# Don't stand so close to Sharpe

Angel León[a]

Lluís Navarro[b]

Belén Nieto[c]

[a] aleon@ua.es, Universidad de Alicante, Address: Ctra. San Vicente SN, E-03080 Alicante, Spain.

[b] lluis.navarro@crfsa.es, Consultora de Riesgos Financieros, S.A., Address: Castellana 93, E-28046 Madrid, Spain.

[c] **Corresponding autor**. belen.nieto@ua.es, Universidad de Alicante, Address: Ctra. San Vicente SN, E-03080 Alicante, Spain. Phone and fax: +34965903621.

**Abstract**

We analyze the use of alternative performance measures for ranking assets. Previous literature on performance evaluation is basically centered on studying the effects of non-normality on rank correlations between measures. We introduce a new approach to compare the effective role of different measures for ranking. We analyze the portfolio composition and the posterior out-of-sample portfolio returns induced by the selection of assets according to each performance measure. The comparison among portfolio returns is performed by applying stochastic dominance tests. The overall empirical findings show that the particular performance measure may influence on both the portfolio composition and its return distribution.

**Highlights**

We analyze the role of a wide range of performance measures as screening rules.
The asset ranking can be very different depending on the measure employed.
The out-of-sample return on the best assets can also behave very differently.
Return differences are verified by using stochastic dominance.

**Keywords:** Investment analysis, performance measures, ranking, returns distribution, stochastic dominance
**JEL:** C10, C40, G11

# Don't stand so close to Sharpe

## 1 Introduction

The securities selection is one important part of the investment process for which the so-called screening rules are useful. These rules aim to restrict the investment universe to a reasonably limited set of assets so as to be analyzed in greater detail by analysts but without specifying assets allocations. Performance measures (PMs, hereafter) are examples of equity screening rules.

Our study tries to analyze the behavior of portfolio returns based on, mainly, daily rebalancing portfolios by using different PMs each period as screening rules. Hence, this method may lead to a different portfolio composition each period depending on the PM we choose. It must be note that we do not implement any portfolio optimization with the selected stocks but we combine them by an equal or a value weighted strategy. This work also tries to summarize the information content of several PMs that belong to the same family by using the principal component analysis (PCA) method. In the same spirit, we can find several works that aim to get together several PMs such as Billio et al. (2012) who construct a performance index, or Hwang and Salmon (2003) who propose a PM combination by using copula functions. Using any of the above techniques as screening rules has the drawback they do not control for the multivariate dependence across asset returns. This problem is left to a future research in which the portfolio optimization framework will be incorporated.[1]

Most PMs are ratios that inform about the risk-reward of the investment. This is a very important concern for risk managers and, thus, these ratios are used for ranking assets. The well-known Sharpe (1966) ratio, that relates the mean return with the standard deviation, has been used as a standard for this aim. It is based on the mean-variance paradigm which requires either elliptical[2] (e.g. Gaussian distribution) returns or quadratic preferences. However, it is well documented that deviations of some financial asset return distributions from normality are statistically significant and then, in such cases, the standard deviation underestimates the total risk and generates biased investment rankings. Therefore, ratios that consider a more general framework such as, among others, one-sided reward and risk

[1] See Bodnar et al. (2015), and the references therein, for the case of portfolio optimization under return predictability.

[2] See Owen et al. (1983).

measures have been proposed.[3] Simultaneously, the debate about the significance in investment applications of these new PMs regarding the Sharpe ratio (SR) is still open.

In particular, the usual way to compare alternative screening rules is based on the Spearman correlations between PM rankings. Although the information provided by each PM could be different, the correlation between two rankings might be large. Therefore, one of the two measures might be redundant as a screening rule. Results from papers that compare PM rank correlations induce controversial conclusions. On the one hand, Eling and Schuhmacher (2007) or Eling (2008), among others, conclude that the PM choice becomes irrelevant since they all produce very similar rankings. However, these papers relay on a small subset of all current available PMs for a sample of hedge funds such that the normality assumption is generally rejected. In the same way, Guo and Xiao (2015).reinforce the previous results. They show that if return distributions belong to the location-scale (LS) family, then the PMs generate identical rank ordering. On the other hand, more recent studies show that rankings can be very different depending on the selected measure. Farinelli et al. (2009) show that measures based on partial moments do a better fit by accommodating to the different investors' risk profile than the Sharpe ratio. Zakamouline (2011) find that severe deviations of normality lead to significant shifts in the rankings for hedge funds. León and Moreno (2015), by assuming the Gram-Charlier distribution (which is not LS) for the stock returns, also agree that the selection of PMs becomes relevant. Caporin and Lisi (2011) find evidence of low rank correlations by using a huge set of different PMs. These authors argue that the results depend on both the sort of assets and the sample period. Additionally, they also show that the rank correlations are time varying and influenced by the sample size. Finally, Magron (2014) shows empirical evidence, with a sample of 24,766 individual investors from a French brokerage, that alternative PMs to the SR, and specifically, the Farinelli-Tibiletti family result in different rankings of investors.

Our paper aims to provide complementary results about the role of different PMs in selecting stocks. Similarly to Caporin and Lisi (2011), we compare a large set of different measures and especially the ones that produce different rankings in previous papers. Our main contribution is just the way of comparison. Instead of analyzing the rank correlations, as in previous papers, we compare the portfolios containing the assets selected by each measure in terms of both the portfolio composition and the out-of-sample (OOS) portfolio returns.

[3] See, for instance, Bacon (2011) and Caporin et al. (2014).

Specifically, we consider 32 PMs computed daily for all stocks in the Standard and Poor (S&P) 500 index by using a 264-days rolling window of past returns. The individual stocks are daily ranked on the basis of each PM and the 20 best stocks are selected to be included in each portfolio. Caporin and Lisi (2011) point out the need of working in a dynamic framework because of the instability of PMs over time. Therefore, the rolling window approach is an additional goal of our paper. Finally, we employ a sufficiently large window to avoid inconsistencies due to the sample size.

Our first analysis consists on comparing the percentage of individual stocks that are simultaneously in two portfolios. It is shown that the portfolio composition is rather similar across portfolios based on measures from the same family. This fact allows for reducing the portfolio number by applying the PCA technique to all portfolios in each family and projecting the OOS returns. In contrast, we find a high distance between the portfolios compositions obtained through PMs from different families.

Our second analysis compares the distribution characteristics of the OOS returns between different portfolios in several ways. First, we analyze the descriptive statistics for the whole OOS period. Second, we implement the dynamic conditional correlation (DCC) model of Engle (2002) between pairs of OOS portfolio returns. Third, we analyze the spread in cumulative returns between each portfolio (and, therefore, PM) and the benchmark (the portfolio based on the Sharpe ratio). Fourth, the empirical distributions of returns are compared by stochastic dominance. All the results indicate that the screening rule influences the subsequent portfolio returns that are, in some portfolios, quite different from the benchmark.

The outline of the work is as follows. The PMs used in our analysis are briefly presented in Section 2. Section 3 contains both the data of individual stocks and the description of the portfolio creation. In Section 4, the empirical results from the comparison of portfolios composition and returns are provided. Section 5 contains two robustness checks. In particular, the effects of either the stock weights on portfolios or the period length for the portfolio rebalancing. Finally, Section 6 summarizes the main conclusions.

## 2 Performance Measures

In this study we consider different PMs into four groups. The first group contains the Sharpe ratio and its extensions. The second one refers to PMs based on partial moments containing both the Kappa and Farinelli-Tibiletti ratios. The third group includes PMs based on quantiles such as the Value-at-Risk (VaR) ratio and the generalized Rachev ratio. Finally, some PMs that do not belong to any other family are considered in the fourth group. Next, we define all the PMs in each group. More details about the specific investment characteristics that these (and other) measures account for can be found in the survey by Caporin et al (2014).

**2.1 The Sharpe ratio and its extensions**

A more generalized version for the original Sharpe ratio (Sharpe, 1966 and 1994) is defined as

$$SR(\theta) = \frac{\mu - \theta}{\sigma},$$

where $\mu$ and $\sigma$ denote the expected return and volatility for the returns distribution, respectively. The parameter $\theta$ is the mean return threshold. In the standard case, $\theta$ is the risk-free rate. An extension of the original $SR$ is the adjusted Sharpe ratio ($ASR$) suggested by Pézier and White (2008), which explicitly adjusts for the skewness, $sk$, and kurtosis, $ku$, of the returns distribution. Hence,

$$ASR(\theta) = SR(\theta)\left[1 + \frac{sk}{6}SR(\theta) - \frac{ku - 3}{24}SR^2(\theta)\right].$$

**2.2 Performance measures based on partial moments**

Lower partial moments (LPM, hereafter) based measures define risk as the negative deviations of the stock returns, $r$, in relation to the mean return threshold, $\theta$, or the minimal acceptable return. Fishburn (1977), among others, defines the LPM of order $m$ as

$$LPM(\theta, m) = \mathbb{E}[(\theta - r)_+^m] = \int_{-\infty}^{\theta} (\theta - r)^m f(r)dr,$$

where $f(\cdot)$ denotes the probability density function (pdf) and $(y)_+ = max(y, 0)$. In contrast to the standard deviation, LPM considers only the negative deviations of returns assuming that investors are especially worried about the losses. The order of the LPM can be interpreted as the investors' risk attitude. Investors with *risk seeking* can be expressed as $0 < m < 1$, $m = 1$ indicates *risk neutrality*, and $m > 1$ indicates *risk aversion*. Thus, for $m > 1$, the higher $m$ the higher the emphasis on the extreme deviations from the threshold $\theta$, whereas the lower the relevance of small deviations from $\theta$. Opposite effects come up with $m < 1$. The LPM of order 0 is the shortfall probability. The LPM of order 1 is related with the expected shortfall.

Equivalently, the upper partial moment (UPM) of order $q$ is defined as

$$UPM(\theta, q) = \mathbb{E}\left[(r-\theta)_+^q\right] = \int_\theta^\infty (r-\theta)^q f(r)dr.$$

So, for a given threshold $\theta$ an analogous reasoning applies to selecting the proper order $q$. We consider two families of PMs within this class: the Kappa and the Farinelli-Tibiletti ratios.

### 2.2.1 Kappa or Sortino-Satchell ratios

Sortino and Satchell (2001) measure the mean excess return per unit of risk by using LPM. Henceforth, we denote this measure as SS. Then,

$$SS(\theta, m) = \frac{\mu-\theta}{\sqrt[m]{LPM(\theta,m)}}. \tag{1}$$

Some popular measures, which are nested in equation (1), are the Sharpe-Omega ratio (see Kaplan and Knowles, 2004) for $m = 1$, the Sortino ratio (see Sortino and Van der Meer, 1991) for $m = 2$, and Kappa 3 (see Kaplan and Knowles, 2004) for $m = 3$. It is verified that $SS(\theta, 1) = \Omega(\theta) - 1$, where $\Omega(\theta)$ denotes the Omega ratio (see Keating and Shadwick, 2002). Note that $\Omega(\theta) = 1$ for $\theta = \mu$. Finally, the Bernardo and Ledoit (2000) ratio is the Omega ratio for $\theta = 0$ and, hence, it represents the gain-to-loss ratio.

We set the following values for the order of the LPM to consider different investors risk attitudes: $m = 10$ (defensive investor), $m = 3$ and 2 (conservative investors), $m = 1.5$ (moderate investor), $m = 0.8$ and 0.5 (aggressive investors).

### 2.2.2 Farinelli-Tibiletti ratios

Farinelli and Tibiletti (2008) propose a ratio (henceforth, FT ratio) that exclusively looks at the upper and lower partial moments by comparing the favorable events and the unfavorable ones:

$$FT(\theta, q, m) = \frac{\sqrt[q]{UPM(\theta,q)}}{\sqrt[m]{LPM(\theta,m)}}, \tag{2}$$

with $q$,$m > 0$ which are called the right and left orders, respectively. The higher the value for $q$, the higher the agent's preference for expected gains. The higher the value for $m$, the higher the investor's dislike for expected losses. Note that equation (2) nests two popular measures: the Omega ratio $\Omega(\theta)$ when $q = m = 1$, and the Upside Potential ratio (see Sortino et al., 1999) when $q = 1$ and $m = 2$.

So, how to choose the proper orders? If $q > m$ the investment strategy aims at no worrying about losses, whereas having huge losses is not desirable when $q < m$. Caporin and Lisi (2011) calibrate the parameters $q$ and $m$ to match them with investors' styles. We follow them and keep the same values for $(q, m)$. Thus, (0.5, 2) for defensive investors; (1.5, 2) for conservative investors; the Omega ratio (1, 1) for moderate investors; (2, 1.5) for investor that are seeking a potential growth in the final wealth; (3, 0.5) for aggressive investors; and the Upside Potential ratio (1, 2).

### 2.3 Performance measures based on quantiles

In this group of measures we present some ratios based on quantiles. The PMs are both the VaR ratios (VaRR) and Generalized Rachev (GR) ratios. To understand better these PMs, we need previously to introduce two downside risk-measures for the returns distribution.

First, the VaR at the $\alpha$ confidence level is the quantity such that the probability that the return will be lower or equal to this quantity is $\alpha$:

$$VaR(\alpha) = -inf\{r|F(r) \geq \alpha\},$$

where $F(r)$ represents the cumulative distribution of returns. Second, the expected shortfall (conditional VaR), which measures the expected value of all returns that are lower or equal to the VaR:

$$CVaR(\alpha) = -\mathbb{E}[r|r \leq -VaR(\alpha)].$$

Note that in this work we take the more common approach of referring VaR (CVaR) as positive numbers.

#### 2.3.1 Generalized Rachev ratio

The Generalized Rachev ratio (Biglova et al., 2004) relates the CVaR to the power $\delta$ applied to the returns lower than the threshold, and the CVaR to the power $\gamma$ applied to the returns higher than the threshold in the symmetric negative VaR. That is,

$$GR(\alpha,\theta,\gamma,\delta) = \frac{\mathbb{E}\left[(r-\theta)_+^{\gamma} \middle| \, r \geq VaR(-r;\alpha)\right]^{\frac{1}{\gamma}}}{\mathbb{E}\left[(\theta-r)_+^{\delta} \middle| \, r \leq -VaR(r;\alpha)\right]^{\frac{1}{\delta}}} \,,$$

where $\gamma, \delta > 0$. As in the VaRR case, we consider 1 and 5 percent confidence levels and denote GR1 and GR5 the corresponding GR ratios. The values for the power parameters $(\gamma, \delta)$ are: (0.8, 0.001), (0.5, 0.8), (0.8, 0.8), (0.5, 1), (0.01, 0.8), and (1, 1). The last case corresponds with the simple Rachev ratio (Biglova et al., 2004).

### 2.3.2 VaR ratio

The VaR ratio is introduced in Caporin and Lisi (2011). With similar foundations than the FT ratio, it relates the positive quantile and the symmetric negative quantile of the distribution:

$$VaRR(\alpha) = \frac{|VaR(-r;\alpha)|}{|VaR(r;\alpha)|} \,,$$

where $|\cdot|$ denotes the absolute value function. We consider two confidence levels: 1 and 5 percent. Hence, we denote VaRR1 and VaRR5 for the 1 and 5 percent confidence levels, respectively.

## 2.4 Other performance measures

The last group of measures includes four ratios representing mean excess return $(\mu - \theta)$ per unit of risk in which the risk in the denominator is approximated by different dispersion measures. The mean absolute deviation (henceforth MAD) ratio proposed by Konno and Yamazaki (1991) uses the mean of $|r-\mu|$ as risk measure. The Minimax (hereafter MM) ratio uses the risk measure proposed by Young (1998): $max(r_{(n)}, -r_{(1)})$, where $r_{(n)}$ and $r_{(1)}$ denote the ordered statistics corresponding to the maximum and the minimum of the return sample with size n, respectively. Last but not least, the Range ratio (hereafter Range) uses $|r_{(n)} - r_{(1)}|$. Finally, we also consider as the risk measure the Gini's mean difference (Shalit and Yitzhaki, 1984) that is defined as

$$\Gamma = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|x - y| f(x) f(y) dx dy,$$

where $x$ and y are a pair of values in realized returns.

Last of all, in our study we implement the corresponding sample estimations for all the PMs exhibited above.

**3 Data and Portfolios Construction**

We get from Bloomberg the end-of-day quotes of the S&P 500 components as well as their market capitalization for the period January 2005-September 2014. For each stock, we compute daily logarithmic returns and set, for instance, a zero mean return as threshold ($\theta = 0$). As expected, these individual returns series are characterized by large deviations from normality. It holds that the Jarque-Bera test rejects the assumption of normally distributed stock returns for all stocks. Therefore, it seems to be an appropriate dataset to compare measures accounting for higher order moments. It is well known that stock indexes composition changes over time. These changes can produce effects on the quantity or quality of the information, or on the demand of an asset because of either its inclusion or exclusion from the index (Pruitt and Wei, 1989; Jain, 1987). In order to avoid potential consequences of delisting on individual stock returns, we finally restrict our sample to the 424 assets that are continuously belonging to the S&P along our sample period.

Our sample period contains 2,453 working days. We employ a window of 264 days (approximately one year) to obtain the estimations of the PMs in Section 2 to the series of individual returns. Specifically, we compute 32 PMs to each individual stock: SR, ASR, six measures in the SS family for different values of order $m$, the FT measures for six alternative combinations of orders $m$ and $q$, the VaRR for both 1 and 5 percent levels, the GR1 and GR5 (also for 1 and 5 percent levels) for six combinations of the power parameters $\gamma$ and $\delta$, respectively, and the MAD, MM, Range and Gini measures. Then, we sort the 424 stocks by each PM and select the best 20 ones to be included in the corresponding PM portfolio. This amount means selecting approximately the best 5% of the whole set of individual assets.

Lastly, we compute the first out-of-sample (OOS) return for each portfolio just the day after the window as the equally-weighted return of the individual stocks in the portfolio. We set this procedure in order to avoid any arbitrary decision about the optimal allocation. Any other asset allocation criterion may be controversial with the PM in which the asset selection is based on. Additionally, De Miguel et al. (2009) find statistical equivalence between the performance of an equally-weighted portfolio and a Markowitz-optimized portfolio. In any case, Section 5 analyzes the robustness of the results for the value-weighted case. By rolling the window each day, we generate series of 2,189 OOS returns.

## 4 Results

We study whether the use of different selecting criteria have consequences in the performance of the resulting portfolios. Therefore, now we compare the 32 portfolios in several ways. First, we analyze the portfolio compositions. Second, we look at the OOS portfolio returns in both static and dynamic frameworks. Third, we also compare the different portfolio returns by testing both first and second stochastic dominance.

### 4.1 Portfolio Composition

For each of the 2,189 days in which 32 portfolios are constructed with the best 20 individual assets selected, we compare the number of assets that are selected in the same day by using two different measures. Table 1 displays the median value of the percentage of the coincident assets when comparing two alternative measures along the 2,189 days.

Comparing the Sharpe ratio with each one of the other measures, we can observe that the median percentage of coincidences is 100 percent for the adjusted Sharpe ratio, it is relatively high (more than 70 percent) for the SS ratios, MAD, Range and Gini, whereas it is high for the MM (around 60 percent). In contrast, the selected assets are in general different by using other measures. That is, the percentage of coincidences goes from 15 and 25 percent when comparing the Sharpe ratio with all the generalized Rachev ratios and the two VaR ratios. The extreme case is the family of FT measures for which the coincident assets are zero on median, indicating a clear different selection criterion from Sharpe ratio. Moreover, the assets selected by FT measures are also different than the assets selected by any other PM, especially the SS ratios for which the number of coincident assets is again zero in all days. Comparing pairs of measures within the same family, as expected, the number of coincident assets tends to be high. However, this is not the case for the FT family. The percentage of coincidences can be as small as 15 percent for different parameter values, suggesting the importance on the investment style. Therefore, we conclude that assets selected by alternative measures can be quite different.

### 4.2 Portfolio Returns: Static Analysis

According to Table 1 the PM selection can produce different portfolio compositions. In this section we aim to study harder the above result by looking at the OOS returns each portfolio generates. In order to reduce the quantity of information, we apply the PCA technique to summarize the information content of all portfolio returns coming from the same performance measure. As a result, we compare returns among 12 instead of 32 portfolios.

Standard descriptive statistics on the portfolio returns computed using the whole OOS period (January 2006 to September 2014) are displayed in Table 2. First panel of Table 2 provides the mean, maximum, minimum, standard deviation, beta, skewness, and excess kurtosis.[4] The mean, maximum, minimum, and standard deviation are in percent numbers. Complementary information about risk is provided by the VaR and CVaR at 1 and 5 percent levels, and the maximum drawdown (MDD). The relative drawdown at time $t$ represents the percentage loss that the investor has suffered from the previous peak of price until time $t$. That is,

$$DD_t = \frac{M_t - P_t}{M_t} \geq 0,$$

where $P_t$ is the price at time $t$ and $M_t$ is the running maximum price in the time interval $[0, t]$. Hence, $M_t = \max_{\tau \in [0,t]} P_\tau$. Then, the relative maximum drawdown is defined as the largest drawdown for the entire sample:

$$MDD = \max_{t \in [0,T]} DD_t.$$

The FT portfolio exhibits the largest mean return followed by the portfolios based on VaRR5, GR1, and MAD. The FT portfolio return also shows the largest range (distance between the maximum and the minimum in Table 2), standard deviation and beta. However, the minimum values are similar among portfolios, while the maximum values are much larger for the FT portfolio. All portfolio returns are negatively skewed (except for FT with practically zero skewness) and show high values for excess kurtosis (highlighting FT). Important differences between portfolios in terms of VaR(5%), CVaR(5%) or MDD are not observed. Only for 1% level, both VaR and CVaR values are larger for the FT portfolio than the rest. Note

[4] Market betas have been computed by employing the S&P 500 index as a proxy for the market portfolio and using monthly frequency.

also, for instance, how the OOS portfolio return with the lowest VaR (1%) is not VaRR1 but GR1.

In order to visualize whether the largest mean return for FT portfolio is compensated by its largest risk, Figure 1 represents mean-risk ratios based on the risk measures provided in Table 2. As shown, FT portfolio exhibits the largest mean-risk compensation. The difference between FT portfolio and the remaining ones is remarkably high. All reward-to-variability ratios obtained from FT portfolio are more than 2.5 times the respective ratios for the VaRR(5%) portfolio, the second best portfolio in terms of risk compensation.

The second panel of Table 2 displays the correlation between the different portfolio returns. Comparing to SR portfolio, the lowest correlation corresponds to FT (65.9 percent). Note also that FT portfolio shows the lowest levels of return correlation with the other portfolios. GR1, GR5, VaRR1 and VaRR5 OOS returns show around 80 percent of correlation with SR. On the contrary, the remaining measures lead to OOS portfolio returns which are highly correlated with SR returns.

### 4.3 Portfolio Returns: Dynamic Analysis

We study here the behavior of PMs in a dynamic framework. First, we start estimating Dynamic Conditional Correlations (DCC) by Engle (2002) between SR portfolio returns and each of the other portfolio returns. Both graphs of Figure 2 display these DCC for two subsets of six (top graph) and five alternative measures (bottom graph). The daily conditional correlation series between SR and ASR, MAD, Gini, SS, or MM returns show high levels and are relatively stable along time. For the remaining PMs, the conditional correlation series are lower and highly time varying. In general, the correlations are positive but decrease considerably during the crisis period for some pairs. In particular, the cases of FT, GR1, GR5, VaRR1 and VaRR5. Therefore, summing up all the results upon this point, we can conclude that very different patterns in the portfolio return distribution are possible for different PMs. Next, we analyze more deeply these differences.

Second, Figure 3 represents the spread between cumulative returns on each portfolio and SR portfolio along the OOS period. The top graph displays the portfolio cumulative returns with the highest spreads regarding SR while the bottom graph shows the spreads for the portfolios with the most similar cumulative returns to SR. This visual analysis indicates that the VaRR5, GR1, MM, and specifically the FT cumulative returns are systematically over the SR ones. Moreover, the FT cumulative return is also substantially larger than the other returns ending with a cumulative value two times larger than the rest. However, this portfolio also shows a sharp drop during the recession. VaRR1 and GR5 portfolios produce returns higher than SR in the pre-crisis period but lower after the crisis. The bottom graph displays spreads of size considerably lower (see scale in vertical axis) and also showing both positive and negative values. In this case, only the MAD portfolio produces higher cumulative returns than SR for almost the whole sample period.

Third, Figure 4 compares the time series of drawdowns for the different portfolios. Remember that the drawdown represents the maximum loss and thus, the higher the drawdown the worse performance. Consistently with Figure 3, the top graph shows that the drawdown is considerably lower for FT than for the remaining portfolios, and mainly, just after the crisis period. It seems that the drawdown is higher for SR than for the others before 2008 but it becomes lower after 2009.

**4.4 Stochastic Dominance Analysis**

In order to compare more rigorously the cumulative portfolio return distributions generated by the different PMs, we employ the stochastic dominance tool. Taking into account the weak assumptions of non-satiation and risk aversion for investor preferences, we can compare sample cumulative distribution functions (CDFs) for each pair of portfolios cumulative returns in terms of first-order and second-order stochastic dominances. We use the Davidson and Duclos (DD) (2000) test. Suppose a random sample of N independent drawings of observations ($y_{i,A}$, $y_{i,B}$), i=1,…,N, from two populations A and B, the statistic for testing the null that neither A nor B dominates at s-order each other has the following form:

$$T^S(x) = \frac{\widehat{D}_A^s(x) - \widehat{D}_B^s(x)}{\sqrt{\widehat{V}^s(x)}},$$

such that

$$\widehat{D}_j^s(x) = \frac{1}{N(s-1)!}\sum_{i=1}^{N}\left(x - y_{i,j}\right)^{s-1} I\left(y_{i,j} \leq x\right) = \frac{1}{N(s-1)!}\sum_{i=1}^{N}\left(x - y_{i,j}\right)_+^{s-1}, \;\; j = A, B$$

where $I(.)$ is an indicator function equal to 1 when the argument is true and 0 otherwise. Note that for s=1, $\widehat{D}$ is simply the empirical distribution function that estimates the population CDF.

And for s>1, $\widehat{D}$ is simply a sum of iid variables. Finally, $\widehat{V}$ is an estimate for the variance of the numerator of the T-statistic:

$$\widehat{V}^s = \widehat{V}_A^s + \widehat{V}_B^s - 2\widehat{V}_{AB}^s,$$

with

$$\widehat{V}_j^s(x) = \frac{1}{N}\left[\left(\frac{1}{N\big((s-1)!\big)^2}\sum_{i=1}^{N}\big(x - y_{i,j}\big)_+^{2(s-1)}\right) - \widehat{D}_j^s(x)^2\right],\ \ j = A, B$$

$$\widehat{V}_{AB}^s(x) = \frac{1}{N}\left[\left(\frac{1}{N\big((s-1)!\big)^2}\sum_{i=1}^{N}\big(x - y_{i,A}\big)_+^{2(s-1)}\big(x - y_{i,B}\big)_+^{2(s-1)}\right) - \widehat{D}_A^s(x)\widehat{D}_B^s(x)\right].$$

Under the null, $T$ is asymptotically distributed as a standard normal variable. We implement the DD test for first and second-order dominances. We set 100 different values for x, which are obtained by dividing the whole possible range (max($y_{i,j}$)-min($y_{i,j}$), for j= A, B) into 100 grids. To control for the joint test size, inference is based on the Studentized Maximum Modulus distribution with a 5 percent critical value of 3.254.[5] Therefore, we compute the number of significant positive statistics (T>3.254) and the number of significant negative statistics (T<-3.254). We denote the numbers of the total of significant positive and negative statistics as DD+ and DD-, respectively. We reject the null and conclude that B (A) dominates A (B) when the number of significant positive (negative) statistics DD+ is equal or higher than 50.

Stochastic dominance results for first-order (FSD) and second-order (SSD) are in Tables 3.a and 3.b, respectively. Since we are comparing the distribution of cumulative returns, then DD+ (DD-) higher than 50 implies that the portfolio indicated in the first row (column) dominates the portfolio shown in the first column (row). Confirming the visual conclusion from Figure 3, we find that FT, VaRR5, MAD, MM and additionally ASR are better than SR in terms of both marginal utility (FSD) and risk (SSD). The maximum number of significant DD+ is for MAD (96 out of 100), indicating than the CDF for SR cumulative returns is practically on the left of CDF for MAD returns. This fact can be observed on the top graph of Figure 5. The percentage of significant DD+ is lower in the FT case (74 or 70, for FSD or SSD). As the middle graph of Figure 5 displays, there is a higher range of values for which Sharpe

[5] The Studentized maximum modulus is the maximum absolute value of a set of independent unit normal variables which is studentized by the standard deviation. Critical values for infinite degrees of freedom are tabulated by Stoline and Ury (1979).

CDF is on the right of FT one. However, the distance between both CDFs is considerably larger than the distance between MAD and SR. This is supported by the average DD statistic value of 9.23 for the SR-MAD comparison, whereas 33.01 for the SR-FT comparison. Tables 3.a and 3.b also show that the SR cumulative return distribution dominates SS, GR1, GR5, and Range distributions. As an example, the bottom graph of Figure 5 represents the empirical CDFs for SR and SS, now the average DD statistic value is equal to -12.15.

Other comparisons show, in general, that SS and GR5 are worse than the other PMs, while FT is the best. In addition, VaRR5, MAD and MM are better than the other measures but worse than FT.

## 5 Robustness Checks

Note that some decisions to obtain the OOS portfolio returns in the previous empirical analysis have been taken arbitrarily. The aim of this section is to check whether the results in Section 4 may be invariant to some changes in the way of computing the OOS portfolio returns. Specifically, we analyze the effects of the individual weights on the portfolio and the period for the portfolio rebalancing.

### 5.1 Value Weighted Portfolios

In this case, the composition of the 32 initial portfolios is the same as in Section 4 but the weight on each individual asset is given by its market capitalization. This is a more realistic case since the maintenance costs are much lower than in the equally-weighted case.

Comparing Figures 3 and 6, we can see that the value-weighted strategy produces portfolio returns closer each other; the large distance between the spread for FT portfolio and any other exhibited in Figure 3 does not hold now. However, the FT portfolio shows again the highest cumulative spread regarding SR during almost the entire sample period. In contrast to the equally-weighted case, there are no many strategies showing higher returns than SR. Only MM (top graph) and Range (bottom graph) produce higher cumulative returns than SR during the whole sample period.

Tables 4.a and 4.b compare cumulative returns in terms of FSD and SSD, respectively. Now, most portfolios are worse than SR. Only FT, MM, and Range dominate SR in both first and second orders. Consistently with Tables 3.a and 3.b, FT strategy is the best one. Now SS is worse than some portfolios but better than others. However, VaRR1 and VaRR5 become the worst strategies.

**5.2 Period of Portfolio Rebalancing**

Daily rebalancing of portfolios can produce important transaction costs. We repeat the same analysis but now carrying out the asset selection only one time per month and so, keeping a buy-and-hold portfolio strategy every day within a month. In this case, the portfolio return is computed equally weighting the individual stock returns.

The graphs in Figure 7 show that spreads regarding SR are lower than in the case of daily rebalancing but the portfolio ordering is similar to the case of equally-weighted and daily rebalancing. As in the previous sections, the performance of the FT portfolio still stands out except for the financial crisis period, which drops sharply reaching large negative values.

Now many selecting criteria seem to do better than SR. This evidence is provided for Tables 5.a and 5.b. The distributions of cumulative returns on SS, FT, GR1, GR5, VaRR1, VaRR5, MAD, MM, Range and Gini portfolios dominate SR. Generally speaking, FT joint to GR1 and GR5 are the best portfolios while SS and Gini are the worst.

**6 Conclusions**

This paper goes on the use of performance measures (PMs) for selecting assets. The comparison between different performance measures as screening rules has previously addressed according to rank correlations. Our contribution is to compare the role of different measures by not looking at the individual asset rankings but the composition and the subsequent out-of-sample return of the portfolio that contains the selected assets. Therefore, our method of comparison complements some previous empirical evidence, already mentioned in the Introduction, since we analyze the consequences of different rankings on the investment results.

We work with 32 different PMs, sort individual assets by each PM, and select the best 20 ones to be included in a portfolio. The first comparison is in terms of portfolio composition. We find a very low percentage of coincident assets when comparing portfolios based on Sharpe ratio (SR) and Generalized Rachev (GR) ratios or VaR ratios (VaRR). The extreme case is the family of Farinelli-Tibiletti (FT) measures that clearly select assets which are not selected by other PMs.

The second comparison analyses the out-of-sample (OOS) portfolio returns. The return correlation between portfolios confirms that not only the composition but also the portfolio resulting returns are different especially between SR and FT, GR or VaRR. Descriptive statistics for the whole OOS period indicate that the mean return is considerably larger for the FT portfolio, followed by the portfolios based on VaRR5, GR1, and mean absolute deviation (MAD). In general, the FT portfolio return also shows the largest total risk (standard deviation, VaR and CVaR at 1%) and systematic risk. However, its reward-risk ratio is the best in terms of Sharpe and Treynor ratios. It is also remarkable the case of VaRR5 with the second largest values for the mean return, Sharpe and Treynor ratios.

The third comparison is about a dynamic analysis of OOS returns. We start looking at DCC correlations, and continue comparing the distributions of cumulative returns by a drawdown analysis and testing the stochastic dominance. In general, FT dominates the rest of portfolios when comparing both cumulative returns and drawdowns. Our dynamic analysis also concludes that the good differential performance of the FT portfolio is observed after the negative shock associated to the recent crisis. Precisely, during the recession period the correlation between SR and FT portfolio returns becomes much lower.

The final analysis deals with the effects of changing the weights on the portfolios (keeping the composition) and the period length for portfolio rebalancing. Results are similar in the second case but we find important differences in portfolio returns when individual stocks are value weighted. However, and generally speaking, the dominance of the FT portfolio is still verified for cumulative returns.

## References


Bacon, C.R., (2011), Practical portfolio performance measurement and attribution, Wiley, second edition.

Bernardo, A.E., Ledoit, O., (2000), Gain, Loss and asset pricing, *Journal of Political Economy* 108(1), 144-172.

Biglova, A., Ortobelli, S., Rachev, S., Stoyanov, S., (2004), Different approaches to risk estimation in portfolio theory, *Journal of Portfolio Management* 3, 103-112.

Billio, M., Caporin, M., Costola, M. (2012), Backward/forward optimal combination of performance measures for equity screening, Ca' Foscari University of Venice, WP No. 13.

Bodnar, T., Parolya, N., Schmid, W., (2015), On the exact solution of the multi-period portfolio choice problem for an exponential utility under return predictability, *Europan Journal of Operational Research* 246, 528-542.

Caporin, M., Jannin G.M., Lisi F., Maillet B. B., (2014). A survey on the four families of performance measures, *Journal of Economic Surveys* 28 (5), 917-942.

Caporin, M., Lisi, F., (2011), Comparing and selecting performance measures using rank correlations, *Economics: The open-access, open-assessment E-Journal* 5, 1-34.

Davidson, R., Duclos, J.Y., (2000), Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality, *Econometrica* 68(6), 1435-1464.

De Miguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naïve diversification: how inefficient is the 1/N portfolio strategy, *Review of Financial Studies*, 22: 1915-1953.

Eling, M., Schuhmacher, F., (2007), Does the choice of performance measure influence the evaluation of hedge funds? , *Journal of Banking & Finance* 31, 2632-2647.

Eling, M., (2008), Does the Measure Matter in the Mutual Fund Industry?, *Financial Analysts Journal*, 64(3), 54-66.

Engle, R. F., (2002), Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroscedasticity models, *Journal of Business & Economic Statistics* 20, 339-50.

Farinelli, S., Ferreira, M., Rossello, D., Thoeny, M., Tibiletti, L., (2009), Optimal asset allocation aid system: From "one-size" vs "tailor-made" performance ratio, *European Journal of Operational Research* 192, 209-215.

Farinelli, S., Tibiletti, L., (2008), Sharpe thinking in asset ranking with one-sided measures, *European Journal of Operational Research* 185, 1542-1547.

Fishburn, P.C., (1977), Mean-risk analysis with risk associated with below-target returns, *The American Economic Review* 67, 116-126.

Guo, B., Xiao, Y., (2015), Why doesn't the choice of performance measure matter? Finance Research Letters, forthcoming.

Hwang, S., Salmon, M., (2003), An analysis of performance measures using copulae, in Knight, J., and Satchell, S. (eds), Performance measurement in finance: firms, funds and managers, Butterworth-Heinemann Finance, Quantitative Finance Series.

Jain, P.C (1987), The Effect on Stock Price of Inclusion in or Exclusion from the S&P 500, *Financial Analysts Journal*, January-February, 58-65.

Kaplan, P.D., Knowles, J.A., (2004), Kappa: A generalized downside risk adjusted performance measure, *Journal of Performance Measurement* 8 (3), 42-54.

Keating, C., Shadwick, W.F., (2002), A Universal Performance Measure, *Journal of Performance Measurement* 6(3), 59-84.

Konno, H., Yamazaki, H., (1991), Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market, *Management Science* 37, 519-531.

León, A., Moreno, M., (2015), One-sided performance measures under Gram-Charlier distribution, Working paper.

Magron, C. (2014), Are individual investors such poor portfolio managers?, Working paper.

Owen, J., Rabinovitch, R. (1983), On the class of elliptical distributions and their applications to the theory of portfolio choice, *Journal of Finance* 38(3), 745-752.

Pézier, J., White, A. (2008), The Relative Merits of Alternative Investments in Passive Portfolios, *The Journal of Alternative Investments* 10(4), 37-49.

Pruitt S. K.C. Wei (1989), Institutional Ownership and Changes in the S&P 500, *Journal of Finance*, June, 509-514.

Shalit, H., Yitzhaki, S., (1984), Mean-Gini, portfolio theory, and the pricing of risky assets, *Journal of Finance* 39, 1449-1468.

Sharpe, W.F., (1994), The Sharpe ratio, *Journal of Portfolio Management* 21, 49-58.

Sharpe, W.F., (1966), Mutual Fund Performance, *Journal of Business* 39, 119-138.

Sortino, F.A., van der Meer, R., (1991), Downside Risk, *Journal of Portfolio Management* 17 (Spring), 27-31.

Sortino, F.A., Van der Meer, R., Platinga, A. (1999), The Dutch triangle, *Journal of Portfolio Management* 26 (Fall), 50-58.

Sortino, F.A., Satchell, S. (2001), *Managing downside risk in financial markets: Theory, Practice and Implementation*, Butterworth Heinemann, Oxford.

Stoline, M.R., H.K. Ury (1979). Tables of the Studentized Maximum Modulus Distribution and an application to multiple comparisons among means. *Technometrics* 21, 87-93.

Young, M.R., (1998), A MiniMax portfolio selection rule with linear programming solution, *Management Science* 44, 673-683.

Zakamouline, V., (2011), The performance measure you choose influences the evaluation of hedge funds, *Journal of Performance Measurement* 15(3), 48-64.

## Acknowledgements

Authors acknowledge the financial support from Generalitat Valenciana through grant PROMETEOII/2013/015. We thank Enrique Sentana for his helpful discussion in the XXIII Finance Forum at the Universidad Pontificia de Comillas in Madrid and seminar participants at the Universidad CEU San Pablo in Elche for useful comments on the paper. We assume full responsibility for any remaining errors.

**Table 1. Percentage of coincident assets comparing pairs of portfolios based on different measures. Median values.**

| | | Kappa | | | | | | Farinelli-Tibiletti | | | | | | Generalized Rachev 1% | | | | | | Generalized Rachev 5% | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 |
| 1 | 1 | 0.75 | 0.85 | 0.9 | 0.9 | 0.9 | 0.85 | 0 | 0 | 0 | 0 | 0 | 0 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.2 | 0.25 | 0.2 | 0.25 | 0.25 | 0.2 | 0.15 | 0.25 | 0.9 | 0.6 | 0.7 | 0.95 |
| 2 | | 0.75 | 0.9 | 0.9 | 0.9 | 0.9 | 0.85 | 0 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.7 | 0.15 | 0.25 | 0.9 | 0.6 | 0.7 |
| 3 | | | 0.85 | 0.8 | 0.75 | 0.7 | 0.65 | 0 | 0 | 0 | 0 | 0 | 0 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.3 | 0.3 | 0.35 | 0.35 | 0.3 | 0.2 | 0.2 | 0.75 | 0.55 | 0.7 | 0.7 |
| 4 | | | | 0.95 | 0.9 | 0.85 | 0.8 | 0 | 0 | 0 | 0 | 0 | 0 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.3 | 0.35 | 0.3 | 0.35 | 0.35 | 0.3 | 0.2 | 0.25 | 0.8 | 0.55 | 0.65 | 0.85 |
| 5 | | | | | 0.95 | 0.9 | 0.85 | 0 | 0 | 0 | 0 | 0 | 0 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.3 | 0.3 | 0.3 | 0.35 | 0.3 | 0.3 | 0.2 | 0.25 | 0.8 | 0.55 | 0.65 | 0.9 |
| 6 | | | | | | 0.95 | 0.9 | 0 | 0 | 0 | 0 | 0 | 0 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.2 | 0.25 | 0.8 | 0.5 | 0.6 | 0.95 |
| 7 | | | | | | | 0.95 | 0 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.25 | 0.3 | 0.25 | 0.3 | 0.3 | 0.25 | 0.2 | 0.25 | 0.8 | 0.5 | 0.6 | 0.95 |
| 8 | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.2 | 0.25 | 0.75 | 0.5 | 0.6 | 0.9 |
| 9 | | | | | | | | | 0.65 | 0.4 | 0.6 | 0.2 | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | | | | | | | | | | 0.65 | 0.65 | 0.4 | 0.25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | | | | | | | | | | | 0.55 | 0.7 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | | | | | | | | | | | | 0.425 | 0.35 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | | | | | | | | | | | | | 0.75 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | | | | | | | | | | | | | | | 0.95 | 0.95 | 0.95 | 0.9 | 0.95 | 0.65 | 0.6 | 0.65 | 0.6 | 0.5 | 0.7 | 0.45 | 0.1 | 0.1 | 0 | 0.05 | 0.2 |
| 16 | | | | | | | | | | | | | | | | 0.95 | 1 | 0.95 | 0.95 | 0.6 | 0.6 | 0.65 | 0.6 | 0.55 | 0.7 | 0.45 | 0.1 | 0.1 | 0 | 0.05 | 0.2 |
| 17 | | | | | | | | | | | | | | | | | 0.95 | 0.9 | 1 | 0.6 | 0.6 | 0.65 | 0.6 | 0.5 | 0.7 | 0.45 | 0.1 | 0.1 | 0 | 0.05 | 0.2 |
| 18 | | | | | | | | | | | | | | | | | | 0.95 | 0.95 | 0.6 | 0.6 | 0.65 | 0.6 | 0.55 | 0.7 | 0.45 | 0.1 | 0.1 | 0 | 0.05 | 0.2 |
| 19 | | | | | | | | | | | | | | | | | | | 0.9 | 0.6 | 0.6 | 0.65 | 0.6 | 0.55 | 0.7 | 0.5 | 0.1 | 0.1 | 0 | 0.05 | 0.2 |
| 20 | | | | | | | | | | | | | | | | | | | | 0.6 | 0.6 | 0.65 | 0.6 | 0.5 | 0.7 | 0.45 | 0.1 | 0.1 | 0 | 0.05 | 0.2 |
| 21 | | | | | | | | | | | | | | | | | | | | | 0.85 | 0.9 | 0.85 | 0.75 | 0.9 | 0.55 | 0.3 | 0.15 | 0.05 | 0.1 | 0.25 |
| 22 | | | | | | | | | | | | | | | | | | | | | | 0.95 | 1 | 0.9 | 0.9 | 0.5 | 0.3 | 0.2 | 0.1 | 0.15 | 0.25 |
| 23 | | | | | | | | | | | | | | | | | | | | | | | 0.9 | 0.85 | 0.95 | 0.5 | 0.25 | 0.2 | 0.05 | 0.1 | 0.25 |
| 24 | | | | | | | | | | | | | | | | | | | | | | | | 0.9 | 0.9 | 0.5 | 0.3 | 0.2 | 0.1 | 0.15 | 0.25 |
| 25 | | | | | | | | | | | | | | | | | | | | | | | | | 0.8 | 0.5 | 0.3 | 0.2 | 0.1 | 0.15 | 0.25 |
| 26 | | | | | | | | | | | | | | | | | | | | | | | | | | 0.5 | 0.25 | 0.15 | 0.05 | 0.1 | 0.25 |
| 27 | | | | | | | | | | | | | | | | | | | | | | | | | | | 0.15 | 0.1 | 0.05 | 0.05 | 0.2 |
| 28 | | | | | | | | | | | | | | | | | | | | | | | | | | | | 0.2 | 0.15 | 0.2 | 0.25 |
| 29 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 0.7 | 0.75 | 0.85 |
| 30 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 0.85 | 0.55 |
| 31 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 0.65 |

Measures 1 and 2 are the Sharpe and Adjusted Sharpe ratio, respectively. Measures from 3 to 8 are Kappa ratios with the following values for the parameter: 10, 3, 2, 1.5, 0.8, and 0.5. Measures from 9 to 14 are Farinelli-Tibilitti ratios with the following pair of values for the parameters: (0.5, 2), (1.5, 2), (1, 1), (2, 1.5), (3, 0.5), and (1, 2). Measures from 15 to 20 and from 21 to 26 are Generalized Rachev ratios evaluated at 1% and 5% levels respectively. In both cases the values for each pair of parameters are: (0.8, 0.001), (0.5, 0.8), (0.8, 0.8), (0.5, 1), (0.01, 0.8), and (1, 1). 27 and 28 are VaR ratios at 1% and 5% level respectively; 29 is the mean absolute deviation; 30 is the distance between the maximum and the minimum; 31 is the Range ratio which is the absolute value of the distance between the maximum and the minimum; and 32 is the Mean-Gini ratio.

**Table 2. Descriptive statistics of out-of-sample portfolio returns.**

| | SR | ASR | SS | FT | GR1 | GR5 | VaRR1 | VaRR5 | MAD | MM | Range | Gini |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean | 0.0039 | 0.0064 | 0.0061 | 0.0446 | 0.0083 | 0.0023 | 0.0035 | 0.0129 | 0.0080 | 0.0069 | 0.0025 | 0.0069 |
| Max | 7.756 | 7.756 | 7.756 | 17.526 | 9.776 | 9.800 | 10.298 | 10.302 | 7.802 | 7.756 | 7.761 | 7.756 |
| Min | -9.934 | -9.934 | -9.983 | -12.495 | -9.709 | -11.735 | -9.801 | -11.706 | -9.934 | -9.758 | -9.758 | -9.934 |
| StD | 1.426 | 1.427 | 1.413 | 1.961 | 1.407 | 1.439 | 1.392 | 1.525 | 1.442 | 1.462 | 1.463 | 1.414 |
| Beta | 0.951 | 0.949 | 0.968 | 1.584 | 1.098 | 1.097 | 1.146 | 1.173 | 0.961 | 0.994 | 1.015 | 0.929 |
| MDD | 0.591 | 0.595 | 0.609 | 0.680 | 0.635 | 0.653 | 0.694 | 0.675 | 0.605 | 0.624 | 0.630 | 0.599 |
| Skewness | -0.561 | -0.565 | -0.576 | 0.069 | -0.491 | -0.802 | -0.592 | -0.655 | -0.577 | -0.740 | -0.739 | -0.564 |
| Exc.Kurtosis | 4.533 | 4.541 | 4.839 | 13.789 | 6.439 | 8.078 | 7.059 | 7.303 | 4.480 | 5.186 | 5.124 | 4.665 |
| VaR(1%) | 0.042 | 0.043 | 0.042 | 0.075 | 0.039 | 0.043 | 0.041 | 0.049 | 0.044 | 0.045 | 0.045 | 0.043 |
| VaR(5%) | 0.024 | 0.024 | 0.024 | 0.027 | 0.023 | 0.023 | 0.022 | 0.025 | 0.025 | 0.025 | 0.025 | 0.024 |
| CVaR(1%) | 0.057 | 0.057 | 0.057 | 0.094 | 0.059 | 0.064 | 0.060 | 0.068 | 0.057 | 0.061 | 0.061 | 0.057 |
| CVaR(5%) | 0.037 | 0.037 | 0.036 | 0.049 | 0.035 | 0.036 | 0.035 | 0.039 | 0.037 | 0.038 | 0.038 | 0.036 |
| Correlations | | ASR | SS | FT | GR1 | GR5 | VaRR1 | VaRR5 | MAD | MM | Range | Gini |
| SR | | 0.999 | 0.993 | 0.659 | 0.813 | 0.830 | 0.801 | 0.833 | 0.993 | 0.969 | 0.976 | 0.994 |
| ASR | | | 0.993 | 0,658 | 0.815 | 0.832 | 0.803 | 0.835 | 0.992 | 0.968 | 0.975 | 0.995 |
| SS | | | | 0.669 | 0.829 | 0.846 | 0.815 | 0.844 | 0.986 | 0.964 | 0.971 | 0.994 |
| FT | | | | | 0.815 | 0.811 | 0.815 | 0.834 | 0.650 | 0.651 | 0.650 | 0.663 |
| GR1 | | | | | | 0.965 | 0.949 | 0.896 | 0.798 | 0.781 | 0.787 | 0.820 |
| GR5 | | | | | | | 0.953 | 0.928 | 0.816 | 0.799 | 0.805 | 0.838 |
| VaRR1 | | | | | | | | 0.895 | 0.788 | 0.775 | 0.780 | 0.808 |
| VaRR5 | | | | | | | | | 0.824 | 0.814 | 0.815 | 0.838 |
| MAD | | | | | | | | | | 0.977 | 0.983 | 0.987 |
| MM | | | | | | | | | | | 0.992 | 0.963 |
| Range | | | | | | | | | | | | 0.970 |

Descriptive statistics on the portfolio returns for the whole out-of-sample period: daily returns between January 2006 and September 2014. The portfolios are constructing equally weighting the best 20 stocks based on a pre-ranking that uses different performance measures: SR and ASR are the standard and the adjusted Sharpe ratios; SS represents Sortino-Satchell ratios; FT represents Farinelli-Tibiletti ratios; GR1 and GR5 are generalized Rachev ratios with 1 and 5 confidence levels; VaRR1 and VaRR5 are value-at-risk ratios at 1 and 5 confidence levels; MAD, MM, Range and Gini are risk-reward ratios that employ the mean absolute deviation, the distance between the maximum and minimum, the absolute distance between the maximum and minimum and the Shalit-Yitzhaki difference, respectively, as the risk measure. First panel provides standard return statistics: mean, maximum, minimum, standard deviation, market beta (Beta), the maximum drawdown (MDD), skewness, excess kurtosis, and value-at-risk (VaR) and conditional value-at-risk (CVaR) at 1 and 5 percent levels. Second panel displays the sample correlation between pairs of portfolio returns.

**Table 3.a. First-order stochastic dominance. Equally-weighted portfolios.**

| | | ASR | SS | FT | GR1 | GR5 | VaRR1 | VaRR5 | MAD | MM | Range | Gini |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SR | DD+ | 60 | 18 | 74 | 37 | 14 | 24 | 92 | 96 | 94 | 4 | 52 |
| | DD- | 2 | 77 | 14 | 50 | 74 | 39 | 0 | 0 | 0 | 75 | 29 |
| ASR | DD+ | | 17 | 74 | 35 | 14 | 15 | 87 | 79 | 69 | 0 | 11 |
| | DD- | | 78 | 14 | 51 | 75 | 56 | 0 | 0 | 10 | 85 | 35 |
| SS | DD+ | | | 98 | 96 | 46 | 80 | 87 | 80 | 82 | 76 | 77 |
| | DD- | | | 0 | 0 | 18 | 13 | 7 | 14 | 12 | 17 | 17 |
| FT | DD+ | | | | 0 | 0 | 13 | 16 | 15 | 17 | 14 | 14 |
| | DD- | | | | 91 | 95 | 74 | 72 | 73 | 71 | 74 | 74 |
| GR1 | DD+ | | | | | 0 | 49 | 53 | 51 | 52 | 46 | 50 |
| | DD- | | | | | 92 | 37 | 29 | 33 | 36 | 41 | 35 |
| GR5 | DD+ | | | | | | 78 | 88 | 77 | 77 | 73 | 75 |
| | DD- | | | | | | 11 | 1 | 11 | 9 | 15 | 14 |
| VaRR1 | DD+ | | | | | | | 93 | 74 | 78 | 11 | 48 |
| | DD- | | | | | | | 0 | 2 | 0 | 68 | 19 |
| VaRR5 | DD+ | | | | | | | | 0 | 15 | 0 | 0 |
| | DD- | | | | | | | | 80 | 61 | 94 | 87 |
| MAD | DD+ | | | | | | | | | 49 | 0 | 0 |
| | DD- | | | | | | | | | 24 | 96 | 80 |
| MM | DD+ | | | | | | | | | | 0 | 9 |
| | DD- | | | | | | | | | | 94 | 65 |
| Range | DD+ | | | | | | | | | | | 70 |
| | DD- | | | | | | | | | | | 0 |

This table displays the results from the first-order stochastic dominance comparison between cumulative returns on pairs of equally-weighted portfolios. The stocks included in each portfolio have been selected using a specific performance measure indicated in the first row. See notes in Table 2 for the definitions. Returns are daily and for the period between January 2006 and September 2014. DD+ (DD-) refers to the percentage of the Davidson and Duclos (2000) significant positive (negative) test statistics. See Section 4.4 for details.

**Table 3.b. Second-order stochastic dominance. Equally-weighted portfolios.**

| | | ASR | SS | FT | GR1 | GR5 | VaRR1 | VaRR5 | MAD | MM | Range | Gini |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SR | DD+ | 52 | 0 | 70 | 18 | 0 | 13 | 91 | 96 | 96 | 0 | 37 |
| | DD- | 39 | 95 | 14 | 65 | 89 | 72 | 0 | 0 | 0 | 96 | 51 |
| ASR | DD+ | | 0 | 70 | 15 | 0 | 0 | 92 | 95 | 96 | 0 | 0 |
| | DD- | | 95 | 14 | 67 | 89 | 90 | 0 | 0 | 0 | 97 | 96 |
| SS | DD+ | | | 98 | 96 | 84 | 95 | 95 | 95 | 95 | 95 | 95 |
| | DD- | | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| FT | DD+ | | | | 0 | 0 | 13 | 17 | 15 | 17 | 13 | 13 |
| | DD- | | | | 92 | 95 | 71 | 67 | 69 | 67 | 71 | 71 |
| GR1 | DD+ | | | | | 0 | 67 | 87 | 71 | 72 | 61 | 67 |
| | DD- | | | | | 90 | 17 | 0 | 11 | 7 | 22 | 16 |
| GR5 | DD+ | | | | | | 89 | 89 | 89 | 89 | 89 | 89 |
| | DD- | | | | | | 0 | 0 | 0 | 0 | 0 | 0 |
| VaRR1 | DD+ | | | | | | | 91 | 90 | 91 | 14 | 59 |
| | DD- | | | | | | | 0 | 0 | 0 | 54 | 0 |
| VaRR5 | DD+ | | | | | | | | 4 | 26 | 0 | 0 |
| | DD- | | | | | | | | 82 | 44 | 94 | 94 |
| MAD | DD+ | | | | | | | | | 96 | 0 | 0 |
| | DD- | | | | | | | | | 0 | 96 | 95 |
| MM | DD+ | | | | | | | | | | 0 | 0 |
| | DD- | | | | | | | | | | 96 | 96 |
| Range | DD+ | | | | | | | | | | | 83 |
| | DD- | | | | | | | | | | | 4 |

This table displays the results from the second-order stochastic dominance comparison between cumulative returns on pairs of equally-weighted portfolios. The stocks included in each portfolio have been selected using a specific performance measure indicated in the first row. See notes in Table 2 for the definitions. Returns are daily and for the period between January 2006 and September 2014. DD+ (DD-) refers to the percentage of the Davidson and Duclos (2000) significant positive (negative) test statistics. See Section 4.4 for details.

**Table 4.a. First-order stochastic dominance. Value-weighted portfolios.**

| | | ASR | SS | FT | GR1 | GR5 | VaRR1 | VaRR5 | MAD | MM | Range | Gini |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SR | DD+ | 0 | 1 | 92 | 29 | 0 | 0 | 11 | 44 | 96 | 95 | 17 |
| | DD- | 82 | 74 | 0 | 42 | 81 | 81 | 69 | 43 | 0 | 0 | 64 |
| ASR | DD+ | | 13 | 95 | 39 | 0 | 5 | 14 | 63 | 96 | 95 | 35 |
| | DD- | | 49 | 0 | 36 | 79 | 81 | 65 | 6 | 0 | 0 | 21 |
| SS | DD+ | | | 96 | 40 | 0 | 5 | 13 | 77 | 96 | 95 | 63 |
| | DD- | | | 0 | 40 | 78 | 79 | 56 | 0 | 0 | 0 | 0 |
| FT | DD+ | | | | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 |
| | DD- | | | | 91 | 94 | 94 | 91 | 94 | 78 | 83 | 96 |
| GR1 | DD+ | | | | | 0 | 0 | 13 | 46 | 91 | 89 | 43 |
| | DD- | | | | | 91 | 88 | 71 | 19 | 0 | 0 | 36 |
| GR5 | DD+ | | | | | | 20 | 86 | 89 | 93 | 93 | 80 |
| | DD- | | | | | | 55 | 0 | 0 | 0 | 0 | 0 |
| VaRR1 | DD+ | | | | | | | 91 | 86 | 93 | 93 | 81 |
| | DD- | | | | | | | 0 | 0 | 0 | 0 | 4 |
| VaRR5 | DD+ | | | | | | | | 76 | 91 | 84 | 64 |
| | DD- | | | | | | | | 6 | 0 | 0 | 11 |
| MAD | DD+ | | | | | | | | | 96 | 81 | 12 |
| | DD- | | | | | | | | | 0 | 11 | 60 |
| MM | DD+ | | | | | | | | | | 0 | 0 |
| | DD- | | | | | | | | | | 95 | 96 |
| Range | DD+ | | | | | | | | | | | 0 |
| | DD- | | | | | | | | | | | 96 |

This table displays the results from the first-order stochastic dominance comparison between cumulative returns on pairs of value-weighted portfolios. The stocks included in each portfolio have been selected using a specific performance measure indicated in the first row. See notes in Table 2 for the definitions. Returns are daily and for the period between January 2006 and September 2014. DD+ (DD-) refers to the percentage of the Davidson and Duclos (2000) significant positive (negative) test statistics. See Section 4.4 for details.

**Table 4.b. Second-order stochastic dominance.Value-weighted portfolios.**

| | | ASR | SS | FT | GR1 | GR5 | VaRR1 | VaRR5 | MAD | MM | Range | Gini |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SR | DD+ | 0 | 0 | 91 | 20 | 0 | 0 | 0 | 13 | 96 | 95 | 0 |
| | DD- | 96 | 95 | 0 | 58 | 91 | 91 | 88 | 73 | 0 | 0 | 95 |
| ASR | DD+ | | 0 | 95 | 34 | 0 | 0 | 0 | 96 | 96 | 95 | 37 |
| | DD- | | 96 | 0 | 38 | 91 | 91 | 88 | 0 | 0 | 0 | 42 |
| SS | DD+ | | | 96 | 39 | 0 | 0 | 0 | 96 | 96 | 95 | 96 |
| | DD- | | | 0 | 32 | 91 | 91 | 88 | 0 | 0 | 0 | 0 |
| FT | DD+ | | | | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 |
| | DD- | | | | 92 | 93 | 93 | 90 | 93 | 74 | 81 | 96 |
| GR1 | DD+ | | | | | 0 | 0 | 0 | 55 | 91 | 90 | 53 |
| | DD- | | | | | 91 | 91 | 88 | 29 | 0 | 0 | 35 |
| GR5 | DD+ | | | | | | 0 | 90 | 91 | 92 | 91 | 91 |
| | DD- | | | | | | 91 | 0 | 0 | 0 | 0 | 0 |
| VaRR1 | DD+ | | | | | | | 91 | 91 | 92 | 91 | 91 |
| | DD- | | | | | | | 0 | 0 | 0 | 0 | 0 |
| VaRR5 | DD+ | | | | | | | | 89 | 89 | 88 | 88 |
| | DD- | | | | | | | | 0 | 0 | 0 | 0 |
| MAD | DD+ | | | | | | | | | 96 | 95 | 0 |
| | DD- | | | | | | | | | 0 | 0 | 87 |
| MM | DD+ | | | | | | | | | | 0 | 0 |
| | DD- | | | | | | | | | | 96 | 96 |
| Range | DD+ | | | | | | | | | | | 0 |
| | DD- | | | | | | | | | | | 95 |

This table displays the results from the second-order stochastic dominance comparison between cumulative returns on pairs of value-weighted portfolios. The stocks included in each portfolio have been selected using a specific performance measure indicated in the first row. See notes in Table 2 for the definitions. Returns are daily and for the period between January 2006 and September 2014. DD+ (DD-) refers to the percentage of the Davidson and Duclos (2000) significant positive (negative) test statistics. See Section 4.4 for details.

**Table 5.a. First-order stochastic dominance. Monthly portfolio rebalancing.**

| | | ASR | SS | FT | GR1 | GR5 | VaRR1 | VaRR5 | MAD | MM | Range | Gini |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SR | DD+ | 27 | 93 | 75 | 96 | 95 | 95 | 92 | 95 | 94 | 94 | 60 |
| | DD- | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 16 |
| ASR | DD+ | | 93 | 75 | 95 | 95 | 95 | 91 | 95 | 93 | 92 | 48 |
| | DD- | | 0 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 23 |
| SS | DD+ | | | 72 | 94 | 95 | 84 | 86 | 71 | 68 | 54 | 9 |
| | DD- | | | 14 | 0 | 0 | 3 | 0 | 9 | 10 | 10 | 59 |
| FT | DD+ | | | | 25 | 27 | 24 | 18 | 16 | 20 | 17 | 15 |
| | DD- | | | | 59 | 57 | 59 | 64 | 70 | 66 | 69 | 71 |
| GR1 | DD+ | | | | | 11 | 0 | 5 | 0 | 0 | 0 | 0 |
| | DD- | | | | | 55 | 75 | 78 | 94 | 89 | 93 | 94 |
| GR5 | DD+ | | | | | | 0 | 28 | 0 | 0 | 0 | 0 |
| | DD- | | | | | | 83 | 56 | 95 | 92 | 95 | 95 |
| VaRR1 | DD+ | | | | | | | 51 | 0 | 11 | 2 | 0 |
| | DD- | | | | | | | 26 | 81 | 63 | 76 | 93 |
| VaRR5 | DD+ | | | | | | | | 2 | 11 | 4 | 0 |
| | DD- | | | | | | | | 83 | 69 | 82 | 85 |
| MAD | DD+ | | | | | | | | | 51 | 27 | 0 |
| | DD- | | | | | | | | | 33 | 45 | 85 |
| MM | DD+ | | | | | | | | | | 15 | 0 |
| | DD- | | | | | | | | | | 65 | 94 |
| Range | DD+ | | | | | | | | | | | 0 |
| | DD- | | | | | | | | | | | 92 |

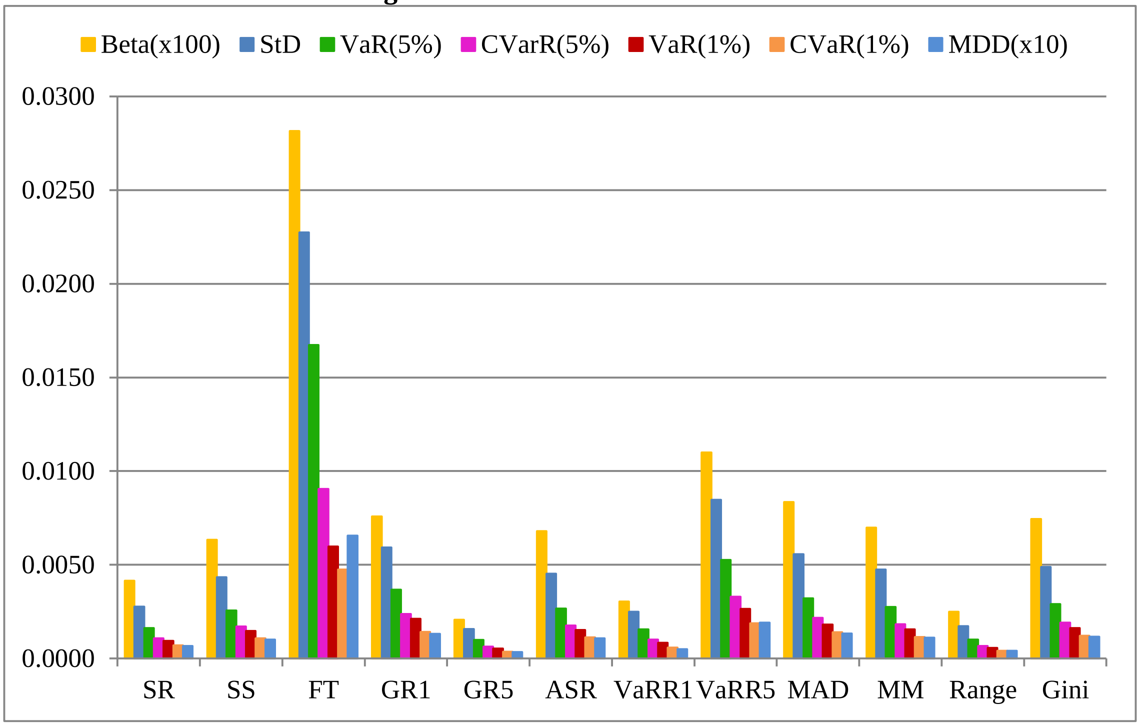
This table displays the results from the first-order stochastic dominance comparison between cumulative returns on pairs of equally-weighted portfolios. The stocks included in each portfolio have been selected using a specific performance measure indicated in the first row and the portfolios are rebalanced monthly. See notes in Table 2 for the definitions. Returns are daily and for the period between January 2006 and September 2014. DD+ (DD-) refers to the percentage of the Davidson and Duclos (2000) significant positive (negative) test statistics. See Section 4.4 for details.

**Table 5.b. Second-orders stochastic dominance. Monthly portfolio rebalancing.**

| | | ASR | SS | FT | GR1 | GR5 | VaRR1 | VaRR5 | MAD | MM | Range | Gini |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SR | DD+ | 95 | 94 | 70 | 95 | 94 | 94 | 92 | 94 | 94 | 94 | 94 |
| | DD- | 0 | 0 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ASR | DD+ | | 94 | 69 | 95 | 94 | 94 | 91 | 94 | 94 | 94 | 94 |
| | DD- | | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| SS | DD+ | | | 64 | 94 | 94 | 93 | 84 | 95 | 94 | 94 | 44 |
| | DD- | | | 16 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 48 |
| FT | DD+ | | | | 34 | 36 | 30 | 23 | 18 | 24 | 20 | 17 |
| | DD- | | | | 45 | 45 | 52 | 54 | 62 | 57 | 60 | 63 |
| GR1 | DD+ | | | | | 40 | 0 | 0 | 0 | 0 | 0 | 0 |
| | DD- | | | | | 44 | 91 | 90 | 94 | 88 | 93 | 94 |
| GR5 | DD+ | | | | | | 0 | 0 | 0 | 0 | 0 | 0 |
| | DD- | | | | | | 89 | 90 | 94 | 94 | 94 | 94 |
| VaRR1 | DD+ | | | | | | | 37 | 0 | 0 | 0 | 0 |
| | DD- | | | | | | | 49 | 92 | 82 | 89 | 93 |
| VaRR5 | DD+ | | | | | | | | 5 | 21 | 8 | 4 |
| | DD- | | | | | | | | 81 | 63 | 76 | 82 |
| MAD | DD+ | | | | | | | | | 53 | 45 | 0 |
| | DD- | | | | | | | | | 27 | 43 | 95 |
| MM | DD+ | | | | | | | | | | 0 | 0 |
| | DD- | | | | | | | | | | 94 | 94 |
| Range | DD+ | | | | | | | | | | | 0 |
| | DD- | | | | | | | | | | | 94 |

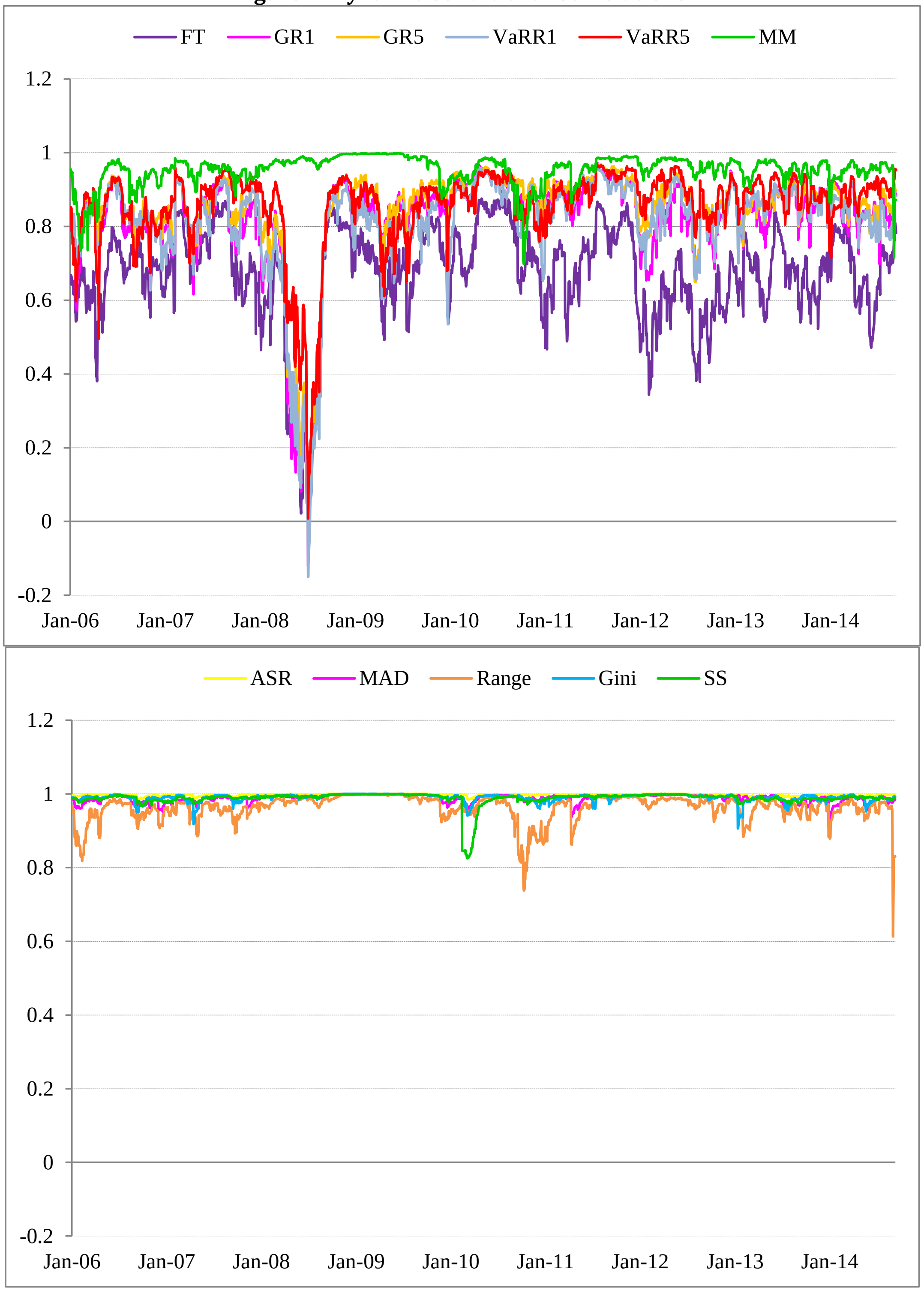
This table displays the results from the second-order stochastic dominance comparison between cumulative returns on pairs of equally-weighted portfolios. The stocks included in each portfolio have been selected using a specific performance measure indicated in the first row and the portfolios are rebalanced monthly. See notes in Table 2 for the definitions. Returns are daily and for the period between January 2006 and September 2014. DD+ (DD-) refers to the percentage of the Davidson and Duclos (2000) significant positive (negative) test statistics. See Section 4.4 for details.

**Figure 1. Mean Return-Risk Ratios**



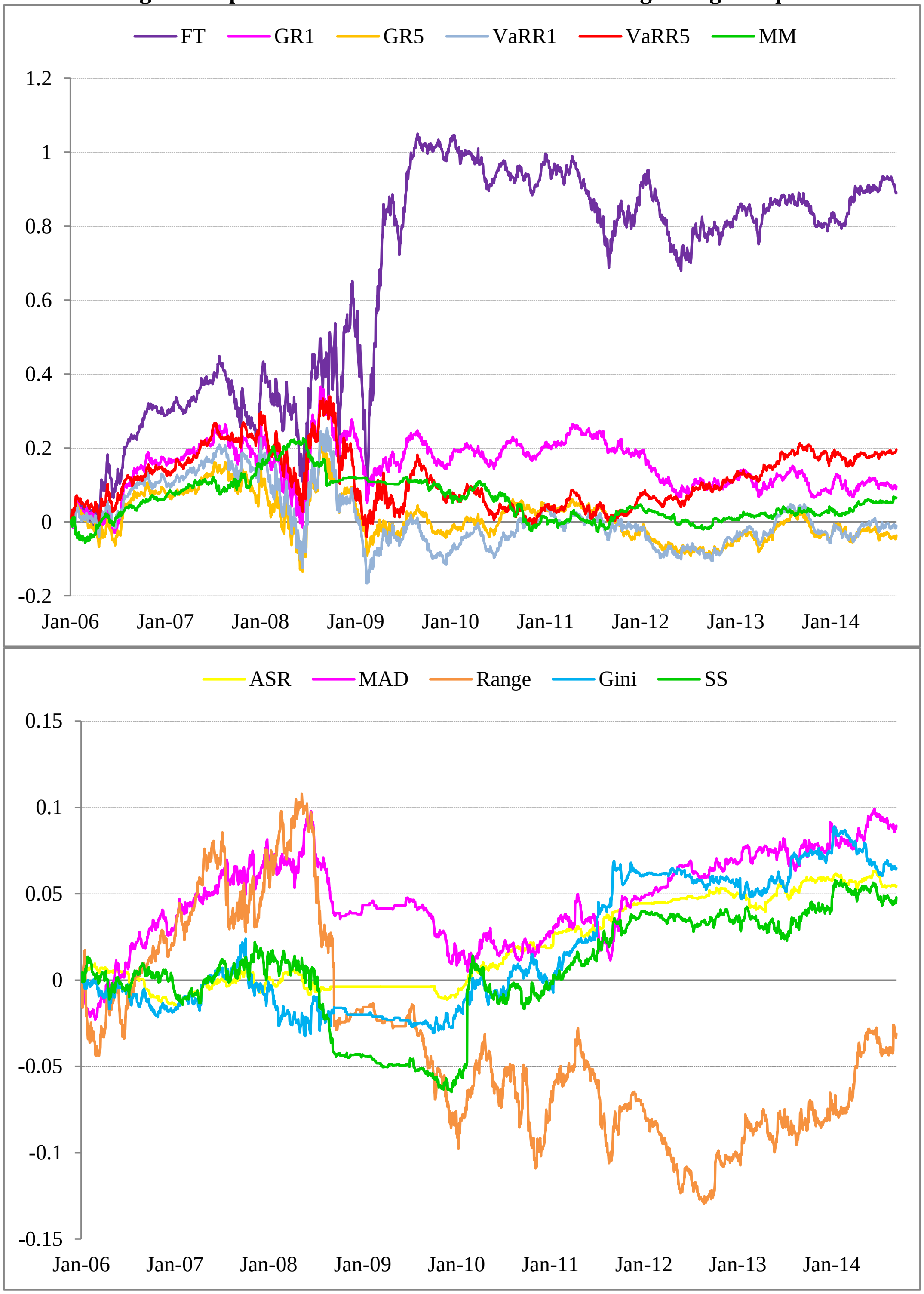

This figure displays the value of some mean return-risk ratios computed from daily return on 12 portfolios during the period between January 2006 and September 2014. The risk measure differs for each ratio and is indicated at the top: market beta (Beta), standard deviation (StD), value-at-risk (VaR), conditional value-at-risk (CVaR), and maximum Drowdawn (MDD). Numbers in parenthesis indicate the confidence level. The portfolios are constructing equally weighting the best 20 stocks based on a pre-ranking that uses different performance measures: SR and ASR are the standard and the adjusted Sharpe ratios; SS represents Sortino-Satchell ratios; FT represents Farinelli-Tibiletti ratios; GR1 and GR5 are generalized Rachev ratios with 1 and 5 confidence levels; VaRR1 and VaRR5 are value-at-risk ratios at 1 and 5 confidence levels; MAD, MM, Range and Gini are risk-reward ratios that employ the mean absolute deviation, the distance between the maximum and minimum, the absolute distance between the maximum and minimum and the Shalit-Yitzhaki difference, respectively, as the risk measure.

**Figure 2. Dynamic Conditional Correlations**



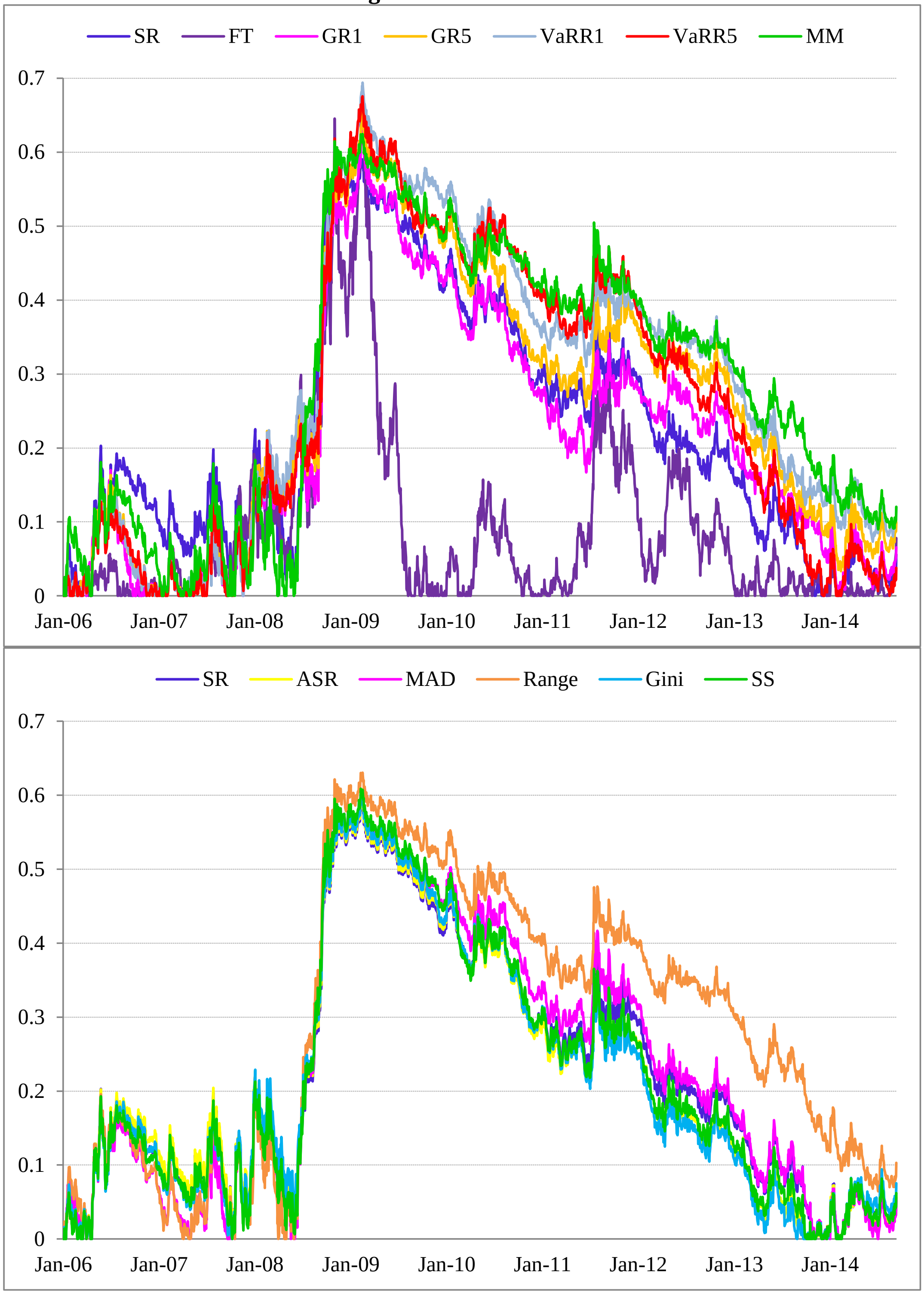

These figures display the dynamic conditional correlations between the returns on the Sharpe portfolio and each of the other portfolios indicated on the top. See notes in Figure 1 for the definitions. Returns are daily and refer to the period between January 2006 and September 2014.

**Figure 3. Spreads between Cumulative Returns regarding Sharpe**



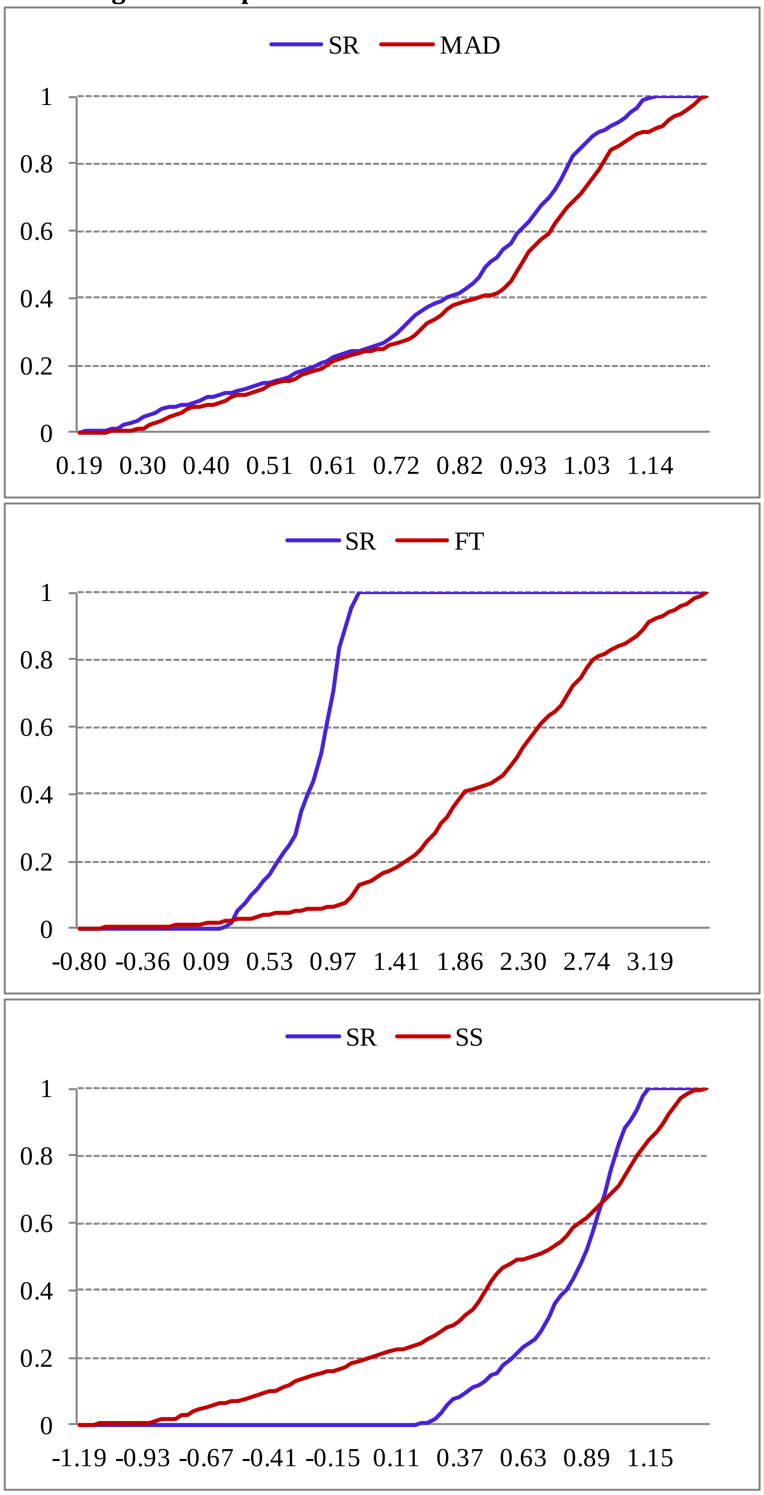

These figures display the difference between the cumulative return on each of the portfolios indicated at the top and on the Sharpe portfolio. See notes in Figure 1 for the definitions. Returns are daily and refer to the period between January 2006 and September 2014.

**Figure 4. Drawdowns**



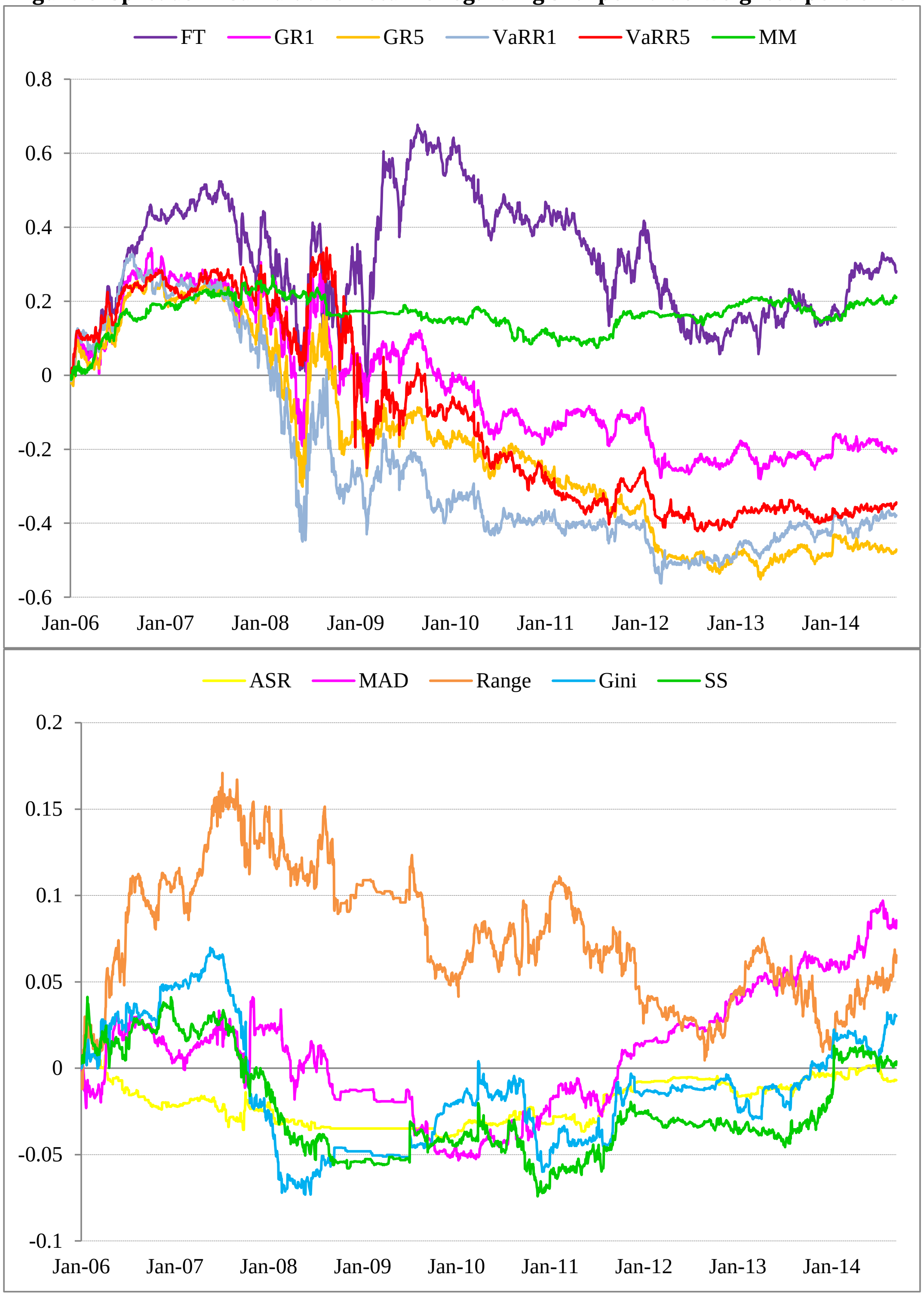

These figures display the time series of drowdawns for each of the portfolios indicated at the top. See notes in Figure 1 for the definitions. Returns are daily and refer to the period between January 2006 and September 2014.

**Figure 5. Empirical CDFs for cumulative returns**



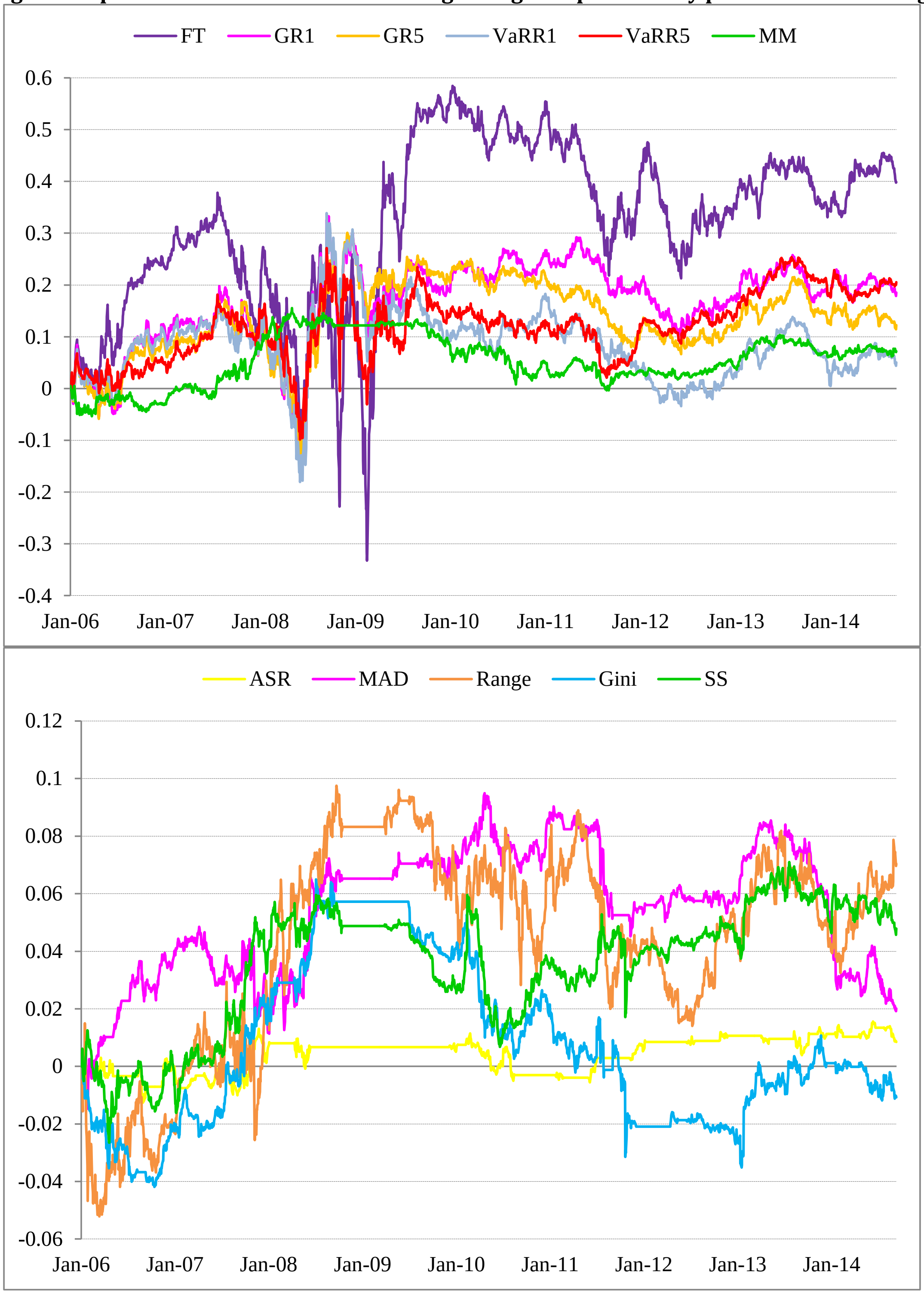

These figures compare the cumulative distribution function for returns on the Sharpe portfolio and MAD (up), FT (middle), and SS (down) portfolios. See Figure 1 for the definitions.

**Figure 6. Spreads in Cumulative Returns regarding Sharpe. Value-weighted portfolios**




These figures display the difference between the cumulative return on each of the portfolios indicated at the top and on the Sharpe portfolio. See notes in Figure 1 for the definitions. In this case, the portfolios are constructed value-weighting the individual stocks. Returns are daily and refer to the period between January 2006 and September 2014.

**Figure 7. Spreads in Cumulative Returns regarding Sharpe. Monthly portfolio rebalancing**




These figures display the difference between the cumulative return on each of the portfolios indicated at the top and on the Sharpe portfolio. See notes in Figure 1 for the definitions. The portfolios are rebalanced monthly. Returns are daily and refer to the period between January 2006 and September 2014.