our rejection of the Fama and French (1993) model cannot be attributed to these biases.

[4] We discuss the potential power problems resulting from low dispersion of the factor loadings in Section 4.1 of Daniel and Titman (1997).

References

Ball, Ray, 1978, Anomalies in relationships between securities' yields and yield-surrogates, *Journal of Financial Economics* 6, 103–126.

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

——— , 1998, Sorting out sorts, unpublished manuscript.

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

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

Douglas, George W., 1969, Risk in the equity markets: An empirical appraisal of market efficiency, *Yale Economic Essays* 9, 3–45.

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

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

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

Miller, Merton H., and Myron Scholes, 1972, Rate of return in relation to risk: A re-examination of some recent findings, in Michael C. Jensen, ed.: *Studies in the Theory of Capital Markets* . pp. 47–78 (Praeger Publishers: New York).