A net beta test of asset pricing models

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

• This article was published in the International Review of Financial Analysis [© Elsevier Inc.] and the definitive version is available at: http://dx.doi.org/10.1016/j.irfa.2009.09.008

Metadata Record: [https://dspace.lboro.ac.uk/2134/15029](https://dspace.lboro.ac.uk/2134/15029)

Version: Accepted for publication

Publisher: © Elsevier Inc.

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
A net beta test of asset pricing models

Cherif Guermat\textsuperscript{a}, Mark C. Freeman\textsuperscript{a, b}

\textsuperscript{a} Bristol Business School, University of the West of England, Frenchay Campus, Coldharbour Lane, Bristol, BS16 1QY, United Kingdom.

\textsuperscript{b} Bradford University School of Management, Emm Lane, Bradford, West Yorkshire, BD9 4JL, United Kingdom. Tel: +44-1274-234363, Fax: +44-1274-235680.

Abstract

While many recent empirical studies of the CAPM have used conditional beta tests, this technique has recently been shown to have several weaknesses. Here we introduce a new, more robust, net beta test which shares a number of characteristics with conditional beta tests. The method is extended to the multi-factor case when there are mimicking portfolios of assets for the underlying factors, including the Fama-French three-factor model. We demonstrate theoretically, by simulation and using market data that the net beta estimators have lower standard errors than those generated by the standard Fama-MacBeth test.

Key words: Factor models, Capital asset pricing, Conditional beta tests.

JEL classification: G12; C12; C15

1. Introduction

Since Fama and MacBeth (1973, hereafter FM), a standard technique for testing asset pricing models involves a two-pass cross-sectional approach. The continued importance of this method in empirical asset pricing has recently been stressed by Shanken and Zhou (2007) amongst others. However, despite its ongoing popularity, the FM test has very low statistical power. For example, more than two decades of empirical tests of
the Capital Asset Pricing Model (CAPM) have been unable to find a statistically significant correlation between stock returns and market betas (see, for example, Reinganum, 1982; Lakonishok and Shapiro, 1986; Ritter and Chopra, 1989; Fama and French, 1992). A question then raised by Chan and Lakonishok (1993) is whether the null hypothesis is true or whether the tests just lack the power to reject it in finite samples. They show that, for realistic levels of market return volatility, the FM test is likely to have low power in samples of the size typically employed in practice. The test cannot reject the null that beta is not priced but it also cannot reject the null of the CAPM assuming any plausible value for the expected return to the market. A researcher who has strong prior beliefs in the CAPM would not be compelled to infer from the point estimate and associated standard error that the CAPM is false.

In an attempt to address this problem, Pettengill et al. (1995) provide a simple adjustment to the FM test. This involves conditioning the test on whether the market is rising or falling at the time and is known as the conditional beta test (CBT). This technique has been widely used in the international asset pricing literature. For example, broad international studies using the method have been conducted by Fletcher (2000) and Tang and Shum (2003a, 2003b, 2008). Within a European context, more specific studies for the UK include Fletcher (1997), Fraser et al. (2004) and Hung et al. (2004), for Germany Elsas et al. (2003), for France Lilti and Montagner (1998), for Turkey Karacabey and Karatepe (2004) and for Poland Zhang and Wihlborg (2004). There have also been a number of studies that have used this technique specifically on Asian data, including Hodoshima et al. (2000) for Japan, Lam (2001), Ho et al. (2006) and Tang and Shum (2006) for Hong Kong, Sheu et al. (1998) for Taiwan and Tang and Shum (2004) for Singapore. Girarda et al. (2003) investigate nine Asian markets using the CBT methodology augmented to allow for conditional heteroskedasticity in stock returns. For the rest of the world, Faff (2001) presents results for Australia, Sandoval and Saens (2004) for Latin America and Huang and Hueng (2008) for the US when betas are time-varying. Basher and Sadorsky (2006) use the CBT methodology to examine the impact of oil prices on emerging market stock returns.
The problem that arises is that, as shown by Freeman and Guermat (2006), it is not possible to draw accurate inferences from the CBT without making significant adjustments, after which it has negligible statistical advantage over the original FM test. Further limitations of the test have been discussed by Koch and Westheide (2009) and Cooper (2009).

This paper describes a simple alternative way of providing estimators with low standard errors in the spirit of the CBT approach but that avoids its limitations. The new technique can be applied to multi-factor linear asset pricing models where the factors have perfectly identified mimicking portfolios of traded assets. This means that, in addition to testing the CAPM, the method can also be used to investigate the three-factor model of Fama and French (1993, 1996) and the exact APT with mimicking portfolios (see, for example, Huberman et al., 1987).

Our modified test (hereafter, the “net beta test”) relies on the observation that the low power of the standard FM test has a simple explanation. Under both the null and alternate hypotheses, the slope coefficient in each monthly cross-section regression is a function of the realized factor returns. The high volatility of this yields an excessively noisy estimated series of monthly slope coefficients. It is this noise that is responsible for the low power of the FM test.

The net beta test involves subtracting (netting out) the realized returns for the mimicking portfolio each period from the estimated FM slope coefficient for each factor. The standard error of the test statistic is then significantly reduced. This is demonstrated analytically, by simulation and by using market data.

A natural consequence of subtracting the observed factor returns from the estimated FM slope coefficient is that the asset pricing model becomes the null hypothesis of the net beta test — a characteristic that it shares in common with the CBT. Any statistically significant non-zero parameter estimate is sufficient to reject the specified asset pricing model. This provides a further improvement on the standard FM test, where, if a coefficient is found to be statistically different from zero, it is not immediately clear that the factor is priced. Instead, the non-zero coefficient may simply reflect the
The structure of the paper is as follows. In the following section, we outline cross-sectional regression methods. We introduce the net beta test and compare it with the FM test and the CBT. Section 3 reports the results of simulation experiments to ascertain the size and power of both the FM test and the net beta test in the single factor case. In section 4, we run the FM and net beta tests on US stock market data. We test the CAPM using the FM and net beta tests over several subperiods since 1973. These confirm that the net beta test statistic has lower standard error than the FM test statistic, although it is still not possible in many cases to reject the CAPM for beta ranked portfolios. We extend our empirical exercise to the Fama-French three-factor model in subsection 4.2. Section 5 offers some concluding remarks.

2. Cross-sectional regression methods

The focus of this paper is on cross-sectional methods for testing the CAPM and other factor models with mimicking portfolios for the factors. A researcher is, of course, not limited to these techniques for such models and may, instead, prefer time-series methods, such as the multivariate extension of Black, Jensen and Scholes (1972) by Gibbons et al. (1989). Our aim in this paper is not to compare the relative merits of testing intercepts against testing risk premia. Most commentators, including Cochrane (2001), Campbell et al. (1997) and Fama and French (2004), present the two techniques as being complementary. In certain situations, though, cross-sectional techniques have clear advantages, including when there are a large number of test assets or when betas are time-varying. It is, therefore, not surprising that these methods remain highly used in practice; see the references in Shanken and Zhou (2007) for FM and the introduction to this paper for the CBT. ¹ Providing methodological improvements to cross-sectional

¹It is also briefly worth noting that there is a clear relationship between our net beta test and standard time-series tests. As we restrict the factors to be identified by mimicking portfolios of assets and because our results extend to generalized least squares we can invoke the result that: “If the factor
techniques remains highly important.

The results that we present in this paper are robust to whether the model is estimated using ordinary, generalized or weighted least squares. While there are alternative approaches for cross-sectional techniques including generalized method of moments (GMM; Hansen, 1982), maximum likelihood (Gibbons, 1982) and semi-parametric methods (Hodgson et al., 2002), our focus is not on how the model is estimated. Despite their theoretical appeal, there is little compelling evidence at present that the more sophisticated estimation techniques actually outperform simple least squares. GMM based methods do not generally lead to fully efficient estimates (Vorkink, 2003). In addition, Harvey and Zhou (1993) find little difference between OLS and GMM based tests, while Ferson and Foerster (1994) provide evidence suggesting that GMM methods lead to asset pricing tests with aberrant properties. The semi-parametric method of Hodgson et al. (2002) seems promising, but the evidence is scant and is based only on a small-scale simulation study (Vorkink, 2003). Shanken and Zhou (2007) express no strong preference for more sophisticated estimation techniques over least squares methods in small sample sizes.

2.1. *The net beta method*

Consider an economy where there exists a one-period risk-free asset, which is usually proxied by Treasury bills, with simple return between $t - 1$ and $t$ of $r_{ft}$. There are also $N$ risky securities, with the simple return to asset $i \in \{1, ..., N\}$ between times $t - 1$ and $t$ being denoted by $r_{it}$. $R_{it} = r_{it} - r_{ft}$ is the excess return for this period for security $i$.

In the case of the CAPM, assume that the beta of each asset is known and fixed, $\beta_i$. Then the standard FM procedure is, at any time $t$ in the test period, to estimate

\begin{itemize}
\item is a return, the GLS cross-sectional regression, including the factor as a test asset, is identical to the time series regression \footnote{Cochrane (2001, p. 243, his emphasis).}
\item A major advantage of cross-sectional approaches is that they can be adapted to time-varying betas. This is equally true of the net beta technique, but we only formally present results with fixed betas for ease of exposition.
\end{itemize}
parameters, $\lambda_{0t}, \lambda_{1t}$ through the cross-sectional regression:

$$R_{it} = \lambda_{0t} + \lambda_{1t}\beta_i + v_{it}$$

(1)

The FM procedure calculates time-series averages and standard errors of the estimate $\hat{\lambda}_{1t}$:

$$\hat{\lambda} = \frac{1}{T} \sum_{t=1}^{T} \hat{\lambda}_{1t}$$

(2)

The null hypothesis for FM is that beta risk is not priced; $E[\hat{\lambda}] = 0$. The FM test statistic for this is:

$$t_{FM} = \frac{\hat{\lambda}}{SE(\hat{\lambda})}$$

(3)

where $SE(\hat{\lambda}) = SD(\hat{\lambda}_{1t})/\sqrt{T-1}$ and $SD(\hat{\lambda}_{1t})$ is the sample standard deviation of the $T$ cross-section regressions estimates $\hat{\lambda}_{1t}$. The problem that we address here is that the FM statistic is known to have low statistical power even in the absence of beta estimation error as $SE(\hat{\lambda})$ is high.

In an attempt to overcome this problem, Pettengill et al. (1995) note that it is natural to expect $\gamma_{1t}$ to be positive (negative) when the ex-post excess return to the market is positive (negative). To account for this, they run the modified regression:

$$R_{it} = \lambda_{0t} + \lambda_{1t}^+D_t\beta_i + \lambda_{1t}^-(1-D_t)\beta_i + v_{it}$$

(4)

where $D_t$ is a dummy variable that equals 1 if the realized excess return to the market is positive at time $t$ and zero otherwise. Denote the time-series average of estimates $\hat{\lambda}_{1t}^+$ and $\hat{\lambda}_{1t}^-$ by $\hat{\lambda}^+$ and $\hat{\lambda}^-$ respectively. When the average excess return to the market is positive in the test period and $\hat{\lambda}^+ + \hat{\lambda}^- = 0$, then they contend the CAPM cannot be rejected. Note that, in contrast to the FM test, the CAPM now becomes the null, rather than part of the alternate, hypothesis.

Freeman and Guermat (2006) prove, though, that the CBT method must be modi-
fied to accurately test the CAPM. This is because the requirement for positive average excess returns ensures non-symmetry around zero. As a consequence, \( \hat{\lambda}^+ + \hat{\lambda}^- = c_1 \) when the CAPM holds for some constant \( c_1 \neq 0 \). As \( c_1 \) requires estimation, and so has an associated standard error, this reduces the statistical power of the CBT. Freeman and Guermat (2006) show that, after making the necessary modifications, the CBT technique has negligible statistical advantage over the original FM test.

The method we present here is very much in the spirit of the CBT. A simple adjustment to the FM test is also made for ex-post returns to the market and the CAPM is again the null hypothesis for the new test. Rather than using dummy variables, though, the modified statistic, \( \hat{\gamma}_{1t} \), is constructed from the standard cross-sectional regression by subtracting, or netting out, the realized excess return to the market:

\[
\hat{\gamma}_{1t} = \hat{\lambda}_{1t} - R_{mt}
\] (5)

The procedure then just follows FM; a time series average of \( \hat{\gamma}_{1t} \) is estimated, \( \hat{\gamma} \), and a t-statistic calculated in the usual way. Given that \( E[\hat{\lambda}_{1t}] = E[R_{mt}] \) if the CAPM is true for the standard FM test, it follows immediately that, if \( \hat{\gamma} \) is not statistically different from zero, then the CAPM cannot be rejected by the data under this test. This is formally proved in the appendix.

The reason for proposing this statistic is that, as shown in the appendix, \( SE(\hat{\lambda}) \) is an increasing function of the ex-post variability of market returns under both the null and alternate hypotheses. By netting out the realized return to the market, the standard error of the estimated coefficient is reduced. The appendix proves the result, which we examine empirically in subsequent sections of this paper, that:

\[
Var(\hat{\lambda}) - Var(\hat{\gamma}) = Var(R_m)
\] (6)

The extension of this test to multi-factors is straightforward when there are mimicking portfolios for the factors. Such cases include the Fama-French three-factor model and exact APT models with mimicking portfolios. While it is perhaps more common
to test the three-factor model using time series techniques, following Fama and French (1993, 1996), there is no *a priori* reason why cross-sectional methods should be discounted. Shanken and Zhou (2007) run simulations on the Fama-French model to test the small sample properties of such tests. Further, the CBT method has been used in the context of the three-factor model by Pettengill et al. (2002), Hung et al. (2004) and Morelli (2007). Pettengill et al. (2002) find that, when allowing for size and book-to-market, the symmetry between \( \tilde{\lambda}^+ \) and \( \tilde{\lambda}^- \) breaks down, suggesting that “investors may not receive adequate compensation for systematic risk as measured by beta” (*ibid.*, p.132). By contrast, when using UK data, Morelli (2007) is unable to reject the null hypothesis that \( \tilde{\lambda}^+ + \tilde{\lambda}^- = 0 \). However, as noted by Hung et al. (2004), the properties of the CBT are not well understood in this context and “it is not clear however, how the Fama French factors should react in up or down markets” (*ibid.*, footnote 4). Koch and Westheide (2009) extend the Freeman and Guermat (2006) critique of conditional beta tests to the three-factor case and confirm that this method adds little to the standard Fama-MacBeth procedure.

For the test presented here, these problems do not arise. Let \( \beta_{ik} \) represent the known beta of asset \( i \) for factor \( k \in [1, K] \). Then, as with FM, the cross-sectional regression

\[
R_{it} = \lambda_{0t} + \sum_{k=1}^{K} \lambda_{kt} \beta_{ik} + v_{it}
\]

(7)

provides estimates \( \tilde{\lambda}_{kt} \) and the modified statistic is given by

\[
\tilde{\gamma}_{kt} = \tilde{\lambda}_{kt} - f_{kt}
\]

(8)

where \( f_{kt} \) is the realized excess return on the mimicking portfolio for the factor at that time. Time series averages \( \tilde{\gamma}_{k} \) and t-statistics are calculated in the usual way. As proved in the appendix, if the model is true, then we should be unable to reject the
null hypothesis that $\hat{\gamma}_k = 0$ for all $k$. The advantage of our test, that

$$Var(\hat{\lambda}_k) - Var(\hat{\gamma}_k) = Var(f_k)$$

(9)

continues through to the multi-factor case, meaning that the decrease in standard errors is not only restricted to coefficients associated with the market portfolio.

2.2. Necessary and sufficient conditions

There is a further feature that distinguishes the net beta and FM tests. The standard FM procedure involves testing whether $\hat{\lambda}$ is statistically different from zero. Even ignoring issues of statistical power, this is not equivalent to testing the null hypothesis that the price of risk is equal to the equity premium, which is the theoretical prediction of the model. Instead, a (statistically significant) non-zero estimate $\hat{\lambda}$ will be viewed as evidence that systematic market risk influences expected return. This is often interpreted as being at least partial evidence in support of the CAPM and something that is empirically useful. In fact, finding $\hat{\lambda} \neq 0$ is neither necessary nor sufficient to establish that market risk is priced. This is formally proved in the appendix but an illustration is provided here.

Assume that, in the real economy, both market and size are priced with respective risk premia $F_m$ and $F_s$. The sensitivities of asset $i$ to these two factors are $b_{im}$ and $b_{is}$ respectively. In error, though, the econometrician thinks that only market risk is priced. She uses a univariate estimate of sensitivity to the market, $\beta_i$, to run the FM test. Using equation 21 in the appendix, the econometrician’s estimate of $\hat{\lambda}$ will be a function of both $F_m$ and $F_s$ as $b_{im} \neq \beta_i$ when size and market effects are correlated. We test this using 1,000 random stocks and find that, on this sample, $E[\hat{\lambda}] = 0.88F_m + 0.53F_s$.$^3$ Even if market risk is not priced, $E[\hat{\lambda}] > 0$ provided that size is positively priced. This, effectively, is the point made by Fama and French (1992); apparent market risk premia are really proxies for other, related, effects. Similarly, if

$^3$The data used here is described in subsection 3.1.
\( F_m = -0.53F_s/0.88 \neq 0 \) then \( E[\lambda] = 0 \) even though market risk is priced. A negative size premium can cancel out an existing positive market premium.

By contrast, with the net beta test, finding a test statistic that is significantly different from zero is sufficient to reject the asset pricing model. Again, the formal proof is given in the appendix. The result of the previous study for the net beta test, using equation 22 in the appendix, is now \( E[\gamma] = -0.12F_m + 0.53F_s \). General non-zero values of \( F_m \) and \( F_s \) will make \( E[\gamma] \neq 0 \), correctly identifying that the economist’s model is false. It should, though, be emphasized that the net beta test does not provide a necessary condition for rejecting the CAPM. It is possible that \( F_m = 0.53F_s/0.12 \), in which case \( E[\gamma] = 0 \) even though the CAPM is false. The extension of these results to the multi-factor case are straightforward. In general, the net beta test gives sufficient grounds to reject a proposed asset pricing model, which cannot be achieved by the standard FM test.

3. Simulation Experiments

We now turn to the evaluation of the small sample properties of both the FM test and the net beta test using simulated data for the single factor CAPM model. We start by running simulations under cases that reflect the null hypothesis of each test: (i) the cost of beta being equal to the equity risk premium and (ii) the cost of beta being zero. We then consider the robustness of the simulation results to different return generating processes.

3.1. Testing the CAPM

We generate data under two distinct hypotheses. Under the first hypothesis, the CAPM holds, while under the second hypothesis, expected returns and beta are uncorrelated. That is, we generate data for the null of the FM test and the net beta test in the case where only market risk is potentially priced. The two models used to generate the simulated data are calibrated using realized monthly returns on 1,000 stocks.
randomly drawn from stocks in the CRSP database that had as least 24 observations over the period January 1968 to December 2000.

For each randomly drawn stock, \( i \), we use the available time series data from January 1968 to December 2000 to estimate \( \beta_i \) as the slope of the time series regression of the monthly excess return of each stock on the stock market excess return. The variance of the residuals from this regression \( \hat{\sigma}_i^2 \) is saved. The excess market return is the monthly CRSP market return minus the US Treasury bill rate (IMF series USI60C).

For the first model, under which the CAPM holds, returns, \( R_{it} \), are generated as follows. At each date, \( t \), excess market returns, \( R_{mt} \), are drawn from a normal distribution with mean \( E(R_{mt}) \) and variance \( \sigma_m^2 \). We use \( E(R_{mt}) = 0.007 \) and \( \sigma_m = 0.055 \). These values closely match the sample moments of the CRSP data and are similar to those used in Jagannathan and Wang (2002). In order to examine the impact of the volatility of the market return on the small sample properties of the two tests, we also use half the standard deviation (\( \sigma_m = 0.0275 \)). Four time-series sample sizes, \( T \), for the test period are considered (60, 120, 240, and 360 months). These time horizons are commonly used in testing the CAPM (see, for example, Fama and French, 1992, and Jagannathan and Wang, 1996). We also consider some cases with a large sample size (\( T = 1,000 \)).

Individual excess returns under the null hypothesis of the CAPM are generated by

\[
R_{i,1,t} = \beta_i R_{mt} + v_{it} \quad v_{it} \sim N(0, \hat{\sigma}_i^2)
\]  

(10)

where \( \beta_i \) is randomly drawn, with replacement, from the vector of 1,000 estimates of \( \beta_i \) and \( v_{it} \) is randomly drawn from a normal distribution with variance \( \hat{\sigma}_i^2 \).

For the second model, where beta risk is not priced, individual excess returns are generated by

\[
R_{i,2,t} = E(R_{mt}) + \beta_i (R_{mt} - E(R_{mt})) + v_{it} \quad v_{it} \sim N(0, \hat{\sigma}_i^2)
\]

(11)

The simulation exercise is based on 1,000 replications. In each replication, equations
10 and 11 are used to generate individual stock return data. In each replication, and for each of the return-generating models, \( j = 1, 2 \), we estimate a series of cross-section regressions 

\[ R_{i,j,t} = \lambda_{0,j,t} + \lambda_{1,j,t}\beta_i + e_{i,j,t}, \quad t = 1, \ldots, T, \]

yielding two series of estimated slope coefficients, \( \hat{\lambda}_{1,j,t}, j = 1, 2, t = 1, \ldots, T \). For each of the two models, the FM statistic is the \( t \)-statistic of the sample mean of \( \hat{\lambda}_{1,j,t} \) given by equation 3, while the net beta statistic is the \( t \)-statistic of the sample mean of \( \tilde{\gamma}_{1,j,t} = \hat{\lambda}_{1,j,t} - R_{mt} \). The FM statistic is then used to test the null hypothesis that expected returns and beta are uncorrelated against the alternative hypothesis that beta is priced, and the net beta test is used to test the null hypothesis that the CAPM holds against the alternative hypothesis that the economist’s model is not identical to the true asset pricing model.

In order to compute the size of the FM and net beta tests, we calculate the proportion of times each test rejects its respective null hypothesis when the respective null hypothesis is true. To compute the power of the two tests, we calculate the proportion of times each test rejects its respective null hypothesis when the null hypothesis of the alternative model is true. The tests are conducted at the 5% and 10% two-sided significance levels.

The simulation results for the FM and net beta tests are reported in Table 1. The two tests have alternate null hypotheses and are therefore not directly comparable. However, the main message of the simulation exercise is to demonstrate how each test performs in identifying the CAPM when it is true and when it is “absolutely” false. Thus, one only needs to bear in mind that for the FM the power is related to its ability to reject an absolute non-CAPM, while the power of the net beta test is related to its ability to reject the CAPM.

For both significance levels, the empirical size of both tests under their respective null hypotheses is close to the nominal significance level. However, the net beta test clearly has greater ability to discriminate between the two hypotheses. Even at the smallest sample size, the power of the net beta test is very close to unity regardless of the volatility of the market return. For the FM test, however, the power is around 20 percent for the smallest sample size when \( \sigma_m = 0.055 \). Even when the variance of the market return is halved, the power is still low, at most 56 percent at the 10%
significance level. To achieve reasonable power, the FM test requires at least 1,000 months of data.

[Insert Table 1 around here]

Some interesting properties of the two tests are shown in Table 2, which reports the mean, standard deviation, skewness and excess kurtosis coefficients of the simulated FM and net beta t-statistics. The two tests show little difference under their respective null hypotheses. But the difference is striking under the null of the alternate test. The means of the two statistics have the expected opposite sign, but the difference in their variance is substantial. Owing to its reduced variance, the mean of the net beta t-statistic is around four and a half times greater than the mean of the FM t-statistic at all sample sizes. Under the alternative, the distribution of the net beta t-statistic is to the left of zero as suggested by its 5th and 95th percentiles. The FM t-statistic is far closer to its null hypothesis value of zero, with the 5th percentile becoming negative at lower sample sizes.

[Insert Table 2 around here]

3.2. Robustness Check

In the previous simulation we contrasted the case where the CAPM held against a situation where the CAPM was absolutely false. It is therefore possible that the simulation results are driven by the selection of the extreme alternative model. To establish the robustness of our simulation experiments, we consider how the FM and net beta tests perform against some alternative data generating processes for returns. As an alternative to the CAPM, Jagannathan and Wang (2002) suggest a model of the form

\[ R_{it} = \alpha_i + \beta_i R_{mt} + v_{it} \]  

(12)

where \( \alpha_i = E(R_{it}) - \beta_i E(R_{mt}) \). This model allows for other firm-specific determinants of expected returns, such as market capitalization and the ratio of book to market value
of equity. In the simulation, $\alpha_i$ is calibrated using the 1,000 randomly drawn stocks from the CRSP database and is calculated as $\hat{\alpha}_i = \overline{R_i} - \hat{\beta}_i \overline{R_m}$. Kan and Zhou (1999) suggest an alternative model that adds noise to the market return. In particular, one of their suggestions is to use

$$R_{it} = \beta_i f_t + v_{it} \quad v_{it} \sim N(0, \sigma_i^2)$$

where $f_t = (R_{mt} + n_t) \sqrt{1 + \sigma_n^2}, n_t$ is a zero mean measurement error with finite variance $\sigma_n^2$, and which is uncorrelated with $v_{it}$. Since, $E(f_t) = E(R_{mt}) \sqrt{1 + \sigma_n^2}$, this is approximately consistent with the CAPM for low $\sigma_n^2$. In the simulation we use equation 13 to generate the data, using $n_t \sim N(0, 0.01)$.

Table 3 presents the rejection rates when the data are generated by the Kan and Zhou (1999) and Jagannathan and Wang (2002) models.

A well specified test should identify the CAPM with a high probability in the case of Kan and Zhou (1999). Panels A and B of the table show that this noisy CAPM is mostly undetected by the FM test. The power of the FM test is extremely low, even for very large samples. The net beta test rejects the CAPM with the expected probability (size) for both levels of market volatility and at both levels of significance.

Panels C and D of Table 3 present results for the case of Jagannathan and Wang (2002). For this data generation process, beta is priced but there are additional pricing errors. Here we would expect both the FM and net beta test to reject the null hypothesis. This is because beta is priced (leading to rejection of the FM null) but the price of risk is not the equity premium (leading to rejection of the net beta null). As can be seen, the net beta test has uniformly greater ability to reject the null than the FM test.

[Insert Table 3 around here]
4. Testing factor models

In this section, we test the CAPM and Fama & French three-factor model on US data using the FM and net beta tests. From this we show that, in general, the standard errors of the net beta test are substantially below those of the FM test. We are also able to draw inferences about the ability of each model to explain cross-sectional differences in expected return.

4.1. Testing the CAPM

In the first set of tests, we randomly selected 9,000 stocks from the CRSP dataset of 22,716 companies. Stocks with less than 60 consecutive monthly observations throughout the period January 1968 to December 2000 were removed from the sample. This left slightly more than five thousand stocks available, of which the first five thousand were selected for the test.

In the first stage, the estimated beta of each stock, \( \hat{\beta}_i \), \( i = 1, ..., 5000 \), was calculated. Following standard procedures, these were calculated using a time series regression over all the available observations for each stock, \( R_{it} = \alpha_i + \beta_i R_{mt} + e_{it} \). On the basis of these estimated individual stock betas, 100 beta-ranked portfolios were formed with 50 stocks in each. Again, following standard Fama-MacBeth procedure, the beta of each portfolio was re-estimated using the 60 months of data prior to the test period in order to remove potential measurement error bias. This provided estimates of portfolio betas \( \hat{\beta}_p \), \( p = 1, ..., 100 \). Then \( \hat{\lambda}_{1t} \) was estimated from the cross section regression, \( R_{pt} = \lambda_0 + \lambda_{1t} \hat{\beta}_p + v_{pt} \). The FM and net beta tests are then simply \( t \)-tests on the sample mean of the time series \( \hat{\lambda}_{1t} \) and \( (\hat{\lambda}_{1t} - R_{mt}) \), respectively. The results, shown in Table 4, contrast the two tests for a selection of starting dates and sample sizes, \( T \).

[Insert Table 4 around here]

As expected, the net beta test statistic has a substantially lower standard error than
the FM test. According to equation 6:

\[
\sigma_m^2 = T \left[ SE^2(\hat{\lambda}) - SE^2(\hat{\gamma}) \right]
\]  

(14)

Using this equation to estimate \(\sigma_m^2\) from the standard errors in Table 4, the average estimated annualized standard deviation of the market is 14.6% with a minimum and maximum estimate of 11.6% and 16.9% respectively, which is broadly consistent with observed returns.

Despite its good power, the net beta test statistic is statistically significant at the 10% level only twice. From Table 1, it has been established that the null hypothesis that this test statistic is equal to zero is rejected almost 100% of the time if beta is unpriced and if market returns in the test period are normally distributed with a mean of 0.7% per month. These results do, then, give some support to the view that beta can help explain cross-sectional variations in average returns. This is backed by the FM test statistics, which are always positive, of similar value to the estimated equity premium and where, despite the test’s low power, the \(p\)-value is less than 10% in three out of the nine cases.

4.2. Testing the three-factor model

We now turn our attention to the three factor model of Fama and French (1993, 1996). Returns data were obtained from Professor French’s website, where a detailed description of the data is available.\(^4\) We use 100 value weighted portfolio monthly returns. These portfolios were constructed by intersecting 10 portfolios formed on size (market equity) and 10 portfolios formed on the ratio of book equity to market equity (BE/ME). The Fama-French factors, SMB and HML, are also obtained from their website. We focus on the period between 1973 and 2006. As we use the previous 60 months to estimate beta, the dataset used runs from January 1968 to February 2006.

The results are presented in Table 5. The \(p\)-values shown in the table are for

\(^{4}\)http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
two sided alternatives, except for the RM coefficient in the FM test. For the FM test, the p-value for the coefficient of market (RM) is obtained as the area to the right of the t-statistic. This is because the risk premium is positive and hence the alternative hypothesis is to the right. The coefficients of SMB and HML are assumed to be positive but we have no theoretical basis to impose a one sided test. This, however, does not affect the results as the negative statistics values are associated with very low t-statistics.

For the SMB and HML factors, the net beta test statistic has a much lower standard error than for the FM test statistic in all cases. This is consistent with the theoretical arguments given above. Interestingly, though, the reduction in standard error for the market factor by using the net beta test rather than FM test is much lower than suggested either by the theory, simulations or empirical results presented in Table 4. We can only speculate as to why this is the case, but suggest that it may be caused by inaccurate estimation of market betas in the multifactor case.

The FM test fails to reject the null that $E(\gamma_{(RM)\mu}) = 0$ in all cases. The FM test also points strongly to the lack of correlation between the size beta (SMB) and asset returns. The coefficient is significant in only one out of 11 cases. On the other hand, the significance of HML is strongly backed by the FM test. The slope associated with the value beta is highly significant in most cases. None of the three factors are found to be significant for the sample beginning in January 1988.

The net beta test supports the rejection of the 3 factor model by rejecting the null that the market risk premium is equal to the equity premium. The null is rejected at the 10% level in all cases. Given the negative and large value of the average of $\bar{\gamma}_{(RM)\mu}$, the net beta model lends support to the findings of the FM test in that the market beta is not priced. The net beta test is never able to reject the null hypothesis that the price of the HML factor is equal to the expected excess return on the mimicking portfolio. However, the FM conclusions about SMB are reversed as, in 8 of the 11 cases, the test...
is unable to reject the null that the size risk premium is equal to the expected return on the SMB mimicking portfolio at the 10% level.

5. Conclusion

It is known from a large number of studies in empirical asset pricing that it is very difficult to reject the null hypothesis that the correlation between market beta and expected returns is zero. Given the low power of the Fama and MacBeth (FM) test and the short test periods employed, the inability to reject the null would be unsurprising even if the factor model held perfectly.

In this paper, we note that the key influence that lies behind the low power of the standard FM test is the volatility of realized factor returns. These are a common component of the test statistic under both the null and alternate hypotheses. We therefore construct an alternate test, the net beta, where the realized return to the factor in each period is netted out from the FM statistic. It is shown theoretically and by simulation that the net beta statistic has a lower standard error than the FM statistic. In principle, therefore, the net beta model is better able to distinguish between the economist’s believed asset pricing model and alternative hypotheses. This is of considerable practical importance. Asset pricing models such as the CAPM and the Fama-French three factor model are often the preferred choice for estimating expected returns (Bartholdy and Peare, 2003, 2005). These estimates, in turn, are crucial inputs for capital budgeting, portfolio construction and performance evaluation. Consequently, spurious inference may lead practitioners to make suboptimal investment and appraisal decisions.

We conduct FM and net beta tests on various samples of US data. Using beta ranked portfolios to test the CAPM, the decreased standard errors of the net beta test is clearly demonstrated and is of the expected order of magnitude. It is interesting to see in these results that, when the modified test with substantially enhanced power is applied, in most cases it is not possible to reject the null hypothesis that the CAPM is true. As an illustration of a multi-factor model, we estimate the Fama and French
three factor model using 100 value weighted portfolios. The FM test suggests that both
the market and SMB are insignificant, while the net beta test suggests that only the
market is insignificant. Also, the net beta test results appear to be more consistent
across sample periods.

We recommend two main areas for future research. The first involves further
empirical tests of asset pricing models. The conditional beta test (CBT) has previously
been extensively applied to both the CAPM and Fama-French three factor model using
a wide variety of international data. However, the work of Freeman and Guermat
(2006), Koch and Westheide (2009) and Cooper (2009) has shown that it is difficult
to draw accurate inferences about the validity of asset pricing models based on these
results. The net beta test, developed here, shares many of the advantages of the
conditional beta test, particularly the low standard errors of the test statistics, but has
the clear benefit of being much easier to interpret. In addition, time-series based tests
of the Fama-French three factor model rely on the assumption that betas are constant.
In the wake of the current financial crisis, this assumption is going to be increasingly
difficult to defend as it is well known that market betas vary between bull and bear
market conditions; see, for example, Howden and Peterson (1998) and the discussion
therein. We therefore recommend that the net beta test be applied in future when
empirically examining factor asset pricing models.

The second area for future research is to make further methodological improvements.
While we have shown that the net beta test significantly improves on existing CBT
techniques, there are a number of areas where it might be further developed. First, it
can only be applied when the factors can be proxied by known mimicking portfolios;
generalizing this to all factors would allow for models such as the APT and Consumption
CAPM to be tested using this technique. Second, while the net beta test can be
directly compared against the CBT, it is not directly comparable with FM as the null
hypotheses are different. Freeman and Guermat (2006) show how the CBT can be
adapted to enable direct comparison of that method with FM; a similar adjustment to
the net beta test that retains the low standard errors associated with the test statistics.
would constitute a major advance. Finally, the current version of the test has not been extended to the zero-beta case where the zero-beta rate is unknown. We conjecture that such an extension is possible and would be based on the null hypothesis that $\hat{\gamma}_0 + \hat{\gamma}_k = 0$ for all $k$ when the cross-sectional regression is run on gross, rather than excess, returns. This, though, introduces additional econometric issues and is left for further research.
A. Appendix

In this appendix, a formal analysis of the net beta test is given for the multi-factor case that nests the CAPM as a specific example. It is shown that it is not possible to reject the null hypothesis that $\hat{\gamma}_k = 0$ for any factor $k$ if the asset pricing model is true. It will also be proved that the coefficients have lower standard error than the FM test.

A.1. Exact linear factor asset pricing models

We first specify the multi-factor asset pricing model that the economist believes best reflects the expected returns process and wishes to test. We then contrast this with the true process for expected returns that cannot be observed directly. A successful test of the asset pricing model should, at least asymptotically, be able to reject the economist’s model when it is not the same as the true model, but fail to reject it when they perfectly overlap.

An economist believes that the assets’ expected returns process is best represented by a factor model. She identifies $K$ risk factors, $f_{kt}$, that she believes are priced. Let $R_t$ denote the $N$–dimensional vector $\{R_{1t}, ..., R_{Nt}\}'$ of excess returns. The economist believes that expected returns, $E^e[R]$, are determined by:

$$E^e[R] = \beta F$$ (15)

Here $\beta$ is an $N \times (K + 1)$ matrix $(1_N, \{\beta_{ik}\})$, where $1_N$ is a vector of 1s and $\beta_{ik}$ provides the sensitivity of the $i^{th}$ risky asset to the $k^{th}$ perceived priced risk factor $(k \in \{1, ..., K\})$. $F$ is a $K + 1$–dimensional vector $(0, F_1, ..., F_K)'$ where $F_k$ is the perceived price of risk for the $k^{th}$ risk factor. For ease of exposition, $\beta$ and $F$ are assumed to be constant over time, although the adjustment to make these time-varying is straight forward. The $e$ superscript here denotes the fact that this is the expectation of the economist — a $t$ superscript will be used below to denote the true generating process.

Let $f_t$ denote the $K$–dimensional vector $(f_{1t}, ..., f_{Kt})'$. The economist places an important restriction on $f_t$. She restrains $f_t = WR_t$ where $W$ is a non-stochastic
A  \times N \) vector of weights that do not vary over time and can be perfectly identified. That is, the factors can be perfectly tracked by “mimicking portfolios” of traded assets. Because the factors are assumed to be perfectly mimicked, it is possible to identify \( F \).

Let \( W_{ki} \) represent the weight of the \( i^{th} \) asset in the \( k^{th} \) mimicking portfolio. Then:

\[
E^e[f_k] = \sum_{i=1}^{N} \sum_{\kappa=1}^{K} W_{ki} \beta_{ik} F_\kappa
\]  

(16)

It is well known that factors have a beta of one against themselves and zero against all other factors; that is, \( \sum_i W_{ki} \beta_{ik} = 1 \) if \( k = \kappa \) and zero otherwise. Therefore, \( E^e[f_k] = F_k \). That is, the risk premium for the \( k^{th} \) factor is estimated by the economist to be the expected excess return on the mimicking factor portfolio. This is a well known result.

The theoretical model in equation 15 should be contrasted with the true model. It will be assumed that this also takes an exact linear factor form, which is consistent with the fundamental theorem of asset pricing. There are \( K^* \) true factors \( \{\phi_1, ..., \phi_{K^*}\} \) that may or may not overlap with the \( K \) factors that the economist thinks are priced. The true expected returns process is given by

\[
E^t[R] = B \Lambda
\]

(17)

where \( B \) is an \( N \times (K^* + 1) \) matrix \( (1_N, \{B_{ik}\}) \), \( B_{ik} \) provides the sensitivity of the \( i^{th} \) risky asset to the \( k^{th} \) priced true risk factor. \( \Lambda \) is a \( (K^* + 1) \)-dimensional vector \( (0, \Lambda_1, ..., \Lambda_{K^*})' \) where \( \Lambda_k \) is the true price of risk for the \( k^{th} \) risk factor. \( B \) and \( \Lambda \) are also assumed to be constant over time. There is no assumption here that the \( K^* \) true priced risk factors can be associated with perfectly identified mimicking portfolios of traded assets.

A.2. Cross-sectional tests of linear factor models

So far, we have only discussed the relationship between expected returns and the factors. However, as expected returns are unobservable, all tests of asset pricing models
are conducted using realized returns. It is therefore also necessary to consider the ex-post relationship between asset returns and factor movements. The realized return can be decomposed as follows:

\[ R_t = E^t[R] + \beta(f_t^* - E^t[f_t^*]) + \varepsilon_t \]  

(18)

First, it has its true predictable component, \( E^t[R] \). Then it has deviations from this expectation. This “noise” is regressed, without loss of generality, onto the economist’s factors. \( f_t^* \) is the \( K+1 \)-dimensional vector with elements \( \{0, f_t\} \). Finally, there is the component \( \varepsilon_t \), which is the \( N \)-dimensional vector caused by unpredicted effects that are uncorrelated with the economist’s factors. There is no assumption here that the elements of \( \varepsilon_t \) are cross-sectionally independent. Taking expectations it is clear that \( E^t[\varepsilon_t] = 0 \).

The standard cross-sectional method in this case is to run OLS regressions on beta plus a constant. While betas are measured with error, for ease of exposition we assume throughout that betas are known with certainty. Then the following regression is run:

\[ R_t = \beta \lambda_t + v_t \]  

(19)

\( \lambda_t \) is a \( K + 1 \)-dimensional vector that captures the empirical prices of risk for the economist’s model with elements labelled \( \lambda_0 \) to \( \lambda_K \).

By the properties of OLS regressions, the estimator of \( \lambda_t \), denoted by \( \hat{\lambda}_t \), is given by \( \hat{\lambda}_t = (\beta'\beta)^{-1}\beta'R_t \). By substituting in from equation 18:

\[ \hat{\lambda}_t = (\beta'\beta)^{-1}\beta' \left\{ E^t[R] + \beta(f_t^* - E^t[f_t^*]) + \varepsilon_t^t \right\} \]  

(20)

\[ = f_t^* - E^t[f_t^*] + (\beta'\beta)^{-1}\beta'(BA + \varepsilon_t^t) \]

Let \( \hat{\lambda} \) denote the \( K+1 \) vector with elements \( \frac{1}{T} \sum_{t=1}^{T} \hat{\lambda}_t \) and \( \Sigma \) the constant variance-covariance matrix of \( \varepsilon_t \), then:
\[ E^t[\lambda] = (\beta'\beta)^{-1}\beta'B\Lambda \]  
\[ Var^t[\lambda] = Var[f^*_t] + (\beta'\beta)^{-1}\beta'\Sigma\beta(\beta'\beta)^{-1} \]

We now propose a simple new test of multi-factor models with mimicking portfolios. We modify the usual parameter \( \hat{\lambda}_t \) by subtracting the realized excess return on the mimicking portfolios:

\[ \hat{\lambda}_t = \hat{\lambda}_t - f^*_t \]  
\[ E^t[\hat{\gamma}] = -E^t[f^*_t] + (\beta'\beta)^{-1}\beta'B\Lambda \]  
\[ Var^t[\hat{\gamma}] = (\beta'\beta)^{-1}\beta'\Sigma\beta(\beta'\beta)^{-1} \]

The point estimate and associated test statistic, \( \hat{\gamma}_k \) and \( t_{\hat{\gamma}_k} \) are defined analogously to equation 3. In the case when the model is true \( F = E^t[f_t] = E^t[f^*_t] \), \( B = \beta \) and \( \Lambda = F \). As a consequence, under the null hypothesis that the economists model accurately captures the true model, \( E^t[\hat{\lambda}] = F \) and \( E^t[\hat{\gamma}] = 0 \). Therefore, in common with the standard CBT test, but in contrast to the FM test, failing to reject the null that \( \hat{\gamma} = 0 \) means that the asset pricing model cannot be rejected.

The key insight of this paper is that \( Var^t[\hat{\lambda}] \) increases one-to-one with \( Var[f^*_t] \). By subtracting the ex-post return to the factor mimicking portfolio from the FM statistic, \( Var^t[\hat{\gamma}] \) becomes independent of \( Var[f^*_t] \), significantly reducing the standard error associated with the test statistic.

The problem with applying this test to general multi-factor models without mimicking portfolios is that the value of \( \Lambda \) is not identified. For example, if we take the macroeconomic APT of Chen, Roll and Ross (1986), then the prices of risk for each factor, \( \Lambda_k \), are not revealed. Given this, it is not clear what the appropriate null hypothesis is for the net beta test. The requirement for the factors to be linked to asset returns is a characteristic that this test shares with time-series tests.
The results of this paper extend in a straightforward way to both generalized and weighted least squares. For example, in GLS, the second step estimate of $\hat{\lambda}_t$ is given by

$$\hat{\lambda}_t = (\beta'\Sigma^{-1}\beta)^{-1}\beta'\Sigma^{-1}R_t.$$  

By substituting in from equation 18, it follows immediately that:

$$\hat{\lambda}_t = (\beta'\Sigma^{-1}\beta)^{-1}\beta'\Sigma^{-1}\{E'[R] + \beta(f_t^* - E'[f_t^*]) + \epsilon_t^f\}$$  

(24)

Given this

$$E'[\hat{\lambda}] = (\beta'\Sigma^{-1}\beta)^{-1}\beta'\Sigma^{-1}BA$$  

$$Var'[\hat{\lambda}] = Var'[f_t^*] + (\beta'\Sigma^{-1}\beta)^{-1}$$

$$E'[\hat{\gamma}] = -E'[f_t^*] + (\beta'\Sigma^{-1}\beta)^{-1}\beta'\Sigma^{-1}BA$$  

$$Var'[\hat{\gamma}] = (\beta'\Sigma^{-1}\beta)^{-1}$$

(25)

### A.3. Necessary and sufficient conditions

We now show that finding $\hat{\lambda}_k \neq 0$ is neither a necessary nor sufficient condition for deducing that factor $k$'s risk is priced. In contrast, finding $\hat{\gamma}_k \neq 0$ for any $k$ is sufficient, but not necessary, to reject the hypothesized model.

1. Finding $\hat{\lambda}_k \neq 0$ for some factor $k$ does not imply that the factor is priced. From equation 21, $E'[\hat{\lambda}] = (\beta'\beta)^{-1}\beta'BA$. As, in general, $B \neq \beta$, $(\beta'\beta)^{-1}\beta'B$ may contain non-zero off-diagonal elements. It is therefore possible for $E'[\hat{\lambda}_k] \neq 0$ even if $\Lambda_k = 0$, provided some other factor is priced and the off-diagonal element of $(\beta'\beta)^{-1}\beta'B$ is non-zero for this factor.

2. Even if factor $k$ is priced, this does not imply that $\hat{\lambda}_k \neq 0$ even asymptotically:

   It is possible that $E'[\hat{\lambda}_k] = 0$ even if $\Lambda_k \neq 0$ provided that the non-zero terms of $(\beta'\beta)^{-1}\beta'B\Lambda$ are of different signs and offsetting magnitudes. This makes interpretation of Fama-MacBeth statistics particularly problematic even in very large samples.

3. Finding $\hat{\gamma}_k \neq 0$ for any $k$ is sufficient to establish that the economist’s model is false. If the economist’s model is true then $K = K^*$, $\beta = B$ and the factors can
still be mimicked by the weight $W_{ki}$. Given this:

$$E'[f_k] = \sum_{i=1}^{N} \sum_{\kappa=1}^{K} W_{ki}\beta_{i\kappa}\Lambda_{\kappa} = \Lambda_{k}$$

That $E[\bar{\gamma}_k] = 0$ for all $k$ in this case follows immediately from equation 22.

4. Failing to reject the null that $E[\bar{\gamma}] = 0$ cannot be taken to imply the model is true even asymptotically. For example, it is possible that $B \neq \beta$ but $(\beta'\beta)^{-1}\beta'B\Lambda = E'[f^*_t]$. In this case, $E[\bar{\gamma}] = 0$ even though the theoretical model is false.
References


Table 1: Empirical size and power of the Fama-MacBeth and net beta tests for a selection of sample sizes and market noise.

<table>
<thead>
<tr>
<th>$T$</th>
<th>Fama-MacBeth 5%</th>
<th>Net beta test 5%</th>
<th>Fama-MacBeth 10%</th>
<th>Net beta test 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_m = 2.75%$</td>
<td>Panel A: Size of test</td>
<td>Panel C: Power of test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.043 0.090</td>
<td>0.057 0.111</td>
<td>0.404 0.556</td>
<td>0.988 0.998</td>
</tr>
<tr>
<td>120</td>
<td>0.053 0.106</td>
<td>0.042 0.086</td>
<td>0.711 0.806</td>
<td>1.000 1.000</td>
</tr>
<tr>
<td>240</td>
<td>0.044 0.086</td>
<td>0.054 0.102</td>
<td>0.958 0.981</td>
<td>1.000 1.000</td>
</tr>
<tr>
<td>360</td>
<td>0.049 0.112</td>
<td>0.053 0.098</td>
<td>0.991 0.996</td>
<td>1.000 1.000</td>
</tr>
<tr>
<td>$\sigma_m = 5.50%$</td>
<td>Panel B: Size of test</td>
<td>Panel D: Power of test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.047 0.100</td>
<td>0.045 0.100</td>
<td>0.162 0.244</td>
<td>0.987 0.996</td>
</tr>
<tr>
<td>120</td>
<td>0.049 0.097</td>
<td>0.049 0.093</td>
<td>0.266 0.380</td>
<td>1.000 1.000</td>
</tr>
<tr>
<td>240</td>
<td>0.055 0.093</td>
<td>0.049 0.092</td>
<td>0.498 0.611</td>
<td>1.000 1.000</td>
</tr>
<tr>
<td>360</td>
<td>0.036 0.082</td>
<td>0.058 0.105</td>
<td>0.651 0.762</td>
<td>1.000 1.000</td>
</tr>
<tr>
<td>1000</td>
<td>0.051 0.091</td>
<td>0.044 0.097</td>
<td>0.971 0.987</td>
<td>1.000 1.000</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Variance</td>
<td>5th Percent</td>
<td>95th Percent</td>
</tr>
<tr>
<td>------------------</td>
<td>-------</td>
<td>----------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Fama-MacBeth ($H_0$)</td>
<td>0.020</td>
<td>1.031</td>
<td>-1.647</td>
<td>1.681</td>
</tr>
<tr>
<td>Net beta ($H_0$)</td>
<td>-0.031</td>
<td>1.017</td>
<td>-1.718</td>
<td>1.644</td>
</tr>
<tr>
<td>Fama-MacBeth ($H_A$)</td>
<td>0.995</td>
<td>1.031</td>
<td>-0.666</td>
<td>2.696</td>
</tr>
<tr>
<td>Net beta ($H_A$)</td>
<td>-4.377</td>
<td>1.216</td>
<td>-6.227</td>
<td>-2.630</td>
</tr>
</tbody>
</table>

Table 2: Summary statistics of simulated t-statistics for a selection of sample sizes ($\sigma_m = 5.50\%$). (*) JB is the Jarque-Bera normality test. Means and Variances are for the sample; the Kurtosis figure refers to excess kurtosis.
<table>
<thead>
<tr>
<th>Start (months)</th>
<th>T (months)</th>
<th>Sample Mean (%)</th>
<th>Standard Error (%)</th>
<th>p-value</th>
<th>Sample Mean (%)</th>
<th>Standard Error (%)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-73</td>
<td>120</td>
<td>0.553</td>
<td>0.535</td>
<td>0.303</td>
<td>0.527</td>
<td>0.298</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>240</td>
<td>0.497</td>
<td>0.351</td>
<td>0.158</td>
<td>0.13</td>
<td>0.193</td>
<td>0.503</td>
</tr>
<tr>
<td></td>
<td>336</td>
<td>0.583</td>
<td>0.305</td>
<td>0.056</td>
<td>0.058</td>
<td>0.183</td>
<td>0.75</td>
</tr>
<tr>
<td>Jan-78</td>
<td>120</td>
<td>0.577</td>
<td>0.451</td>
<td>0.204</td>
<td>0.026</td>
<td>0.235</td>
<td>0.911</td>
</tr>
<tr>
<td></td>
<td>240</td>
<td>0.512</td>
<td>0.297</td>
<td>0.085</td>
<td>-0.234</td>
<td>0.168</td>
<td>0.165</td>
</tr>
<tr>
<td>Jan-83</td>
<td>120</td>
<td>0.440</td>
<td>0.457</td>
<td>0.337</td>
<td>-0.267</td>
<td>0.243</td>
<td>0.274</td>
</tr>
<tr>
<td></td>
<td>216</td>
<td>0.600</td>
<td>0.370</td>
<td>0.107</td>
<td>-0.202</td>
<td>0.230</td>
<td>0.383</td>
</tr>
<tr>
<td>Jan-88</td>
<td>120</td>
<td>0.448</td>
<td>0.387</td>
<td>0.249</td>
<td>-0.494</td>
<td>0.239</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>156</td>
<td>0.771</td>
<td>0.449</td>
<td>0.088</td>
<td>-0.09</td>
<td>0.294</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Table 4: CAPM estimation results for the FM and net beta tests using US market data, four starting dates, and a selection of sample sizes.
<table>
<thead>
<tr>
<th>T (months)</th>
<th>Factor</th>
<th>Sample Mean (%)</th>
<th>Standard Error</th>
<th>p-value</th>
<th>Sample Mean(%)</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FM test</td>
<td></td>
<td></td>
<td></td>
<td>Net beta test</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sample beginning January 1973</td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>RM</td>
<td>-0.565</td>
<td>0.367</td>
<td>0.937</td>
<td>-0.645</td>
<td>0.383</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>0.570</td>
<td>0.263</td>
<td>0.032</td>
<td>-0.025</td>
<td>0.070</td>
<td>0.720</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.622</td>
<td>0.258</td>
<td>0.017</td>
<td>-0.023</td>
<td>0.084</td>
<td>0.785</td>
</tr>
<tr>
<td>240</td>
<td>RM</td>
<td>-0.443</td>
<td>0.265</td>
<td>0.952</td>
<td>-0.828</td>
<td>0.268</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>0.160</td>
<td>0.169</td>
<td>0.344</td>
<td>-0.073</td>
<td>0.043</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.520</td>
<td>0.169</td>
<td>0.002</td>
<td>-0.016</td>
<td>0.051</td>
<td>0.754</td>
</tr>
<tr>
<td>360</td>
<td>RM</td>
<td>-0.638</td>
<td>0.228</td>
<td>0.997</td>
<td>-1.030</td>
<td>0.222</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>0.137</td>
<td>0.169</td>
<td>0.418</td>
<td>-0.043</td>
<td>0.041</td>
<td>0.293</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.465</td>
<td>0.167</td>
<td>0.006</td>
<td>-0.020</td>
<td>0.048</td>
<td>0.679</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sample beginning January 1978</td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>RM</td>
<td>-0.637</td>
<td>0.343</td>
<td>0.967</td>
<td>-1.240</td>
<td>0.398</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>0.229</td>
<td>0.213</td>
<td>0.286</td>
<td>-0.082</td>
<td>0.055</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.430</td>
<td>0.242</td>
<td>0.079</td>
<td>-0.026</td>
<td>0.063</td>
<td>0.681</td>
</tr>
<tr>
<td>240</td>
<td>RM</td>
<td>-0.506</td>
<td>0.245</td>
<td>0.980</td>
<td>-1.262</td>
<td>0.245</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>0.045</td>
<td>0.154</td>
<td>0.770</td>
<td>-0.078</td>
<td>0.041</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.344</td>
<td>0.154</td>
<td>0.026</td>
<td>0.007</td>
<td>0.044</td>
<td>0.875</td>
</tr>
<tr>
<td>337</td>
<td>RM</td>
<td>-0.586</td>
<td>0.229</td>
<td>0.994</td>
<td>-1.221</td>
<td>0.215</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>0.170</td>
<td>0.170</td>
<td>0.317</td>
<td>-0.053</td>
<td>0.041</td>
<td>0.192</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.344</td>
<td>0.166</td>
<td>0.039</td>
<td>-0.034</td>
<td>0.046</td>
<td>0.463</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sample beginning January 1983</td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>RM</td>
<td>-0.318</td>
<td>0.380</td>
<td>0.798</td>
<td>-1.033</td>
<td>0.374</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>-0.217</td>
<td>0.207</td>
<td>0.298</td>
<td>-0.109</td>
<td>0.050</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.396</td>
<td>0.216</td>
<td>0.069</td>
<td>-0.019</td>
<td>0.057</td>
<td>0.738</td>
</tr>
<tr>
<td>240</td>
<td>RM</td>
<td>-0.672</td>
<td>0.287</td>
<td>0.990</td>
<td>-1.233</td>
<td>0.270</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>-0.064</td>
<td>0.215</td>
<td>0.767</td>
<td>-0.046</td>
<td>0.050</td>
<td>0.361</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.375</td>
<td>0.214</td>
<td>0.082</td>
<td>-0.023</td>
<td>0.057</td>
<td>0.689</td>
</tr>
<tr>
<td>277</td>
<td>RM</td>
<td>-0.605</td>
<td>0.254</td>
<td>0.991</td>
<td>-1.272</td>
<td>0.239</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>0.026</td>
<td>0.191</td>
<td>0.892</td>
<td>-0.056</td>
<td>0.046</td>
<td>0.223</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.381</td>
<td>0.188</td>
<td>0.044</td>
<td>-0.039</td>
<td>0.053</td>
<td>0.460</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sample beginning January 1988</td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>RM</td>
<td>-0.426</td>
<td>0.351</td>
<td>0.890</td>
<td>-1.365</td>
<td>0.295</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>-0.143</td>
<td>0.224</td>
<td>0.524</td>
<td>-0.071</td>
<td>0.059</td>
<td>0.233</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.290</td>
<td>0.190</td>
<td>0.130</td>
<td>0.031</td>
<td>0.055</td>
<td>0.572</td>
</tr>
<tr>
<td>217</td>
<td>RM</td>
<td>-0.585</td>
<td>0.301</td>
<td>0.973</td>
<td>-1.256</td>
<td>0.254</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>SMB</td>
<td>0.134</td>
<td>0.235</td>
<td>0.569</td>
<td>-0.036</td>
<td>0.055</td>
<td>0.515</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td>0.314</td>
<td>0.221</td>
<td>0.156</td>
<td>-0.042</td>
<td>0.062</td>
<td>0.496</td>
</tr>
</tbody>
</table>

Table 5: Fama-French three factor model estimation results for the FM and net beta tests using US market data, four starting dates, and a selection of sample sizes.