NBER WORKING PAPER SERIES

PRICING WITHOUT MISPRICING

Tobias J. Moskowitz
Robert F. Stambaugh

Working Paper 29016
http://www.nber.org/papers/w29016

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2021, Revised September 2025

We gratefully acknowledge the contribution of Jianan Liu, who was an earlier coauthor but bears no responsibility for any shortcomings of the current paper. We are also grateful for comments from Nick Barberis, Martijn Boons, Gene Fama, Peter Nyberg, Lasse Pedersen, and seminar participants at AQR, IDC, TU Munich, Wharton, and Yale. AQR Capital is a global investment firm that may or may not use the insights in this paper. The views expressed here are those of the authors and not necessarily those of AQR Capital or the National Bureau of Economic Research.

Pricing Without Mispricing
Tobias J. Moskowitz and Robert F. Stambaugh
NBER Working Paper No. 29016
July 2021, Revised September 2025
JEL No. G12, G14, G40

ABSTRACT

We test whether a model could price assets if the market were efficient. Specifically, we test whether a model assigns zero alpha to a strategy that uses only decade-old information, which even an inefficient market would correctly price. Persistence in the strategy's multifactor betas gives our test power. Multifactor betas can help capture mispricing, but persistence in those betas then leads the multifactor model to distort expected returns well after that information gets priced correctly. The CAPM passes our test, but prominent multifactor models do not.

Tobias J. Moskowitz
Yale University
Yale School of Management
and NBER
tobias.moskowitz@yale.edu

Robert F. Stambaugh
University of Pennsylvania
The Wharton School
Finance Department
and NBER
stambaugh@wharton.upenn.edu

1. Introduction

The joint-hypothesis problem posed by Fama (1970) is central to empirical investigations of market efficiency. Fama (1991) expresses the inherent challenge as follows:

> *We can only test whether information is properly reflected in prices in the context of a pricing model that defines the meaning of "properly."*

In this study, we test whether a model can be the one correctly defining "properly." Such a model offers the best benchmark for assessing market efficiency and gauging mispricing. Models that capture expected returns empirically could be identifying mispricing in an inefficient market or compensated risks in an efficient market.[1] Which pricing models, among prominent candidates, could not describe expected stock returns if there were no mispricing? That is the question we address in this paper.

To address the joint hypothesis problem, tests of market efficiency traditionally seek to control for the equilibrium model of expected returns or identify situations where the underlying equilibrium model differences out (such as examining deviations from the law of one price, e.g., Lamont and Thaler, 2003, Du, Tepper, and Verdelhan, 2018, Hu, Pan, and Wang, 2013). Conversely, we seek to avoid having market inefficiency confound tests of asset pricing models.

To implement this idea empirically, we attempt to plausibly identify when mispricing is absent. Our key assumption is that any mispricing of currently available information about a stock gets corrected in less than ten years. Regardless of how investors interpret information or how arbitrage is impeded, we assume the forces of learning and arbitrage are strong enough over 10 years to correct any mispricing. In U.S. equity markets, a 10-year horizon seems, if anything, a conservative assumption. Under this assumption, if a long-short spread based on decade-old information produces a significant alpha with respect to a given pricing model, that alpha would not reflect mispricing. Instead it would reveal that the model does not describe expected stock returns in the absence of mispricing. To be clear, our assumption is that decade-old information is not mispriced, but this does not mean a stock could not

[1]Prominent empirical models in asset pricing have received numerous interpretations from both sides of the market efficiency debate. An inexhaustive list of rational and behavioral theories includes Fama and French (1993, 1995, 2015), Gomes, Kogan, and Zhang (2003) , Zhang (2005), Li, Livdan, and Zhang (2009), Belo (2010), Li and Zhang (2010), Liu and Zhang (2008), Berk, Green, and Naik (1999), Johnson (2002), Sagi and Seasholes (2007), Liu, Whited, and Zhang (2009), and Hou, Xue, and Zhang (2015) on the rational, no mispricing side and Lakonishok, Shleifer, and Vishny (1994), Daniel and Titman (1997), Daniel, Hirshleifer, and Subrahmanyam (1998), Barberis, Shleifer, and Vishny (1998), Hong and Stein (1999), and Stambaugh and Yuan (2017) on the behavioral, mispricing side.

be mispriced for 10 years or longer. Such a scenario would simply reflect the mispricing of information that arrived more recently than a decade earlier.

In many ways, finding settings where information must surely be reflected in prices is an easier task than identifying the right pricing model (or finding deviations from the law of one price). While our plausible (and conservative) assumption avoids mispricing, a potential drawback is low power. Tests where we can confidently rule out mispricing may also be tests that have low power to discriminate among pricing models. For instance, while decade-old information should be reflected correctly in today's prices, much of that information may relate little to the current values of multifactor betas that pricing models either omit or include. The betas of long-short spreads based on such information could then be essentially zero, as if stock positions were chosen randomly. Our test would then have little power to discriminate among models. Thus, we aim to find information among a set of test assets that is plausibly not mispriced and is related to the current values of betas in popular asset pricing models.

The decade-old information with which we illustrate our test is a return forecast based on the well-known 3-factor model of Fama and French (1993), hereafter FF3. FF3 includes factors based on size and value, which have long been recognized as fundamental stock characteristics in both academic research and investment practice. The FF3 expected return gives a natural univariate measure incorporating those familiar properties while avoiding arbitrary selections of stock characteristics from the ever-expanding "factor zoo."

For each pricing model, our test's null hypothesis is that the model assigns zero alpha to investment strategies formed using only decade-old information. For a long-short strategy based on stocks' decade-old FF3 predictions, our test does not reject the CAPM but does reject prominent multifactor models: the three-, five- and six-factor models of Fama and French (1993, 2015, 2018), the four-factor model of Hou et al. (2015), and the five-factor model of Hou, Mo, Xue, and Zhang (2021). Not only does the CAPM pass our test, but so do various extensions augmenting that model with a single factor, such as the betting-against-beta (BAB) factor of Frazzini and Pedersen (2014)[2] and the liquidity factor of Pástor and Stambaugh (2003). Even a model with no factors, implying equal expected returns across assets, passes our test. Certainly our approach is not powerful enough to identify a unique model as the "proper" no-mispricing benchmark for tests of market efficiency. Nevertheless,

[2]A positive premium on the latter factor, theoretically motivated, implies positive (negative) CAPM alphas on assets with low (high) betas. Black (1972) and Fama (1976) provide earlier theoretical motivations for augmenting the CAPM with such a beta factor, and Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) provide related evidence.

the overall message seems clear: simpler is better. The prominent models we consider with three or more factors seem unsuitable for the role of equilibrium pricing in an efficient market, at least when even an inefficient market would price decade-old information correctly. Hence, our test can rule out various models as the proper no-mispricing benchmark, but it cannot narrow the field to just one model.

Our results admit a simple explanation. Multifactor betas can help capture stocks' actual expected returns, which can include mispricing. Some of those multifactor betas, however, can persist longer than the stocks' mispricing. If a multifactor beta gets included in a pricing model to help explain expected returns that in part reflect mispricing, then persistence in that beta injects unhelpful distortions into stocks' implied expected returns after the mispricing corrects. When a security's mispricing is more transitory than its multifactor betas, the latter will distort the security's expected return when mispricing is absent. By excluding those multifactor betas, the CAPM and other simple models avoid such distortions.

With this premise in mind, we conduct our tests separately within different firm size segments, for a number of reasons. Consider the largest firms, for example. Our assumption that mispricing disappears within 10 years seems especially plausible for the largest stocks, which are liquid and well-monitored. Large stocks should be those least likely to have persistent market frictions impeding price correction over periods as long as a decade. Correctly pricing large stocks thus seems the purest test for any proper benchmark model to clear. Moreover, our test relies on decade-old information being related to current multifactor betas, and betas seem likely to be more stable over time for the largest firms. We empirically confirm that larger stocks have more persistent and more precise multifactor betas. The type of securities affording our test the most power are those with 1) no mispricing based on 10-year old information and 2) persistent and reliable multifactor betas. The largest firms embody these characteristics.

Consider instead the smallest stocks. Because they are less heavily followed and face greater market frictions, small stocks are probably subject to the greatest misprising. In that sense they present salient targets for multifactor models aiming to fit average short-run returns. Small stocks could therefore offer our test an especially informative sample for distinguishing such models from the proper no-mispricing benchmark. Offsetting this potential advantage, however, is that multifactor betas are less stable and less reliable for small stocks, making the link between decade-old information and current betas weaker. Because determinants of our test's informativeness could differ across firm size, we allow each segment to speak separately in the data.

We find that results differ meaningfully across firm size. When run on just the smallest stocks (below the 20th NYSE percentile), our test generally does not reject any of the models considered. When run on the remaining stocks, even just the largest ones, the test generally rejects the models with three or more factors as being the no-mispricing benchmark. The rejections for even the largest firms provide compelling support for the notion that multifactor models pick up temporary mispricing that then distorts expected returns that should only reflect underlying risk premia. In other words, the largest firms would seem to present the strongest test of what model should be the proper benchmark for addressing questions about market efficiency (Fama, 1991).

Models failing our hurdle for a no-mispricing benchmark can perform well by the usual empirical metrics that assess a pricing model's ability to describe expected returns conditional on any available information. Those models potentially capture mispricing well, and some of them are even cast explicitly in those terms (e.g., Stambaugh and Yuan, 2017). Such models can be quite useful in designing trading strategies, for example, but they are less useful for providing the benchmark model of expected returns to which "properly" applies in the sense of Fama (1991).

Our test and insights are very different from the recent literature on the persistence of stock characteristics and the predictability of asset returns over various horizons (Keloharju, Linnainmaa, and Nyberg, 2020 and Baba-Yara, Boons, and Tamoni, 2024). We are not interested in, nor have anything to say about, the best model for expected long-term returns. Rather, our goal is simply to find models that satisfy the no-mispricing benchmark of Fama (1991), which is a model for short-term expected returns in the absence of mispricing. Even the CAPM, which passes our test for such a model, could be a bad model for long-term expected returns in the absence of mispricing. For example, if market betas converge in the long-run, then the CAPM with current betas could be a bad long-term no-mispricing benchmark, while a no beta/factor model could be a better long-term model. All of this is to say that our objective is very different from the literature on short- versus long-term return predictability. The fact that we use 10-year-old information to implement our test creates an apparent connection to this literature but is fundamentally unrelated. Had we come up with information unrelated to horizon, potentially correlated with multifactor betas, and still reasonably assumed to be correctly priced (we welcome suggestions!), the apparent connection to this literature would disappear.

The rest of the paper is organized as follows. Section 2 motivates our test and discusses its power. Section 3 describes our empirical setting and presents test results for various pricing models. Section 4 concludes.

2. A simple test of pricing without mispricing

In the absence of mispricing, the proper benchmark asset pricing model should deliver zero alpha for any investment strategy. This implication motivates the canonical test of market efficiency, but it is plagued by the joint hypothesis problem identified by Fama (1970): any test of efficiency is a joint test of the pricing model. The proper model for testing efficiency is one that characterizes prices in an equilibrium wherein (public) information is correctly reflected in prices. Such an equilibrium may be just hypothetical in a market with noise traders or other frictions that hinder information from being incorporated immediately into prices. Nevertheless, the "proper" model is the one relevant to the joint hypothesis problem posed by Fama (1970). Under this view, non-zero alphas imply mispricing or an incorrect model, while zero alphas imply no mispricing or, again, an incorrect model. In the latter case, the incorrect model captures mispricing but not expected returns in the absence thereof. In other words, the model is not a useful empirical benchmark for gauging the extent of market inefficiencies.

Rather than investigate potential market inefficiencies, we explore the other side of the joint hypothesis. That is, we ask which asset pricing models could capture expected returns in the absence of mispricing. To do so, we assume that decade-old information is correctly reflected in prices today, however inefficient the market may be in reflecting current information. Specifically, we test the abilities of various models to produce zero alpha for strategies based on decade-old information. Such a test, we argue, is unlikely to be contaminated by information inefficiency.[3]

The test's inherent challenge is achieving power. We seek to discriminate among pricing models that include or omit various factors. For many assets, however, current betas on those factors may have little or no relation to decade-old information. Our test relies on finding assets and information for which the relation between the assets' current betas and lagged information is strong enough to provide power. We expect the greatest power to be

[3] As pointed out earlier, this assumption does not mean a security cannot be mispriced for 10 years or more, but rather any mispricing associated with any security cannot come from information that has been around for at least 10 years. In other words, if a security is persistently mispriced even beyond 10 years, that mispricing must come from more recent information that is less than 10 years old. A security could persistently underreact to its earnings news forever and be consistently mispriced, however that security's mispricing is not related to earnings news from 10 years ago, but rather more recent earnings news.

offered by assets most likely to have stable betas.

To understand better the nature of our test and its potential power, let f_t denote the vector of returns in month t on the factor portfolios associated with a given asset pricing model. Consider the time-series regression,

$$r_t = \alpha + \beta' f_t + \epsilon_t, \tag{1}$$

where r_t is the excess return in month t on an investment strategy whose portfolio weights in month t are determined by sorting a given universe of assets, A, on the values in a vector x of an asset-specific information variable observed (publicly) in month $t - \tau$. The value of α depends on the identity of the information, x, its lag length, τ, the set of factors, f, and the asset universe, A. We represent this dependence via the functional notation,

$$\alpha = \alpha(x, \tau, f, A). \tag{2}$$

Let f^* denote the factors in the proper pricing model that captures expected return in the absence of mispricing. Our test assumes

$$\alpha(x, 120, f^*, A) = 0 \tag{3}$$

for any information x. If the market is efficient, then $\alpha(x, 1, f^*, A) = 0$ for any x. In an inefficient market, various choices of x can make $\alpha(x, 1, f^*, A) \neq 0$.[4]

A model can fail as the no-mispricing benchmark by omitting a relevant factor or by including an irrelevant one. Omitting a relevant factor from a pricing model is a possibility long recognized in the literature. In an inefficient market, however, there is a distinct possibility that models are developed with factors that are irrelevant in the absence of mispricing. This is the key point of our test and one that has not received much attention in the literature. The factors in prominent pricing models are generally returns on portfolios formed using recent information, unavailable before month $t - 1$. In an inefficient market, a set of such factors, $\tilde{f}$, can be useful in making

$$\alpha(x, 1, \tilde{f}, A) \approx 0 \tag{4}$$

for various choices of x. To reject such a model as being the correct no-mispricing benchmark, our test relies on detecting

$$\alpha(\hat{x}, 120, \tilde{f}, \hat{A}) \neq 0, \tag{5}$$

[4]Or, more generally, in an efficient market $\alpha(x, \tau, f^*, A) = 0$ and in an inefficient market various choices of x can make $\alpha(x, \tau < 120, f^*, A) \neq 0$.

for a given asset universe, $A = \hat{A}$, and a given $x = \hat{x}$.

Suppose that the model is not the no-mispricing benchmark because $\tilde{f}$ incorrectly includes or excludes a given factor, f_j. Let $\beta_j(\hat{x}, \tau, \tilde{f}, \hat{A})$ denote the factor's corresponding element of β in equation (1) if $\tilde{f}$ includes f_j. Otherwise, let $\beta_j(\hat{x}, \tau, \tilde{f}, \hat{A})$ denote the factor's slope coefficient in that regression when $\tilde{f}$ is augmented by f_j. A necessary condition for (5) to obtain, and thus for our test to have power, is that that

$$\beta_j(\hat{x}, 120, \tilde{f}, \hat{A}) \neq 0. \tag{6}$$

The nonzero beta in (6), when multiplied by the non-zero mean (factor premium) of f_j, then represents the component of expected return that is incorrectly included or excluded, resulting in the nonzero alpha in (5). The challenge faced by our test, noted earlier, is that for many asset universes and information choices, there may be essentially no relation between individual assets' decade-old values in $\hat{x}$ and their current betas on f_j. In that scenario, a long-short strategy based on those values in $\hat{x}$ is likely to produce $\beta_j(\hat{x}, 120, \tilde{f}, \hat{A}) = 0$, the same as one would expect if the assets in the long and short legs were just randomly selected. Settings more likely to avoid this no-power scenario are those in which $\beta_j(\hat{x}, 1, \tilde{f}, \hat{A}) \neq 0$ and the individual assets' betas on f_j are stable over time. The latter condition, in turn, seems more likely when the selected universe, $\hat{A}$, comprises assets tending to have more stable factor betas.

3. Empirical setting and test results

As explained earlier, for the information vector $\hat{x}$ we use $\hat{E}$, containing assets' return forecasts from the FF3 model. Specifically, for stock i in month t, we compute the decade-old quantity $\hat{E}_{i,t-120} = \bar{f}'_{t-120}\tilde{\beta}_{i,t-120}$, where $\bar{f}_{t-120}$ contains the historical averages of the FF3 factors through month $t-120$, and $\tilde{\beta}_{i,t-120}$ contains ten-year-ahead forecasts of the stock's factor betas. We construct $\tilde{\beta}_{i,t-120}$ as follows. For each factor k and month $t-120$, we first run a cross-sectional multiple regression in which the i-th observation of the dependent variable is $\hat{\beta}_{i,k,t-120}$ and the three independent variables are $\hat{\beta}_{Ind(i),j,t-240}, j = 1, 2, 3$, where $\hat{\beta}_{Ind(i),j,t-240}$ is the average value of $\hat{\beta}_{n,k,t-240}$ for all stocks n whose industry classification in month $t-240$ is the same as that of stock i in month $t-120$. Each $\hat{\beta}_{i,k,s}$ is estimated using months $s-35$ through s, and the 49 industry classifications follow Fama and French (1997). We then compute the three elements of $\tilde{\beta}_{i,t-120}$ by applying the coefficients from the corresponding multiple regressions to the values of $\hat{\beta}_{Ind(i),j,t-120}, j = 1, 2, 3$.

The above procedure for computing the values of $\hat{E}_{i,t-120}$ is applied separately within three segments of market capitalization. Using NYSE market-cap percentiles as break points, and using market values at the end of month $t-120$ (so all information is 10-years old), we form three subsamples consisting of all NYSE, AMEX, and NASDAQ stocks (i) above the 70th percentile, (ii) between the 70th and 20th percentiles, and (iii) below the 20th percentile. For labeling ease, we denote these segments as "large," "medium," and "small," while recognizing that other terms can be applied. For example, one might instead use "microcaps" to label the stocks below the 20th percentile, following Fama and French (2008).[5]

Computing $\hat{E}_{i,t-120}$ using decade-old industry betas, rather than individual-stock betas, allows us to include more stocks in our test portfolios. Sorting in month t on $\hat{E}_{i,t-120}$ can be applied, at best, only to stocks that also existed in month $t-120$, an age requirement that already reduces the sample nontrivially. Using individual-stock betas to compute $\hat{E}_{i,t-120}$ would further reduce the sample, as the stocks eligible to be sorted would then be only those that also existed for an earlier N-month period ending in month $t-120$, where N is the length of the period used to estimate betas (36 months in our case). Computing $\hat{E}_{i,t-120}$ using $\tilde{\beta}_{i,t-120}$, i.e., month-t betas predicted ten years earlier, rather then simply using the industry betas in month $t-120$, allows $\hat{E}_{i,t-120}$ in part to reflect persistence in betas, which is the source of our test's power.

Table 1 reports the average value of R-squared in multiple regressions of month-t individual stock betas on decade-old individual stock betas, in Panel A, and on decade-old industry-average betas, in Panel B. For comparability, the samples in both panels are limited only to stocks for which the results in Panel A can be computed, meaning stocks existing in month t that also have a 36-month sample ending in month $t-120$.

One consistent result in Table 1 is that the R-squared values are higher for large stocks. In part, this result likely reflects small stocks' greater volatility, and thus noisier betas, but the result is also consistent with large stocks having more persistent betas. By virtue of their size, large firms are likely to have greater inertia, taking longer to change course in ways that would significantly impact their characteristics relevant to factor exposures. (Aircraft carriers turn more slowly than destroyers.) Greater persistence in large-cap betas is likely to endow our test with more power within that segment. Our test results, discussed later, are consistent with that interpretation. Table 1 also reveals that using industry betas, which allows our test to include more stocks, appears to sacrifice little, if any, explanatory power

[5]Stocks between the 70th and 20th NYSE percentiles essentially combine categories that are often denoted as "mid-cap" and "small-cap." In that respect, readers should not misinterpret our "small" labeling.

relative to using stock-specific betas. In fact, the largest R-squared (0.139), occurs when forecasting ten-year-ahead HML betas using lagged industry-average FF3 betas.

To conduct our main test, we compute return spreads between the bottom and top 20% of stocks sorted by $\hat{E}_{i,t-120}$, value weighting the stocks in each leg. The value weighting uses market caps as of month $t-120$, so that only decade-old information is used to construct the long-short spread. Of course, we use the information that a stock exists in the most recent month, t, but that information is identical for all stocks in the strategy. The designation of the bottom 20% of $\hat{E}_{i,t-120}$ as the long leg is motivated by the possibility that part of $\hat{E}$ could reflect the FF3 model's ability to capture mispricing of current information. If so, and if that mispriced information eventually gets priced correctly within a decade, then a high (low) value of $\hat{E}_{i,t-120}$ would predict too high (low) a return in month $t+1$. Thus, the long-short strategy would have a positive alpha with respect to FF3 and other models that also help capture whatever mispricing FF3 does.

We apply our test to 11 different pricing models. The first five models, which contain at least three factors, include the three-, five- and six-factor models of Fama and French (1993, 2015, 2018), denoted FF3, FF5 and FF6, the four-factor model of Hou et al. (2015), denoted Q4, and the five-factor model of Hou et al. (2021), denoted Q5. The next five models include the CAPM of Sharpe (1964) and Lintner (1965) plus four models that augment the CAPM with a single factor: the betting-against-beta (BAB) factor of Frazzini and Pedersen (2014), a 12-month momentum factor (UMD from Fama and French (2018)), the traded liquidity factor (LIQ) of Pástor and Stambaugh (2003), and the size factor (SMB) of Fama and French (1993). Finally, we include a model with no factors, which simply equates expected returns across all stocks.

We form spreads separately within the large-, medium-, and small-cap segments as well as within the whole universe and just the largest 200 ("mega-cap") stocks. The first five columns of Table 2 report the alphas and t-statistics with respect to each pricing model across mega-cap, large, medium, small, and all stocks categories. The last column reports the p-value for a test of whether the alphas for the disjoint size segments (large, medium, small) are jointly equal to zero.

The first row of Table 2 reports alphas for the FF3 model across the size groups. The first column reports a 41 basis point (bp) alpha for the mega-cap stocks (largest 200) with a statistically significant 2.62 t-statistic (using White (1980)-corrected errors), indicating that the FF3 model yields significant abnormal returns for a portfolio of mega-cap stocks sorted on 10-year old information, where there should be no mispricing. This non-zero alpha indicates that the FF3 model is not a proper benchmark for equilibrium asset pricing in the sense of Fama (1991), if decade-old information represents the absence of mispricing.

The remaining columns report alphas with respect to the FF3 model for the other size groups. For large-cap stocks, the alpha is a significant 37 bps (t-statistic = 2.70), for medium cap stocks it is a marginally significant 22 bps (t-statistic = 1.71), and for small-cap stocks the alpha is an insignificant -11 bps. These results are consistent with betas being more precise and more persistent among larger stocks: small stocks have noisy and unstable betas, and using 10-year-old information to forecast them results in noise. The portfolio sort then yields no reliable differences in current loadings on the factors. These findings are supported by the results in Table 1 that show little persistence in betas for small caps. The second-to-last column reports the alpha for all stocks in the investment universe, which is 39 bps (t-statistic = 3.17). Finally, the last column reports the p-value of a Gibbons, Ross, and Shanken (1989) F-test (with a White (1980) heteroskedasticity-consistent covariance matrix) of whether the alphas of the large, medium, and small-cap segments are jointly zero. The p-value of 0.0143 indicates that the test for alphas being jointly zero is rejected, confirming that the FF3 model produces significant alphas when mispricing is absent.

The next four rows report the analysis for the FF5, FF6, Q4, and Q5 models, who each produce positive alphas (of 22 to 38 bps) among medium- and large-cap stocks, and zero alphas among small caps. The GRS (1989) joint test of alphas being zero across the size segments rejects for all four models, indicating that each of these multifactor models, which are prominently used in the asset pricing literature, are likely not good candidates for the no-mispricing benchmark of Fama (1991).

The next row reports results for the single factor CAPM. Here, the GRS (1989) test that the alphas are jointly zero fails to reject (p-value = 0.2337). Hence, the CAPM passes our test and could be a viable candidate for the equilibrium asset pricing model in the absence of mispricing. However, our test cannot tell whether the CAPM is the right no-mispricing benchmark, just that it does not fail our test of pricing decade-old information, as the other multifactor models do. This is an important distinction. Our test is useful in ruling out the multifactor models as viable candidates for the model of expected returns in an efficient market with no mispricing. However, our test does not identify what that model is.

To illustrate the point, consider various modifications of the CAPM. The next four rows of Table 2 report results from our test applied to the CAPM plus the addition of one other factor: the betting-against-beta (BAB) factor of Frazzini and Pedersen (2014), a 12-month momentum factor (UMD from Fama and French (2018)), the traded liquidity factor (LIQ) of Pástor and Stambaugh (2003), and the size factor (SMB) of Fama and French (1993). All of these modifications to the CAPM also pass our test and hence are consistent with being the proper benchmark in an efficient market. But, we have no power to detect which of these models best fills that role. Take the CAPM + MOM model, for example. This model passes our test, but a substantial literature links MOM to mispricing, suggesting CAPM + MOM is unlikely to be the no-mispricing benchmark. Our test has no power to reject CAPM + MOM if MOM reflects mispricing, because our test uses decade-old information to implement the no mispricing condition. Decade-old information has little to say about current momentum, which is a relatively short-term characteristic having little predictability beyond a year. A model of the CAPM plus short-term reversals would face the same issue. In fact, the CAPM plus noise would also satisfy our test, but it would not be a good description of equilibrium returns in an efficient market. To distinguish among these models, we would need a test with more power, e.g., a test among securities immune to MOM. In other words, if one could come up with other plausible conditions where mispricing is absent and there is power to detect the potential influences of MOM (or BAB, LIQ, SMB), then we could conduct similar tests and possibly distinguish among these models.

Finally, the last row of Table 2 reports results from our test on a no-factor model. Our test fails to reject this model as well, which is to say that sorting on 10-year old predictions from the FF3 model produces no reliable differences in average raw returns. Despite passing our test, the no factor model has little theoretical appeal as a model of short-run expected return in an efficient market. The no-factor model could, however, be a good description of expected long-horizon returns in a CAPM world if betas are expected to compress in the long run. In general, our test has little to say about long-term expected return models and therefore does not have much relevance for the recent literature on short- versus long-term signals in asset pricing.[6] The fact that we use decade-old information as our condition for ruling out mispricing may seemingly provide a link, but our motivation is wholly different

[6]Having said that, our framework may provide some additional clarity to the results from the literature. For example, Cho and Polk (2020) find that the CAPM describes the cross-section of prices better than it describes expected short-horizon returns. Since price levels of stocks depend more on discount rates applied in the long run, if mispricing gets corrected in the short run, the cross-section of prices will be less affected by mispricing, which is consistent with our results on the CAPM, too. Keloharju, Linnainmaa, and Nyberg (2020) find that persistent differences in firm characteristics do not predict stock returns, a result that they argue is consistent with long-term expected returns not varying across stocks. Our findings suggest that the reason persistent characteristics fail to have long-term return consequences is perhaps because they help identify mispricing, which dissipates faster than the characteristics change. Baba-Yara, Boons, and Tamoni (2024) examine the returns to 56 characteristic-sorted portfolios using the most recent information about a characteristic and its lag (up to five years). They find that 2/3 (1/3) of the characteristics contain more return predictability from newer (older) information and that the CAPM does better at pricing older sorts, while popular multifactor models do better at pricing newer sorts, consistent with our results and the assumption that older information is not mispriced.

and we have little to say about short- versus long-term predictability. Put differently, had we found another condition or assumption, besides decade-old information, where mispricing is plausibly absent, there would be no seeming link to the literature on short- versus long-term expected returns or the persistence of signals.

Our test rejects each of the models with three or more factors as being the no-mispricing benchmark model, and it fails to reject the CAPM and two-factor variations of the CAPM. To understand better the source of these rejections (and lack of rejections), Table 3 decomposes the alphas reported in Table 2. Specifically, the alpha at each lag is decomposed into the average return spread minus the product of the spread's estimated beta times the sample-average factor premium for each factor in the model.

The first set of results in Table 3 pertain to the FF3 model. Rejection of the model for mega- and large-cap stocks is driven by the persistence in market beta and HML beta, that when multiplied by their return premia result in unhelpful variation in average returns that generates the positive alpha leading to rejection. The raw return spread between the long and short legs of the portfolio sorted on decade-old return predictions for mega- and large-cap stocks is 9 bps (statistically insignificant from zero). But when these returns are regressed on the FF3 model, negative betas on the market and HML show up, which when multiplied by the market and HML premia, result in a more positive alpha. Specifically, for mega caps, this exposure adds 31 more bps of returns to the 9 bp spread, resulting in a significant 41 bp alpha that leads to the FF3 model's rejection. For large caps, these residual exposures add 28 bps of return to yield a significant 37 bp alpha. For small stocks, there is no reliable factor exposure, consistent with small stocks having less persistent and less reliable betas, and therefore no power to reject.

The story is nearly identical for the FF5 and FF6 models: persistent exposure to the market and HML among large stocks long after mispricing has abated leads to unhelpful factor exposure a decade later, inflating the alpha and leading to rejection of the model. For the Q4 and Q5 models, which replace the FF factors with ME, IA, ROW and (in the case of Q5) EG, all of the action is driven by the IA factor instead of HML. Persistent exposure to IA a decade later inflates the zero alpha raw return spread by 12 to 16 bps when adjusting returns for these models.

Interestingly, the FF3, 5, and 6 models as well as the Q4 and Q5 models are motivated by and couched in terms of equilibrium risk premia consistent with an efficient market, and hence have been suggested as candidates for the proper benchmark in the sense of Fama (1991). Our test's rejection of these models casts doubt on that interpretation, suggesting that these models are not good descriptions of expected returns in the absence of mispricing. The fact that these models fail to price the largest, most liquid stocks in particular is even further testament to that claim. Such stocks should be subject to the least mispricing, especially based on decade-old information, and they have the most reliable and stable factor exposures, giving our test the most power.

The remaining entries in Table 3 pertain to the CAPM and its variants, where we don't find rejection of the models. Hence, these models pass our test and could be viable candidates for the proper equilibrium benchmark in an efficient market per Fama (1991). However, as discussed previously, this does not necessarily mean they provide good descriptions of equilibrium returns with no mispricing. We would need more powerful tests to find the best model among the set of candidate models that pass our test.

4. Conclusion

In an efficient market, public information is properly reflected in prices, but assessing efficiency rests on having a pricing model that defines "properly" (Fama, 1970, 1991). We investigate whether prominent asset pricing models appear suitable for that role. We assume prices properly reflect at least the information the market has had ten years to evaluate and exploit, whether or not the market is efficient. With this assumption, a model suitable as the no-mispricing benchmark should clear the seemingly modest hurdle of assigning zero alphas to long-short spreads based on decade-old information.

We find a number of prominent asset pricing models fail that test, assigning significant alphas to spreads formed using ten-year lags of FF3 expected return forecasts. Such models include the FF3, FF5, FF6 models and the Q4 and Q5 models of Hou, Xue, and Zhang (2015) and Hou et al. (2021). In contrast, the same long-short spreads do not produce significant alphas with respect to the traditional CAPM of Sharpe (1964) and Lintner (1965) and simple variations of this model. Hence, such models emerge as viable candidates for the no-mispricing benchmark, with more powerful tests required to discriminate among them.

While it seems reasonable that spreads based on decade-old information should receive zero alpha with respect to the no-mispricing benchmark, the main challenge faced by our test is achieving power. For many stocks, current values of factor betas may be unrelated to decade-old information. However, we show that large stocks have the most stable multifactor betas, thus offering our test the most power. Moreover, to have large stocks play the strongest role in our test underscores the economic importance of the result. For these stocks, which are the backbone of the US economy, prominent multifactor models evidently distort expected returns purged of mispricing. The stronger results of our test among large stocks reveal what seems to be an economically significant shortcoming of popular multifactor models. These models fail significantly in capturing expected returns on strategies based on decade-old information. Our assumption that such information should be fully reflected in prices seems especially reasonable for the market's largest and most liquid stocks. The CAPM and its simpler variants, in contrast, fare well in this regard, emerging as better candidates for the no-mispricing benchmark model.

The relatively parsimonious set of models we consider is far from exhaustive. In an initial attempt to compare the abilities of pricing models to serve as the no-mispricing benchmark, we believe the models we consider present a horserace with interesting entrants. We certainly acknowledge that there are other horses out there. Seeking refinements of our approach that potentially offer more power also seems a worthy research objective and may help better identify which models serve best as the proper efficient benchmark.

Of course, another worthy and parallel objective for research in asset pricing is to continue building models that better describe actual expected returns, whether or not the prices determining those expected returns include mispricing. Although such models may be less useful for gauging the extent of market inefficiencies or understanding risk premia, they can be otherwise useful, such as in designing investment strategies.

Table 1

R-squared when regressing individual-stock betas on decade-old individual-stock or industry betas

The table reports the average R-squared value in a multiple regression of the current individual-stock OLS beta with respect to a given FF3 factor on decade-old estimates of betas with respect to the three FF3 factors. The decade-old estimates are either OLS estimates for individual stocks (Panel A) or within-industry averages of those estimates (Panel B). The regressions and the industry averaging are performed separately within three (decade-old) market-cap segments formed using NYSE percentiles as breakpoints: large (above 70th), medium (70th to 20th), and small (below 20th). The R-squared values are averaged over the sample period for current betas, 1/1968 to 12/2024. The beta estimates for a given month use the past 36 months of data.

Dependent variable: individual-stock OLS beta	Market-cap segment		
	Large	Medium	Small
Panel A. Independent variables: decade-old individual-stock OLS betas			
β_{MKT}	0.128	0.096	0.058
β_{SMB}	0.073	0.061	0.024
β_{HML}	0.113	0.054	0.015
Panel B. Independent variables: decade-old industry-average OLS betas			
β_{MKT}	0.114	0.082	0.038
β_{SMB}	0.069	0.066	0.026
β_{HML}	0.139	0.080	0.019

Table 2

Pricing tests using decade-old return predictions

The table reports estimated monthly alphas (in percent) and *t*-statistics (in parentheses) for spreads between value-weighted portfolios of stocks in the bottom and top 20% of stocks sorted by decade-old return predictions from the three-factor model of Fama and French (1993). The return predictions use industry-level betas to predict stocks' betas ten years later. Results are shown for the total stock universe as well as mega-cap stocks (largest 200) and three market-cap segments formed using NYSE percentiles as breakpoints: large (above 70th), medium (70th to 20th), and small (below 20th). The models tested are the three-, five- and six-factor models of Fama and French (1993, 2015, 2018), denoted FF3, FF5 and FF6, the four- and five-factor models of Hou et al. (2015) and Hou et al. (2021), denoted Q4 and Q5, the CAPM of Sharpe (1964) and Lintner (1965), and the latter model augmented by various single factors: the betting-against-beta (BAB) factor of Frazzini and Pedersen (2014), a 12-month momentum factor (UMD from Fama and French (2018)), the traded liquidity factor (LIQ) of Pástor and Stambaugh (2003), and the size factor (SMB) of Fama and French (1993). Also included is a model with no factors. The last column tests joint equality to zero of the three element vector containing the alphas for the large, medium, small size segments, using the test of Gibbons, Ross, and Shanken (1989) but with a White (1980) heteroskedasticity-consistent covariance matrix (also used to compute the *t*-statistics). The sample period for computing alphas is 1/1968 to 12/2024.

	Mega-cap	Large	Medium	Small	All stocks	L/M/S *p*-value
FF3	0.41	0.37	0.22	-0.11	0.39	0.0143
	(2.62)	(2.70)	(1.71)	(-0.77)	(3.17)	
FF5	0.30	0.26	0.38	-0.20	0.28	0.0058
	(1.76)	(1.81)	(2.97)	(-1.31)	(2.25)	
FF6	0.25	0.22	0.34	-0.20	0.24	0.0232
	(1.42)	(1.45)	(2.55)	(-1.23)	(1.78)	
Q4	0.28	0.26	0.34	-0.24	0.29	0.0189
	(1.56)	(1.70)	(2.43)	(-1.37)	(2.07)	
Q5	0.20	0.19	0.34	-0.33	0.24	0.0288
	(1.05)	(1.19)	(2.31)	(-1.77)	(1.66)	
CAPM	0.25	0.23	0.18	0.00	0.24	0.2337
	(1.50)	(1.60)	(1.47)	(0.03)	(1.63)	
CAPM+BAB	0.22	0.19	0.25	-0.30	0.21	0.0701
	(1.17)	(1.18)	(1.80)	(-1.86)	(1.32)	
CAPM+MOM	0.15	0.14	0.15	0.03	0.13	0.5807
	(0.84)	(0.90)	(1.13)	(0.18)	(0.87)	
CAPM+LIQ	0.21	0.20	0.19	0.01	0.21	0.2856
	(1.26)	(1.37)	(1.54)	(0.04)	(1.38)	
CAPM+SMB	0.25	0.23	0.18	0.00	0.24	0.2346
	(1.50)	(1.60)	(1.47)	(0.01)	(1.76)	
No factors	0.09	0.09	0.04	-0.10	0.03	0.8370
	(0.52)	(0.57)	(0.32)	(-0.66)	(0.18)	

Table 3

Alpha components for spreads formed with decade-old return predictions

The table reports the components (in percent) that sum to the estimated monthly alphas for spreads between value-weighted portfolios of stocks in the bottom and top 20% of stocks sorted by decade-old return predictions from the three-factor model of Fama and French (1993). The return predictions use industry-level betas to predict stocks' betas ten years later. Results are shown for the total stock universe as well as mega-cap stocks (largest 200) and three market-cap segments formed using NYSE percentiles as breakpoints: large (above 70th), medium (70th to 20th), and small (below 20th). The models tested are the three-, five- and six-factor models of Fama and French (1993, 2015, 2018), denoted FF3, FF5 and FF6, the four- and five-factor models of Hou et al. (2015) and Hou et al. (2021), denoted Q4 and Q5, the CAPM of Sharpe (1964) and Lintner (1965), and the latter model augmented by various single factors: the betting-against-beta (BAB) factor of Frazzini and Pedersen (2014), a 12-month momentum factor (UMD from Fama and French (2018)), the traded liquidity factor (LIQ) of Pástor and Stambaugh (2003), and the size factor (SMB) of Fama and French (1993). The sample period for computing alphas is 1/1968 to 12/2024.

Model	Alpha components	Mega-cap	Large	Medium	Small	All stocks
FF3						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.19	0.17	0.14	0.05	0.18
	$-\beta_{SMB} \times \overline{SMB}$	0.00	0.00	0.01	0.03	0.06
	$-\beta_{HML} \times \overline{HML}$	0.12	0.11	0.03	-0.09	0.12
	total: α	0.41	0.37	0.22	-0.11	0.39
FF5						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.18	0.16	0.16	0.05	0.16
	$-\beta_{SMB} \times \overline{SMB}$	-0.01	-0.00	0.03	0.02	0.08
	$-\beta_{HML} \times \overline{HML}$	0.14	0.13	-0.02	-0.10	0.12
	$-\beta_{RMW} \times \overline{RMW}$	-0.08	-0.06	0.08	-0.09	-0.04
	$-\beta_{CMA} \times \overline{CMA}$	-0.03	-0.06	0.10	0.02	-0.06
	total: α	0.30	0.26	0.38	-0.20	0.28
FF6						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.17	0.15	0.15	0.05	0.15
	$-\beta_{SMB} \times \overline{SMB}$	-0.01	-0.00	0.03	0.02	0.08
	$-\beta_{HML} \times \overline{HML}$	0.13	0.12	-0.03	-0.10	0.11
	$-\beta_{RMW} \times \overline{RMW}$	-0.07	-0.05	0.08	-0.09	-0.04
	$-\beta_{CMA} \times \overline{CMA}$	-0.03	-0.05	0.10	0.02	-0.05
	$-\beta_{MOM} \times \overline{MOM}$	-0.04	-0.04	-0.03	-0.01	-0.04
	total: α	0.25	0.22	0.34	-0.20	0.24
Q4						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.18	0.16	0.15	0.04	0.16
	$-\beta_{ME} \times \overline{ME}$	-0.01	-0.00	0.04	0.04	0.12
	$-\beta_{IA} \times \overline{IA}$	0.16	0.12	0.08	-0.13	0.08
	$-\beta_{ROE} \times \overline{ROE}$	-0.14	-0.11	0.04	-0.09	-0.11
	total: α	0.28	0.26	0.34	-0.24	0.29

Table 3 (continued)

Alpha components for spreads formed with decade-old return predictions

Model	Alpha components	Mega-cap	Large	Medium	Small	All stocks
Q5						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.18	0.16	0.15	0.03	0.16
	$-\beta_{ME} \times \overline{ME}$	-0.01	-0.01	0.04	0.03	0.12
	$-\beta_{IA} \times \overline{IA}$	0.16	0.13	0.08	-0.12	0.08
	$-\beta_{ROE} \times \overline{ROE}$	-0.12	-0.09	0.04	-0.07	-0.09
	$-\beta_{EG} \times \overline{EG}$	-0.10	-0.08	-0.00	-0.10	-0.05
	total: α	0.20	0.19	0.34	-0.33	0.24
CAPM						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.16	0.15	0.14	0.10	0.21
	total: α	0.25	0.23	0.18	0.00	0.24
CAPM+BAB						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.16	0.14	0.15	0.09	0.21
	$-\beta_{BAB} \times \overline{BAB}$	-0.03	-0.04	0.07	-0.29	-0.03
	total: α	0.22	0.19	0.25	-0.30	0.21
CAPM+MOM						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.15	0.13	0.14	0.11	0.20
	$-\beta_{MOM} \times \overline{MOM}$	-0.09	-0.08	-0.03	0.02	-0.09
	total: α	0.15	0.14	0.15	0.03	0.13
CAPM+LIQ						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.16	0.15	0.14	0.10	0.21
	$-\beta_{LIQ} \times \overline{LIQ}$	-0.04	-0.03	0.01	0.00	-0.03
	total: α	0.21	0.20	0.19	0.01	0.21
CAPM+SMB						
	average spread	0.09	0.09	0.04	-0.10	0.03
	$-\beta_{MKT} \times \overline{MKT}$	0.16	0.15	0.13	0.07	0.15
	$-\beta_{SMB} \times \overline{SMB}$	-0.00	-0.00	0.01	0.03	0.06
	total: α	0.25	0.23	0.18	0.00	0.24

References

Baba-Yara, F., Boons, M., Tamoni, A., 2024. Persistent and transitory components of firm characteristics: Implications for asset pricing. Journal of Financial Economics 154.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

Belo, F., 2010. Production-based measures of risk for asset pricing. Journal of Monetary Economics 57, 146–163.

Berk, J. B., Green, R. C., Naik, V., 1999. Optimal investment, growth options, and security returns. Journal of finance 54, 1553–1607.

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

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

Cho, T., Polk, C., 2020. Asset pricing with price levels. Working paper, London School of Economics .

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under-and overreactions. Journal of Finance 53, 1839–1885.

Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. Journal of Finance 52, 1–33.

Du, W., Tepper, A., Verdelhan, A., 2018. Deviations from covered interest rate parity. Journal of Finance 73, 915–957.

Fama, E. F., 1970. Efficient capital markets: A review of theory and empirical work. Journal of Finance 25, 383–417.

Fama, E. F., 1976. Foundations of Finance. Basic Books, New York.

Fama, E. F., 1991. Efficient capital markets: II. Journal of Finance 46, 1575–1617.

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

Fama, E. F., French, K. R., 1995. Size and book-to-market factors in earnings and returns. Journal of Finance 50, 131–155.

Fama, E. F., French, K. R., 1997. Industry costs of equity. Journal of Financial Economics 43, 153–193.

Fama, E. F., French, K. R., 2008. Dissecting anomalies. Journal of Finance 63, 1653–1678.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E. F., French, K. R., 2018. Choosing factors. Journal of Financial Economics 128, 234–252.

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

Frazzini, A., Pedersen, L. H., 2014. Betting against beta. Journal of Financial Economics 111, 1–25.

Gibbons, M. R., Ross, S. A., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica pp. 1121–1152.

Gomes, J., Kogan, L., Zhang, L., 2003. Equilibrium cross section of returns. Journal of Political Economy 111, 693–732.

Hong, H., Stein, J. C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54, 2143–2184.

Hou, K., Mo, H., Xue, C., Zhang, L., 2021. An augmented q-factor model with expected growth. Review of Finance 25, 1–41.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: an investment approach. Review of Financial Studies 28, 650–705.

Hu, G. X., Pan, J., Wang, J., 2013. Noise as information for illiquidity. Journal of Finance 68, 2341–2382.

Johnson, T. C., 2002. Rational momentum effects. Journal of Finance 57, 585–608.

Keloharju, M., Linnainmaa, J. T., Nyberg, P., 2020. Long-term discount rates do not vary across firms. Journal of Financial Economics, forthcoming.

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

Lamont, O. A., Thaler, R. H., 2003. Can the market add and subtract? Mispricing in tech stock carve-outs. Journal of Political Economy 111, 227–268.

Li, D., Zhang, L., 2010. Does q-theory with investment frictions explain anomalies in the cross section of returns? Journal of Financial Economics 98, 297–314.

Li, E. X., Livdan, D., Zhang, L., 2009. Anomalies. Review of Financial Studies 22, 4301–4334.

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

Liu, L. X., Whited, T. M., Zhang, L., 2009. Investment-based expected stock returns. Journal of Political Economy 117, 1105–1139.

Liu, L. X., Zhang, L., 2008. Momentum profits, factor pricing, and macroeconomic risk. Review of Financial Studies 21, 2417–2448.

Pástor, L., Stambaugh, R. F., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Sagi, J. S., Seasholes, M. S., 2007. Firm-specific attributes and the cross-section of momentum. Journal of Financial Economics 84, 389–434.

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

Stambaugh, R. F., Yuan, Y., 2017. Mispricing factors. Review of Financial Studies 30, 1270–1315.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heterockedasticity. Econometrica 8, 817–838.

Zhang, L., 2005. The value premium. Journal of Finance 60, 67–103.