

NBER WORKING PAPER SERIES

WHICH ALPHA?

Francisco Barillas  
Jay Shanken

Working Paper 21698  
<http://www.nber.org/papers/w21698>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
November 2015

Thanks to Ken French for comments on an earlier draft. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Francisco Barillas and Jay Shanken. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Which Alpha?  
Francisco Barillas and Jay Shanken  
NBER Working Paper No. 21698  
November 2015  
JEL No. G11,G12

**ABSTRACT**

A common approach to comparing asset pricing models with traded factors involves a competition between models in pricing test-asset returns. We find that such practice, while seemingly reasonable, cannot be relied on to determine which is the superior model for several widely accepted criteria including statistical likelihood, Sharpe ratios and a modified HJ distance. All that matters for model comparison is the extent to which each model is able to price the factors in the other model. Given this information, test assets are actually irrelevant, whether the models are nested or non-nested.

Francisco Barillas  
Goizueta Business School  
Emory University  
1866 Brockton Glen NE  
Atlanta, GA 30329  
francisco\_barillas@bus.emory.edu

Jay Shanken  
Goizueta Business School  
Emory University  
1300 Clifton Road  
Atlanta, GA 30322  
and NBER  
jay.shanken@emory.edu

Many papers in the empirical literature compare the performance of different models in pricing test assets using metrics like the average absolute “alpha” or test results for the zero-alpha restriction. These papers often leave the impression that the model with smaller alphas is preferred. But can we necessarily conclude that the data favor such a model? This question, though clearly one of great interest for asset pricing has not, to our knowledge, been thoroughly analyzed in the traditional time-series alpha-based framework. Surprisingly, for many important metrics, the answer is no. Simply comparing pricing performance for test assets, while it may provide some useful descriptive information, cannot serve to identify a superior model and can even be misleading in this regard.

Metrics for which this conclusion holds include classic ones like the general statistical criterion of model likelihood, as well as fundamental asset-pricing measures based on relative factor-portfolio efficiency (Sharpe ratio) or a related notion of distance for stochastic discount factors (HJ distance). In fact, we prove that *test assets tell us nothing about model comparison in these cases, beyond what we learn by examining the extent to which each model prices the factors in the other model*. This is the main message of our paper. Most of our colleagues have, at first, found this counter-intuitive, as did we. But the logical argument is straightforward once some simple algebraic pricing results are established and it applies to nested as well as non-nested models, the latter being the case in which each model contains factors not included in the other.

The classic “alpha” of investment analysis is the intercept in the time-series regression of an asset’s excess returns on those of the market portfolio (Mkt). The early paper by Jensen (1968) recognized the interpretation of alpha as an asset’s deviation from the capital asset pricing model (CAPM) expected return relation of Sharpe (1964) and Lintner (1965). Over the decades, other benchmark factors have also been included in asset pricing models, notably the SMB (small minus big) size and HML (high minus low) value factors of the Fama-French (1993) three-factor model (FF3). With such “traded factors,” i.e., portfolio excess returns or spread portfolios, the time-series intercept can still be viewed as an asset’s deviation from the model.

Alpha also plays a fundamental role in analyzing portfolio performance, as measured by the Sharpe ratio, a portfolio’s expected excess return divided by its standard deviation. A non-zero alpha indicates that the Sharpe ratio can be improved and a more efficient portfolio obtained by complementing investment in the benchmark portfolios with a position in the given asset.<sup>2</sup> If the

---

<sup>2</sup> Strictly speaking, the square root of the squared Sharpe ratio can be increased in this case. It is possible that the corresponding tangency portfolio will be inefficient and have a negative Sharpe ratio. This case is rarely encountered in practice, however, and so we assume henceforth that the Sharpe ratio is positive.

factors already span the tangency portfolio, however, such improvement is not possible and alpha is zero. Gibbons, Ross and Shanken (1989), henceforth GRS, examine the case of multiple assets and develop an F-test of the joint hypothesis that the alphas are all zero.

Here, we address a different, but related question – the one posed in our initial discussion. We suppose there are two pricing models of interest to be compared and the better model identified. In addition to the two sets of factors, which may overlap, there is a set of test-asset returns (securities or portfolios) that can be used in evaluating the models. Ideally, we want a model that will “price,” i.e., produce zero alphas for all of these investments, both test-asset and factor returns. Equivalently, the factors in the model should span the tangency portfolio for the full investment universe and, therefore, maximize the Sharpe ratio. The classic asset-pricing challenge is then to identify a small number of *observable* factors, as few as possible, that meet these equivalent objectives. We do not presume that any of the models under consideration is perfect, however. Instead, we explore several notions of which model is the “better” one.

A number of papers provide empirical results on the pricing of test assets across different models, e.g., Fama and French (1993, 1996, 2015a,b), Avramov and Chao (2006), Hou, Karolyi and Kho (2011) and Hou, Xue and Zhang (2015a,b). The recent papers also provide some evidence on each model’s ability to price the factors in the other model (also see Asness and Frazzini (2013)). But the issue of how to combine the different kinds of evidence in comparing models is not addressed. And while a given model might perform well from both perspectives, this need not be so. We provide several examples below, in which the model that does a better job of pricing test assets has more trouble pricing the excluded factors and, therefore, is not the better model. As we will see, a traditional likelihood ratio finds such a model inferior.

A related paper by Fama (1998) explores the pricing of state variables in an intertemporal CAPM context. Fama argues that it is not necessary to test alternative ICAPMs on all assets for the purpose of identifying the state variables of concern to investors. Cochrane (2005) also has a related result for *nested* models in the stochastic discount factor framework, but does not discuss its relevance for the comparison of standard alpha-based models with traded factors, nor does he consider non-nested models.<sup>3</sup> Thus, while the pricing relations that our conclusions are based on have antecedents in

---

<sup>3</sup> Cochrane (Section 13.4) shows that “if you want to know whether factor  $i$  helps to price other assets, look at  $b_i$ ,” the coefficient on that factor in the linear specification for the SDF. As he shows, the  $b$  vector equals the inverse of the factor covariance matrix times the price of risk vector  $\lambda$  in the expected-return/beta relation (also see Kan, Robotti and Shanken (2013)). Test-asset irrelevance is not discussed by Cochrane and, indeed, is not immediately apparent since, in

the theoretical literature, we extend those results along several dimensions. But, more importantly, we explore the implications for empirical analysis involving model comparison, as these issues do not seem to be recognized in the empirical asset pricing literature.

Before continuing, it deserves emphasis that by “model comparison,” we mean here the determination of which model is the superior one according to a given metric. A researcher may, nonetheless, be interested in exploring how various models price particular assets and this is certainly a form of comparison, as that term is used more generally. But, as we demonstrate, it is not the same as identifying the better model based on well-established criteria. In fact, the ability of models to price sets of test assets is irrelevant in this regard. Thus, our paper highlights the differences between these notions of “comparison” and, hopefully, will contribute to a better understanding of the relation between the contrasting research objectives.

We consider first the case of one model nested in a larger model. While it may seem that additional factors can only improve on pricing and produce smaller test-asset alphas, this need not be true in general. An expanded model can yield expected-return predictions that are farther from the actual test-asset means than the predictions of the nested model. Although this is not a new insight, it deserves emphasis in the context of our analysis. As a simple illustration, suppose we want to compare CAPM to a two-factor model with factors Mkt and SMB. There is just one test asset, the loser decile portfolio based on past-year returns. For 1963-2013, the annualized CAPM alpha for this portfolio is -11.06% (the momentum effect of Jegadeesh and Titman (1993)), but the two-factor loser alpha, -11.79%, is even larger in magnitude. This occurs because losers load positively on SMB (many losers are small firms). Given the positive SMB alpha (the size effect) relative to the Mkt, this raises the two-factor model’s prediction for the expected loser return. Consequently, the loser alpha, the difference between the actual and predicted values, is lowered.

Thus, one set of restrictions favors CAPM (the loser test-asset restrictions) while another favors the two-factor model (the SMB excluded-factor restriction). How do we choose between models in such a case? We could think about specifying some metric that aggregates the various alphas; consistent estimators for each model and the difference might then provide the basis for a formal statistical comparison. One concern, however, is that an ad hoc metric of this sort might be influenced by the number and choice of test assets in a somewhat arbitrary fashion. The issue of conflicting asset-

---

general, identification of  $\lambda$  requires test assets, not just factors. However, when the factors are traded,  $\lambda$  is the vector of expected factor returns and  $b$  depends on the factor moments only. In fact,  $b$  is the vector of weights in the tangency portfolio based on the factors in this case. Thus, factor  $i$  is needed if its tangency portfolio weight is non-zero. We derive an equivalent result in terms of alphas later in the paper.

pricing restrictions is by no means an artifact of our simple example. Fama and French (2015b) encounter this phenomenon as well. The three-factor model cannot price the additional investment and profitability factors of their expanded five-factor model (FF5). But while FF5 does about as well or better than FF3 in accounting for many anomalies, FF3 is better at pricing portfolios formed by independent sorts on size and accruals.

These examples relate to a basic property of nested models. Assume the nested model,  $M_1$ , contains the factors  $f_1$  and the expanded model,  $M$ , contains the factors  $f_1$  and  $f_2$ .  $R$  is a vector of test-asset returns. Let  $\alpha_{R1}$  denote the alphas of the test assets relative to  $M_1$ ,  $\alpha_{21}$  the alphas of the additional factors on  $f_1$ , and  $\alpha_R$  the test-asset alphas relative to the larger model,  $M$ . Suppose that  $M_1$  improves on the pricing of test assets, in the sense that the magnitude of the elements of  $\alpha_{R1}$  is smaller than that of  $\alpha_R$  in some metric. Ironically, it then follows that  $M_1$  *must* fail to price some or all of the excluded factors, i.e.,  $\alpha_{21} \neq 0$ . In fact, the same conclusion holds if there is *any* difference in the test-asset alphas.

The examples and observations above highlight the fact that, while it may be of some empirical interest to separately examine the pricing of test assets and excluded factors, it is essential to *jointly* consider both types of evidence in the ultimate evaluation of a model. This can easily be carried out by letting test assets and factors serve simultaneously as “left-hand-side assets” in a standard GRS test. Comparing test statistics across models is problematic, however, and subject to the limitation that power will generally differ, making an evaluation based on p-values difficult to interpret. In particular, a model with many factors and relatively low residual variance of left-hand-side returns might produce a more extreme test statistic and a lower p-value even if it does a better job of pricing (smaller alphas). This point has been made previously by Fama and French (1993). So the question remains - how should we compare two models when evaluation of some restrictions points to the first model and evaluation of others points to the second model.

Fortunately, a simple resolution is possible. This follows directly from a basic equivalence that we establish: the nested model  $M_1$  holds in the usual sense that  $\alpha_{21} = 0$  and  $\alpha_{R1} = 0$  if and only if  $\alpha_{21} = 0$  and  $\alpha_R = 0$ .<sup>4</sup> Thus, in characterizing  $M_1$ , the test-asset restrictions can be formulated either

---

<sup>4</sup> This equivalence is related to a result in Sharpe (1984), who provides a multibeta interpretation of the CAPM. Sharpe describes his result as an approximation. However, one direction of our equivalence - that the usual CAPM restrictions imply the multifactor pricing model with restricted prices of risk - can be obtained using a variation on his

relative to  $f_1$  or in terms of the expanded set of factors,  $f_1$  and  $f_2$ , as long as the excluded-factor constraints are also imposed. For example, if  $M_1$  is CAPM and  $M$  is FF3, CAPM holds if and only if the one-factor alphas for SMB and HML are zero and the *three-factor* test-asset alphas are zero as well. Hence, CAPM is seen to be a restricted version of FF3 in this representation and so comparing the models amounts to evaluating whether the excluded-factor CAPM restrictions hold. The test assets are *irrelevant* in this regard since those restrictions are common to both models.

If  $\alpha_{21} = 0$ , the excluded-factor premia  $E(f_2)$  are given by the betas on  $f_1$  times  $E(f_1)$ . We can also show that  $\alpha_{R1} = \alpha_R$  in this case, so the same pricing of factors and test assets is achieved with the more parsimonious model  $M_1$ . The nested model is the “better” model in this sense. Of course, if  $\alpha_{21} \neq 0$ , the expanded model  $M$  is the superior model since it leaves  $E(f_2)$  unrestricted, rather than imposing a constraint that is violated. This is true even if the nested model  $M_1$  improves on the pricing of test assets, as compared to the larger model  $M$ . From a portfolio perspective,  $\alpha_{21} \neq 0$  means that the additional factors  $f_2$  allow for the attainment of a higher Sharpe ratio than would be possible based on  $f_1$  alone. This simple efficiency criterion based solely on the factors is shown here to be valid for nested-model comparison, despite the availability of test-asset information; it is not an implicit *assumption*, as in some earlier work.

With these observations in mind, consider again the common procedure of comparing models on the basis of their relative success in pricing a set of test assets. In the simple two-factor loser example discussed earlier, CAPM would be judged superior since the loser portfolio CAPM alpha is smaller in magnitude than the loser two-factor alpha. However, this conclusion would be contrary to that based on a comparison of Sharpe ratios or, equivalently, a version of the Hansen-Jagannathan (1997) distance. These measures will favor the expanded model (higher Sharpe ratio, lower distance) whenever a nested model fails to price the excluded factors, regardless of how the addition of those factors affects the pricing of test assets.

Although we have focused on nested models in the discussion thus far, similar conclusions about the role of test assets hold for a pair of *non-nested* models  $M_{1a}$  and  $M_{1b}$ . Here, we can characterize each model in terms of its excluded-factor restrictions ( $\alpha_{21a} = 0$  and  $\alpha_{21b} = 0$ , respectively), together with the same test-asset restrictions ( $\alpha_R = 0$ ). The latter alphas are defined

---

argument. To see this, consider the case in which all factors are traded and the market is one of the factors. Then the factors will obviously span the market portfolio, so Sharpe’s equation (4) holds exactly. Equation (5) then gives the desired result. Assuming the factors are multifactor minimum-variance (MMV) portfolios in the sense of Fama (1996), Fama (1998) provides an alternative proof in this same direction. We do not impose the MMV assumption.

relative to the combined model,  $M$ , that includes all of the factors and so, as earlier, the test assets are irrelevant for model comparison. In other words, how well each model prices the excluded factors is all that matters in deciding which model is the better one – what we call a relative test. Moreover, the logic would be the same even if the set of test-asset returns, together with the factors, spanned the entire investment universe. Of course, the evidence on test assets *is* relevant for evaluating the validity of each model (an absolute test); it is just not needed for comparing the models.

Later, we examine a non-nested models example from several perspectives. The models are FF3 and a four-factor model that substitutes an alternative more timely value factor for HML and adds in a momentum factor UMD, along with Mkt and SMB. As in past studies, UMD has a huge FF3 alpha (nearly 11% annualized) and, consequently, the four-factor model does a much better job of pricing excluded factors. In contrast, FF3 performs better in pricing 6 of 10 sets of test-asset portfolios, by several commonly-used criteria. Nonetheless, we show that a standard likelihood ratio clearly favors the four-factor model, consistent with our basic message.

When evaluating models in terms of statistical likelihood, test-asset irrelevance amounts to the observation that the likelihood ratio for a pair of models is independent of the test-asset returns. Interestingly, this is true regardless of the model parameterization, e.g., whether CAPM is represented in the usual way, in terms of restrictions on CAPM alphas only, or an equivalent combination of restrictions on CAPM alphas (excluded factors) and FF3 alphas (test assets). Barillas and Shanken (2015) build on this observation, showing how to formally aggregate excluded-factor intercept evidence across several models in a Bayesian framework, permitting both nested and non-nested comparisons.

The rest of the paper is organized as follows. Section 1 establishes the equivalence result for nested models and discusses the implied test-asset irrelevance for both nested and non-nested models. Section 2 presents an example that illustrates the conflicting conclusions that can be obtained in comparing non-nested models based on test-asset alphas, instead of the excluded-factor alphas. Test-asset irrelevance in terms of model likelihoods is explored under estimation uncertainty in Section 3. Section 4 extends our results to model comparison based on a modified version of the HJ distance and Section 5 concludes.

## **1. Comparing Asset Pricing Models**

In this section, we develop our key result, that test assets are irrelevant for model comparison, given the factor returns. We first address the evaluation of a factor-pricing model  $M_1$  relative to



another model  $M$  that contains all the factors in  $M_1$  as well as some additional factors. For example, CAPM is *nested* in FF3 in this sense.

We begin by laying out the model notation and assumptions. The factor model  $M$  is a multivariate linear regression with  $N$  excess returns,  $\mathbf{R}$ , and  $K$  factor returns  $\mathbf{f}$ :

$$\mathbf{R} = \alpha_{\mathbf{R}} + \beta \mathbf{f} + \varepsilon, \quad (1)$$

where  $\mathbf{R}$ ,  $\varepsilon$  and  $\alpha$  are  $N \times 1$ ,  $\beta$  is  $N \times K$  and  $\mathbf{f}$  is  $K \times 1$ . The disturbance has zero mean and covariance matrix  $\Sigma$ . As shown in Gibbons, Ross and Shanken (1989), the following relation holds:

$$\alpha_{\mathbf{R}}' \Sigma^{-1} \alpha_{\mathbf{R}} = \text{Sh}(\mathbf{f}, \mathbf{R})^2 - \text{Sh}(\mathbf{f})^2, \quad (2)$$

where  $\text{Sh}(\cdot)^2$  denotes the maximum squared Sharpe ratio (mean excess return over standard deviation) obtainable from portfolios of the given returns.<sup>5</sup> This Sharpe ratio corresponds to the tangency portfolio determined by the test-asset and factor returns, along with the risk-free rate.

We now establish a basic proposition that will greatly simplify the task of comparing nested models. In particular, think about CAPM as nested in the Fama-French (1993) three-factor model (henceforth FF3). In this case, the proposition shows that the alpha restriction of the single-factor CAPM can be reformulated in terms of the usual one-factor alpha restriction for the excluded-factor returns (HML and SMB), together with the FF3 alpha restriction for the test-asset returns. Since the models differ only with respect to the excluded-factor restrictions, we have the surprising (to us) conclusion that the test-asset returns do not play any role in comparing CAPM and FF3. This conclusion holds under the assumption, which we maintain throughout, that there is no interaction in the chosen metric between the excluded-factor alphas and the expanded-model test-asset alphas. We later show that this is true, in particular, for the standard measure of a model's statistical likelihood, a basis for both classical and Bayesian modes of inference. It is also true for several familiar asset-pricing metrics.

One can see the equivalence between zero-alpha restrictions as a direct consequence of standard portfolio algebra.  $M$  is the pricing model with factors  $\mathbf{f} = (\mathbf{f}_1, \mathbf{f}_2)$ , where  $\mathbf{f}_1$  consists of the factors in the nested model  $M_1$ . The key observation is that the tangency portfolio  $\tau(\mathbf{f}_1, \mathbf{f}_2, \mathbf{R})$  equals the tangency  $\tau(\mathbf{f}_1)$  if and only if  $\tau(\mathbf{f}_1) = \tau(\mathbf{f}_1, \mathbf{f}_2)$  and  $\tau(\mathbf{f}_1, \mathbf{f}_2) = \tau(\mathbf{f}_1, \mathbf{f}_2, \mathbf{R})$ . In words, if the factors  $\mathbf{f}_1$

---

<sup>5</sup> Also see related work by Jobson and Korkie (1982)

already span the tangency portfolio for the investment universe consisting of all the factors and test-asset returns, then the additional  $f_2$  factors will not improve on this tangency portfolio, nor will adding test assets to the factors. More formally, if the nested model  $M_1$  holds for the factors  $f_2$  ( $\alpha_{21} = 0$ ), then  $\text{Sh}(f)^2 = \text{Sh}(f_1)^2$  by (2). Therefore,  $\text{Sh}(f, R)^2 = \text{Sh}(f_1)^2$  if and only if  $\text{Sh}(f, R)^2 = \text{Sh}(f)^2$  in this case and so, appealing to (2) again, we have:

*Proposition 1.* The nested pricing model  $M_1$  with factors  $f_1$  holds for both the test asset returns  $R$  and the excluded-factor returns  $f_2$  ( $\alpha_{R1} = 0$  and  $\alpha_{21} = 0$ ) if and only if those excluded-factor returns satisfy  $M_1$  ( $\alpha_{21} = 0$ ) and the larger model  $M$  holds for the test asset returns ( $\alpha_R = 0$ ).

Further insight can be obtained by deriving the proposition in terms of the statistical relation between the parameters in the regression models for  $M$  and  $M_1$ . Partitioning  $\beta = (\beta_1, \beta_2)$  to conform with the factor partition  $f = (f_1, f_2)$ , we have

$$R = \alpha_R + \beta_1 f_1 + \beta_2 f_2 + \varepsilon. \quad (3)$$

We are interested in the relation between these parameters and those in the regression of  $R$  on a constant and  $f_1$  only:

$$R = \alpha_{R1} + b f_1 + e. \quad (4)$$

This relation depends on the parameters in the ‘‘auxiliary regression’’:

$$f_2 = \alpha_{21} + d f_1 + u. \quad (5)$$

Substituting this expression for  $f_2$  in (3) gives

$$R = \alpha_R + \beta_1 f_1 + \beta_2 (\alpha_{21} + d f_1 + u) + \varepsilon = (\alpha_R + \beta_2 \alpha_{21}) + (\beta_1 + \beta_2 d) f_1 + (\beta_2 u + \varepsilon).$$

It follows from the standard regression orthogonality conditions that  $u$  and  $\varepsilon$  have mean 0 and are uncorrelated with  $f_1$ . Therefore, the regression on  $f_1$  satisfies:

$$\alpha_{R1} = \alpha_R + \beta_2 \alpha_{21}, \quad b = \beta_1 + \beta_2 d \quad \text{and} \quad e = \beta_2 u + \varepsilon. \quad (6)$$

Now suppose  $M_1$  holds for both the test asset returns and the excluded-factor returns, i.e.,  $\alpha_{R1} = 0$  and  $\alpha_{21} = 0$ . Then  $\alpha_R = 0$  by (6). Conversely, if  $\alpha_R = 0$  and  $\alpha_{21} = 0$  then  $\alpha_{R1} = 0$ , again establishing the proposition.<sup>6</sup>

Thus, a model  $M_1$  that is nested in a larger model  $M$ , in the sense that its factors are all included in  $M$ , is nested in the *statistical* sense that  $M_1$  may be obtained by imposing restrictions on  $M$ . As noted above, it follows that the only condition relevant in distinguishing between  $M_1$  and  $M$  is the requirement that  $M_1$  hold for the excluded-factor returns  $f_2$  ( $\alpha_{21} = 0$ ). The test-asset restriction ( $\alpha_R = 0$ ) is common to both models and, therefore, cannot help in deciding which model performs better. Hence, the test-asset returns are not relevant in *comparing*  $M_1$  and  $M$ , though they are, of course, important in assessing model validity.

If  $\alpha_{21} = 0$ , the model predictions for the expected test-asset returns are identical for  $M_1$  and  $M$ . To see this, note that under  $M_1$ , the prediction is  $bE(f_1)$ , while under  $M$ , it is  $\beta_1 E(f_1) + \beta_2 E(f_2)$ . Now  $\alpha_{21} = 0$  implies that  $E(f_2) = dE(f_1)$ , so test-asset expected returns under  $M$  equal  $\beta_1 E(f_1) + \beta_2 dE(f_1) = (\beta_1 + \beta_2 d)E(f_1)$ , and this equals  $bE(f_1)$  by the middle relation in (6). Therefore, the more parsimonious  $M_1$  is favored in this case. On the other hand,  $\alpha_{21} \neq 0$  implies that  $M$  is the better model since it does not impose the false restriction on  $\alpha_{21}$ . In this case, the factor-based tangency portfolio places some weight on the factors in  $f_2$ .

Equation (6) also yields the interesting implication that if  $\alpha_{21} = 0$  then  $\alpha_R = \alpha_{R1}$ , i.e., the test-asset alphas on  $f_1$  and  $f_2$  equal those on  $f_1$ , whether the models hold or not. Thus, any difference between  $\alpha_R$  and  $\alpha_{R1}$  requires that  $\alpha_{21} \neq 0$ , which in turn implies that  $M$  is the better model. This is true, even if the elements of  $\alpha_{R1}$  are smaller in magnitude than those of  $\alpha_R$  in some metric. Thus, by focusing on test assets in isolation, a false inference about model comparison can be obtained. These observations imply that it is essential to consider test-asset pricing jointly with the pricing of excluded factors in the evaluation of a model. But if we do that, the test assets end up dropping out in the model comparison.

As noted earlier, a metric that allows for interaction between the test-asset deviations from  $M$  and the excluded-factor deviations from  $M_1$  need not satisfy test-asset irrelevance. From equation (6),

---

<sup>6</sup> See Pastor and Stambaugh (2002) for a very different application of (6) to the estimation of mutual fund alphas.

we see that focusing on test-asset deviations from  $M_1$  involves such an interaction since the corresponding squared alphas ( $\alpha_{R_1}^2$ ) depend on products of elements of  $\alpha_{21}$  and  $\alpha_R$  when . Thus, test assets need not drop out in a comparison based on *model-specific* test-asset alphas. But such a comparison would be inconsistent with the likelihood principle and asset-pricing metrics that will be discussed later.

Under standard simplifying regression assumptions: i) constant parameters in the linear regression of  $f_2$  on  $f_1$  and ii) disturbances that are independent and identically normally distributed over time, the standard GRS F-test can be used to evaluate the hypothesis  $\alpha_{21} = 0$ . In independent work, Fama and French (2015b) note that FF5 performs better than FF3 for almost all portfolio sorts examined and consider whether the differences are statistically reliable. They report that the GRS test for the investment and profitability factors regressed on the FF3 factors is highly significant and assert (without proof) that if some stocks have nonzero exposures to the additional factors, then those factors “add information about expected returns to the three-factor model.” The first relation in equation (6) above formalizes this idea, showing that if we assume  $\beta_2 \neq 0$ , then a finding that  $\alpha_{21} \neq 0$  implies that the test-asset alpha vectors  $\alpha_{R_1}$  and  $\alpha_R$  differ. However, as we noted earlier,  $\alpha_{21} \neq 0$  does *not* ensure that the superior model  $M$  with factors  $(f_1, f_2)$  will do a better job of pricing the test assets; i.e., the added information can be detrimental to pricing.

Now consider a pair of non-nested models  $M_{1a}$  and  $M_{1b}$ , with some factor overlap possible, and let  $M$  be the model corresponding to the union of the factors from these models. For example, {Mkt SMB} and {Mkt HML} are non-nested in this sense and  $M$  is {Mkt SMB HML}. In general,  $M_{1a}$  and  $M_{1b}$  are both nested in  $M$ , so Proposition 1 can be invoked with each model playing the role of  $M_1$ . Thus,  $M_{1a}$  holds if and only if it prices the factors in  $M_{1b}$  and  $M$  prices the test asset returns,  $R$ . Similarly,  $M_{1b}$  holds if and only if it prices the factors  $M_{1a}$  and  $M$  prices  $R$ . Once again, the test-asset restrictions ( $\alpha_R = 0$ ) are the same for each model and, therefore, the pricing of test-asset returns is irrelevant for comparing  $M_{1a}$  and  $M_{1b}$ . We need only consider the pricing of the factors in each model relative to the factors in the other model (the common factors will automatically be priced). For example, we need to ask how well {Mkt SMB} prices HML and {Mkt HML} prices SMB.

Let us represent the excluded-factor restrictions as  $\alpha_{21a} = 0$  and  $\alpha_{21b} = 0$ . Suppose  $\alpha_{21a} = 0$  is rejected, but we fail to reject  $\alpha_{21b} = 0$ . Power issues aside, this evidence would be consistent with a scenario in which  $M_{1a}$  is the better model - equivalent to the larger model  $M$  (which necessarily prices

all the factors), while  $M_{1b}$  is not. But more generally, both  $\alpha_{21a}$  and  $\alpha_{21b}$  could be nonzero, in which case the preferred model would not be obvious. Thus, model comparison is less straightforward here than in the nested scenario, where  $\alpha_{21}$  is either zero or not, leaving no ambiguity about the ranking of models (apart from estimation issues). Nonetheless, in *both* the nested and non-nested cases, test assets are irrelevant.

Asness and Frazzini (2013), Fama and French (2015a,b) and Hou, Xue and Zhang (2015a,b) report classical tests for the pricing of some factors relative to other factors. Thus, they recognize the relevance of such procedures for model comparison. They do not formalize this as we do in Proposition 1, though, nor do they discuss the irrelevance of test assets in comparing models. To be clear, we are not questioning their empirical evidence, but rather seek to provide a more rigorous basis for *interpreting* such results. Our analysis highlights important conditions under which model comparison should be based solely on excluded-factor alphas and not on test-asset alphas.

We have already presented a simple nested-model example illustrating the possible pitfalls in comparing test-asset findings across models. In the next section, we provide a non-nested example. The goal here is to further convey what can go wrong, not to identify the definitive *empirical* model.

## 2. A Non-Nested Example

We now consider a comparison of the non-nested models,  $FF3 = \{\text{Mkt SMB HML}\}$  and a four-factor model  $4FM = \{\text{Mkt SMB HML}^m \text{UMD}\}$ .  $\text{HML}^m$  is an alternative version of the value factor due to Asness and Frazzini (2013), which is based on book-to-market rankings that use the most recent monthly stock price in the denominator. This is in contrast to Fama and French (1993), who use annually updated lagged prices in constructing HML. The up-minus-down factor UMD is the momentum factor of Carhart (1997), motivated by the work of Jegadeesh and Titman (1993).

As in the previous section, we nest both FF3 and 4FM in  $M = \{\text{Mkt SMB HML HML}^m \text{UMD}\}$ . Then, using Proposition 1, FF3 can be characterized in terms of zero restrictions on the three-factor alphas of the excluded factors,  $\text{HML}^m$  and UMD, and the five-factor alphas of the test assets. Similarly, 4FM amounts to zero restrictions on HML's four-factor alpha and the five-factor alphas of the test assets. Since the test-asset restrictions are the same for each model, that evidence is irrelevant for comparing the models.

Over the period July 1963 to December 2013, the relevant annualized alpha estimates are 0.56%, 10.85% and 2.20% for  $\text{HML}^m$ , UMD and HML, respectively. The alphas are, of course, zero

for the factors included in each model. Therefore, the average absolute alpha over all five factors is  $(.56 + 10.85)/5 = 2.28\%$  for FF3 and  $2.2/5 = 0.44\%$  for 4FM. Alternatively, the square root of the average squared factor alpha is 4.86% for FF3 and 0.44% for 4FM. These metrics clearly favor 4FM over FF3 (we ignore sampling variation). On the other hand, Table 1 shows measures of performance on test assets for each model and FF3 provides the better fit for 6 of the 10 sets. For example, the annualized average absolute alpha across 17 industry portfolios is 2.22% for FF3 and 2.40% for 4FM. Similarly, the GRS test statistic with these test assets, another basis of comparison that has been used previously, is 3.41 for FF3 and 4.96 for 4FM.

Not surprisingly, when challenged to price the 25 portfolios based on independent size and momentum sorts, the 4FM that includes a momentum factor performs better, with an average absolute alpha of 1.30% compared to 3.90% for FF3. Similar results are obtained with the alphas scaled by average deviations from the cross-sectional mean return, or if the GRS statistic is used. Here, the inability of FF3 to explain momentum drives the excluded-factor alpha evidence as well as the size-momentum test-asset evidence. But for many other portfolios, the test-asset alphas and excluded-factor alphas point to different conclusions. Proposition 1 cuts through these sometimes confusing *impressions*, implying that only the excluded-factor evidence matters for a *relative* test of model comparison.

To reflect common practice in the literature, our example has compared test-asset results relative to the factors in each model. In contrast, the argument for test-asset irrelevance focuses on test-asset restrictions for the *combined* factor model M ( $\alpha_R = 0$ ). As we have emphasized, these restrictions are common to both models. What then causes test-asset information relative to the factors in one model to sometimes provide a misleading indication of its performance relative to another model? The answer can be found in equation (6), specifically the relation  $\alpha_{R1} = \alpha_R + \beta_2 \alpha_{21}$ . This equation shows, for example, that test-asset alphas relative to FF3 equal those for the combined model {Mkt SMB HML HML<sup>m</sup> UMD}, plus a multiple of the alphas of the excluded factors HML<sup>m</sup> and UMD on FF3.<sup>7</sup>

Thus, two kinds of information end up being mixed together in  $\alpha_{r1}$ : *relevant* evidence about the pricing of excluded factors and *irrelevant* (for model comparison) information about the pricing of test

---

<sup>7</sup> The multiple depends on the elements of  $\beta_2$ , which can be positive or negative and can make  $\alpha_{r1}$  larger or smaller than  $\alpha_r$ . If  $\beta_2 = 0$ , the information about  $\alpha_{21}$  is lost.

assets by the set of both included and excluded factors. Examining information that pools what truly does matter with information that does not, ends up providing a “noisy signal” that ultimately can obscure the comparison, rather than clarifying it. For the purpose of evaluating a single model in an absolute test, however, whether we focus on  $\alpha_{21}$  and  $\alpha_R$  or  $\alpha_{21}$  and  $\alpha_{R1}$  does not matter (apart from convenience) since the corresponding restrictions are equivalent by Proposition 1.

### 3. Likelihoods and Estimation Uncertainty

A natural approach to model comparison in the presence of estimation uncertainty relates the likelihoods for each model, given the observed data. This notion plays a fundamental role in classical as well as Bayesian inference. While the development of a detailed methodology of either sort is beyond the scope of this paper, we now apply the well-known Akaike information criterion (AIC), which is often used as a heuristic in model selection, though not in a formal hypothesis test. Let  $L$  be the maximized value of the likelihood function and  $m$  be the number of parameters in a model. To avoid overfitting with too many parameters, AIC incorporates a penalty that increases with  $m$ :

$$\text{AIC} = -2\ln(L) + 2m,$$

where lower values of AIC provide more support for a model. Accordingly, given several candidate models, nested or non-nested, the criterion favors the model with the lowest AIC value (highest adjusted likelihood).

The likelihood function is the joint density for the data, viewed as a function of the model parameters. In our earlier notation for model  $M_1$  with factors  $f_1$  and excluded factors  $f_2$ , we can express that joint density for the factor and test-asset returns as a product of three densities: the unrestricted density for  $f_1$ , the restricted ( $\alpha_{21} = 0$ ) conditional density for  $f_2$  given  $f_1$ , and the restricted ( $\alpha_r = 0$ ) conditional density for the test asset returns  $R$  given  $f_1$  and  $f_2$ . All densities are taken to be multivariate normal in this example, although normality is not required for test-asset irrelevance. It is important to note that this representation of the joint density parallels the alternative characterization of  $M_1$  in Proposition 1, with the restrictions imposed on  $\alpha_{21}$  and  $\alpha_R$ . The corresponding likelihood under  $M_1$  can then be written as

$$L = L_1 \times L_{21} \times L_R.$$

Now let us return to the non-nested example of Section 2, comparing the models FF3 and 4FM. In computing the likelihood for FF3,  $f_1$  is {Mkt SMB HML} and  $f_2$  is {HML<sup>m</sup> UMD}. For 4FM,  $f_1$  is

{Mkt SMB HML<sup>m</sup> UMD} and  $f_2$  is {HML}. The test assets for this illustration are taken to be the 25 portfolios formed on size and momentum. Thus, we have

$$\ln(L^{\text{FF3}}) = 4253.0 + 3258.9 + 42251.5 = 49763.4$$

and

$$\ln(L^{\text{4FM}}) = 5564.5 + 1977.0 + 42251.5 = 49793.0$$

(7)

Note that, for both models, the test-asset portion of the log-likelihood is the same. Therefore,  $L_R$  cancels out in the likelihood ratio for the two models, so it has no effect on the AIC comparison either. This is just our earlier test-asset irrelevance conclusion viewed in terms of the likelihood function.

The number of parameters is 18 for FF3 (3 factor means, 6 factor variances/covariances, 3 residual variances/covariances and 6 betas) and, similarly, 19 for 4FM. The one additional parameter in the latter case reflects the fact that there is one less alpha constraint. Empirically, the AIC values for FF3 and 4FM are -99490.9 and -99548.1, respectively, with a difference of 57.2. The relative (adjusted) likelihood of 4FM to FF3 is, therefore,  $\exp(57.2/2)$ , overwhelmingly in favor of 4FM, largely due to the great difficulty FF3 has in pricing momentum.

Now *suppose we were unaware of the equivalence in Proposition 1 and its implication for test-asset irrelevance*, but still wanted to compare the models above. How would we go about that and would the outcome be the same or different? In this case, we would again calculate the maximum likelihood values, only now employing the usual parameterization for each model and the associated factorization of the joint density of returns. As emphasized earlier, it is essential to evaluate each model's ability to simultaneously price the test assets and any factors excluded from the model. Hence, we start with the unrestricted density for  $f_1$ , as before, but now multiply by the restricted ( $\alpha_{21} = 0$  and  $\alpha_R = 0$ ) joint conditional density for  $(f_2, R)$  given  $f_1$ . The corresponding likelihood under model  $M_1$  can then be written as

$$L = L_1 \times L_{2R} .$$

The associated log-likelihoods are

$$\ln(L^{\text{FF3}}) = 4253.0 + 45510.4 = 49763.4$$

and

$$\ln(L^{\text{4FM}}) = 5564.5 + 44228.5 = 49793.0$$

(8)



identical to the values in (7). This was inevitable, given the well-known invariance of the likelihood function to one-to-one transformation of the parameter space which, in this case, is induced by the different factorizations of the joint density of factor and test-asset returns.

The conclusion for model comparison is that the same outcome is obtained, regardless of which model representation is employed. In the end, both approaches evaluate the likelihood that each model is true, given *all* the returns. Interestingly, though, the fact that test assets have no impact on this model comparison would not be apparent from (8) and the usual way of thinking about asset-pricing restrictions. The reason is that, in this calculation, we only observe the *joint* impact of test assets and the different factors excluded from each model in the term  $L_{2R}$ .

We have illustrated these ideas using the AIC criterion based on *maximum* likelihood estimates. However, similar conclusions apply for model likelihoods calculated with fixed values of the parameters that are not restricted under the given model. Barillas and Shanken (2015) develop this observation further in a Bayesian approach to inference about model comparison.

#### **4. Model Comparison with the Modified HJ Distance**

Using the stochastic-discount-factor (SDF) representation of an asset pricing model, Hansen and Jagannathan (1997) suggest a measure of model misspecification that involves the distance in the standard mean-squared norm between the proposed SDF and the set of valid SDFs. This measure can be used to compare models.

As shown by Kan and Zhou (2004), the HJ distance is closely related to the cross-sectional regression measure introduced in Shanken (1985), except for the manner in which the zero-beta rate is selected. In the excess-return context adopted here, the zero-beta rate is constrained to be the risk-free rate. In this case, Kan and Robotti (2008) suggest a modification to the HJ distance, explaining that this “amounts to requiring all competing SDFs to assign the same price to the risk-free asset, so that we only compare their performances based on their pricing errors on excess returns.” When the factors are traded assets, they further note (footnote 9) that the GRS measure in equation (2) above is a version of the modified HJ ( $HJ_m$ ) distance that imposes the restriction that the traded factors are priced without error. This constraint parallels the standard assumption in the alpha-based framework that factor premia equal the expected factor values.

We have argued that all models should be compared in terms of their ability to price the *same* set of investment returns, both test assets and traded factors. As before, let  $M_1$  be a model with factors  $f_1$ , and let  $f_2$  be the factors excluded from the model. The test-asset returns are denoted by  $R$ . The alpha vector for evaluating the pricing of all these investments under  $M_1$  can be written as  $\alpha_{M_1} = (0, \alpha_{21}, \alpha_{R1})$ , where the alphas of  $f_1$  on  $f_1$  are necessarily 0,  $\alpha_{21}$  refers to the alphas of  $f_2$  on  $f_1$  and  $\alpha_{R1}$  the alphas of  $R$  on  $f_1$ . Using the Kan and Robotti results, the corresponding squared  $HJ_m$  distance is then

$$\alpha_{M_1}' V^{-1} \alpha_{M_1} = (\alpha_{21}, \alpha_{R1})' \Sigma_{M_1}^{-1} (\alpha_{21}, \alpha_{R1}), \quad (9)$$

where  $V$  is the covariance matrix for  $(f_1, f_2, R)$  and  $\Sigma_{M_1}$  is the residual covariance matrix of  $(f_2, R)$  on  $f_1$ . By (2), the right-hand-side of (9) is the difference between the squared Sharpe measure for all the investments and that for just the factors in  $M_1$ ,  $Sh^2(f_1, f_2, R) - Sh^2(f_1)$ .

Now consider another model corresponding to a different subset of the same factors,  $(f_1, f_2)$ . We can derive a similar expression for its  $HJ_m$  distance and the equivalent difference of Sharpe ratios. Note, however, that the weighting matrix  $V$  and the Sharpe measure for all the investments will be unchanged (the set of factors is the same, regardless of which are included or excluded from the models). It thus follows that selecting the model with the lower  $HJ_m$  distance is equivalent to choosing on the basis of the higher model Sharpe measure. Thus, the  $HJ_m$  criterion turns out to be identical to the simple “efficiency criterion” discussed earlier. It follows, of course, that test assets are again irrelevant for model comparison, as can be shown directly by a similar argument.<sup>8</sup>

## 5. Conclusions

It is natural to examine the performance of a given model in pricing test assets and traded factors that are not included in the model. However, the question that we have addressed is how this information should be interpreted for the specific purpose of model comparison. We have obtained a simple, yet unexpected answer to this question. In comparing models on the basis of statistical likelihoods or commonly employed asset-pricing metrics, test assets tell us nothing beyond what we learn from the evidence on the pricing of factors excluded from each model. In particular, when test-

---

<sup>8</sup> The squared GRS measure for  $M_1$  in (9) is the sum of the quadratic form for adding the factors  $f_2$  to  $f_1$  plus that for adding the test assets  $R$  to  $(f_1, f_2)$ . Therefore, the latter quadratic form drops out in comparing two models. In other words, the GRS measure is a sum of separate functions of these excluded-factor and test-asset alphas, so there is no interaction of the sort discussed earlier. Therefore, test assets are irrelevant.

asset performance favors one model, but the excluded-factor evidence favors the other, it is only the latter that is relevant in identifying the better model. The point is that a proper evaluation of a model must consider the *totality* of the test-asset and factor-pricing evidence. But when this is done, the test assets drop out in the model comparison.

## References

- Avramov, Doron and John Chao, 2006, An exact Bayes Test of Asset Pricing Models with Application to International Markets, *Journal of Business* 79, 293-323.
- Barillas, Francisco and Jay Shanken, 2015, Comparing Asset Pricing Models, Working paper, Emory University.
- Black, Fisher, Michael C. Jensen and Myron Scholes, 1972, The Capital Asset Pricing Model: Some Empirical Tests, in *Studies in the Theory of Capital markets*. Michael C. Jensen, ed. New York: Praeger, pp. 79-121
- Carhart, Mark, 1997, On Persistence in Mutual Fund Performance, *Journal of Finance* 52, 57–82.
- Cochrane, John, 2005, Asset Pricing, Princeton University press.
- Fama, Eugene F., 1996, Multifactor Portfolio Efficiency and Multifactor Asset Pricing, *Journal of Financial and Quantitative Analysis* 31, 441-65.
- Fama, Eugene F., 1998, Determining the Number of Priced State Variables in the ICAPM, *Journal of Financial and Quantitative Analysis* 33, 217-31.
- 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 Kenneth R. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance* 51, 55-84.
- Fama, Eugene F., And Kenneth R. French, 2015a, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1-22.
- Fama, Eugene F., and Kenneth R. French, 2015b, Dissecting anomalies with a five-factor model, Manuscript, Booth School of business, University of Chicago.
- Gibbons, M., Ross, S., Shanken, J., 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121-1152.
- Hansen, L. P., and R. Jagannathan, 1997, Assessing Specification Errors in Stochastic Discount Factor Models." *Journal of Finance*, 52, 557-590.
- Kewei Hou, G. Andrew Karolyi and Bong-Chan Kho, 2011, What Factors Drive Global Stock Returns? *Review of Financial Studies* 24, 2527-2574.

- Hou, Kewei, Xue Chen, and Zhang, Lu, 2015, Digesting Anomalies: An Investment Approach, *Review of Financial Studies* 28, 650-705.
- Hou, Kewei, Xue Chen, and Zhang, Lu, 2015, A Comparison of New Factor Models, Working paper.
- Jegadeesh, Narasimhan and Sheridan Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance*, 48, 65-91.
- Jensen, Michael C., 1968, The Performance of Mutual Funds in the Period 1945-1964, *Journal of Finance* 23, 389-416.
- Jobson, D. and R. Korkie, 1982, Potential performance and tests of portfolio efficiency, *Journal of Financial Economics* 10, 433-466.
- Kan, R., and C. Robotti, 2008, Specification Tests of Asset Pricing Models Using Excess Returns." *Journal of Empirical Finance* 15, 816-838.
- Kan, R., and G. Zhou, 2004, Hansen-Jagannathan Distance: Geometry and Exact Distribution." Working paper, University of Toronto and Washington University in St. Louis.
- Kan, Raymond, Cesare Robotti and Jay Shanken, 2013, Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology, *Journal of Finance*, 68, 2617-2649.
- Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics*, 47, 13-27.
- Pastor, Lubos and Robert Stambaugh, 2002, Mutual Fund Performance and Seemingly Unrelated Assets, *Journal of Financial Economics* 63, 315–349.
- Shanken, Jay, 1985. Multivariate tests of the zero-beta CAPM, *Journal of Financial Economics*, 14, 327–348.
- Sharpe, F. William, 1964, Capital asset prices: a theory of market equilibrium under conditions of risk, *Journal of Finance*, 19, 425-442.
- Sharpe, F. William, 1984, Factor models, CAPMs, and the APT, *Journal of Portfolio Management*, 11.1, 21-25.

Table 1

**Statistics for FF3 and 4F models: July 1963 to December 2013.**

This table presents statistics for the three-factor Fama-French (1993) model versus a four-factor model that adds the momentum factor UMD and replaces HML with a more timely version (HML<sup>m</sup>). The table shows (i) the factors in each regression model, (ii) the annualized average absolute value of the intercepts  $A|a_i|$ , (iii)  $A|a_i|/A|r_i|$ , the average absolute value of the intercepts over the average absolute value of the average return deviation, which is the average return on portfolio  $i$  minus the cross-sectional average of the time-series average portfolio returns, (iv) the GRS F-statistic testing whether the true intercepts are zero and (v) the p-value for the GRS statistic.

<b>Model Factors</b>	<b><math>A a_i </math></b>	<b><math>A a_i /A r_i </math></b>	<b>GRS</b>	<b>p(GRS)</b>
<b>25 Size-B/M</b>				
Mkt SMB HML	1.21	0.53	3.61	0.0000
Mkt SMB HML <sup>m</sup> UMD	1.30	0.57	4.00	0.0000
<b>17 Industries</b>				
Mkt SMB HML	2.22	2.18	3.41	0.0000
Mkt SMB HML <sup>m</sup> UMD	2.40	2.37	4.96	0.0000
<b>25 Size-UMD</b>				
Mkt SMB HML	3.90	1.13	4.72	0.0000
Mkt SMB HML <sup>m</sup> UMD	1.30	0.41	3.78	0.0000
<b>25 Size-Inv</b>				
Mkt SMB HML	1.31	0.63	4.48	0.0000
Mkt SMB HML <sup>m</sup> UMD	0.94	0.44	3.35	0.0000
<b>25 BM-Inv</b>				
Mkt SMB HML	1.31	0.83	1.96	0.0037
Mkt SMB HML <sup>m</sup> UMD	1.82	1.17	2.37	0.0002
<b>25 Size-OP</b>				
Mkt SMB HML	1.31	0.68	2.40	0.0002
Mkt SMB HML <sup>m</sup> UMD	1.36	0.71	2.94	0.0000
<b>32 Size-B/M-OP</b>				
Mkt SMB HML	1.84	0.61	2.52	0.0000
Mkt SMB HML <sup>m</sup> UMD	1.96	0.65	3.36	0.0000
<b>32 Size-B/M-INV</b>				
Mkt SMB HML	1.54	0.63	2.75	0.0000
Mkt SMB HML <sup>m</sup> UMD	1.74	0.72	3.28	0.0000
<b>32 Size-Prof-INV</b>				
Mkt SMB HML	2.20	0.80	4.36	0.0000
Mkt SMB HML <sup>m</sup> UMD	1.90	0.69	3.49	0.0000
<b>16 Size-B/M-Prof-INV</b>				
Mkt SMB HML	1.66	0.62	4.88	0.0000
Mkt SMB HML <sup>m</sup> UMD	1.61	0.60	4.96	0.0000