ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

ECONOMETRICS & OPERATIONS RESEARCH
BACHELOR THESIS

Testing the low-volatility anomaly using monotonic relation tests

Author:
Daan OPSCHOOR
410188

Supervisor:
Dr. Erik KOLE

Second assessor:
Xuan LENG

Final version: July 2, 2017

Abstract

In this thesis we test the low-volatility anomaly using monotonic relation tests. We use both portfolios based on all-cap and large cap stocks which are sorted on some characteristic, such as the volatility, to test whether there is a significant relation between the expected or risk-adjusted return and the characteristic. We show that top-minus-bottom tests can give the wrong conclusion about the presence of a monotonic relation. Also, we conclude that there is indeed an anomaly present in the risk-return relation, although we generally find that this relation is not monotonic.

Contents

1 Introduction

In this thesis we test the low-volatility anomaly using monotonic relation tests. This anomaly that low volatility stocks earn higher risk-adjusted returns than high volatility stocks contradicts the well-known capital asset pricing model (CAPM), since this model implies that there is a positive linear relation between the expected return and the risk of an asset. Empirical studies done by Jensen et al. (1972), Fama and MacBeth (1973) and Haugen and Heins (1975) already show that the expected returns of low-beta (high-beta) assets are higher (lower) than the CAPM model suggests, which implies that the risk-return relation is flatter than the CAPM predicts. Twenty years later, Fama and French (1992) find similar results even when they control for size effects.

More recently, Ang et al. (2006) show that this phenomenon is not only observed between beta en return, but also between volatility and return. Specifically, they show that stocks with high volatility have abnormally low returns even when they account for size, value, momentum and liquidity effects. In fact, Blitz and van Vliet (2007) find that low volatility stocks are more attractive in terms of return than high volatility stocks, which suggests that there is even a negative relation between risk and return. They observe this relation on global and regional stock markets, whereas Blitz et al. (2013) find that this volatility effect is also present in emerging equity markets and that it cannot be explained by size, value and momentum effects. Blitz et al. (2013) also show that the low-volatility anomaly is robust to using other risk measures than variance for the volatility, such as the CAPM beta and mean absolute deviation. To summarize, Baker and Haugen (2012) show that the low-volatility anomaly is observed in all testable equity markets.

Ang et al. (2006), Blitz and van Vliet (2007) and Blitz et al. (2013) all look at the significance of the top-minus-bottom return differentials between low-volatility stocks and high-volatility stocks to test whether there is a relation between volatility and return. However, Patton and Timmermann (2010) show in their simulation experiment that using this top-minus-bottom differential is not always a good way to test for a relation as it does not take into account intermediate observations. Therefore, Patton and Timmermann (2010) propose a new monotonic relation test , which allows us to test whether there is a significant monotonic relation between volatility and return.

The test of Patton and Timmermann (2010) (henceforth MR test) is nonparametric, easy to implement and has the monotonic relation under the alternative hypothesis such that we only find a significant relation when there is enough evidence in the data to support it. Alongside the MR test, we also implement simple top-minus-bottom tests, the multivariate inequality test of Wolak (1989) and the Bonferonni test of Fama (1984), where the latter two tests have the (weakly) monotonic relation under the null hypothesis such that we only find a significant relation when there is not enough evidence in the data against it. Since the MR test is nonparametric, we do not need to assume a specific relation or distribution. The Wolak test, however, needs to assume that the return differentials are normally distributed.

As a first approach, we use portfolio sorts based on all-cap stocks. First, we show that there is significant evidence for a monotonic increasing relation between book-to-market ratio and return and somewhat weaker evidence for a relation between momentum and return. For the portfolios sorted on variance, market beta and size, however, we can not find such significant evidence. We do show that the relation between risk-adjusted return and risk seems to be more decreasing than for the expected return, where the risk adjusted return measures are based on a constant of the CAPM model (also known as Jensen's alpha developed by Jensen (1967)), three factor model of Fama and French (1993) and four factor model of Carhart (1997). This indicates the presence of the low-volatility anomaly, however, the relation is not significantly monotonically decreasing.

Furthermore, we see that the highest decile portfolio sorted on variance, has a very low return compared to the rest which is consistent with the findings of Ang et al. (2006). As a consequence, the top-minus-bottom test rejects the null hypothesis of equal returns for the bottom and top ranked asset, which implies that there is a negative relation. However, according to the MR test this conclusion is wrong, which shows that the top-minus-bottom test is not always a good way to test for a monotonic relation as Patton and Timmermann (2010) already stated. The MR test of Patton and Timmermann (2010) is easily generalized to deal with two-way sorted portfolios, such that we can test for a monotonic relation across two variables. We show that there seems to be a significant relation between variance and return for small-cap stocks, but that this relation is not present for large-cap stocks.

Next, we use the returns of large-cap stocks which are components of the S&P500 market index to construct stocks and portfolios which we sort on variance and idiosyncratic volatility. Beside using monotonic relation tests, we also use a regression analysis which shows that their is a positive relation between the average return and volatility, but a flat relation between risk-adjusted return and volatility. For regression analysis, however, we need to assume that the relation is indeed linear, which makes it less robust. The monotonic relation tests show that the pattern between variance and (risk-adjusted) return for large cap stocks, in contrary to the portfolios based on all-cap stocks, seems to be increasing, although the MR test does not find significant evidence for this. For idiosyncratic volatility, we find similar results, also when we use the Sharpe ratio proposed by Sharpe (1966, 1994) as risk-adjusted return measure. In addition, we also construct two-way sorted portfolios formed on market beta and (idiosyncratic) volatility with the large cap stock returns. We show that their is a significant relation between return and both volatility as idiosyncratic volatility across the market betas. However, their is no joint relation between these risk measures and the return.

Our contribution to the existing literature, is that we show that top-minus-bottom tests can give wrong conclusions about the presence of a monotonic relation when we compare their conclusions with that of the MR test. Also, we conclude that there is indeed an anomaly present in the risk-return relation, although we generally find that this relation is not monotonic.

2 Methodology

2.1 Notation

Let $\{R_{i,t},\ t = 1, \ldots, T;\ i = 0, \ldots, N\}$ be the returns of $N+1$ assets over T time periods, where the assets are ranked on some characteristic (e.g. market beta) such that asset 0 has the lowest value of the characteristic and asset N the highest. Since we are also interested in risk-adjusted returns, we use the constant of a factor model and the Sharpe ratio developed by Sharpe (1966, 1994) as risk-adjusted return measures. We can denote the general factor model as

$$R_{i,t} - R_{f,t} = \alpha_i + \boldsymbol{\beta}_i' \boldsymbol{F}_t + \varepsilon_{i,t}, \quad t = 0, 1, \ldots T \text{ for each } i \tag{1}$$

where $R_{i,t}$ is the return of asset i at time t, $R_{f,t}$ is the risk free rate of return at time t, the vector $\boldsymbol{\beta}_i$ contains the betas of the corresponding factors in $\boldsymbol{F}_t$ and $\varepsilon_{i,t}$ is the residual for asset i at time t.

We use three different factor models and therefore obtain three different estimates for α_i. First, we use the CAPM model where $\boldsymbol{F}_t = R_{M,t} - R_{f,t}$ is the excess market return. The constant in the CAPM model is known as Jensen's alpha proposed by Jensen (1967), but we refer to it as the one-factor alpha. The second model is the three factor model of Fama and French (1993) which uses $\boldsymbol{F}_t = (R_{M,t} - R_{f,t}, SMB_t, HML_t)'$ where SMB_t and HML_t represent the size and value factor, respectively. We call the constant, α_i, in this model the three-factor alpha. The last model is the four-factor model of Carhart (1997) where $\boldsymbol{F}_t = (R_{M,t} - R_{f,t}, SMB_t, HML_t, MOM_t)'$ and MOM_t represents the momentum factor. The corresponding α_i in this model is called the four-factor alpha. The Sharpe ratio of asset i is defined as

$$SR_i = \frac{\mathrm{E}(R_{i,t} - R_{f,t})}{\sqrt{\mathrm{var}(R_{i,t} - R_{f,t})}}. \tag{2}$$

We can determine $\hat{\alpha}_i$ as the ordinary least square (OLS) estimate of the constant for each model. The Sharpe ratio can be calculated using the sample mean and sample variance of $R_{i,t} - R_{f,t}$ such that we can say that $\hat{\alpha}_i = \hat{SR}_i$. If we are interested in the average return, we use a regression of $R_{i,t}$ on a constant which results in $\hat{\alpha}_i = \frac{1}{T}\sum_{t=1}^{T} R_{i,t}$. We denote the vector with the average or risk-adjusted returns of the assets as $\boldsymbol{\alpha} = (\alpha_0, \alpha_2, \ldots, \alpha_N)'$. The vector with the corresponding return differentials is denoted as $\boldsymbol{\Delta} = (\Delta_1, \ldots, \Delta_N)'$ where $\Delta_i = \alpha_i - \alpha_{i-1}$. To test whether there is an increasing relation in the average or risk-adjusted returns, we need to test whether

$$\Delta_i > 0 \text{ for } i = 1, \ldots, N. \tag{3}$$

holds. If we want to test for a decreasing relation, we simply multiply $\boldsymbol{\Delta}$ with -1 and test for an increasing relation.

2.2 Monotonic relation tests

2.2.1 Top-minus-bottom tests

The first test we discuss is the top-minus-bottom test. The null and alternative hypotheses of this test are denoted as

$$H_0 : \alpha_0 = \alpha_N \text{ versus } H_1 : \alpha_0 < \alpha_N, \tag{4}$$

such that we test whether the expected return of the bottom ranked asset is lower than the expected return of the top ranked asset. We can test this with a regression of $R_{N,t} - R_{0,t}$ on a constant. Naturally, the obtained constant in the regression is equal to the mean of the top-minus-bottom differential, $R_{N,t} - R_{0,t}$. Using this estimate and its standard error we can compute the t-statistic and test whether the differential is significantly different from zero. For the risk-adjusted return measures based on the factor models, we simply regress $R_{N,t} - R_{0,t}$ on a constant and the factors of interest. Again, we use the t-statistic of the constant to test whether the differential is zero or not. We use the heteroskedasticity and autocorrelation consistent (HAC) standard errors of Newey and West (1987) in the construction of the t-statistics to correct for possible heteroskedasticity or autocorrelation.

For the Sharpe ratio, however, we cannot use this approach. Instead we use the Jobson and Korkie (1981) test with the Memmel (2003) correction (henceforth JKM test) to test whether two Sharpe ratios are significantly different. The test statistic of the JKM test is

$$Z = \frac{SR_N - SR_0}{\sqrt{\frac{1}{T}\left(2(1-\rho_{0,N}) + \frac{1}{2}\left(SR_0^2 + SR_N^2 - SR_0 SR_N (1+\rho_{0,N}^2)\right)\right)}} \tag{5}$$

where $\rho_{0,N}$ is the correlation between asset 0 and N. Asymptotically, Z follows a standard normal distribution such that we can easily test the significance of $SR_N - SR_0$.

The top-minus-bottom tests do not take into account intermediate observations, and therefore have limited purpose to test for a monotonic relation as is shown in the simulation experiment in Patton and Timmermann (2010). On the other hand, it is an easy way to find an indication of the sign of the relation.

2.2.2 Bonferroni test

The second test that we discuss summarizes multiple t-tests into one test in terms of a Bonferroni bound as is proposed by Fama (1984). The null and alternative hypotheses of this Bonferroni test are denoted as

$$H_0 : \boldsymbol{\Delta} \geq \mathbf{0} \text{ versus } H_1 : \boldsymbol{\Delta} \text{ unrestricted.} \tag{6}$$

Under the null hypothesis we have a (weakly) monotonically increasing relation and under the alternative hypothesis we have a non-monotonic relation. To test this we regress $R_{i,t} - R_{i-1,t}$ on a constant and eventual factors and obtain the t-statistic of the constant for $i = 1, \ldots, N$. Again, we use the HAC standard errors of Newey and West (1987) to construct the t-statistics. Using the Bonferroni inequality as it is used in Fama (1984) and Patton and Timmermann (2010) , we calculate the p-value as

$$p\text{-value} = \min((N-1) \cdot \Phi(t_{1:N}), 1) \tag{7}$$

where $t_{1:N}$ is the smallest t-statistic on $\hat{\Delta}_i, i = 1, \ldots, N$ and Φ is the cumulative distribution function of a standard normal distribution. We take the minimum of $(N-1) \cdot \Phi(t_{1:N})$ and one to make sure that the p-value is between zero and one as $(N-1) \cdot \Phi(t_{1:N})$ can become bigger than one. This test is simple to implement, but according to the simulation experiment in Patton and Timmermann (2010) it still seems to be a conservative test as it has a low power.

2.2.3 Wolak test

Here we talk in more detail about the multivariate inequality test of Wolak (1989). The test of Wolak (1989) has the same null and alternative hypothesis as in equation (6). The test statistic of this test is based on a comparison between an unrestricted estimate of $\boldsymbol{\Delta}$ and a restricted estimate of $\boldsymbol{\Delta}$, where the restriction imposes a weakly monotonic relation. The unrestricted estimate $\hat{\boldsymbol{\Delta}}$ is just the sample return differential $(\hat{\alpha}_1 - \hat{\alpha}_0, \ldots, \hat{\alpha}_N - \hat{\alpha}_{N-1})'$. The restricted estimate $\tilde{\boldsymbol{\Delta}}$ is obtained from the minimization

$$\begin{aligned} &\min_{\boldsymbol{\Delta}} (\hat{\boldsymbol{\Delta}} - \boldsymbol{\Delta})' \hat{\boldsymbol{\Omega}}^{-1} (\hat{\boldsymbol{\Delta}} - \boldsymbol{\Delta}) \\ &\text{subject to } \boldsymbol{\Delta} \geq \mathbf{0} \end{aligned} \tag{8}$$

where $\hat{\boldsymbol{\Omega}}$ is the HAC estimator of the covariance matrix of $\hat{\boldsymbol{\Delta}}$ which can be obtained using the results in Newey and West (1987). Following the results in Wolak (1989), we calculate the Wald test statistic of this test as

$$W = (\hat{\boldsymbol{\Delta}} - \tilde{\boldsymbol{\Delta}})' \hat{\boldsymbol{\Omega}}^{-1} (\hat{\boldsymbol{\Delta}} - \tilde{\boldsymbol{\Delta}}). \tag{9}$$

Wolak (1987, 1989) show that if we assume that $\boldsymbol{\Delta}$ is normally distributed, that the distribution of W is a weighted sum of chi-squared distributions ranging from one to N degrees of freedom. To obtain the weights, Wolak (1989) proposes to perform a Monte Carlo simulation where we take S draws from a multivariate normal distribution with mean zero and covariance matrix $\hat{\boldsymbol{\Omega}}$. We use draw $\hat{\boldsymbol{\Delta}}^s$ to obtain a restricted estimate $\tilde{\boldsymbol{\Delta}}^s$ for $s = 1, \ldots, S$ and we count how many elements in $\tilde{\boldsymbol{\Delta}}^s$ are bigger than zero. The weights $w(N, k, \hat{\boldsymbol{\Omega}})$ are obtained as the proportion of the S draws that $\tilde{\boldsymbol{\Delta}}^s$ has k elements bigger than zero. The p-value of this test is equal to

$$p\text{-value} = \sum_{k=1}^{N} w(N, k, \hat{\boldsymbol{\Omega}}) \Pr(\chi^2_k \geq W) \tag{10}$$

In a simulation experiment, Patton and Timmermann (2010) find that this test often performs better in detecting a monotonic relation than the top-minus-bottom t-test and Bonferroni test. However, a disadvantage of this test is that the Monte Carlo simulation takes a lot of time, especially for large N, as we need to perform S minimization procedures to find N unknown parameters. Also, Patton and Timmermann (2010) note that the test has limited power as it does not reject the null hypothesis of a (weakly) monotonic relation when there is not enough evidence in the data against it.

2.2.4 One-way MR test

In this subsection we discuss the monotonic relation test of Patton and Timmermann (2010) based on one-way sorts. This test has, in contrary to the Bonferroni and Wolak test, a flat or weakly declining relation under the null and a strictly increasing relation under the alternative. Therefore, this test only rejects the null hypothesis if there is enough evidence in the data in favor of a monotonic relation. Mathematically, the null hypothesis and alternative hypothesis of the MR test are denoted as

$$H_0 : \boldsymbol{\Delta} \leq \mathbf{0} \text{ versus } H_1 : \boldsymbol{\Delta} > \mathbf{0}. \tag{11}$$

which can also be written as

$$H_0 : \boldsymbol{\Delta} \leq \mathbf{0} \text{ versus } H_1 : \min_{i=1,\ldots,N} \Delta_i > 0. \tag{12}$$

These hypotheses motivate the test statistic that is used for this test, namely

$$J_T = \min_{i=1,\ldots,N} \hat{\Delta}_i. \tag{13}$$

Under standard conditions which can be found in Patton and Timmermann (2010), $\hat{\boldsymbol{\Delta}} = (\hat{\Delta}_1, \ldots, \hat{\Delta}_N)'$ is asymptotically normally distributed for $T \to \infty$, that is

$$\sqrt{T}\big((\hat{\Delta}_1, \ldots, \hat{\Delta}_N)' - (\Delta_1, \ldots, \Delta_N)'\big) \overset{a}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}). \tag{14}$$

Since the critical value for the minimum value of $\boldsymbol{\Delta}$ would depend on the entire covariance matrix $\boldsymbol{\Omega}$, which is difficult to estimate for large numbers of N, Patton and Timmermann (2010) propose a bootstrap method to find the critical value of the MR test.

First, they use the stationary bootstrap method of Politis and Romano (1994) to draw (with replacement) a new random sample of returns $\{\tilde{R}^{(b)}_{i,\tau(t)}, \tau(1), \ldots, \tau(T); \quad i = 1, \ldots, N\}$, where $\tau(t)$ is a random draw from the original time indices $\{1, \ldots, T\}$ and $b = 1, \ldots, B$ denotes the bootstrap number. To preserve the time series dependencies in the returns, we draw the data in blocks where both the starting point and the length of the blocks are random. We draw the block length from a geometric distribution with parameter θ which denotes the average length of each block. For the average returns we use the bootstrap mean of $\tilde{R}^{(b)}_{i,\tau(t)}$ for $\alpha^{(b)}_i$ and for the risk-adjusted return measures we use the bootstrap regression

$$\tilde{R}^{(b)}_{i,\tau(t)} - \tilde{R}^{(b)}_{f,\tau(t)} = \alpha^{(b)}_i + \boldsymbol{\beta}^{(b)\prime}_i \boldsymbol{F}^{(b)}_{i,\tau(t)} + e^{(b)}_{i,\tau(t)}, \quad i = 0, 1, \ldots N \tag{15}$$

to find an estimate of $\alpha^{(b)}_i$ where $\tilde{R}^{(b)}_{f,\tau(t)}$ is the bootstrap risk free rate of return, $\alpha^{(b)}_i$ represents the expected risk-adjusted return, $\boldsymbol{\beta}^{(b)}_i$ is a vector of the betas of the specified bootstrap factors in $\boldsymbol{F}^{(b)}_{i,\tau(t)}$, which are constructed with the same bootstrap time indices as the returns, and $e^{(b)}_{i,\tau(t)}$ are the residuals for asset i and bootstrap b. For the Sharpe ratio, we just use the excess bootstrap returns, $\tilde{R}^{(b)}_{i,\tau(t)} - \tilde{R}^{(b)}_{f,\tau(t)}$. Using the estimates of $\alpha^{(b)}_i$, we can calculate the bootstrap return differentials as $\hat{\Delta}^{(b)}_i = \hat{\alpha}^{(b)}_i - \hat{\alpha}^{(b)}_{i-1}$. The bootstrap p-value of this test is equal to

$$p\text{-value} = \frac{1}{B} \sum_{b=1}^{B} \mathbb{1}(J^{(b)}_T > J_T) \tag{16}$$

where

$$J^{(b)}_T = \min_{i=1,\ldots,N} (\hat{\Delta}^{(b)}_i - \hat{\Delta}_i), \quad b = 1, 2, \ldots, B. \tag{17}$$

Furthermore, Patton and Timmermann (2010) implement a studentized version of this test to eliminate possible cross-sectional heteroskedasticity effects. In this version they divide $\hat{\Delta}_i$ in (13) and $\hat{\Delta}^{(b)}_i - \hat{\Delta}_i$ in (17) with the stationary bootstrap estimate of the long-run standard deviation of differential i which is defined in Politis and Romano (1994). The estimate of the long-run N-by-N covariance matrix is calculated in a similar fasion as the HAC estimator of the covariance matrix of Newey and West (1987) and is equal to

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{d}_t' \boldsymbol{d}_t + \frac{1}{T} \sum_{m=1}^{\min(T,N)-1} \left(1 - \frac{m}{N}\right)\left(1 - \frac{1}{\theta}\right)^m \left(\sum_{t=1}^{T} (\boldsymbol{d}_{t+m}' \boldsymbol{d}_t + \boldsymbol{d}_t' \boldsymbol{d}_{t+m}) \right) \tag{18}$$

where $\boldsymbol{d}_t = (d_{1,t}, \ldots, d_{N,t})$, $d_{i,t} = R_{i,t} - R_{i-1,t} - \hat{\Delta}_i$ and θ is the average block length in the stationary bootstrap. The estimate of the long-run standard deviation of differential i is equal to the square root of the i-th diagonal element of $\hat{\boldsymbol{\Sigma}}$. Since the stationary bootstrap makes use of circularity in the time-series, we can say that $d_{i,T+m} = d_{i,m}$ as this assumption is needed in the calculation of the last sum in (18). The upper bound of the second sum in (18), $\min(T, N) - 1$, comes from the fact that we want to deal with cases where $N > T$ as we need this for stock sorts in subsection 2.3.2

Patton and Timmermann (2010) also provide two other measures, namely the Up and Down statistic, which are denoted as

$$J^+_T = \sum_{i=1}^{N} |\hat{\Delta}_i| \mathbb{1}(\hat{\Delta}_i > 0), \tag{19}$$

$$J^-_T = \sum_{i=1}^{N} |\hat{\Delta}_i| \mathbb{1}(\hat{\Delta}_i < 0), \tag{20}$$

respectively. These measures account for the direction, frequency and the magnitude of deviations from a flat pattern and therefore provide information whether there is at least a part of the pattern monotonically increasing or decreasing. Again, the critical values of these tests can be obtained using bootstrap techniques and it is also possible to perform the studentized version by dividing $\hat{\Delta}_i$ in (19) and (20) with the long-run standard deviation of differential i.

An advantage of the MR test compared to the Wolak test is that we do not need to assume a distribution for Δ_i. Also, the MR test is easily applicable for large N where the Wolak test is not. Since the alternative hypothesis of the MR test is the one that contains the monotonic relation, we only reject the null hypothesis in favour of a monotonic relation when there is enough evidence in the data to support it. This in contrary to the Wolak and Bonferroni test, which do not reject the null hypothesis of a (weakly) monotonic relation when there is not enough evidence in the data against it.

Patton and Timmermann (2010) find that the Wolak test and Bonferroni test do not always reject the null hypothesis when there is in fact no monotonic relation. Conversely, the MR test sometimes does not reject the null hypothesis when there is in fact a monotonic relation. Since we cannot directly compare the size and power of the Wolak and Bonferroni test with the MR test as their hypothesis are different, Patton and Timmermann (2010) claim that it depends on the research question or economic framework which test we should use. In our research, though, we use all aforementioned tests where we rely the most on the MR test.

2.2.5 Two-way MR test

Patton and Timmermann (2010) show that their one-way monotonic relation test is easily generalized to deal with two-way sorted assets. We refer to this generalization as the two-way MR test. If we report the results of the two-way sorts in a $(N+1) \times (N+1)$ table, where we sort one variable across rows and the other variable across columns, we can denote the row and column return differentials as $\Delta^r_{i,j} = \alpha_{i,j} - \alpha_{i-1,j}$ and $\Delta^c_{i,j} = \alpha_{i,j} - \alpha_{i,j-1}$ respectively. The null and alternative hypothesis are

$$H_0 : \Delta^r_{i,j} \leq 0,\ \Delta^c_{i,j} \leq 0 \text{ for all } i,j \quad \text{versus} \quad H_1 : \Delta^r_{i,j} > 0 \text{ and } \Delta^c_{i,j} > 0 \text{ for all } i,j. \tag{21}$$

We can rewrite (21) as

$$H_0 : \Delta^r_{i,j} \leq 0,\ \Delta^c_{i,j} \leq 0 \text{ for all } i,j \quad \text{versus} \quad H_1 : \min_{i,j=1,\ldots,N} (\Delta^r_{i,j}, \Delta^c_{ij}). \tag{22}$$

In a similar fashion as the one-way MR test, the test statistic is denoted as

$$J_T = \min_{i,j=1,\ldots,N} (\hat{\Delta}^r_{i,j}, \hat{\Delta}^c_{i,j}). \tag{23}$$

We can use the same stationary bootstrap method and studentized version as for the one-way MR test, but now with $(\hat{\Delta}^r_{i,j}, \hat{\Delta}^c_{i,j})$ instead of $\hat{\Delta}_i$.

2.3 Application on the low-volatility anomaly

2.3.1 Portfolio sorts based on all-cap stocks

For this first approach we use decile portfolios sorted on variance, market beta, market equity, book-to-market ratio and momentum. The monthly returns of these portfolios can be found on the website of Kenneth R. French[1] and are based on all the NYSE, AMEX and NASDAQ stocks which have available data on the required sorting characteristic. We use the monthly returns from July 1963 to March 2017, which gives us 645 observations.

The portfolios formed on variance are based on the variance of daily returns using 60 days (about 3 months) of data. The portfolios formed on market beta are based on an estimate of beta using 60 monthly returns, where beta is estimated with the method of Scholes and Williams (1977) to account for nonsynchronous trading. Using the monotonic relation tests, we are able to test whether there is a significant negative monotonic relation between (risk-adjusted) return and volatility. The monthly factors needed for the factor models to determine the risk-adjusted returns can also be found on the website of Kenneth R. French.

Beside the variance and market beta, we are also interested whether there is a relation in the expected returns of decile portfolios that are sorted on market equity, book-to-market ratio and momentum, since these factors are used in the three-factor and four-factor models. For both portfolios sorted on book-to-market ratio and momentum, we test for an increasing relation in the expected returns, whereas for market equity we test for a negative relation.

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Next, we look at two-way sorted portfolios formed on market equity and market beta and portfolios sorted on market equity and variance. Fama and French (1992) already showed that the low-beta anomaly is unrelated to size. However, by using two-way sorted portfolios we are able to get more insight in the overall relation between size, risk and return. We use portfolios which are the intersections of five portfolios formed on market equity and five portfolios formed on variance or market beta. Again, monthly returns of these portfolios can be found on the website of Kenneth R. French. Using the one-way MR test and two-way MR test, we can test for conditional monotonicity (by keeping one variable fixed) and joint monotonicity in the returns. Both across rows and columns, we test for a decreasing relation.

Lastly, we use $S = 1,000$ simulation runs for the Wolak test to determine the weights. For the MR test, we use $B = 1,000$ bootstrap replications with an average block length of $\theta = 10$. Furthermore, we use the studentized version for the one-way MR test, two-way MR test, Up test and Down test.

2.3.2 Stock and portfolio sorts based on large cap stocks

In the second approach we look at individual stock returns of large-cap companies. We use the adjusted closing stock prices of S&P500 companies (stated on 27 May 2017) which can be found on Yahoo Finance and are downloaded using the Multiple Stock Quote Downloader of Samir Khan[2]. The selection of these stocks that are in the S&P500 market index entails a bias as we know that these stocks are performing well over the past, but we would not know this in the beginning of the sample period. However, in the context of our research this is not a problem as we are only interested in their mean, standard deviation and the stock and portfolio sorts that we make with them. We use the stocks which have stock prices available from 10 July 2001 till 28 April 2017 (3976 observations) and do not have more than one consecutive missing values. This selection procedure leaves us with 377 S&P500 companies. For the stocks which have a missing value, we use linear interpolation between the prior and subsequent stock price to find an estimate of this missing value. The return of stock i at day t can be calculated as

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \tag{24}$$

where $P_{i,t}$ is the stock price of company i at day t. Instead of only 10 portfolios in the previous approach, we now have return data for 377 stocks which gives us the possibility to do a regression between the mean and standard deviation of the returns, this is we do the regression

$$\alpha_i = \phi_i + \gamma_i s_i + \varepsilon_i, \quad i = 1, 2, \ldots, 377 \tag{25}$$

where α_i is the mean of the returns of stock i, ϕ_i is the constant, γ_i is the slope parameter, s_i is the standard deviation of the returns of stock i and ε_i is the error term of stock i. We denote the cross-sectional OLS estimates as $\hat{\phi}_{i,\text{OLS}}$ and $\hat{\gamma}_{i,\text{OLS}}$. To correct for possible heteroskedasticity and/or autocorrelation when the error terms are not white noise, we use the HAC standard errors of Newey and West (1987).

Beside the HAC standard errors, we also use the Fama-MacBeth (FM) procedure suggested by Fama and MacBeth (1973) to estimate standard errors which are corrected for cross-sectional correlation between the different stocks. In this procedure we first run a regression of the returns at each time t on a constant and the standard deviation of stock i , this is do the regression

$$R_{i,t} = \phi_t + \gamma_t s_i + \varepsilon_{i,t}, \quad i = 1, 2, \ldots, 377 \text{ for each } t. \tag{26}$$

Next, Fama and MacBeth (1973) propose to estimate the parameters and residuals as

$$\hat{\phi}_{i,\text{FM}} = \frac{1}{T}\sum_{t=1}^{T}\hat{\phi}_t, \hat{\gamma}_{i,\text{FM}} = \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_t \text{ and } \hat{\varepsilon}_{i,\text{FM}} = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{i,t} \tag{27}$$

Cochrane (2005) shows that the cross-sectional OLS estimates are equivalent to the FM estimates, this is $\hat{\phi}_i = \hat{\phi}_{i,\text{FM}} = \hat{\phi}_{i,\text{OLS}}$ and $\hat{\gamma}_i = \hat{\gamma}_{i,\text{FM}} = \hat{\gamma}_{i,\text{OLS}}$. We can now estimate the standard errors corrected for cross-sectional correlation as

$$\sigma^2(\hat{\phi}_{i,\text{FM}}) = \frac{1}{T^2}\sum_{t=1}^{T}(\hat{\phi}_t - \hat{\phi}_{i,\text{FM}})^2 \text{ and } \sigma^2(\hat{\gamma}_{i,\text{FM}}) = \frac{1}{T^2}\sum_{t=1}^{T}(\hat{\gamma}_t - \hat{\gamma}_{i,\text{FM}})^2. \tag{28}$$

[2] http://investexcel.net/multiple-stock-quote-downloader-for-excel/

By using the estimated slope parameter and both the HAC and FM standard errors, we can test whether the slope parameter, γ_i, is significantly different from zero. We also look at the risk-adjusted returns where we set α_i as the constant of one of the used factor models. For the FM procedure we first calculate $R_{i,t} - R_{f,t} - \hat{\boldsymbol{\beta}}_i{}' \boldsymbol{F}_{i,t}$, where $\boldsymbol{F}_{i,t}$ contains the eventual daily factors and $\hat{\boldsymbol{\beta}}_i$ contains the corresponding OLS estimates, and then use this term instead of $R_{i,t}$ in (26).

Since we need to assume that there is a linear relation between the return and volatility with the regression methods, we also want to look at the situation where we do not need to assume a specific relation, namely by using the monotonic relation test. To use these tests, we need to construct new return series which are sorted on variance. First, we take a month (20 trading days) of data for each stock and use this to calculate the sample standard deviation. Next, we sort all the stocks from smallest standard deviation to highest. Then we calculate the monthly return for each stock and save this in the sorted return series. After that, we continue to the next month and do exactly the same such that we get the sorted monthly returns of the month after the first. Finally, we continue this process till we have our new 377 return series which are sorted on variance and based on the returns of 377 stocks. In the end, the sorted return series consist of $T = \lfloor \frac{3975-20}{20} \rfloor = 197$ monthly returns. We use the floor function to remove the remaining days which are than 20 days and therefore do not have enough days to determine another monthly return.

Beside sorting on volatility, we also sort on idiosyncratic volatility which means that we sort on unsystematic risk that cannot be captured by certain market factors. To find this idiosyncratic volatility we use one of the three factor models discussed in 2.1 and use the standard deviation of the residuals as the idiosyncratic volatility. The daily factors needed in the estimation of the factor models can be found on the website of Kenneth French. The sorting procedure is the same as with the regular volatility, though, we now use 3 months (60 trading days) to determine the idiosyncratic volatility since we need enough observations to do the regression. This procedure gives us sorted return series of $\lfloor \frac{3975-60}{20} \rfloor = 195$ monthly returns.

Alongside the stock sorts, we also construct equally-weighted decile portfolios sorted on variance or idiosyncratic volatility based on the 377 stocks. The construction of these decile portfolios is the same as for the stock sorts, except that we now take the average of the monthly returns of each block of $\lceil \frac{377}{10} \rceil = 38$ sorted stocks except for the last three decile portfolios where we use only 37 stocks as we do have not enough stocks for an equal division. Using the stock and portfolio sorts, we now have the possibility to test whether there is a significant relation between the (idiosyncratic) volatility and expected returns. However, since N is 377 for the stock sorts it is not possible to use the Wolak test, since this test is computationally to intensive for such a large N. Therefore, we only focus on the top-minus-bottom tests and the MR test since these test are still usable for large N in this case. For the portfolio sorts, we are still able to use all monotonic relation tests.

For the portfolios sorted on variance, we use the constant of one of the three factor models as risk-adjusted return measures. However, to find these we need to estimate the factors which correspond with the constructed monthly returns. Therefore, we take the sum of the daily factor returns of the 20 trading days and use this as the corresponding factor. For the idiosyncratic volatility, however, we use the Sharpe ratio as risk adjusted return measure where the monthly risk free rate of return is created the same way as the factors but now by taking the sum of the daily risk free rate of returns for 60 trading days. Furthermore, we construct two-way sorted portfolios where we sort on (idiosyncratic) volatility and market beta. We first sort on variance or idiosyncratic volatility and then we sort the stocks on the market beta, where we use the CAPM beta as an estimate of the market beta. We make intersections of quantile portfolios for both characteristics, such that we get 25 portfolios. Every monthly portfolio return is based on $\lfloor \frac{377}{25} \rfloor = 15$ stocks, except for the lowest and highest (idiosyncratic) volatility portfolios with the lowest market beta which are based on 16 stocks. We use 60 trading days to determine the standard deviaton and market beta. By using the two-way MR test, we are able to jointly test for a monotonic relation across market beta and (idiosyncratic) volatility.

We take $B = 10,000$ to make sure we have enough bootstrap replications as we have a large N for the stock sorts and we set $\theta = 10$ as the average block length. For the portfolio sorts, we use $S = 1,000$ simulation runs for the Wolak test. Again, for all tests we use the standardized version of the MR test where we just take the same estimate of the long-run bootstrap variance for the Sharpe ratio as for the expected returns. For all tests, we test for an increasing relation.

3 Results

3.1 Preliminary results

Before we discuss the results of the monotonic relation tests, we first discuss some summary statistics of the portfolio and stock data. Table 1 shows that the average returns of portfolios sorted on variance and momentum have a lower minimum average return than the other portfolios. Therefore, the standard deviation of their average returns is higher. The standard deviation of the standard deviations of the portfolio returns are, however, lower for size and value sorts than for the variance, market beta and momentum sorts. This also becomes clear from the minimums and maximums of standard deviation of the the size and value , which are indeed further away from each other than for the other characteristics.

For the stock data, in the last five columns of table 1, we naturally see that the daily stock returns have a much lower standard deviation of average returns than the monthly returns of the sorted stocks. Since the statistics are based on a large number of individual stocks, their is always a return series which has a negative average return. This explains the negative minimums of the average stock returns. Lastly, we see that the statistics for the stocks sorted on variance and idiosyncratic volatility are all very close to each other which could indicate that they have quite similar return series.

Table 1: Summary statistics of portfolio and stock data

		Portfolios					Stocks				
		VA	BE	ME	BE-ME	MOM	All	VA	IV1	IV3	IV4
Mean	Std.	0.22	0.06	0.10	0.15	0.32	0.03	0.69	0.74	0.73	0.74
	Min.	0.37	0.90	0.85	0.82	0.21	-0.02	-0.60	-0.59	-0.56	-0.43
	Max.	1.15	1.04	1.19	1.30	1.49	0.15	5.33	7.12	6.66	7.21
Std.	Std.	1.73	1.37	0.71	0.50	1.19	0.64	3.05	3.13	3.11	3.10
	Min.	3.28	3.45	4.19	4.33	4.31	1.11	4.24	4.37	4.40	4.33
	Max.	8.81	7.87	6.31	6.08	8.03	4.65	30.22	28.23	27.93	31.27

Note: This table shows the standard deviation (Std.), minimum (Min.) and maximum (Max.) of the average returns (Mean) and standard deviations of the portfolio ans stock returns. The sorting variables of the portfolios are variance (VA), market beta (BE), market-equity (ME), book-to-market value (BE-ME) and momentum (MOM). For the stocks, we have a column that contains information on the daily returns of the large-cap stocks which are not sorted. Also, we show the statistics of monthly returns of stocks sorted on variance and idiosyncratic volatility based on the CAPM model (IV1), three factor model (IV3) and four factor model (IV4). The average returns and standard deviations are in percentages. The sample period for the portfolio data is July 1963 - March 2017 and for the stock data 11 July 2001 till 28 April 2017f for the daily returns, August 2001 till April 2017 for the stocks sorted on variance and October 2001 till April 2017 for the stocks sorted on idiosyncatic volatility.

3.2 Portfolio sorts based on all-cap stocks

3.2.1 Average returns

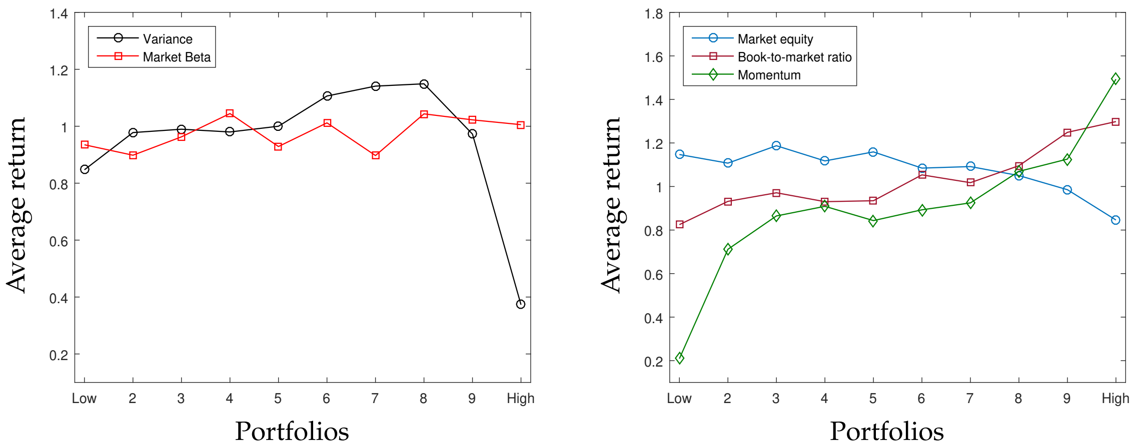
We now discuss the results of the tests for monotonicity in the average returns. The first five columns of table 2 show that the top-minus-bottom t-test only finds a significant top-minus-bottom spread for book-to-market ratio (p-value of 0.013) and momentum (p-value of 0.000). Furthermore, we see that the Wolak and Bonferroni test fail to reject the null of a (weakly) monotonic relation for all sorting variables, where the p-values are the highest for book-to-market ratio and momentum. The MR test only rejects the null hypothesis in favour of a monotonic relation for book-to-market ratio with a p-value of 0.039, whereas the Up test also finds that at least a part of the pattern for momentum is monotonically increasing with a p-value of 0.032. The last three columns of table 2 correspond to results of Patton and Timmermann (2010) based on the same portfolios, but with a different sample. We see that the p-values are quite similar despite the different sample such that we can draw the same conclusions as Patton and Timmermann (2010).

From figure 1b we indeed observe that the relation between book-to-market ratio and return seems to be monotonically increasing. The same holds for momentum, although this pattern has a more flat part in the middle which could explain why the MR test does not reject the null hypothesis in favour of a monotonic relation here. The pattern of market equity looks slowly decreasing or even flat. The relation between market beta and average return in figure 1a also looks flat. In contrast, the pattern between variance and returns seems to be slightly increasing, except for the highest two decile portfolios which have very low returns compared to the rest. This observation is consistent with the finding of Ang et al. (2006) that high volatility stocks have abnormally low returns.

Table 2: p-values of monotonic relation tests in average returns

	07/1963 - 03/2017					07/1963 - 12/2006		
	VA	BE	ME	BE-ME*	MOM*	ME	BE-ME*	MOM*
t-test	0.080	0.599	0.084	0.013	0.000	0.062	0.005	0.000
Bonferroni test	0.481	0.606	0.661	1.000	1.000	1.000	1.000	1.000
Wolak test	0.410	0.309	0.552	0.962	0.897	0.736	1.000	0.873
MR test	0.638	0.564	0.552	0.039	0.166	0.274	0.000	0.291
Up test	0.380	0.270	0.621	0.032	0.000	0.737	0.045	0.000
Down test	0.115	0.442	0.068	0.968	0.964	0.051	1.000	0.954

Note: This table contains the p-values of monotonic relation tests of portfolios sorted on variance (VA), market beta (BE), market equity (ME), book-to-market ratio (BE-ME) and momentum (MOM). For the variables with a star, we test for an increasing relation and for all other variables for a decreasing relation. We use $S = 1,000$ for the Wolak test and $B = 1,000$ and $\theta = 10$ for the MR, Up and Down test. Also, we use the studentized version of these latter three tests.



(a) Sorted on variance and market beta

(b) Sorted on ME, BE-ME and MOM

Figure 1: Average monthly returns on decile portfolios sorted on variance, market beta, market equity (ME), book-to-market ratio (BE-ME) and momentum (MOM) from July 1963 to March 2017

3.2.2 Risk-adjusted returns

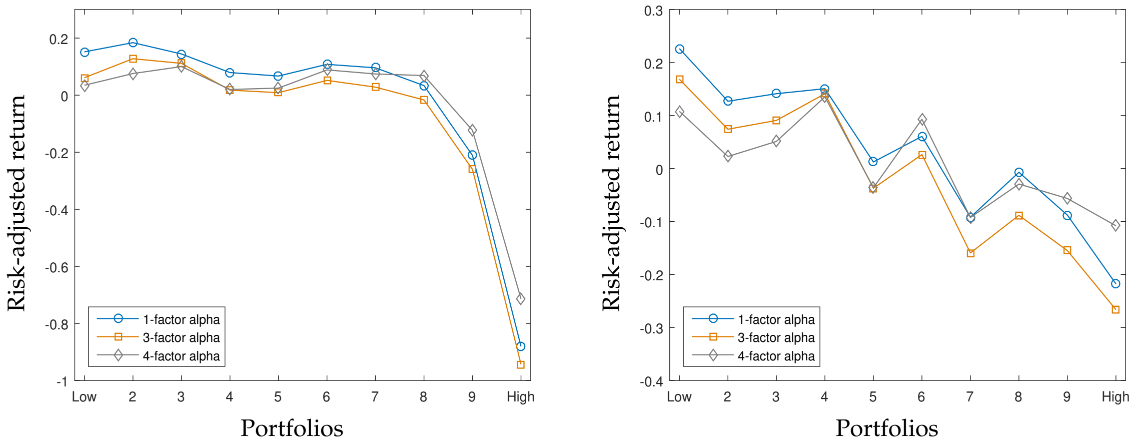
Instead of average returns, we now look for monotonicity in risk-adjusted returns. Table 3 shows that the top-minus-bottom differentials are significantly smaller than zero for all risk-adjusted return measures, except for the 4-factor α of the portfolios sorted on market beta (p-value of 0.083). The Wolak and Bonferroni tests are again unable to reject the null hypothesis of a (weakly) decreasing monotonic relation for both variance and market beta. The MR test, however, does not reject the null in favour of a decreasing relation. The conclusions in terms of rejections are mostly the same as with the average returns, but the p-values of the MR test (Wolak and Bonferroni tests) for risk-adjusted returns are way smaller (higher) than for the average returns. This suggests that there is more evidence for a decreasing relation in the risk-adjusted returns than for the average returns.

This also becomes clear when we look at the Down test. We find that for all risk-adjusted return measures and sorting variables at least a part of the pattern is monotonically decreasing with p-values ranging from 0.001 to 0.045. Figure 2 also supports this, especially for the three highest decile portfolios. However, the patterns between variance and risk-adjusted returns in figure 2a seems to be more flat for the lower decile portfolios, which could explain that the MR test does not reject its null hypothesis. We also find that the highest portfolio has abnormally low risk-adjusted returns compared to the rest, which causes the t-tests to reject an equal top-minus-bottom return differential. The overall patterns between market beta and risk-adjusted return in figure 2b seems to be decreasing, although we cannot say that this relation is monotonic as there are lots of reversals of the relation between the lowest and highest decile portfolio.

Table 3: p-values of monotonic relation tests in risk-adjusted returns, July 1963 - March 2017

	Variance			Market beta		
	1-factor α	3-factor α	4-factor α	1-factor α	3-factor α	4-factor α
t-test	0.000	0.000	0.000	0.018	0.003	0.083
Bonferroni test	1.000	1.000	1.000	1.000	1.000	0.209
Wolak test	0.943	0.891	0.878	0.898	0.865	0.428
MR test	0.050	0.190	0.117	0.143	0.148	0.726
Up test	0.932	0.914	0.849	0.833	0.777	0.357
Down test	0.004	0.001	0.012	0.039	0.009	0.045

Note: This table contains the p-values of monotonic relation tests of portfolios sorted on variance and market beta. We test for a decreasing relation of variance and market beta with the risk-adjusted return measures 1-factor alpha, 3-factor alpha and 4-factor alpha. We use $S = 1,000$ for the Wolak test and $B = 1,000$ and $\theta = 10$ for the MR, Up and Down test. Also, we use the studentized version of these latter three tests.



(a) Sorted on variance

(b) Sorted on market beta

Figure 2: Average monthly risk-adjusted returns (based on the CAPM model, three-factor model and four-factor model) of decile portfolios sorted on variance and market beta from July 1963 to March 2017

3.2.3 Two-way sorts

We now discuss the results when we test for monotonicitiy in two-way sorted portfolios. Table 4 indicates that we cannot reject the null hypothesis in favour of a joint monotonic relation (p-value of 0.963). However, we see that there is a significant relation between size and return for the smallest three variance portfolios with p-values of 0.001, 0.003 and 0.043. The finding that there is no overall significant relation between size and return (p-value of 0.988) across all variance portfolios, is due to the fact that the high variance and small market equity portfolio has a very low return of 0.200 %. Figure 3a shows this more clearly. We see that for the low variance portfolios there is a clear relation between the size and return, but for the high variance portfolio this relation disappears. Furthermore, the conditional MR tests do not reject the null hypothesis in favor of a monotonic relation between variance and return, since there is only a decreasing relation for the small market equity and high variance portfolios. This suggests that volatile small-cap stocks often perform worse in terms of their return than less volatile small-cap stocks.

Table 5 shows that across all market beta portfolios, we reject the null hypothesis in favour of a decreasing relation between size and return with a p-value of 0.030. Figure 3a also shows that this relation is more present than in the case of the variance in figure 3b. However, we cannot reject the null hypothesis of a (weakly) increasing relation between market beta and return, since the patterns seems to be parabolic for the small market equity stocks and flat for the big market equity stocks. As a consequence, we cannot find evidence for a joint monotonic relation across both market beta and market equity (p-value of 0.750). Since the size effects are unable to explain the strange relation between market beta and return, we can say that the size is indeed unrelated to the low-beta anomaly which is consistent with the findings of Fama and French (1992).

Table 4: Conditional and joint monotonicity tests for market equity and variance, July 1963 to March 2017

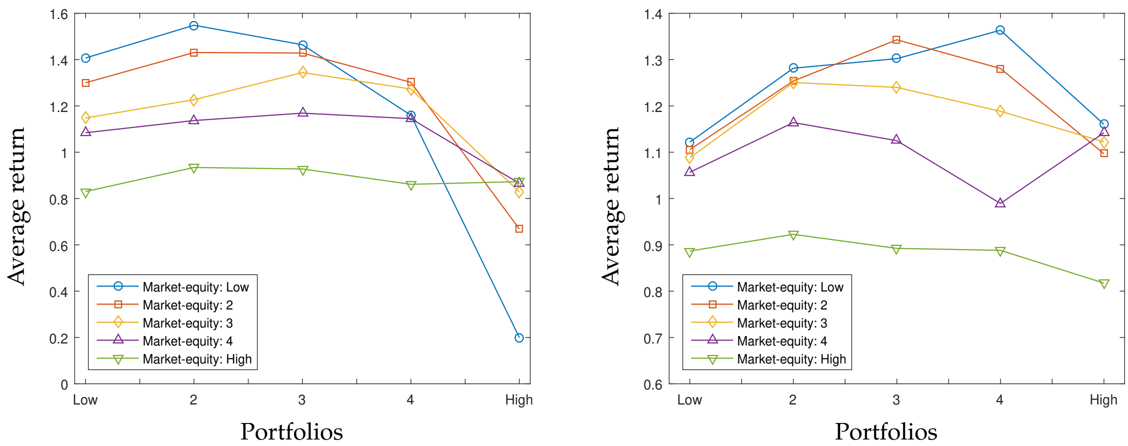
	Variance					MR p-value	Joint MR p-value
	Low	2	3	4	High		
Market equity							
Small	1.407	1.548	1.464	1.160	0.200	0.855	
2	1.300	1.431	1.428	1.302	0.668	0.812	
3	1.148	1.227	1.344	1.273	0.829	0.880	0.723
4	1.084	1.137	1.169	1.145	0.866	0.401	
Big	0.830	0.934	0.928	0.861	0.873	0.645	
MR p-value	0.001	0.003	0.043	0.644	0.995		
Joint MR p-value			0.988				0.963

Note: This table contains the average returns (in %) of the two-way sorted portfolios on market equity and variance and the p-values of the corresponding joint and conditional monotonicity tests. Both across rows and columns, we test for a decreasing relation. We use $B = 1,000$ and $\theta = 10$ and the studentized version of the MR tests.

Table 5: Conditional and joint monotonicity tests for market equity and market beta, July 1963 to March 2017

	Market beta					MR p-value	Joint MR p-value
	Low	2	3	4	High		
Market equity							
Low	1.122	1.281	1.302	1.363	1.160	0.958	
2	1.106	1.254	1.343	1.281	1.097	0.921	
3	1.088	1.250	1.240	1.188	1.122	0.899	0.852
4	1.057	1.164	1.125	0.989	1.142	0.643	
High	0.887	0.923	0.892	0.888	0.817	0.236	
MR p-value	0.054	0.109	0.194	0.001	0.201		
Joint MR p-value			0.030				0.750

Note: This table contains the average returns (in %) of the two-way sorted portfolios on market equity and market beta and the p-values of the corresponding joint and conditional monotonicity tests. Both across rows and columns, we test for a decreasing relation. We use $B = 1,000$ and $\theta = 10$ and the studentized version of the MR tests.



(a) Sorted on variance

(b) Sorted on market beta

Figure 3: Average monthly returns of two-way sorted portfolios for the combinations market equity with variance and market beta from July 1963 to March2017

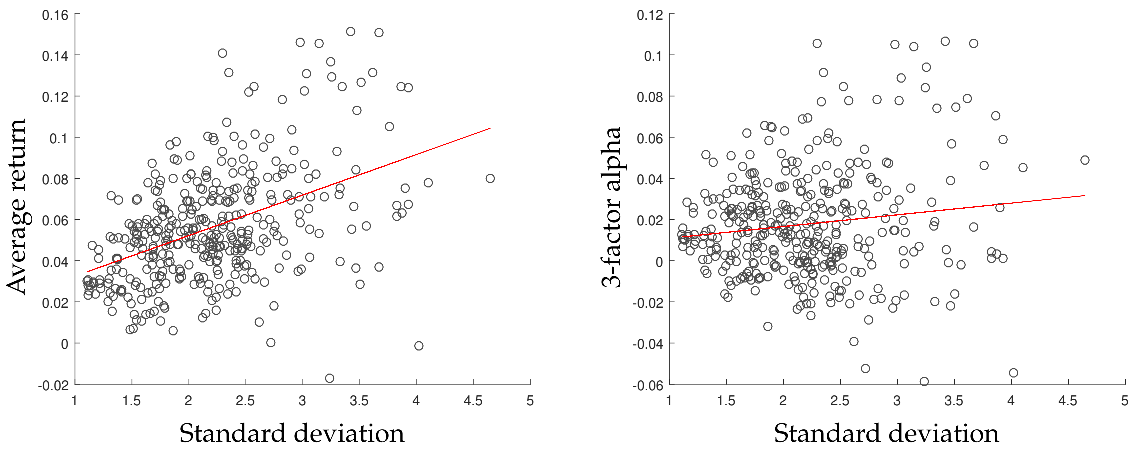
3.3 Stock and portfolio sorts based on large cap stocks

3.3.1 Regression analysis

Here we discuss the results of regressing the stock returns on a constant and the standard deviation of the stock returns. Figure 4 shows that the regression line of the average returns seems to be more positive than the line of the risk-adjusted return measure, 3-factor alpha. This is conform the expectation, since the relation between risk-adjusted return and risk is expected to be flat. We also observe that the returns are more spread out for a higher standard deviation which suggests the presence of heteroskedasticity, so it is good thing that we use the HAC standard errors instead of the normal OLS standard errors since these are biased in case of heteroskedasticity.

In table 6 we see indeed that the the estimate of the slope parameter is bigger than the estimates for the risk-adjusted returns. By using the HAC standard errors for the t-test, we find that the slope parameters of the fitted line for both the average return as the alphas are significantly different from zero. However, when we use the FM standard errors we find that the slope parameters are not significantly different from zero for the risk-adjusted return measures. This is due to the fact that the FM standard errors are higher than the HAC standard errors, which suggests that there is a strong correlation between the different stock returns, since the FM standard errors correct for this correlation.

In the last column of 6 we find the R-squared of the regressions. We see that they are very low, especially for the risk-adjusted returns, which means that the model does only explain a small part of the variation in the returns. Also, when we reject a significant relation with this model it could by due to the fact that we made a wrong assumption, namely that the relation is linear. Therefore, we need a more robust way to test for a monotonic relation, namely with the monotonic relation tests, as they do not need to make an assumption about a specific relation.



(a) Average returns (b) Risk-adjusted returns

Figure 4: Scatter plots of average returns and 3-factor alphas versus standard deviation of 377 S&P500 stocks based on daily returns from 11 July 2001 till 28 April 2017

Table 6: Estimation output of regressions based on daily returns of 377 stocks for the period 11 July 2001 till 28 April 2017

	$\hat{\gamma}_i$	HAC		FM		R^2
		S.E.	p-value	S.E.	p-value	
Average return	0.020	0.002	0.000	0.010	0.023	0.204
1-factor alpha	0.008	0.003	0.001	0.007	0.106	0.039
3-factor alpha	0.006	0.002	0.010	0.006	0.177	0.019
4-factor alpha	0.008	0.002	0.000	0.006	0.075	0.041

Note: This table contains the estimates of γ_i and the HAC and FM standard errors (S.E.) with the corresponding p-values of the t-tests. The last column contains the R-squareds of the regressions.

3.3.2 One-way sorts

In this subsection we first discuss the results of the monotonic relation tests when we apply them on the stock sorts formed on variance. Table 7 shows that the top-minus-bottom t-test finds a significant increasing relation for all return measures based on stock sorts. Figure 5a shows that this is due to the the very high return for the highest variance stock. The MR test, however, does not even once reject the null hypothesis in favour of a monotonic relation with p-values ranging from 0.836 to 0.921. The overall pattern, namely, seems to be more flat when we look at figure 5a. This is a clear example that the top-minus-bottom can draw controversial conclusions, since it only looks only at the lowest and highest ranked assets. The Up and Down test are both unable to find support for only a part in the pattern which is monotonically increasing or decreasing.

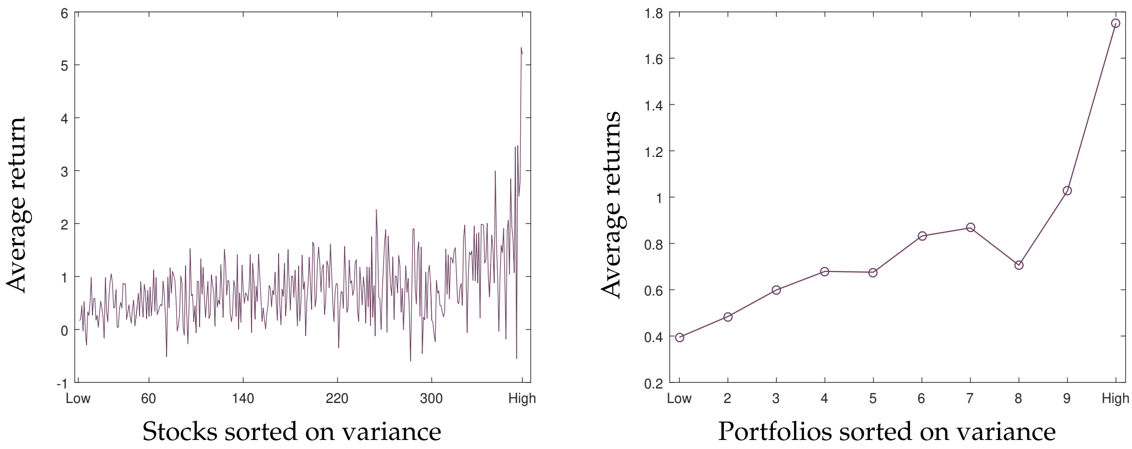
In the last 4 columns of table 7, we find the p-values of the monotonic relation tests applied on portfolios sorted on variance. Again, the top-minus-bottom t-tests find evidence for an increasing relation, except for the 3-factor alpha (p-value of 0.070). The Wolak and Bonferroni test do not reject the null hypothesis of a (weakly) increasing relation. The MR test, though, does not find a significant relation between variance and return. On the other hand, the Up test shows that there is at least a part of the pattern which is monotonically increasing for the average return and 4-factor alpha with p-values of 0.021 and 0.023 respectively. We can also see this in figure 5b, where the pattern, especially for the lower variance portfolios, seems to be monotonically increasing.

The finding that there is some evidence for an increasing relation between variance and return, whereas we found evidence for a decreasing relation in table3, is due to the fact that we now use large cap stocks instead of all-cap stocks. In table 4 we already showed that the relation between variance and returns seems to be more decreasing for small cap stocks and more flat or increasing for large cap stocks. Instead of observing it in the two-way sorted portfolios, we now observe it more generally by looking at the differences between large cap and all-cap stock returns.

Table 7: p-values of monotonic relation tests based on stocks and portfolios sorted on variance, August 2001 to April 2017

	Stocks				Portfolios			
	AVR	Risk-adjusted returns			AVR	Risk-adjusted returns		
		1F α	3F α	4F α		1F α	3F α	4F α
t-test	0.000	0.002	0.002	0.002	0.008	0.044	0.070	0.003
MR test	0.908	0.836	0.882	0.921	0.342	0.605	0.706	0.546
Up Test	0.164	0.201	0.206	0.212	0.021	0.190	0.219	0.023
Down test	0.217	0.242	0.242	0.257	0.917	0.786	0.665	0.776
Bonferroni test	-	-	-	-	1.000	0.413	0.300	0.417
Wolak test	-	-	-	-	0.752	0.550	0.412	0.501

Note: This table contains the p-values of monotonic relation tests of stocks and portfolios sorted on variance. We look at both the average returns (AVR) and the risk-adjusted return measures 1-factor alpha (1F α), 3-factor alpha (3F α) and 4-factor alpha (4F α). We use $S = 1,000$ for the Wolak test and $B = 10,000$ for the MR, Up and Down test. We use the studentized version of the MR, Up and Down test. We test for an increasing relation.



(a) Stocks

(b) Portfolios

Figure 5: Average monthly returns of stocks and portfolios sorted on variance for the period August 2001 to April 2017

We now discuss the monotonic relation tests on the portfolios and stocks sorted on idiosyncratic volatility. Panel A in table 8 shows that the p-values for stocks sorted on idiosyncratic volatility are similar to the p-values in table 7 based on stocks sorted on variance. However, we see that the JKM test does not reject equal top-minus-bottom Sharpe ratios for idiosyncratic volatility based on the CAPM model. Also, the p-values of the Up and Down test for the stock sorts are higher in table 8 than they are in table 7, which suggests that there is less support for a part of the pattern that is monotonically decreasing for stocks sorted on idiosyncratic volatility.

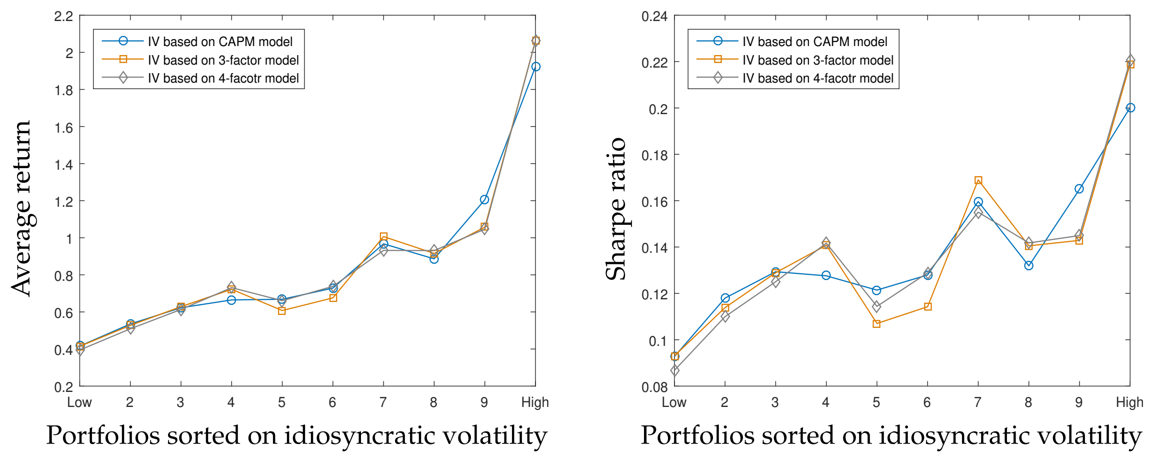
Panel B in table 8 shows that the top-minus-bottom tests reject equal return differentials. The Wolak and Bonferroni tests also find significant evidence for an increasing relation by not rejecting the null hypothesis. The MR test still does not reject the null hypothesis in favour of an increasing, although the p-values are somewhat smaller than they are for the portfolios sorted on variance, especially for the average returns of the portfolios sorted on the variance of the residuals of the CAPM and four facor model with p-values of 0.062 and 0.135. We also observe this in figure 6a where the relation between idiosyncratic volatility and return seems to be increasing. When we compare figure 6a with figure 5a, we see that the patterns are pretty similar.

Moreover, the Up test rejects the null hypothesis in favour of a partly monotonic relation for almost all return measures (p-values ranging from 0.002 to 0.039), except for the Sharpe ratios of the portfolios based on idiosyncratic volatility of the CAPM and four factor model. Figure 6b shows that this is the case for the lower four idiosyncratic volatility portfolios. We see that the Sharpe ratios of the portfolios based on the three factor model has a slightly stronger increasing relation than for the CAPM and four factor model, which explains the p-value of 0.039 for the Up test instead of the p-values of 0.116 and 0.056. The overall pattern between Sharpe ratio and idiosyncratic volatility does not seems to be monotonically increasing which explains the higher p-values of the MR test than those of the MR test based on average returns.

Table 8: p-values of monotonic relation tests based on stocks and portfolios sorted on idiosyncratic volatiliy, October 2001 to April 2017

Panel A:	Stocks					
	IV1		IV3		IV4	
	AVR	SR	AVR	SR	AVR	SR
t-test	0.000	0.065	0.000	0.029	0.000	0.027
MR test	0.770	0.396	0.701	0.934	0.995	0.998
Up Test	0.198	0.233	0.255	0.275	0.277	0.280
Down test	0.332	0.331	0.401	0.349	0.300	0.280
Panel B:	**Portfolios**					
	IV1		IV3		IV4	
	AVR	SR	AVR	SR	AVR	SR
t-test	0.003	0.022	0.002	0.012	0.001	0.008
MR test	0.062	0.153	0.574	0.775	0.135	0.549
Up Test	0.004	0.116	0.002	0.039	0.003	0.056
Down test	0.974	0.933	0.711	0.533	0.956	0.811
Bonferroni test	1.000	-	0.649	-	1.000	-
Wolak test	0.948	-	0.597	-	0.904	-

Note: This table contains the p-values of monotonic relation tests of stocks (Panel A) and portfolios (Panel B) sorted on idiosyncratic volatility based on the CAPM model (IV1), three factor model (IV3) and four factor model (IV4). We look at both the average return (AVR) and the Sharpe ratio (SR). The p-values of the top-minus-bottom tests in the colums of the Sharpe ratios (SR) are based on the JKM test and of the average returns on the t-tests. We use $S = 1,000$ for the Wolak test and $B = 10,000$ for the MR, Up and Down test. We use the studentized version of the MR, Up and Down test.



(a) Average returns

(b) Risk-adjusted returns

Figure 6: Average monthly returns and Sharpe ratios of stocks and portfolios sorted on idiosyncratic volatility (IV) based on the CAPM model, three factor model and four factor model for the period October 2001 to April 2017

3.3.3 Two-way sorts

Lastly, we discuss the results of two-way sorted portfolios on (idiosyncratic) volatility and market beta. Table 9 shows that there is an significant increasing relation between variance and return across all market betas (p-value of 0.010) and that this relation is also present for the conditional MR tests, except the high market beta portfolio, with p-values ranging from 0.008 to 0.02. Figure 7a also shows this clear monotonic pattern for all market betas between variance and return. The relation between market beta and return, however, is not significant across all variance portfolios where the joint p-value is 0.448. Figure 7a shows, namely, that the pattern seems to reverse. This is, for the low variance portfolio does the highest market beta portfolio have the lowest average return, whereas for the high variance portfolio the highest market beta portfolio has the highest average return. This phenomenon also happen for the fourth quantile portfolio of the market beta, except than reversed.

For the portfolios sorted on idiosyncratic volatility and market beta, we find similar results as we can see in table 10. We find a significant increasing relation between idiosyncratic volatility and return with a p-value is 0.004. This relation seems to be the strongest for the highest two market beta portfolios with p-values of 0.000 and 0.006. The relation between market beta and return is still insignificant though. We can also observe this in figure 7b. Only the low market beta portfolio has significantly lower average returns than the rest, but all the other portfolios are mixed and therefore do not show a clear pattern.

Table 9: Conditional and joint monotonicity tests for variance and market beta, October 2001 till April 2017

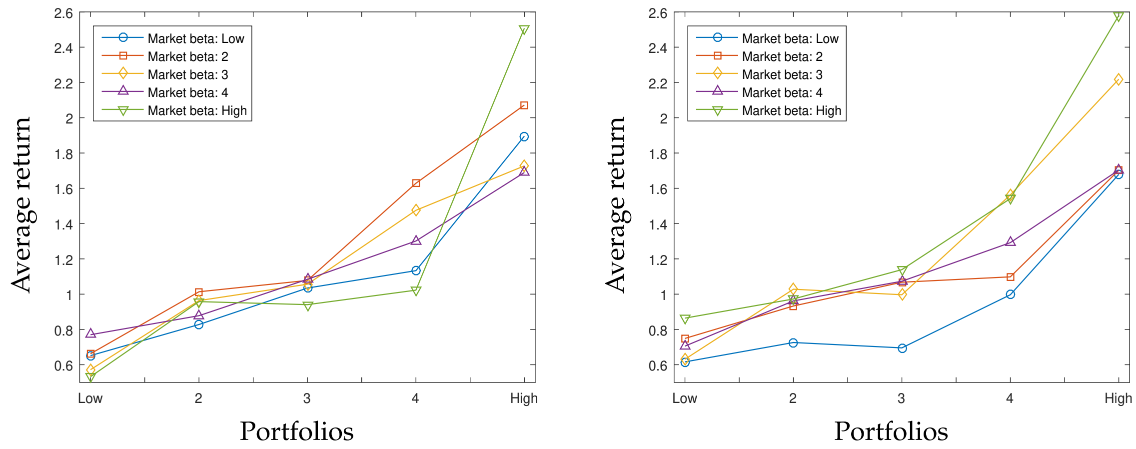
	Market beta						
	Low	2	3	4	High	MR p-value	Joint MR p-value
Variance							
Low	0.652	0.662	0.572	0.770	0.532	0.797	
2	0.828	1.014	0.964	0.878	0.958	0.210	
3	1.035	1.078	1.056	1.084	0.941	0.349	0.448
4	1.133	1.628	1.475	1.301	1.023	0.827	
High	1.896	2.072	1.727	1.690	2.507	0.741	
MR p-value	0.020	0.011	0.008	0.009	0.093		
Joint MR p-value			0.010				0.360

Note: This table contains the average returns (in %) of the two-way sorted portfolios on variance and market beta and the p-values of the corresponding joint and conditional monotonicity tests. We use $B = 1,000$ and the studentized version of the MR tests.

Table 10: Conditional and joint monotonicity tests for idiosyncratic volatility and market beta, October 2001 till April 2017

	Market beta						
	Low	2	3	4	High	MR p-value	Joint MR p-value
Idiosyncratic volatility							
Low	0.616	0.749	0.631	0.705	0.864	0.582	
2	0.726	0.933	1.028	0.962	0.973	0.187	
3	0.695	1.068	0.997	1.073	1.139	0.114	0.609
4	0.997	1.098	1.562	1.292	1.542	0.613	
High	1.679	1.702	2.219	1.704	2.580	0.862	
MR p-value	0.134	0.027	0.105	0.000	0.006		
Joint MR p-value			0.004				0.501

Note: This table contains the average returns (in %) of the two-way sorted portfolios on idiosyncratic volatility and market beta and the p-values of the corresponding joint and conditional monotonicity tests. We use $B = 1,000$ and the studentized version of the MR tests.



(a) Sorted on variance

(b) Sorted on idiosyncratic volatility

Figure 7: Average monthly returns of two-way sorted portfolios formed on market beta and (idiosyncratic) volatility from October 2001 to April 2017

4 Conclusion

We found significant evidence for a monotonic relation between book-to-market ratio and return and somewhat weaker evidence for a relation between momentum and return. For the portfolios sorted on variance, market beta and size, however, we can not find such significant evidence. We do show that for the portfolios based on all-cap stocks, the relation between risk-adjusted return and risk seems to be more decreasing than for the expected return, which indicates the presence of the low-volatility anomaly, although, the relation is not significantly monotonically decreasing.

Furthermore, we saw that the highest decile portfolio sorted on variance, has a very low return compared to the rest which is consistent with the findings of Ang et al. (2006). The consequence of this is that the top-minus-bottom test finds a significant relation, whereas the MR test shows that this conclusion is wrong, which shows that the top-minus-bottom test is not always a good way to test for a monotonic relation as Patton and Timmermann (2010) already stated. Also, we showed that there seems to be a significant relation between variance and return for small-cap stocks, but that this relation is not present for large-cap stocks.

We saw with regression analysis that their seems to be a positive relation between the average return and volatility, but a flat relation between risk-adjusted return and volatility. However, this conclusion relies on the assumption that the relation between return and volatility is linear, which suggest that this method is not a very robust way of testing for a significant relation. Also, we found that the relation between (idiosyncratic) volatility and return seems to be increasing for large cap stocks, whereas it was decreasing for all-cap stocks, although both relations are not significant. Lastly, we showed that their is a significant relation between return and both volatility as idiosyncratic volatility when we control for different market betas.

To summarize, we showed that the top-minus-bottom tests can give wrong conclusions about the presence of a monotonic relation. Also, we conclude that there is indeed an anomaly present in the risk-return relation, although we generally find that this relation is not monotonic.

References

Ang, A., Hodrick, R., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299.

Baker, N. and Haugen, R. (2012). Low risk stocks outperform within all observable markets of the world. *SSRN Working paper*, page no. 2055431.

Blitz, D., Pang, J., and van Vliet, P. (2013). The volatility effect in emerging markets. *Emerging Markets Review*, 16:31–45.

Blitz, D. and van Vliet, P. (2007). The volatility effect: Lower risk without lower return. *Journal of Portfolio Management*, 34:102–113.

Carhart, M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Cochrane, J. (2005). Asset pricing. chapter 12, pages 245–251. Princeton University Press, revised edition.

Fama, E. (1984). Term premiums in bond returns. *Journal of Financial Economics*, 13(4):529–546.

Fama, E. and French, K. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

Fama, E. and French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. and MacBeth, J. (1973). Risk, return, and equilibrium: empirical tests. *The Journal of Political Economy*, 81(3):607–636.

Haugen, R. and Heins, A. (1975). Risk and the rate of return on financial assets: Some old wine in new bottles. *The Journal of Financial and Quantitative Analysis,*, 10(5):775–784.

Jensen, M. (1967). The performance of mutual funds in the period 1945-1964. *Journal of Finance*, 23(2):389–416.

Jensen, M., Black, F., and Scholes, M. (1972). The capital asset pricing model: some empirical tests. *Studies in the theory of capital markets*. Praeger Publisher Inc.

Jobson, J. and Korkie, B. (1981). Hypothesis testing with the sharpe and treynor measures. *The Journal of Finance*, 36(4):889–908.

Memmel, C. (2003). Performance hypothesis testing with the sharpe ratio. *Finance Letters*, 1(1):21–23.

Newey, W. and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Patton, A. and Timmermann, A. (2010). Monotonicity in asset returns: New tests with applications to the term structure, the capm, and portfolio sorts. *Journal of Financial Economics*, 98(3):602–625.

Politis, D. and Romano, J. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313.

Scholes, M. and Williams, J. (1977). Estimating betas from nonsynchronous data. *Journal of Financial Economics*, 5(3):309–327.

Sharpe, W. (1966). Mutual fund performance. *The Journal of Business*, 39(1):119–138.

Sharpe, W. (1994). The sharpe ratio. *The Journal of Portfolio Management*, 23(1):49–58.

Wolak, F. (1987). An exact test for multiple inequality and equality constraints in the linear regression model. *Journal of the American Statistical Association*, 82(399):782–793.

Wolak, F. (1989). Testing inequality constraints in linear econometric models. *Journal of Econometrics*, 41(2):205–235.