A Closer Look at Return Predictability of US Stock Market: Evidence from Fama-French Portfolio and Panel Variance Ratio Test

Jae H. Kim
Department of Finance
La Trobe School of Business
La Trobe University
VIC 3086, Australia

Abul Shamsuddin
Newcastle Business School
University of Newcastle
Callaghan, NSW 2308
NSW 2300 Australia

Abstract
This paper examines return predictability of the US stock market using Fama-French stock portfolios. We consider a wide range of portfolios classified by risk factors such as size, book-to-market ratios, and momentum. We also consider general industry portfolios, as well as those classified into tradable vs. non-tradable; and hi-tech vs. non-hi-tech. We evaluate time-varying return predictability using a new panel variance ratio test proposed in this paper. We conduct extensive Monte Carlo experiment to find that these tests exhibit desirable small sample properties, with correct size and power that substantially increases with the number of cross-sectional units.

At the aggregate level, it is found that the portfolio returns have been highly predictable from 1964 to 1997, except for the timing of 1987 stock market crash and its aftermath. After 1997, the U.S stock portfolio returns have been unpredictable overall, apart from the period of the global financial crisis. At the disaggregated level, the large-cap portfolio return and hi-tech industry portfolio returns exhibit different patterns from the general industry portfolio, showing less degree of return predictability over time. Contrary to the general belief that the U.S. stock market has been weak-form efficient, this paper finds that the weak-from efficiency is prevalent only from 1997. It is also found that large-cap and hi-tech industries stocks are found to be more efficient than the other sections of the market.

1 Corresponding author: Tel +61 3 94796616, Fax +61 3 94791654, Email: J.Kim@latrobe.edu.au
1. Introduction

Under the efficient market hypothesis (Fama, 1970), stock returns are purely unpredictable since the stock prices adjust instantly to the desired level fully reflecting all available information. With the weak-form efficiency in which the information set is limited to the past prices and returns, the stock returns cannot be predicted by exploiting past price information. The hypothesis has been tested extensively for decades, with rather mixed empirical results (see, for example, Yen and Lee; 2008; Park and Irwin; 2007). However, a generally held view is that, while the stock market is efficient most of times, it can deviate from efficiency from time to time, depending on prevailing market situations (Malkiel, 2003; Timmermann, 2008). Lo’s (2005) adaptive markets hypothesis asserts that market efficiency is highly context dependent and dynamic; and that return predictability can arises from time to time as the market participants adapt to different market conditions. Recent empirical studies in support of the adaptive markets hypothesis include Kim et al. (2011) in the context of the stock market; and Neely et al. (2009) and Charles et al., (2012) for foreign exchange markets. In particular, Kim et al. (2011), exploiting a century-long Dow-Jones industrial average index, report strong evidence that the US stock return predictability has been changing over time, depending largely on market conditions such as political crises, market crashes, and economic bubbles.

In this paper, we examine the return predictability of U.S. stock portfolios at a highly disaggregated level. This is in contrast with recent and past studies where index returns are exclusively examined. For this purpose, the portfolios constructed by Fama and French (1993) are employed, in which stock portfolios are classified by industry; and by a range of risk factors. The latter include size, book-to-market ratio, momentum, short-term reversal, and long-term reversal. Fama and French (1993) have found that the returns from
these risk factors (particularly size and book-to-market ratio) have shown high explanatory power for portfolio returns, when they are added to the CAPM regression as additional explanatory variables. In addition, it is possible that industry structure plays a role in return predictability. For example, Shykevich (2012) notes that small-cap and hi-tech industries represent growth industries whose stock returns can be more predictable than the others. Griffin and Karolyi (1998) pay their attention to traded vs. non-traded industries, and argue that the two are fundamentally different in their exposure to exchange rates and sensitivity to prices, which can affect their profitability.

The purpose of this paper is to examine how return predictability of US stock portfolios changes over time, in relation to these Fama-French risk factors and industry factors. The results will have strong implications as to how these risk and industry factors contribute to the degree of weak-form efficiency of the US market at highly disaggregated levels. We examine Fama-French portfolio returns, daily from 1964 to 2011, classified by industry and a range of risk factors. The industry portfolios are further classified into tradable vs non-tradable; and hi-tech vs. non-hi-tech industry portfolios. To the best of our knowledge, an investigation of stock return predictability at such a disaggregated level has not been conducted extensively in the literature.

To measure the return predictability of Fama-French portfolios at a disaggregated level, we employ a variance ratio test for a cross-section of portfolio returns from 1964 to 2011. This requires a reliable and powerful test which can be applied to a panel of portfolio returns. To this end, we propose a panel variance ratio test, as an extension of the automatic variance ratio test of Kim (2009) and Choi (1999). Through an extensive Monte Carlo experiment, we find that these new panel variance ratio tests show desirable size and
power properties in small samples, under a wide range of models subject to unknown forms of conditional heteroskedasticity. We apply the panel variance ratio test to the daily data, using 1-year subsample windows moving every 6 month. This enables us to estimate time-varying return predictability, also providing effective guard against possible data snooping bias (see Hsu and Kuan, 2005).

The main findings of the paper can be summarized as follows. The U.S industry portfolios have been highly predictable from 1964 to 1997, except for the timing of the 1987 stock market crash and its aftermath. From 1997, the returns have been unpredictable apart from the period of global financial crisis (2007-2008). Among the Fama-French risk factors, only the size is found to make difference in return predictability. Large-cap portfolios exhibit different pattern in return predictability from that of the general industry portfolios. Other risk factors such as book-to-market ratio, momentum, long-term reversal, and short-term are found to make little differences in return predictability. It is also found that returns form hi-tech industry portfolios have been more efficient than non-hi-tech portfolios, while return predictability of the latter follow the general pattern of the industry portfolios. The portfolios from the tradable industries also show little difference from non-tradable industries. In the next section, we present the panel variance ratio tests and report the results of the Monte Carlo experiment for its small sample properties. Section 3 provides the data details. Section 4 presents the empirical results, and Section 5 concludes the paper.

2. Panel Variance Ratio Test
The variance ratio test has been widely used in empirical finance literature as a tool to test for the presence of return predictability in asset returns or weak-form efficiency of the financial market, since the seminal work of Lo and MacKinlay (1988). A number of improved versions have been proposed; Charles and Darne (2009) provide a detailed survey of recent developments in the variance ratio tests. Notable recent contribution is the automatic variance ratio test of Kim (2009), which improves the small sample properties of the original version of Choi (1999) with the use of wild bootstrapping. Kim’s (2009) test selects the optimal holding period automatically and delivers desirable small sample properties under conditional heteroskedasticity: see, for details, Charles et al. (2011).

All tests developed so far, however, are univariate tests applicable to a single time series. When there are a number of cross-sectional units are available, it is possible that the use of panel test can substantially improve the power of the test. To this end, we propose a panel test version of the automatic variance ratio test in this paper. To the best of our knowledge, the panel variance ratio test has not been explored in the literature, although a recent study of Okui (2010) considers estimation of autocorrelations using panel data.

**Panel Variance Ratio Test**

Let $Y_{it}$ be an asset return at time $t$ ($t = 1, \ldots, T$) for a cross-sectional unit $i$ ($i = 1, \ldots, N$). It is assumed that asset returns of different cross-sectional units ($Y_{it}$ and $Y_{jt}$) are correlated contemporaneously. We consider the automatic variance ratio test of the form

$$VR_{i}(k) = 1 + 2 \sum_{j=1}^{T-1} m(j/k) \hat{\rho}(j)$$

(1)
where \( \hat{\rho}_j(j) = \frac{\sum_{i=1}^{T-j}(Y_{it} - \hat{\mu})(Y_{it+j} - \hat{\mu})}{\sum_{i=1}^{T} (Y_{it} - \hat{\mu})^2} \) and \( \hat{\mu} = T^{-1} \sum_{i=1}^{T} Y_{it} \), while

\[
m(x) = \frac{25}{12\pi^2 x^2} \left[ \frac{\sin(6\pi x / 5)}{6\pi x / 5} - \cos(6\pi x / 5) \right]
\]

is the quadratic spectral kernel. Choi (1999) stated that \( VR(k) \) in (1) is a consistent estimator for \( 2\pi f_Y(0) \), where \( f_Y(0) \) is the normalized spectral density for \( Y_t \) at the frequency zero. Choi (1999) showed that, under \( H_0^a: Y_{it} \) is serially uncorrelated (or \( H_0^b: 2\pi f_Y(0) = 1 \)),

\[
AVR_i(k) = \sqrt{\frac{k}{T}} VR_i(k) - 1 \sqrt{2} \xrightarrow{d} N(0,1)
\]

as \( k \to \infty, T \to \infty, T/k \to \infty \). Kim (2009) proposes wild bootstrapping of the \( AVR \) statistic given in (1) for substantially improved small sample properties.

The panel VR test is to test for the null that \( Y_{it} \) is serially uncorrelated for all \( i \), against the alternative that at least one \( Y_{it} \) is serially correlated. Assuming the independence of cross sectional units, \( AVR_i = \sqrt{N AVR} \) follow \( N(0,1) \) asymptotically where \( AVR = \frac{1}{N} \sum AVR_i \);

and \( AVR_2 = \sum_{i=1}^{N} AVR_i^2 \) asymptotically follows the chi-squared distribution with \( N \) degrees of freedom, under the null hypothesis. However, it is unlikely that asset returns are independent over the cross-sectional units. We propose the wild bootstrap which can replicate the cross-sectional dependence as well as the unknown forms of the heteroskedasticity in individual return time series.

Let \( Y_t = (Y_{1t}, \ldots, Y_{Nt}) \). The wild bootstrap can be conducted in three stages as follows:
(i) Form a bootstrap sample of $T$ observations $Y_t = \eta_t Y_t \ (t=1, \ldots, T)$ where $\eta_t$ is a random sequence with $E(\eta_t) = 0$ and $E(\eta_t^2) = 1$;

(ii) Calculate the $AVR_1^*$ statistic, which is the $AVR_1$ statistic calculated from $Y_t^*$.

(iii) Repeat (i) and (ii) $B$ times to obtain the bootstrap distribution $\{AVR_1^*(j)\}_{j=1}^B$.

The two-tailed p-value can be obtained as the proportion of the $\{AVR_1^*(j)\}_{j=1}^B$ in absolute values, which are greater than $AVR_1$ statistic in absolute value. The bootstrap distribution $\{AVR_2^*(j)\}_{j=1}^B$ can be obtained in a similar way, and its one-tailed p-value can be obtained as the proportion of the $\{AVR_2^*(j)\}_{j=1}^B$ in absolute values, which are greater than $AVR_2$ statistic.

Without loss of generality, consider the bivariate case where $Y_t = (Y_{1t}, Y_{2t})$ and $Y_t^* = (\eta_t Y_{1t}, \eta_t Y_{2t})$. Conditionally on data, it can be shown that

\[ Var(Y_{it}^*) = Y_{it}^2 \text{ for } i = 1, 2; \]
\[ E(Y_{1t}^* Y_{2t}^*) = Y_{1t} Y_{2t}; \]
\[ Cov(Y_{1t}^{*2}, Y_{2t}^{*2}) = E(\eta_t^4 Y_{1t}^{*2} Y_{2t}^{*2} - Y_{1t}^{*2} Y_{2t}^{*2}). \]

That is, the wild bootstrap can effectively replicate the heteroskedasticity of the individual time series and possible correlations among the mean. To replicate the second-order dependence between the two cross-sectional units, we need to find $\eta_t$ that satisfies $E(\eta_t^4) = 2$ so that $Cov(Y_{1t}^{*2}, Y_{2t}^{*2}) = Y_{1t}^{*2} Y_{2t}^{*2}$. A widely used distribution that meets such a requirement is the two-point distribution of Mammen (1993), which can be written as.
Monte Carlo Design

We consider the sample size $T = 50, 100, 300$ and cross-sectional unit $N = 1, 3, 5, 10, 20$. To evaluate the size properties, we consider GARCH(1,1) process of the form $Y_{it} = u_{it}$;

$$u_{it} = \sqrt{h_u} \varepsilon_{it}; h_u = 0.01 + \alpha_i h_{u,t-1} + \beta_i \varepsilon_{it-1}^2,$$

where $\alpha_i \sim U(0.8, 0.9)$ and $\beta_i = 0.09$. We also consider stochastic volatility (SVOL) process of the form

$$u_{it} = \exp(0.5h_u)\varepsilon_{it}; h_u = \gamma_i h_{u,t-1} + \nu_u,$$

where $\gamma_i \sim U(0.9, 0.95)$. Note that $\varepsilon_{it}$ is generated from a multivariate normal distribution with zero mean and variance-covariance matrix $\Sigma$, where $\Sigma = (\rho_{ij})^{n\times n}$ with $\rho_{ij} = E(\varepsilon_{it} \varepsilon_{jt})$ while $E(\varepsilon_{it} \varepsilon_{js}) = 0$ where $t \neq s$. The form of $\Sigma$ matrix considered is $\rho_{ii} = 1$ and $\rho_{ij} = 0.5$ for $i \neq j$. Note that $\nu_u \sim N(0,0.1)$ independent of $\varepsilon_{it}$.

To evaluate the power, we consider AR(1) models, ARFIMA models, and NDAR (white noise plus the first difference of an AR(1) models. The AR(1) models take the following form:

$$Y_{it} = \alpha_i Y_{it-1} + u_{it},$$

where $\alpha_i \sim U(0.09,0.11)$. We consider two types models for the error term $u_{it}$: a GARCH(1,1) and SVOL model as detailed above. For the ARFIMA model, we consider

$$(1 - B)^{\delta} Y_{it} = u_{it},$$
where $\delta_i \sim U(0.10, 0.15)$, and GARCH(1,1) and SVOL model for the error terms $u_{it}$. The
NDAR model takes the form

$$Y_{it} = \varepsilon_{it} + X_{it} - X_{it-1}; X_{it} = \beta_i X_{it-1} + v_{it},$$

where $\beta_i \sim U(0.8, 0.9)$, while $u_{it}$ and $\varepsilon_{it}$ are independent $N(0,1)$ error terms. Note that $\varepsilon_{it}$’s
have the correlation structure given by $\Sigma$ above.

The models for power evaluation presented above are representatives of the cases where
all cross-sectional units show serial correlation or predictability to a similar degree. In
practice, however, it is often the case that a high degree of cross-sectional heterogeneity
exists in the panel, where some show little predictability while other cross-section units
are highly correlated. In the latter, it is also possible that some show positive serial
correlation while others show negative serial correlation. To evaluate such situation, we
consider the AR(1) models with $\alpha_i \sim U(-0.2, 0.2)$. These models are labeled AR1HET
models. Table 1 presents the realizations for the AR(1) coefficients used for simulation
and their summary statistics for each number of cross-sectional units.

**Monte Carlo Results**

We first present the size and power of $AVR_1$ test only. This is because $AVR_2$ test shows
power always slightly less than those of the $AVR_1$ test in small samples, with no size
distortion. This is true for all models except for AR1HET. For the latter case, the power
properties of both statistics will be presented.

Figure 1 presents the size properties (probability of rejecting the true null hypothesis) of
the $AVR_1$ test. For the GARCH(1,1) model, there is no sign of size distortion for all sample
sizes and the number of cross-sectional units. For the SVOL model, there is tendency of
over-rejecting the null hypothesis when the sample size is 50, but this size distortion quickly disappears as the sample size increases. Hence, the results show that there is no sign of serious size distortion, especially when the sample size is higher than 100, for all numbers of cross-sectional units.

Figure 2 presents the power properties (probability of rejecting the false null hypothesis) for alternative models for the $AVR_1$ test. It is evident that the power increases with the sample size, and also as the number of cross-sectional unit increases. The results so far are obtained under the assumption that the cross-sectional units are moderately correlated with correlation coefficient 0.5. Figure 3 presents the power properties under different values of this correlation for the AR(1) model with GARCH(1,1) errors. When this correlation is high (0.9), the power increases with the sample size but at a much slower rate than when the correlation is moderate. In addition, the power function when the correlation is 0.9 is flat against the number of $N$, indicating that inclusion of extra cross-sectional units does not provide noticeable power improvement. When the cross-sectional units are correlated weakly with correlation of 0.1, the reverse is the case. The power improves more sharply with sample size and with the number of cross-sectional units than when the correlation is 0.5.

Figure 4 presents the power properties of $AVR_1$ and $AVR_2$ tests for AR1HET models under GARCH(1,1) errors with $\rho_y = 0.5$. The summary statistics for AR(1) coefficient for cross-sectional units are reported in Table 1. The former shows little power when $N \geq 3$ even when the sample size is large. As it is based on the sample mean of the $AVR$ statistics from all cross-sectional units, the $AVR_1$ test statistic can be misleading under a high degree of cross-sectional heterogeneity since the individual statistics cancel out when the mean is
calculated. By construction, the $AVR_2$ test statistic should be robust to a high degree of cross-sectional heterogeneity. Indeed, the power of the $AVR_2$ test increases with $N$ and $T$. Hence, when a high degree of cross-sectional heterogeneity is expected, the use of $AVR_2$ test statistic should be preferred.

3. Data and Computation Details

We use portfolio returns available at Ken French’s data library\(^2\), daily from 1964 to 2011, both value-weighted and equal-weighted. All returns are excess returns adjusted with risk-free rate. We use 5, 10, 30 industry portfolios return to examine the return predictability of general portfolios. The observed pattern will be used as the benchmark for the general trend for return predictability. We will also examine whether the degree of disaggregation has any impact on the return predictability. In the context of the panel variance ratio test, the numbers of cross-section units $N$ are 5, 10, and 30 respectively.

We take 25 portfolios formed based on size and book-to-market: we use five-smallest portfolios and five-largest portfolios based on size ($N=5$). Similarly, we take 100 portfolios formed based on size and book-to-market, and use ten-smallest portfolios and 10-largest portfolios based on size ($N=10$). By taking these extreme portfolios, we examine the size effect on the return predictability and weak-form efficiency. We also take ten portfolios formed on momentum factor; ten formed on short-term reversal factor; and ten formed on long-term reversal factor. For each set of portfolio, we use three extreme portfolios to examine the effects of these risk factors on return predictability ($N=3$).

\(^2\) [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)
We classify 10 industry portfolios into tradable and non-tradable industries guided by Griffin and Karolyi (1998); the former include consumer goods (durables and non-durables), manufacturing goods, oil products, and business equipments while the latter telecommunications, shops (retail and wholesale) and services, healthcare, and utilities. We also classify and hi-tech and non-hi-tech industries as in Shynkevich (2012): the former includes consumer durables (cars, TV’s, furniture, and household appliances), business equipments (computers, software, electronic equipments), and telecommunication (telephone and TV transmissions); while the latter for all other industries. We also considered the same classifications (tradable vs. non-tradable; hi-tech vs. non-hi-tech) from more disaggregated industry portfolios (e.g., 30 industry portfolios), but the empirical results were qualitatively no different.

We take a moving rolling-windows approach in this paper; that is, one-year window moving every 6 month. We take a cross-section of daily portfolio returns for a year of January to December 1964 (T=252 approximately), and then calculate the p-values of the panel variance ratio tests. Then, we move the window six month forward, by taking one-year daily data from July 1964 to June 1965, and calculate the p-values. This process continues until the last window which covers the period from January to December 2011. There are two benefits of this moving sub-sample window approach. First, it enables us to examine the evolution of time-varying return predictability. Second, as Hsu and Kuan (2005; p.608) point out, it is a useful tool to address the problem of data snooping bias, as an effective alternative to the statistical tests of White (2000) and Hansen (2005). It should be noted that the moving sub-sample window approach is not intended for multiple testing, but as a means of measuring the time variation in the degree of return predictability.
4. Empirical Results

For simplicity, we only report $AVR_2$ test statistics for value-weighted portfolios. The results from $AVR_1$ statistic and those from equal-weighted portfolios show qualitatively similar results overall. The p-values of the $AVR_2$ statistics, calculated with moving sub-sample window of 1-year, are plotted over time. If the p-values are less than 0.05 or 0.10, the null hypothesis of no return predictability (weak-form efficiency) is rejected at 5% or 10% level of significance at the particular time.

Figure 5 reports industry portfolios with different levels of aggregation: 5, 10, and 30 industry portfolios. The panel VR tests applied to these portfolios show virtually identical pattern over time, regardless of the level of disaggregation. That is, the portfolios returns have been highly predictable since mid-sixties until the period of 1987-1989, which represents the timing of the 1987 stock market crash and its aftermath. The returns become predictable again until 1997. From 1997, the portfolio returns become unpredictable overall, apart from the brief period of the global financial crisis (2007-2008). The findings are consistent with those of Kim et al. (2011), who reported no predictability during the market crashes but predictability during the economic crises.

Figure 6 plots p-values of the test for the size-sorted portfolios. The first graph corresponds to the 10 smallest of the 100 size-sorted portfolios; the second to the 10 middle-sized portfolios; and the third 10 biggest of the 100 size-sorted portfolios. From the first graph, it can be seen that the small-sized portfolios have been highly predictable until 2006, with an exception of the timing of the 1987 stock market crash. The middle-sized portfolios show similar pattern, but they become unpredictable from 1999. These small-cap and mid-cap portfolios show the same pattern in return predictability as the
benchmark industry portfolios as observed in Figure 5. However, the large-sized portfolios show different pattern, exhibiting no return predictability from early 1980’s. They show three brief episodes return predictability after 1987 crash: mid-90’s, in 2003; and in 2007-2009 during the global financial crisis. This observation is in broad agreement with that of Kim et al. (2011, Figure 2), where the returns from the Dow-Jones index (a portfolio of large stocks) show predictability until eighties and show a period predictability during the period of 2007 global financial crisis.

Figure 7 plots the p-values of the test on the returns formed based on book-to-market ratio. The first graph plots the p-values of the portfolio on three lowest deciles and the second those from three highest deciles. The former are chosen to represent growth stocks and the latter value stocks. The result shows no noticeable differences between the two groups, indicating that these sectors of the market show similar degree of weak-from market efficiency. In addition, they show the same pattern as the general industry portfolios observed in Figure 5. Overall, the growth stocks tend to show more episodes no return predictability, while classification based on growth or value stocks based on book-to-market ratio makes little difference in return predictability. Figure 8 plots the p-values of the test on the returns formed based on momentum factors. It again appears that the difference in momentum factors makes little contribution to the return predictability.

Figure 9 plots the case of non-tradable versus tradable goods. Again, there are no noticeable differences in return predictability. For all three cases, the return predictability follows the pattern of those of the general industry portfolios as observed in Figure 5.

Figure 10 presents the p-values for the industry portfolios for hi-tech and non hi-tech industries. It is clear that the former show different pattern form the latter, while the latter
shows similar pattern from general industry portfolios presented in Figure 5. There are a number of prolonged episodes of market efficiency from 60’s to late nineties. Even in 2000’s, the former industry portfolios show smaller number of predictable episodes. During the global financial crisis, the hi-tech industry portfolio shows a prolonged period of no return predictability.

The empirical results suggest that the overall market has been inefficient from mid-60’s to late 90’s, apart from the period of 1987 stock market crash. From late 1990’s, the market has been efficient apart from the episode of the global financial crisis. The results also show that the large-cap stocks and hi-tech stocks has been showing different pattern in return predictability from the others, as they have been more efficient before 1997. Other factors such as the book-to-market ratio and momentum factors contribute make little difference to the evolution of stock return predictability.

5. Conclusion

This paper examines stock return predictability of the U.S. market at a highly disaggregated level, by exploiting the returns from the Fama-French portfolios daily from 1964 to 2011. The latter provides portfolio returns formed based on a range of risk factors such as size, book-to-market ratio, and momentum factor. They also provide highly disaggregated industry portfolios, from which the returns from tradable and non-tradable industry portfolios are obtained. We also obtain the returns from hi-tech and non-high-tech industry portfolios. We consider both value-weighted and equal weighted portfolios, and their returns are adjusted with the risk free rate.
To test and measure the degree of predictability for the panel of portfolio returns, we propose the use of a panel variance ratio test in this paper. It is an extension of the automatic variance ratio test of Kim (2009), which is valid under unknown forms of conditional heteroskedasticity and cross-section correlations. Our Monte Carlo experiment reveals that the test has desirable small sample properties with correct size and high power. It is found that the power of the test increases dramatically with the number of cross-sectional units, when the cross-sectional correlation is low or moderate.

From the analysis of a range of industry portfolios, a general pattern in the predictability of the U.S. market has emerged. The returns have been highly predictable from 1964 to 1997, except for the timing of the 1987 stock market crash and its aftermath. From 1997, the returns have been unpredictable apart from the period of global financial crisis (2007-2008). This finding is consistent with Kim et al. (2011) who reported little predictability during stock market crash and high predictability during crises. Among the Fama-French risk factors, only the size is found to make difference in return predictability. It is found that the returns from small-cap portfolios have been highly predictable, compared to those from large-cap portfolios which have been largely unpredictable from 1980’s. Other risk factors such as book-to-market ratio, momentum, long-term reversal, and short-term are found to make little differences in return predictability. It is also found that returns form hi-tech industry portfolios have been more efficient than non-hi-tech portfolios, while return predictability of the latter follow the general pattern of the industry portfolios. The portfolios from the tradable industries also show little difference from non-tradable industries.
Contrary to the general belief that the U.S. stock market has been weak-form efficient, this paper, with the use of highly powerful and robust panel variance ratio test, finds that the market efficiency is prevalent only from 1997. It is also found that only the small sections of the market (large cap companies; hitech industries) are found to be more efficient than the other sections of the market.
References


Table 1. Realizations of the AR(1) coefficients used for AR1HET models

<table>
<thead>
<tr>
<th>$N$</th>
<th>Min</th>
<th>1$^{st}$ Q</th>
<th>Median</th>
<th>3$^{rd}$ Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.19</td>
<td></td>
<td>-0.19</td>
<td></td>
<td>-0.19</td>
</tr>
<tr>
<td>3</td>
<td>-0.19</td>
<td>0.02</td>
<td></td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.16</td>
<td>-0.12</td>
<td>0.02</td>
<td>0.03</td>
<td>0.16</td>
</tr>
<tr>
<td>10</td>
<td>-0.18</td>
<td>-0.11</td>
<td>0.03</td>
<td>0.11</td>
<td>0.19</td>
</tr>
<tr>
<td>20</td>
<td>-0.19</td>
<td>-0.09</td>
<td>-0.02</td>
<td>0.07</td>
<td>0.19</td>
</tr>
</tbody>
</table>
Figure 1  Size Properties of the Test (Nominal Size 5%)

Note

Y-axis represents the percentage of rejecting the null hypothesis and X-axis the number of cross-sectional units.

95% confidence interval for the population proportion of 5% is (3.65, 6.35)
Figure 2  Power Properties of the Test (Nominal Size 5%)

Note
Y-axis represents the percentage of rejecting the null hypothesis and X-axis the number of cross-sectional units.
Note

Y-axis represents the percentage of rejecting the null hypothesis and X-axis the number of cross-sectional units.
Figure 3  Power Properties of the Test Under Different Contemporaneous Correlations (Nominal Size 5%)

Note
Y-axis represents the percentage of rejecting the null hypothesis and X-axis the number of cross-sectional units
Figure 4  Power Properties of AVR1 and AVR2 Statistics under Cross-Sectional Heterogeneity

Note
Y-axis represents the percentage of rejecting the null hypothesis and X-axis the number of cross-sectional units
Figure 5. P-values of the AVR$_2$ statistic for industry portfolios

The horizontal lines represent 0.05 and 0.10. The AVR2 test is applied to the industry portfolio returns consist of 5, 10, 30 industries (i.e., N = 5, 10, 30).
Figure 6. P-values of the AVR₂ statistic for industry portfolios based on size.

The horizontal lines represent 0.05 and 0.10. The AVR₂ test is applied to the 100 industry portfolio returns sorted on size. The first graph corresponds to 10 smallest; the second 10 middle; and the third 10 largest in size (i.e., N = 10 for each graph).
Figure 7. P-values of the AVR2 statistic for portfolios formed on book-to-market value ratio

The horizontal lines represent 0.05 and 0.10. The AVR2 test is applied to the 10 industry portfolio returns sorted on book-to-market ratio. The first graph corresponds to 3 smallest; and the second 3 largest in book-to-market ratio (i.e., N = 3 for each graph).
Figure 8. P-values of the AVR2 statistic for portfolios formed on momentum

The horizontal lines represent 0.05 and 0.10. The AVR2 test is applied to the 10 industry portfolio returns sorted on momentum. The first graph corresponds to 3 lowest; and the second 3 highest momentum (i.e., N = 3 for each graph).
Figure 9. P-values of the AVR2 statistic for portfolios formed for non-tradable and tradable industries.

The horizontal lines represent 0.05 and 0.10. The AVR2 test is applied to the 10 industry portfolio returns. The first graph corresponds to 5 no-tradable industries; and the second 5 tradable industries. The former include telecommunications, shops (retail and wholesale) and services, healthcare, and utilities; while the latter include consumer goods (durables and non-durables), manufacturing goods, oil products, and business equipments while (i.e., N =5 for each graph).
Figure 10. P-values of the AVR2 statistic for portfolios for hi-tech and non hi-tech industries.

Hi-Tech

Non Hi-Tech

The horizontal lines represent 0.05 and 0.10. The AVR2 test is applied to the 10 industry portfolio returns. The first graph corresponds to 3 hi-tech industries (N=3); and the second 7 non hi-tech industries (N=7). The former includes consumer durables (cars, TV’s, furniture, and household appliances), business equipments (computers, software, electronic equipments), and telecommunication (telephone and TV transmissions), while the latter for all other industries.