NUS
National University of Singapore

SCAPE

Singapore Centre for Applied and Policy Economics

Stochastic Dominance Analysis of iShares

by

Dominic Gasbarro, Wing-Keung Wong and J. Kenton Zumwalt

Department of Economics
SCAPE Working Paper Series
Paper No. 2007/06 – April 2007
http://nt2.fas.nus.edu.sg/ecs/pub/wp-scape/0706.pdf

Stochastic Dominance Analysis of iShares

Dominic Gasbarro
Murdoch Business School
Murdoch University

Wing-Keung Wong
Risk Management Institute and Department of Economics
National University of Singapore

J. Kenton Zumwalt
University of Western Australia
Colorado State University

Abstract

Country indices as represented by iShares exhibit non-normal return distributions with both skewness and kurtosis. Davidson and Duclos (2000) and Memmel (2003**)** provide procedures for determining the statistical significance of stochastic dominance measures and the Sharpe Ratio, respectively. This study uses these refinements to compare the performance of 18 country market indices. The iShares are indistinguishable when using the Sharpe Ratio as no significant differences are found. In contrast, stochastic dominance procedures identify dominant iShares. Although the results vary over time, stochastic dominance appears to be both more robust and discriminating than the CAPM in the ranking of the iShares.

JEL classifications: G11, G15

Keywords: Stochastic dominance; Sharpe ratio; skewness; country index funds

1. Introduction

Contemporary finance advocates the use of the mean-variance model developed by Markowitz (1952) and the capital asset pricing model statistics (CAPM) developed by Sharpe (1964), Treynor (1965) and Jensen (1969) for portfolio construction and performance evaluation. These methodologies depend on normal return distributions and quadratic utility functions and are not appropriate if return distributions are not normal or investors' utility functions are not quadratic.

Stochastic dominance (SD) rules offer superior criteria on which to base investment decisions relative to the traditional mean-variance (MV) model because the assumptions underlying SD are less restrictive than those of the MV. SD incorporates information on the entire return distribution, rather than the first two moments as with MV and requires no precise assessment as to the specific form of the investor's risk preference or utility function. It also allows us to determine if an arbitrage opportunity exists among the investment alternatives so that once an arbitrage opportunity is identified, investors can increase their utilities as well as wealth by setting up zero dollar portfolios to exploit this opportunity.

The SD requirements on investors' utility functions depend on the level of stochastic dominance being examined. They must exhibit non-satiation (more is preferred to less) under first-order SD (FSD); non-satiation and risk aversion under second-order SD (SSD); and non-satiation, risk aversion, and decreasing absolute risk aversion (DARA) under third-order SD (TSD). Jarrow (1986) shows that if the return distribution of investment Y dominates the return distribution of investment Z in the sense of first order stochastic dominance, investors can increase both their wealth and their utilities by shifting from Z to Y. Stochastic dominance procedures allow the identification of these arbitrage opportunities.

These advantages of stochastic dominance have motivated prior studies to use SD techniques to evaluate the performance of mutual funds. Unfortunately, earlier research was unable to determine the statistical significance of stochastic dominance. However, recent advances in stochastic dominance techniques by Davidson and Duclos (2000) (DD)[i] allow differences between any two return cumulative density functions to be tested for statistical significance. The Davidson and Duclos SD procedures allow us to identify the negative and positive regions for FSD, SSD, and TSD and their levels of significance.

An opportunity for applying these innovations emerged with the introduction of country index funds. Standard and Poor's Depository Receipts, (SPDRs or "spiders") track the S&P 500 Index and began trading in January 1993. The acceptance and wide use of SPDRs (ticker symbol: SPY) led to the introduction in March, 1996, of seventeen exchange traded funds (ETFs) known as World Equity Benchmark Shares (WEBS). WEBS, now known as iShares, are investment companies designed to track the Morgan Stanley Capital International (MSCI) foreign stock market indices. These innovations allow investors to continuously trade shares of several well-diversified portfolios.[ii]

Empirically these securities' distributions are non-normal and exhibit both skewness and kurtosis. In addition, shocks to the system cause stock returns to exhibit non-normal behavior and the return distributions may exhibit the "fat tails" associated with extraordinary gains or losses. It is essential that the shocks be correctly modelled. In the MV framework, shocks are only modelled by changes in the mean and variance. However, stochastic dominance considers the entire distribution and shock information is more fully impounded in the evaluation process.

We find that the traditional CAPM measures are ambiguous in their evaluation of the iShares. Specifically, ambiguity is present both between and within measures. For example, although the Sharpe ratio can be used to rank the iShares and the dispersion is relatively wide, we find that none of the differences are statistically significant, including the difference between the highest and lowest ratios. The evaluation problem is compounded because the Treynor and Jensen measures suggest different rankings. These measures use systematic risk in their calculations and these are problematic during periods when markets are volatile.

We use SD procedures that allow us to determine whether statistically significant stochastic dominance occurs among 18 marketable iShares. The procedures allow the determination of whether dominance is due to the positive or negative portion of the return distributions and how they impact a risk-averse investor. We find that over the entire 1996 – 2003 period certain iShares dominate others. Conversely, some do not dominate any other iShares, but they themselves are not dominated by all iShares. We find that SPY dominates most of the other funds while Malaysia (EWM) is dominated by most of the other funds. Spain (EWP) and Japan (EWJ) show the greatest difference in Sharpe ratio, but do not appear different from a SD perspective.

Taylor and Yoder (1999) point out that during periods of extreme market stress increased skewness is observed in the return distribution. So, to examine how stochastic dominance captures the impact of shocks, we divide our sample into three periods based on the Asian financial crisis and the bursting of the technology bubble in the U.S. We find substantial changes occur over time irrespective of evaluation technique. Surprisingly, it appears that Spain (EWP) exhibits FSD over Japan (EWJ) during the period leading up to the Asian financial crisis. On the other hand, U.S. (SPY) continues to dominate the most funds over all three subperiods.

Finally, we examine the efficacy of SD procedures during "up" and "down" markets. A regime shifting procedure is applied to the MSCI index to identify up and down markets for the exchange traded funds, and, again SD procedures are applied. The dominance of the SPY over the entire period and the three subperiods are confirmed with the up-market and down-market results.

2. Literature Review

Early stochastic dominance research falls into two categories. The first type compares efficient frontiers generated by mean-variance models with efficient frontiers generated by stochastic dominance models. Levy and Sarnat (1970) find that the efficient set according to the MV criteria is reasonably similar to a set using a concave utility function. They suggest, however, that stochastic dominance may be used to reduce the number of alternatives via a first screening of the data. Porter (1973) compares the MV frontier with the frontier developed by stochastic dominance procedures. He reports the two efficient frontiers are similar and that discrepancies are minor. Recently, Kjetsaa and Kieff (2003) show how to use stochastic dominance to iteratively reduce a large set of equity mutual funds operating over the period 1985-2000 to a single-digit set of non-dominated funds. They suggest stochastic dominance may be used to identify funds that outperform market indices.

Sharpe (1966) develops the Sharpe Ratio to compare 34 mutual funds with the Dow Jones Industrial Average (the Dow). He reports that over the 1954-1963 period, 11 funds outperformed the Dow while 23 were outperformed by the Dow. Joy and Porter (1974) and Meyer (1977) make the same comparison using stochastic dominance, and both studies report that over the 1954-1963 the Dow stochastically dominated mutual funds.

Porter and Gaumnitz (1972) generate a Markowitz MV efficient set of portfolios from 140 stocks and first, second and third degree stochastic dominance tests are applied. They report that the most significant difference between the MV and stochastic dominance portfolios is the tendency for SD to eliminate low return - low variance portfolios. Although they conclude that the choice between stochastic dominance and mean variance models is not critical, the MV rule can lead highly risk-averse investors to make choices inconsistent with the maximization of expected utility.

In recent years, a number of studies have focused on the skewness of return distributions. Peiro (1999) finds that sample skewness can be used to reject symmetry in eight of nine stock return indices but that the results are sensitive to extreme outliers. He observes that two markets exhibit differences in location between negative and positive excess returns and three markets exhibit different dispersion. The dispersion occurred because of the higher frequency of negative excess returns. Taylor and Yoder (1999) examine load and no-load mutual fund performance during the 1987 market crash using stochastic dominance. They also find that during the period of extreme market stress, variance and beta are not suitable proxies for risk because of increased skewness in the return distribution.

Sun and Yan (2003) report that ex post stock returns are positively skewed, but the skewness is due to individual stocks rather than portfolios. They indicate that earlier studies did not use skewness in the construction of the portfolios. Using a polynomial goal programming method and considering skewness, they report portfolios can be formed with the positive skewness that investors prefer. Prakash, et al (2003) also use polynomial goal programming incorporating skewness to determine optimal portfolios. They report the incorporation of skewness into the analysis results in major changes in the optimal portfolio. They also suggest investors use skewness preferences in their trading activities. Interestingly, Jondeau and Rockinger (2003) find no statistical differences in the extreme tails for the return distributions of 20 international stock market indices.

Post (2003) uses the value-weighted average of all NYSE, AMEX and Nasdaq stocks in a Fama-French (1995) type of analysis using stochastic dominance. He finds the markets are inefficient in second order stochastic dominance and his results indicate that the inefficiency is both economically and statistically significant. However, he suggests return distributions vary over time and indicates his results may be influenced by the particular sampling period and return horizon.

It should be pointed out that many studies use stock market indices that are not actually tradable or marketable. For example, daily returns are used by Peiro (1999) from nine stock indices while Jondeau and Rockinger (2003) utilize 20 international stock market indices. Prakash, et al (2003) employ both weekly and monthly returns from 17 international stock market indices. The introduction of iShares, securities that mimic international indices, and that are tradable, allows an investigation of realizable return distributions and an examination of the shape of the empirical distribution function.

3. Data and methodology

Standard and Poor's Depository Receipts, (SPY) began trading in January 1993. SPYs track the S&P 500 Index, and are created and redeemed via "creation units" of 50,000 shares. In March 1996, seventeen exchange traded funds (ETFs) known as World Equity Benchmark Shares (WEBS) began trading on the American Stock Exchange. These funds, now known as iShares, are investment companies designed to track the Morgan Stanley Capital International (MSCI) foreign stock market indices. A listing of the country market indices, their ticker symbols and inception dates is presented in Table 1.

iShares were developed by Morgan Stanley, and Barclays Global Fund Advisors (BGFA) serves as each fund's investment advisor. BGFA uses either "replication" or "representative sampling" in the construction of portfolios designed to mimic a particular country's index. Replication means the fund contains essentially all of the securities in the relevant country index in relatively the same proportions as that country's index. Alternatively, representative sampling means the portfolio has a similar investment profile but not all securities in the index are included in the iShare. Changes in portfolio values come from both changes in the share prices in the portfolio and in changes in the exchange rate between U.S. dollars and the currency for a particular country. For consistency, foreign currencies are converted at the same time and at the same rate as used in the determination of each of the MSCI indices.

Although technically iShares are open-end index funds, the "creation units" cause their shares to trade in the secondary market just like ordinary shares. These "creation units" are in-kind deposits of portfolios of securities designed to represent a particular MSCI Index. If price differences between the underlying country index and the associated iShare emerge, arbitrage opportunities exist and large investors will quickly eliminate the price differences. These funds are not actively managed, turnover is virtually nonexistent, and operating expenses are low. In addition, only the usual brokerage fee is paid to buy or sell shares, as there are no front-end loads or deferred sales charges.

The daily return data were obtained from the Center for Research in Security Prices (CRSP) for the 17 iShares and SPY for the March 12, 1996 – December 31, 2003 period. Because the markets are impacted by different events and economic conditions, we consider three subperiods with events that are region specific. The first subperiod includes the Asian financial crisis that began in July of 1997 but ends before the Russian devaluation in October 1998. Secondly, a subperiod incorporating the U.S. technology bubble boom and concluding with the bursting of the bubble in March of 2000 is examined. Finally, the third subperiod goes from the bursting of the bubble to the end of 2003.

The 3-month U.S. T-bill rate and the Morgan Stanley Capital International index returns (MSCI) proxy the risk-free rate and the global market index, respectively. Mean-variance measures and several statistics derived from the CAPM including beta, Sharpe Ratio, Treynor's Index and Jensen's Alpha (referred to as CAPM statistics) are used along with stochastic dominance criteria to study the performance of the 18 closed-end funds.

Daily returns are used to compute the descriptive statistics described above for each fund in the sample. As will be seen in the next section, the measures of skewness, kurtosis, and the Jarque-Bera statistic indicate that none of the 18 return distributions are normal. Because the mean-variance criterion and the CAPM statistics are restricted to the first two moments of the data, important information contained in the higher moments is ignored and, hence, may result in inappropriate investment decisions.

To overcome the shortcomings associated with the MV and CAPM models and to investigate the performance of the entire distributions of the returns, we apply the Davidson and Duclos (2000) nonparametric stochastic dominance (DD) statistics to test for the dominance of any pair of the returns series.[iii]

Assume there are two return distributions, Y and Z, with N_y and N_z observations with the corresponding cumulative distribution functions to be F_y and F_z, respectively. Let $D_i^1 = F_i$ for i= y, z and let

$$D_i^k(x) = \int_{-\infty}^{x} D_i^{k-1}(y)dy$$

for any integer k > 1, Y is said to dominate Z stochastically at order k (denoted by $Y \succ_k Z$) if $D_y^k(x_i) \le D_z^k(x_i)$ for all i, with strictly significant inequality for some i. Modified from the Kolmogorov-Smirnov statistic, the DD statistic tests the null hypothesis, H_0, of the equality of $D_y^k(x) = D_z^k(x)$ is:

$$T^k(x) = \frac{\hat{D}_y^k(x) - \hat{D}_z^k(x)}{\sqrt{\hat{V}^k(x)}} \tag{1}$$

where:

$$\hat{V}^k(x) = \hat{V}_y^k(x) + \hat{V}_z^k(x) - 2\hat{V}_{y,z}^k(x)$$

$$\hat{D}_w^k(x) = \frac{1}{N(k-1)!}\sum_{i=1}^{N}(x-w_i)_+^{k-1}$$

$$\hat{V}_w^k(x) = \frac{1}{N}\left[\frac{1}{N((k-1)!)^2}\sum_{i=1}^{N}(x-w_i)_+^{2(k-1)} - \hat{D}_w^k(x)^2\right] \qquad \text{w = y or z,}$$

$$\hat{V}_{y,z}^k(x) = \frac{1}{N}\left[\frac{1}{N((k-1)!)^2}\sum_{i=1}^{N}(x-y_i)_+^{(k-1)}(x-z_i)_+^{(k-1)} - \hat{D}_y^k(x).\hat{D}_z^k(x)\right].$$

Note that $N_y = N_z = N$ as (x, y) are paired observations.

To test for stochastic dominance, H_0 should be examined for the full support, which is empirically impossible. A compromise is to examine a pre-designed finite number of values of x based on adopted multiple comparisons (Bishop, et al 1992). For any fixed values of x_1, x_2, …, x_m and their corresponding statistics $T^k(x_i)$ for k = 1, 2, 3 and i =1,…, m, the following hypotheses are investigated:

$H_0 : D_y^k(x_i) = D_z^k(x_i)$ for all i;

$H_A : D_y^k(x_i) \neq D_z^k(x_i)$ for some i; (2)

$H_{A1} : D_y^k(x_i) \geq D_z^k(x_i)$ for all i, $D_y^k(x_i) > D_z^k(x_i)$ for some i; and

H_{A2} :. $D_y^k(x_i) \leq D_z^k(x_i)$ for all i, $D_y^k(x_i) < D_z^k(x_i)$ for some i.

We note that in the above hypotheses, H_A is set to be exclusive of both H_{A1} and H_{A2}, which means that if the test accepts H_{A1} or H_{A2}, it will not classify to be H_A. Under the null hypothesis, Davidson and Duclos show that $T^k(x)$ is asymptotically distributed as the Studentized Maximum Modulus (SMM) distribution (Richmond 1982) to account for joint test size. For risk averters, the null hypothesis is rejected in favor of the alternative hypothesis A1 (fund Z dominates fund Y) if no DD statistic is significantly negative and at least some DD statistics are significantly positive[iv], or in favor of the alternative hypothesis A2 (fund Y dominates fund Z) if no DD statistic is significantly positive and at least some DD statistics are significantly negative.

The DD test compares the distributions at a finite number of grid points. Various studies examine the choice of grid points. For example, Barrett and Donald (2003) and Tse and Zhang (2004) show that an appropriate choice of 'm' for reasonably large samples ranges from 6 to 15. Too few grids will miss information of the distributions between any two consecutive grids and too many grids will violate the independence assumption required by the SMM distribution. To allow more detailed comparisons without violating the independence assumption, we follow Fong, et al (2005) to create 10 major partitions with 10 minor partitions within any two consecutive major partitions in each comparison, and to make the statistical inference based on the SMM distribution for k=10 and infinite degrees of freedom[v]. This allows the examination of the consistency of the magnitudes and the signs of the DD statistics between two consecutive major partitions.

4. Results

The descriptive statistics for the returns of the 18 closed-end funds for the entire period and three subperiods are reported in Table 2, Panels A through D. As can be seen, the means and standard deviations vary widely across iShares and over time. Though not shown here, two-sample t-tests indicate that some funds have significantly higher mean returns than others, and the F-statistic shows some standard deviations are significantly different at the 1% level. For example, the U.S. SPY exhibits a significantly higher mean and a significantly smaller standard deviation than Malaysia's EWM while Spain EWP exhibits a significantly higher mean but not significantly smaller standard deviation than Japan's EWJ. The characteristics of these four iShares are examined later in more detail.

However, the results also show that the return distributions are non-normal and exhibit both skewness and kurtosis and, hence, the distributions do not satisfy the normality requirements of the traditional CAPM measures. Specifically, for the entire

1996-2003 period, seven of the 18 skewness measures are significant at the 1% level (5 positive and 2 negative), and all kurtosis and Jarque-Bera (JB) measures are significant at the 1% level, highlighting the non-normality of the return distributions. Furthermore, the table highlights the ambiguity that is present both between and within the traditional CAPM measures. Although most betas are near 1.0, the range goes from 0.3325 for Austria (EWO) to 1.3035 for Sweden (EWD). The Sharpe ratios also exhibit wide variation. At the extremes are Spain (EWP) with a value of 0.0312 and Japan (EWJ) at -0.0093. Surprisingly, although their difference is large, the Sharpe Ratio Test shows no statistically significant difference between these funds. In fact, none of the Sharpe Ratios are significantly different.[vi] In addition, the evaluation issue is exacerbated in that the Treynor and Jensen measures suggest different rankings for the 18 closed-end funds.

The three subperiod results are presented in Panels B through D and allow the examination of the performance measures during different economic conditions. Again, we observe substantial differences among the distributions during different time periods. Skewness appears to be reduced over time as the number exhibiting significant skewness at the 1% level decreases from 11, to 8, to 3 (including MSCI) over the three subperiods. Changes in kurtosis and the JB statistic are much less with a maximum of 3 kurtosis measures being not significant in the first subperiod.

Unlike the CAPM criteria, stochastic dominance procedures allow us to determine whether one iShare stochastically dominates another based on the entire empirical return distribution. Table 3 shows the results of the Davidson and Duclos (2000) stochastic dominance procedures for the entire period. The DD test is a pairwise comparison of the iShares. The rows indicate which funds are dominated by the fund in the left hand column while the columns show which funds dominate the fund at the top of the column. For example, for the EWP (Spain) row, the ND under EWJ (Japan) means that EWP does not stochastically dominate EWJ, while the SSD under the EWH (Hong Kong) column indicates EWP dominates EWH in the sense of second order stochastic dominance. For the EWP column, the NDs indicate the 16 funds that do not dominate EWP, while the SSD associated with SPY shows that the SPY dominates EWP. The column on the far right shows the number of funds each individual fund dominates. We see that Spain, EWP, dominates five other funds and SPY dominates the most other funds at 14. Along the bottom row we see the number of funds each individual fund is dominated by. For example, Malaysia, EWM, is dominated by 13 other funds while four funds are not dominated by any other fund. We skip reporting the dominance relationship among each iShare for each of the sub-periods. However, we summarize the number of dominated iShares and number of dominant iShares of the entire period as well as each of the sub-periods in Table 4.

Interestingly, one instance of first order stochastic dominance is found. In the March 1996-June 1998 period before the Asian financial crisis, Spain, EWP, exhibited FSD over Japan, EWJ.[1] Thus, investors could have increased both their wealth and their utility by switching from Japan to Spain. This is an interesting finding, as most, if not all, prior studies find no first order stochastic dominance. However, we must conclude it is time-specific as the relation does not appear in any other period.

Bawa (1978) and Jarrow (1986) point out that if there is no first order stochastic dominance, investors cannot increase their wealth by switching from one fund to another, and no arbitrage opportunity exists. However, by considering second and third order stochastic dominance, we can determine whether investors could increase their utility by switching from one fund to another. In addition, the impact of the positive and negative portions of the return distributions is examined.

Table 3 reveals that SPY dominates 14 other closed-end funds in SSD while EWM, EWK, EWH, EWW and EWS do not stochastically dominate any of the other funds.[vii] Unlike the results from the Sharpe Ratio Test, the DD test reveals that SPY is statistically superior to 14 other funds. Hence, SPY is the best choice or at least one of the best choices among the 18 closed-end funds in the sense of second order stochastic dominance.

Hanoch and Levy (1969) indicate risk-averse investors will increase their utility but not necessarily their wealth by switching portfolios. In the present study, risk-averse

---

[1] As we skip reporting the dominance relationship among the iShares for each of the sub-periods, this finding could only be revealed in Table 6 Panel B.

investors will increase their utility by switching from these 14 other closed-end funds to SPY. The existence of second order stochastic dominance does not imply any arbitrage opportunity, and neither implies the failure of market efficiency nor market rationality. Thus, we conclude that although SPY does not significantly outperform most other funds from a wealth perspective, risk-averse investors prefer SPY as they will increase their utility by switching from 14 other funds to SPY. At the other extreme, all risk-averse investors holding EWM, EWK, EWH, EWW and EWS will increase their utility by switching to some other funds.

Four funds are selected for further analysis. Although the Sharpe Ratio indicates SPY (U.S.) and EWM (Malaysia) are not significantly different, the DD tests show that SPY dominates the most other funds (14) while EWM is dominated by the most other funds (13). In addition, the SSD shows a statistically significant difference between these two funds. For these reasons they are chosen for further analysis and we next examine the sources of the difference.[viii]

The comparison is depicted in Table 5 and Figure 1. From the figure, we see that for the entire period from 1996 to 2003, the empirical cumulative density function (CDF) of EWM is either greater than or equal to that of SPY in the entire negative return region, but the reverse is true over the positive return region. The first-order DD test (T1)[ix] shown in Table 5 and Figure 1 reveals that 27% of the return differences are significantly negative in the downside return region while 37% are significantly positive in the upside return region. Thus, SPY dominates EWM in the downside return region and EWM dominates SPY in the up-side return region in the sense of FSD. Second-order DD test (T2) and the third-order DD test (T3) are also examined. As expected, risk-averse investors fear the large negative returns associated with EWM more than they value the large positive returns. As shown, all SSD (T2) and TSD (T3) values are negative with 37% and 75% significant, respectively, over the entire period. Hence, we confirm that risk-averse investors will unambiguously prefer SPY to EWM.

When subperiods are considered, the first two subperiods are similar to the entire period, but the third subperiod (March 2002-December 2003) is different. A closer examination of Table 2, Panels B through D, reveals that SPY possesses a significantly higher mean in the first sub-period,[x] but EWM possesses a significantly higher mean in the second sub-period and a higher but insignificant mean in the third sub-period. On the other hand, SPY possesses a significantly smaller standard deviation in the first and second sub-periods, but an insignificantly smaller standard deviation in the third sub-period. In addition, SPY obtains a higher Sharpe Ratio in sub-periods 1 and 3 but not in sub-period 2.

Table 5 shows the stochastic dominance results between SPY and EWM does not change sign, and basically all three sub-periods draw a similar conclusion, that EWM dominates SPY in the up-side returns while SPY dominates to EWM in the downside returns in the sense of FSD. The dominance of SPY over EWM in SSD still remains the same as that in the entire period. The difference is that the stochastic dominance of SPY is significant in the first two sub-periods but becomes insignificant in the third sub-period. In this period, the DD stochastic dominance tests indicate SPY does not dominate EWM nor does EWM dominate SPY.

The second set of funds considered further is Spain EWP and Japan EWJ. Interestingly, the DD test indicates first order stochastic dominance of EWP over EWJ in the first subperiod. The FSD finding is important as most past studies find no evidence of first order dominance. The results for these two funds are presented in Table 6 and Figures 5 through 8. For the entire period, the CDFs and the first-order DD statistics (T1) show that EWP first-order dominates EWJ in the negative return region with marginal significance (2%), but EWJ dominates EWP (but not significant) in the positive region. As we use a conservative 10% cutoff point for the proportion of t-statistics to minimize type II error of finding dominance, and to avoid Almost Stochastic Dominance (Leshno and Levy 2002), 2% dominance implies that we cannot conclude first-order dominance of EWP over EWJ over the entire period.

However, in the first sub-period, Table 6 reveals that there are 11% (41% and 70%) first (second and third respectively)-order dominance of EWP over EWJ and Figure 6 confirms that these dominances are in the negative region. The first order dominance (though marginal) of EWP over EWJ implies that all investors with increasing utility will prefer EWP than EWJ. There is an arbitrage opportunity between EWP and EWJ such that all investors will increase both their wealth and their utility if they shift their investments from EWJ to EWP. However, Table 6 and Figures 7 and 8 show that the dominances (first, second and third orders) disappear in the following two sub-periods. This could be due to its exploitation after investors realize this arbitrage opportunity.

In addition we examine stochastic dominance in up-markets and down-markets by applying a regime shifting technique (see for example, Hamilton (1994). We use the MSCI index to classify the up-market and down-market regimes and to estimate the likelihood of being in an up-market or down-market on each day by applying the Hamiltonian regime switching approach. By using the regime switching technique, we find days with low likelihood of being a down market prevail in our first two sub-periods, while those with high likelihood of being down market are more pronounced within the third sub-period. Figure 9 A shows the long upward trend in the MSCI until the bursting of the technology bubble in early 2000, followed by the substantial decline until the end of the period. Figure 9 B shows the probability of a down market over the entire period. We see that early in the period, the data are dominated by a lower probability of a down market, while later in the period there is a high probability of a down market. Because the results of the up- and down-market regime changes are similar to the original analyses of the sub-periods, we do not report the results in this paper.

5. Conclusion

Recent improvements in statistical procedures allow a more rigorous assessment of different return distributions. Memmel (2003) modifies the procedure developed by Jobson and Korkie (1981) to test the statistical significance of the Sharpe Ratio. Davidson and Duclos (2000) provide a procedure for determining the statistical significance of stochastic dominance measures. The present study uses these refinements to compare the performance of 18 country market indices represented by iShares. Our examination considers their returns from their inception through 2003, and we find that, empirically, iShare returns exhibit both skewness and kurtosis and are not normally distributed.

The results show that even though there appear to be large differences among the Sharpe Ratios, the Memmel test indicates none of the differences are statistically significant. Thus, the ratios are indistinguishable on the first two moments, possibly due to the existence of skewness and kurtosis. Furthermore, the Treynor and Jensen measures provide conflicting rankings. This may be caused by the use of betas that may be biased due to volatile markets and non-normal return distributions. Variations in these measures over the subperiods support this view.

Previous stochastic dominance tests provide an overall assessment of return distributions. We extend the research by using the Davidson and Duclos (2000) stochastic dominance tests of significance and isolate the regions of statistical significance. Specifically, using four different iShares and three subperiods, we identify the existence for first-, second- and third-order stochastic dominance and the levels of significance. Although the results vary over time, stochastic dominance appears to be more robust than the CAPM in the ranking of the iShares.

Further research could include the impact of differing volatilities on stochastic dominance tests. This is an important issue and there are some studies on the impact of the volatilities on stochastic dominance (see, for example, Theorem 4 in Hadar and Russell (1971), Theorem 1' in Tesfatsion (1976) and Theorem 8 in Li and Wong (1999)) but so far, there is no study on the impact on stochastic dominance tests.

References

Anderson, G.., 1996. Nonparametric tests of stochastic dominance in income distributions. *Econometrica* 64, 1183 – 1193.

Anderson, G.J., 2004. Toward an empirical analysis of polarization. *Journal of Econometrics* 122, 1-26.

Barrett, G.. and Donald, S., 2003. Consistent tests for stochastic dominance. *Econometrica* 71, 71-104.

Bawa, V. S., 1978. Safety-first, stochastic dominance, and optimal portfolio choice. *Journal of Financial and Quantitative Analysis* 13, 255-271.

Bishop, J.A., Formly, J,P., and Thistle, P.D., 1992. Convergence of the south and non-south income distributions. *American Economic Review* 82(1), 262-272.

Davidson, R. and Duclos, J-Y., 2000. Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica* 68, 1435-1464.

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

Fong, W.M., Wong, W.K. and Lean, H.H., 2005. Stochastic dominance and the rationality of the momentum effect across markets. *Journal of Financial Markets* 8, 89-109.

Hadar J. and Russell W.R. 1971. Stochastic dominance and diversification. *Journal of Economic Theory* 3, 288-305.

Hamilton, J.D., 1994, *Time Series Analysis* (Princeton University Press, Princeton, NJ).

Hanoch, G. and Levy, H., 1969. The efficiency analysis of choices involving risk. *Review of Economic Studies* 36, 335-346.

Jarrow, R., 1986. The relationship between arbitrage and first order stochastic dominance. *Journal of Finance* 41, 915-921.

Jensen, M.C., 1969. Risk, the pricing of capital assets and the evaluation of investment portfolios. *Journal of Business* 42, 167-247.

Jobson, J.D. and Korkie, B., 1981. Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance* 36, 889-908.

Jondeau, E. and Rockinger, M., 2003. Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance* 10, 559-581.

Joy, O. M. and Porter, R.B., 1974. Stochastic dominance and mutual fund performance. *Journal of Financial and Quantitative Analysis* 9, 25-31.

Kaur, A., Rao, B.L.S.P. and Singh, H., 1994, Testing for second order stochastic dominance of two distributions. *Econometric Theory* 10, 849 – 866.

Kjetsaa, R. and Kieff, M., 2003. Stochastic dominance analysis of equity mutual fund performance. *American Business Review* 21, 1-8.

Lean, H.H., Wong, W.K. and Zhang, X.B., 2004. Size and power tests of various stochastic dominance tests. Working Paper, National University of Singapore.

Leshno, M. and Levy, H., 2002. Preferred by "All" and preferred by "Most" decision makers: almost stochastic dominance. *Management Science* 48, 1074-1085.

Levy, H. and Sarnat, M., 1970. Alternative efficiency criteria: an empirical analysis. *Journal of Finance* 25, 1153-1158.

Li, C.K. and Wong, W.K. 1999. A note on stochastic dominance for risk averters and risk takers. *RAIRO Recherche Operationnelle* 33, 509-524.

Linton, O., Maasoumi, E. and Whang, Y. J., 2002, Consistent testing for stochastic dominance: a subsampling approach, Working Paper, London School of Economics.

Markowitz, H. M., 1952. Portfolio selection. *Journal of Finance* 7, 77-91.

Memmel, C., 2003. Performance hypothesis testing with the Sharpe Ratio. *Finance Letters* 1, 21-23.

Meyer, J., 1977. Further applications of stochastic dominance to mutual fund performance. *Journal of Financial and Quantitative Analysis* 12, 235-242.

Peiro, A., 1999. Skewness in financial returns. *Journal of Banking and Finance* 23, 847-862.

Porter, R.B., 1973. An empirical comparison of stochastic dominance and mean-variance portfolio choice criteria. *Journal of Financial and Quantitative Analysis* 8, 587-608.

Porter, R.B. and Gaumnitz, J.E., 1972. Stochastic dominance vs. mean-variance portfolio analysis: an empirical evaluation. *American Economic Review* 62, 438-446.

Post, T., 2003. Empirical tests for stochastic dominance efficiency. *Journal of Finance* 58, 1905-1931.

Prakash, A.J., Chang, C.H. and Pactwa, T.E., 2003. Selecting a portfolio with skewness: recent evidence from US, European, and Latin American equity markets. *Journal of Banking and Finance* 27, 1375-1390.

Richmond, J., 1982. A general method for constructing simultaneous confidence Intervals. *Journal of the American Statistical Association* 77, 455-460.

Sharpe, W.F., 1964. Capital asset prices: theory of market equilibrium under conditions of risk. *Journal of Finance* 19, 452-442.

Sharpe, W.F., 1966. Mutual fund performance. *Journal of Business 39*, 119-138.

Stoline, M.R. and Ury, H.K., 1979. Tables of the studentized maximum modulus distribution and an application to multiple comparison among means. *Technometrics* 21, 87-93.

Sun, Q. and Yan, Y., 2003. Skewness persistence with optimal portfolio selection. *Journal of Banking and Finance* 27, 1111-1121.

Taylor, W.R.L. and Yoder, J.A., 1999. Load and no-load mutual fund dynamics during the 1987 Market Crash: A Stochastic dominance analysis. *Journal of Economics and Finance* 23, 255-265.

Tesfatsion, L. 1976. Stochastic dominance and maximization of expected utility. *Review of Economic Studies* 43, 301-315.

Treynor, J.L., 1965. How to rate management of investment funds. *Harvard Business*

Review 43, 63-75.

Tse, Y.K. and Zhang, X.B., 2004. A Monte Carlo investigation of some tests for stochastic dominance. *Journal of Statistical Computation and Simulation* 74, 361-378.

Table 1
Ticker symbols and the date of inception for the 17 iShares and the U.S. SPY.

Country Fund	Symbol	Inception Date
U.S. SPY	SPY	January 1993
Australia	EWA	March 1996
Austria	EWO	March 1996
Belgium	EWK	March 1996
Canada	EWC	March 1996
France	EWQ	March 1996
Germany	EWG	March 1996
Hong Kong	EWH	March 1996
Italy	EWI	March 1996
Japan	EWJ	March 1996
Malaysia	EWM	March 1996
Mexico	EWW	March 1996
Netherlands	EWN	March 1996
Singapore	EWS	March 1996
Spain	EWP	March 1996
Sweden	EWD	March 1996
Switzerland	EWL	March 1996
United Kingdom	EWU	March 1996

Subsequently, the following iShares have been added: South Korea EWY (May 2000), Taiwan EWT (June 2000), European Monetary Union EZU (July 2000), Brazil EWZ (July 2000), Pacific ex-Japan EPP (October 2001), South Africa EZA (February 2003), and Emerging Markets EEM (April 2003).

Table 2
Summary statistics, results of normality tests, and statistics derived from the CAPM using daily returns for the 17 iShares and SPY. The CAPM statistics use as the risk-free asset the 3-month U.S. T-bill rate and market portfolio is the MSCI World Index.

Panel A: March 1996 – December 2003

iShare	Daily Returns		Normality Tests			CAPM Statistics			
	mean	σ	skewness	kurtosis	Jarque-Bera	beta	Sharpe	Treynor	Jensen
EWA	0.0004	0.0155	-0.2443**	3.4273**	979.2866**	0.6212	0.0172	0.0004	0.0002
EWO	0.0003	0.0150	0.0674	3.6263**	1075.963**	0.3325	0.0116	0.0005	0.0001
EWK	0.0004	0.0211	4.1341**	175.0082**	2508137**	0.7959	0.0136	0.0004	0.0002
EWC	0.0005	0.0146	-0.2288**	2.3689**	475.6367**	0.8510	0.0260	0.0004	0.0003
EWQ	0.0004	0.0158	-0.0600	1.7160**	241.7922**	1.1317	0.0211	0.0003	0.0002
EWG	0.0003	0.0172	-0.0252	2.3339**	445.2621**	1.2541	0.0136	0.0002	0.0000
EWH	0.0002	0.0226	0.6324**	7.1731**	4334.903**	1.2758	0.0058	0.0001	-0.0000
EWI	0.0005	0.0162	-0.0553	2.0185**	333.8914**	0.9809	0.0252	0.0004	0.0003
EWJ	-0.0000	0.0175	0.5818**	3.9784**	1403.916**	0.9534	-0.0093	-0.0002	-0.0003
EWM	0.0000	0.0269	0.8762**	7.0148**	4271.62**	0.8396	-0.0017	-0.0000	-0.0001
EWW	0.0006	0.0223	0.1386*	7.8962**	5100.811**	1.2731	0.0217	0.0004	0.0003
EWN	0.0003	0.0163	-0.0710	1.7916**	263.9275**	1.1042	0.0100	0.0001	0.0000
EWS	-0.0000	0.0230	0.4436**	4.4773**	1702.231**	1.1365	-0.0065	-0.0001	-0.0003
EWP	0.0006	0.0163	0.1039	1.7431**	251.7784**	1.0149	0.0312	0.0005	0.0004
EWD	0.0005	0.0208	0.0870	2.1628**	384.6655**	1.3035	0.0182	0.0003	0.0002
EWL	0.0003	0.0160	0.0322	1.5554**	198.0173**	0.8414	0.0105	0.0002	0.0000
EWU	0.0004	0.0149	0.0451	1.5294**	191.775**	0.9595	0.0179	0.0003	0.0002
SPY	0.0004	0.0131	0.0105	2.2838**	426.202**	1.1288	0.0232	0.0003	0.0002
MSCI	0.0002	0.0099	-0.1106	2.6039**	558.0178**				

The risk-free asset is 3-month T-bill in US and market return is from the MSCI World Index. * means the statistics are significant at 5% level, and ** means the statistics are significant at 1% level.

Panel B: From date of inception to the Asian financial crisis: March 1996 – June 1998									
	mean	σ	skewness	kurtosis	Jarque-Bera	beta	Sharpe	Treynor	Jensen
EWA	0.000109	0.014976	-0.1288	4.6472**	520.8011**	0.9013	-0.0022	-0.0000	-0.0006
EWO	0.0004	0.0130	0.4274**	13.2296**	4225.384**	0.4683	0.0172	0.0005	-0.0000
EWK	0.0009	0.0120	0.0842	1.2921**	40.8212**	0.6922	0.0603	0.0010	0.0003
EWC	0.0007	0.0112	-0.6594**	3.6840**	368.1015**	0.8952	0.0494	0.0006	-0.0000
EWQ	0.0011	0.0124	-0.3335**	1.6118**	73.1585**	0.9466	0.0752	0.0010	0.0003
EWG	0.0011	0.0117	-0.3224**	2.0141**	107.5278**	0.9569	0.0806	0.0010	0.0003
EWH	-0.0005	0.0233	0.6716**	13.1656**	4210.561**	1.5173	-0.0271	-0.0004	-0.0016
EWI	0.0013	0.0155	-0.2220	2.2981**	131.7062**	1.0881	0.0756	0.0011	0.0005
EWJ	-0.0006	0.0168	0.6004**	4.9460**	622.7988**	1.1316	-0.0441	-0.0007	-0.0015
EWM	-0.0019	0.0290	0.8864**	7.3297**	1367.185**	1.5652	-0.0718	-0.0013	-0.0031
EWW	0.0006	0.0210	-0.6949**	10.1663**	2531.216**	1.6758	0.0211	0.0003	-0.0007
EWN	0.0011	0.0121	-0.0599	1.8191**	79.9022**	1.0034	0.0799	0.0010	0.0003
EWS	-0.0015	0.0238	0.6688**	6.5985**	1089.784**	1.2565	-0.0693	-0.0013	-0.0025
EWP	0.0016	0.0135	-0.0240	1.4135**	48.0912**	0.9650	0.1081	0.0015	0.0008
EWD	0.0011	0.0137	-0.0596	0.3304	2.9667	1.0321	0.0708	0.0009	0.0003
EWL	0.0008	0.0131	0.0370	0.4603*	5.224399	0.7720	0.0511	0.0009	0.0002
EWU	0.0011	0.0132	-0.2099	1.1262**	34.72782**	0.8095	0.0730	0.0012	0.0004
SPY	0.0011	0.010513	-0.5309**	5.3458	714.144**	1.1876	0.0887	0.0008	0.0002
MSCI	0.0008	0.006954	-0.6393**	4.5885	545.4772**				
The risk-free asset is 3-month T-bill in US and market return is from the MSCI World Index. * means the statistics are significant at 5% level, and ** means the statistics are significant at 1% level.									

Panel C: From the Asian financial crisis to the bursting of the technology bubble: July 1998 – February 2002									
	mean	σ	skewness	kurtosis	Jarque-Bera	beta	Sharpe	Treynor	Jensen
EWA	0.0003	0.0171	-0.2597**	2.9809**	350.965**	0.6279	0.0078	0.0002	0.0002
EWO	-0.0003	0.0169	0.0812	1.3457**	70.4254**	0.4125	-0.0245	-0.001	-0.0004
EWK	-0.0001	0.0259	4.7396**	164.7029**	1043314**	0.6564	-0.0091	-0.0004	-0.0002
EWC	0.0002	0.0169	-0.2707**	1.4784**	95.0148**	0.9722	0.0061	0.0001	0.0002
EWQ	0.0000	0.0162	-0.0160	1.8020**	124.5172**	1.0838	-0.0073	-0.0001	0.0000
EWG	-0.0002	0.0181	0.0693	2.4256**	226.2636**	1.2045	-0.0193	-0.0003	-0.0002
EWH	0.0006	0.0245	0.6765**	4.0064**	685.4874**	1.4764	0.01745	0.0003	0.0006
EWI	-0.0000	0.0169	0.0161	1.9629**	147.7375**	1.0177	-0.0130	-0.0002	-0.0000
EWJ	-0.0000	0.0186	0.7238**	4.3711**	812.7283**	1.0031	-0.011	-0.0002	-0.0000
EWM	0.0011	0.0305	0.8132**	4.5983**	911.9368**	0.9828	0.0316	0.0010	0.0011
EWW	0.0007	0.0258	0.3838**	6.1967**	1494.536**	1.4651	0.0231	0.0004	0.0008
EWN	-0.0002	0.0165	-0.0555	1.4196**	77.71904**	1.0011	-0.0224	-0.0004	-0.0002
EWS	0.0006	0.0243	0.4024**	3.6661**	540.0522**	1.3650	0.0209	0.0004	0.0007
EWP	-0.0002	0.0176	0.2155*	1.5923**	104.3126**	1.0521	-0.0154	-0.0003	-0.0001
EWD	-0.0000	0.0231	0.0918	1.7679**	121.0995**	1.4561	-0.0070	-0.0001	0.0000
EWL	-0.0002	0.0166	0.1339	1.7873**	125.2061**	0.7993	-0.0203	-0.0004	-0.0002
EWU	-0.0001	0.0155	0.1607	1.2145**	60.5004**	0.9688	-0.0170	-0.0003	-0.0001
SPY	0.0001	0.0140	0.0155	1.7705**	120.1994**	1.1524	-0.0004	-0.0000	0.0002
MSCI	-0.0000	0.0104	-0.2245*	1.9834**	158.5253**				
The risk-free asset is 3-month T-bill in US and market return is from the MSCI World Index. * means the statistics are significant at 5% level, and ** means the statistics are significant at 1% level.									

Panel D: From the technology bubble to end of data: March 2002 – December 2003									
	mean	σ	Skewness	kurtosis	Jarque-Bera	beta	Sharpe	Treynor	Jensen
EWA	0.0009	0.0127	-0.3004*	0.7485**	17.8094**	0.4968	0.0712	0.0018	0.0009
EWO	0.0013	0.0130	-0.2650*	1.8117**	68.8885**	0.1545	0.0986	0.0083	0.0013
EWK	0.0008	0.0194	0.3648**	2.5151**	132.5816**	1.0505	0.0401	0.0007	0.0009
EWC	0.0007	0.0134	0.3883**	2.5026**	132.7427**	0.6508	0.0529	0.0011	0.0008
EWQ	0.0005	0.0187	0.0222	0.8499**	14.0038**	1.2826	0.0260	0.0004	0.0006
EWG	0.0005	0.0207	-0.0276	0.8548**	14.1860**	1.4544	0.0245	0.0003	0.0006
EWH	0.0005	0.0169	0.1193	0.4976	5.8876	0.8801	0.0289	0.0006	0.0006
EWI	0.0007	0.0155	0.0126	1.8035**	62.8957**	0.8780	0.0451	0.0008	0.0008
EWJ	0.0007	0.0160	0.1212	0.7198**	11.1530**	0.8102	0.0402	0.0008	0.0007
EWM	0.0005	0.0127	-0.0562	0.8307**	13.5855**	0.3332	0.0382	0.0015	0.0005
EWW	0.0004	0.0156	0.2980*	1.6540**	59.75627**	0.8174	0.0203	0.0004	0.0004
EWN	0.0003	0.0200	-0.0345	1.0851**	22.85539**	1.3014	0.0111	0.0002	0.0003
EWS	0.0004	0.0188	0.0513	0.4361	3.8805	0.7546	0.0216	0.0005	0.0005
EWP	0.0009	0.0165	0.0289	1.7030**	56.1329**	0.9766	0.0525	0.0009	0.0009
EWD	0.0007	0.0231	0.1674	1.3921**	39.6317**	1.1879	0.0308	0.0006	0.0008
EWL	0.0006	0.0178	-0.1044	1.0956**	24.0513**	0.9338	0.0302	0.0006	0.0006
EWU	0.0005	0.0157	0.0587	2.1914**	93.1069**	1.0074	0.0283	0.0004	0.0005
SPY	0.0002	0.0141	0.3574**	1.3689**	46.1044**	1.0680	0.0096	0.0001	0.0002
MSCI	-0.0000	0.0119	0.2752*	1.6737**	60.0161**				
The risk-free asset is 3-month T-bill in US and market return is from the MSCI World Index. * means the statistics are significant at 5% level, and ** means the statistics are significant at 1% level.									

Table 3
Pairwise results of the Davidson-Duclos (DD) tests between iShares March 1996 – December 2003

	EWA	EWO	EWK	EWC	EWQ	EWG	EWH	EWI	EWJ	EWM	EWW	EWN	EWS	EWP	EWD	EWL	EWU	SPY	Dominates
EWA		ND	ND	ND	ND	ND	SSD	ND	ND	SSD	SSD	ND	SSD	ND	SSD	ND	ND	ND	5
EWO	ND		ND	ND	ND	ND	SSD	ND	ND	SSD	SSD	ND	SSD	ND	SSD	ND	ND	ND	5
EWK	ND	ND		ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	0
EWC	ND	ND	ND		ND	SSD	SSD	ND	SSD	SSD	SSD	ND	SSD	ND	SSD	ND	ND	ND	7
EWQ	ND	ND	ND	ND		ND	SSD	ND	ND	SSD	SSD	ND	SSD	ND	SSD	ND	ND	ND	5
EWG	ND	ND	ND	ND	ND		SSD	ND	ND	SSD	SSD	ND	SSD	ND	SSD	ND	ND	ND	5
EWH	ND	ND	ND	ND	ND	ND		ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	0
EWI	ND	ND	ND	ND	ND	ND	SSD		ND	SSD	SSD	ND	SSD	ND	SSD	ND	ND	ND	5
EWJ	ND	ND	ND	ND	ND	ND	SSD	ND		SSD	SSD	ND	SSD	ND	SSD	ND	ND	ND	5
EWM	ND	ND	ND	ND	ND	ND	ND	ND	ND		ND	ND	ND	ND	ND	ND	ND	ND	0
EWW	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND		ND	ND	ND	ND	ND	ND	ND	0
EWN	ND	ND	ND	ND	ND	ND	SSD	ND	ND	SSD	ND		SSD	ND	SSD	ND	ND	ND	4
EWS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND		ND	ND	ND	ND	ND	0
EWP	ND	ND	ND	ND	ND	ND	SSD	ND	ND	SSD	SSD	ND	SSD		SSD	ND	ND	ND	5
EWD	ND	ND	ND	ND	ND	ND	ND	ND	ND	SSD	ND	ND	ND	ND		ND	ND	ND	1
EWL	ND	ND	ND	ND	ND	ND	SSD	ND	ND	SSD	SSD	ND	SSD	ND	SSD		ND	ND	5
EWU	ND	ND	ND	ND	ND	SSD	SSD	ND	SSD	SSD	SSD	ND	SSD	ND	SSD	ND		ND	7
SPY	SSD	ND	ND	ND	SSD	SSD	SSD	SSD	SSD	SSD	SSD	SSD	SSD	SSD	SSD	SSD	SSD		14
Dominated by	1	0	0	0	1	3	12	1	3	13	11	1	12	1	12	1	1	0	

The results in this Table are read based on row versus column. For example, the first row EWA and the second column EWO means that EWA does not stochastically dominate EWO while the second row EWO and the first column EWA means that EWO does not stochastically dominate EWA. Alternatively, reading along the row SPY it can be seen that SPY dominates 14 other funds while reading down the SPY column shows that SPY is not dominated by any other fund.

Table 4
Country fund, ticker symbol and a summary of the stochastic dominance results

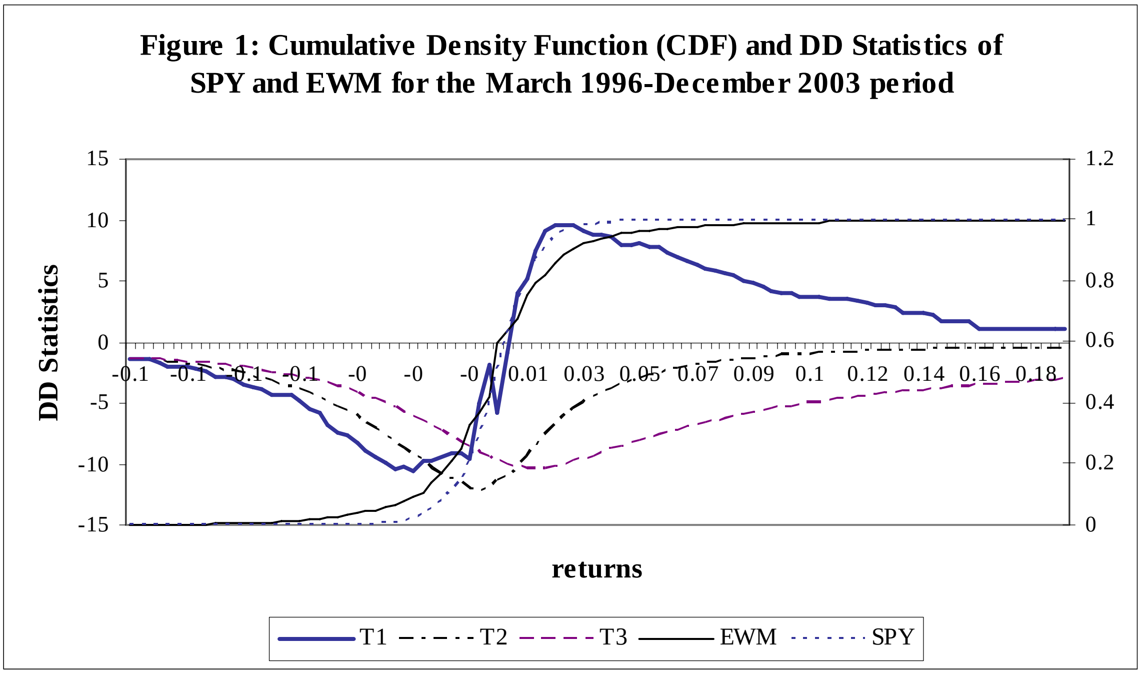
		Entire Period		**March 1996 – June 1998**		**July 1998 - February 2002**		**March 2002 – December 2003**	
Country Fund	**Symbol**	**Dominates**	**Dominated By**	**Dominates**	**Dominated By**	**Dominates**	**Dominated By**	**Dominates**	**Dominated By**
Australia	EWA	5	1	3	4	4	0	6	0
Austria	EWO	5	0	4	0	1	1	9	0
Belgium	EWK	0	0	5	0	4	0	0	0
Canada	EWC	7	0	6	0	4	0	4	0
France	EWQ	5	1	5	0	4	0	1	2
Germany	EWG	5	3	6	0	1	1	1	6
Hong Kong	EWH	0	12	0	13	1	11	1	8
Italy	EWI	5	1	3	1	4	0	5	0
Japan	EWJ	5	3	2	11*	3	0	3	2
Malaysia	EWM	0	13	0	15	0	16	3	0
Mexico	EWW	0	11	1	10	0	11	1	4
Netherlands	EWN	4	1	6	0	4	0	1	5
Singapore	EWS	0	12	0	14	1	10	0	9
Spain	EWP	5	1	5*	0	4	1	2	2
Sweden	EWD	1	12	5	0	1	0	0	15
Switzerland	EWL	5	1	5	0	4	1	1	1
United Kingdom	EWU	7	1	5	0	4	0	6	0
U.S. SPY	SPY	14	0	7	0	8	0	10	0

The values indicate the number of funds for each fund dominates and the number of funds that it is dominated by. For example, for the entire period the Australian fund dominates five other funds and is dominated by one other fund. * indicates significance of FSD between EWP and EWJ.

Table 5
Results of the Davidson-Duclos (DD) tests comparing SPY with EWM (SPY – EWM)

Panel A: March 1996 – December 2003			
	FSD	**SSD**	**TSD**
%DD+	37%	0	0
%DD-	27%	37%	75%
Panel B: From date of inception to the Asian financial crisis: March 1996 – June 1998			
%DD+	22%	0	0
%DD-	24%	50%	79%
Panel C: From the Asian financial crisis to the bursting of the technology bubble: July 1998 – February 2002			
%DD+	30%	0	0
%DD-	20%	28%	43%
Panel D: From the technology bubble to end of data: March 2002 – December 2003			
%DD+	0	0	0
%DD-	0	0	0

Note that %DD+ (%DD-) for FSD, SSD and TSD indicate that the percentage of the DD test as shown in (1) for k = 1, 2 and 3 respective to be (5%) significantly positive (negative). For example, the first two entries in the second columns are 37% and 27% which indicate that 37% of the first order DD test is significantly positive and 27% of the first order DD test is significantly negative.



Note: EWM is CDF of EWM, SPY is CPF of SPY, Tk is the k[th] order DD test statistic as shown in (1) for k = 1, 2 and 3.

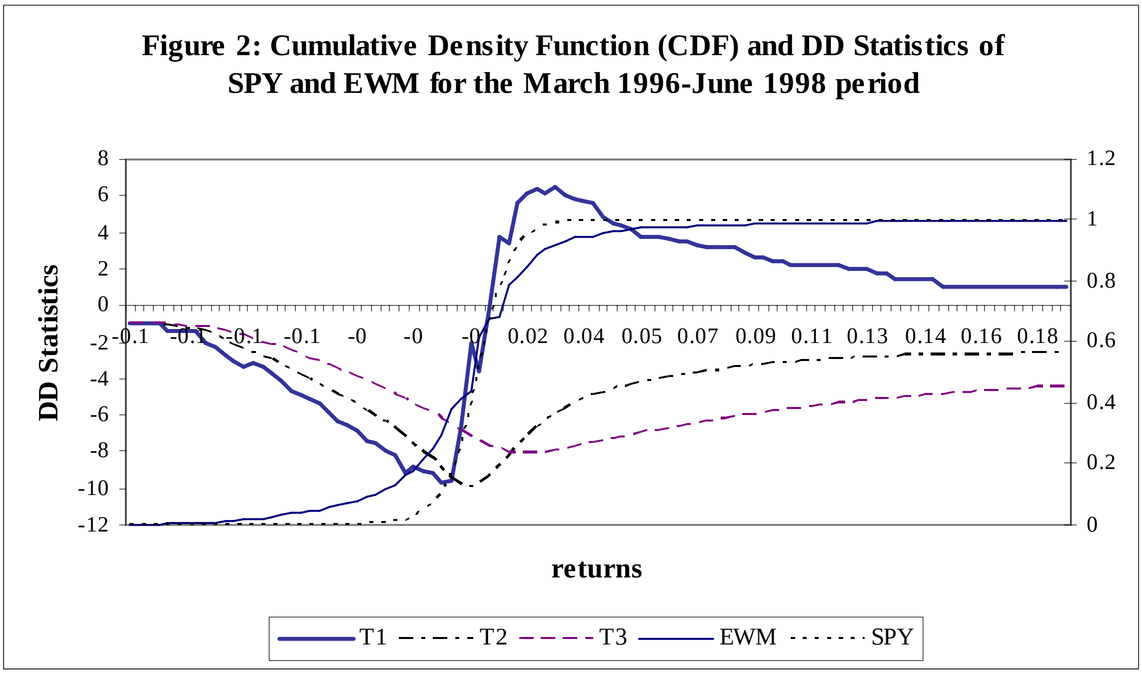


Figure 3: Cumulative Density Function (CDF) and DD Statistics of SPY and EWM for the July 1998--February 2002 period

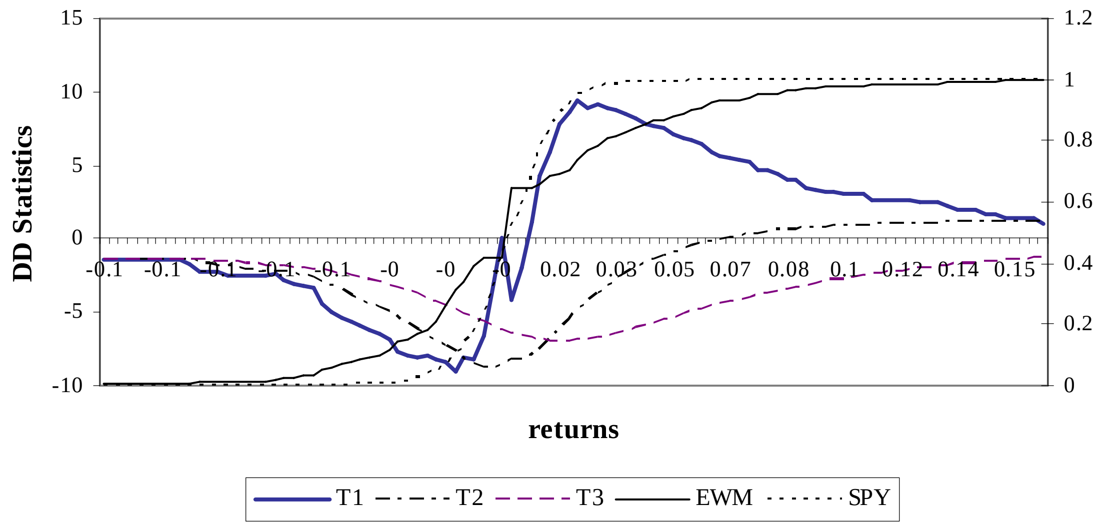


Figure 4: Cumulative Density Function (CDF) and DD Statistics of SPY and EWM for the March 2002--December 2003 period

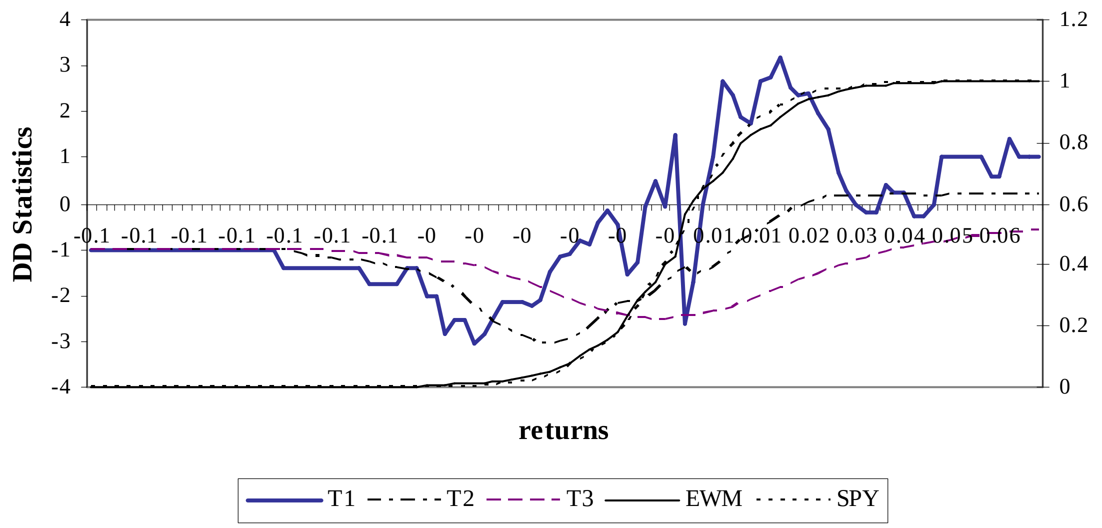


Table 6

Results of the Davidson-Duclos (DD) tests comparing EWP with EWJ (EWP – EWJ)

Panel A: March 1996 – December 2003			
	FSD	**SSD**	**TSD**
%DD+	0	0	0
%DD-	2%	0	0
Panel B: From date of inception to the Asian financial crisis: March 1996 – June 1998			
%DD+	0	0	0
%DD-	11%	41%	70%
Panel C: From the Asian financial crisis to the bursting of the technology bubble: July 1998 – February 2002			
%DD+	0	0	0
%DD-	0	0	0
Panel D: From the technology bubble to end of data: March 2002 – December 2003			
%DD+	0	0	0
%DD-	0	0	0

Note that %DD+ (%DD-) for FSD, SSD and TSD indicate that the percentage of the DD test as shown in (1) for k = 1, 2 and 3 respective to be (5%) significantly positive (negative). For example, the first two entries in the second columns are 0% and 2% which indicate that 0% of the first order DD test is significantly positive and 2% of the first order DD test is significantly negative.

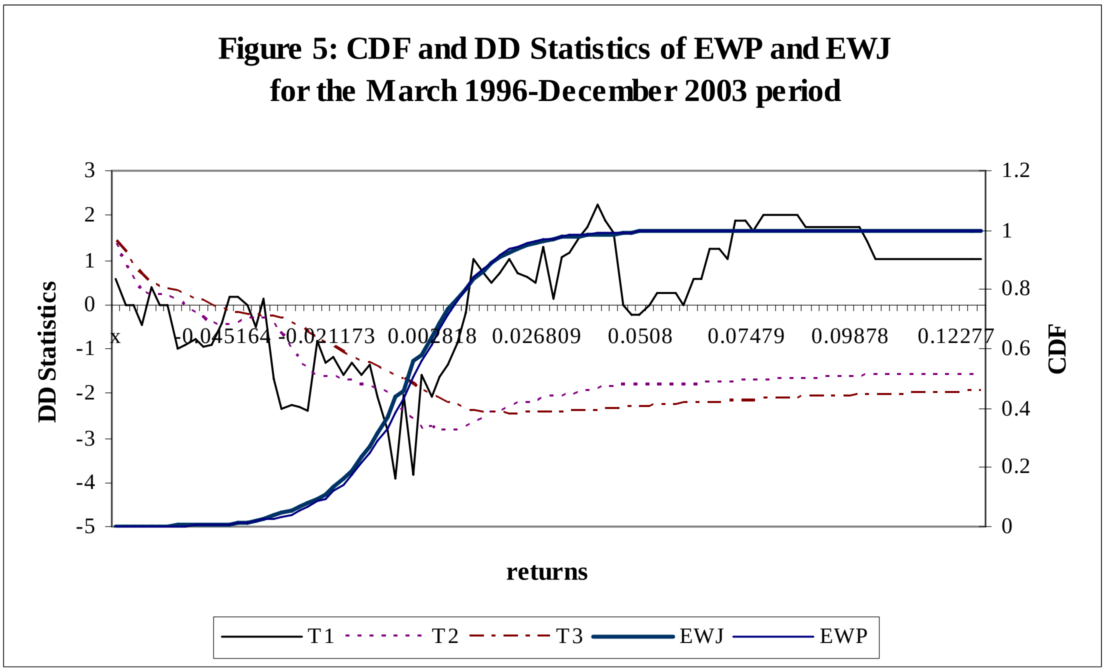


Figure 5: CDF and DD Statistics of EWP and EWJ
for the March 1996-December 2003 period
DD Statistics
CDF
returns
x
-0.045164
-0.021173
0.002818
0.026809
0.0508
0.07479
0.09878
0.12277
T1
T2
T3
EWJ
EWP

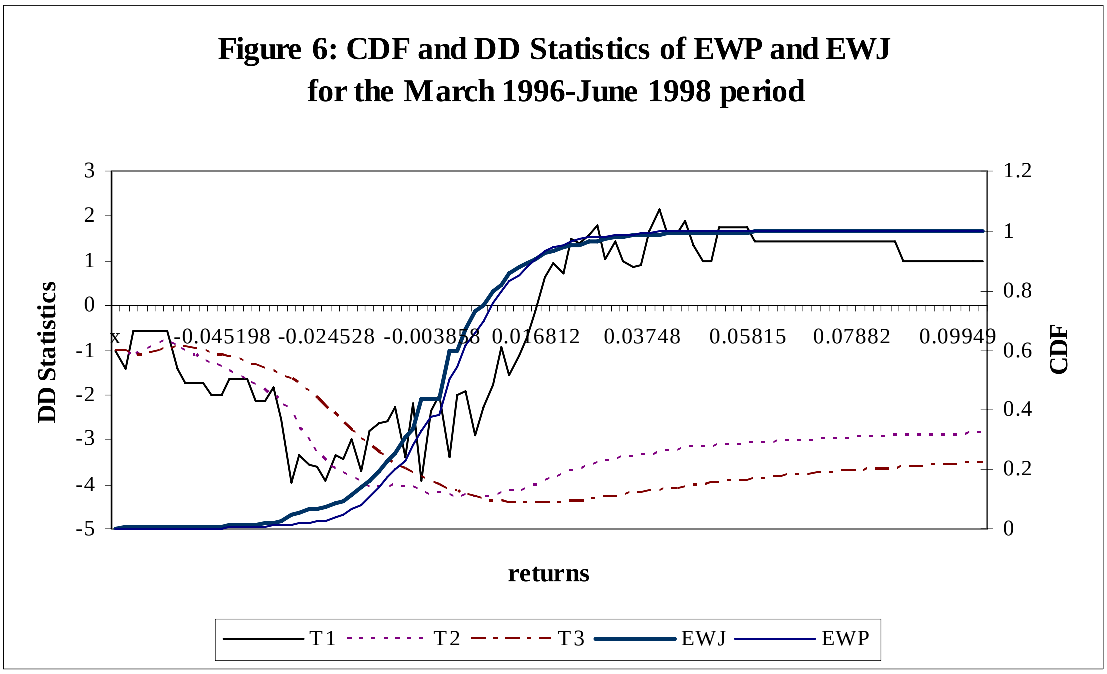


Figure 6: CDF and DD Statistics of EWP and EWJ
for the March 1996-June 1998 period
DD Statistics
CDF
returns
x
-0.045198
-0.024528
-0.003858
0.016812
0.03748
0.05815
0.07882
0.09949
T1
T2
T3
EWJ
EWP

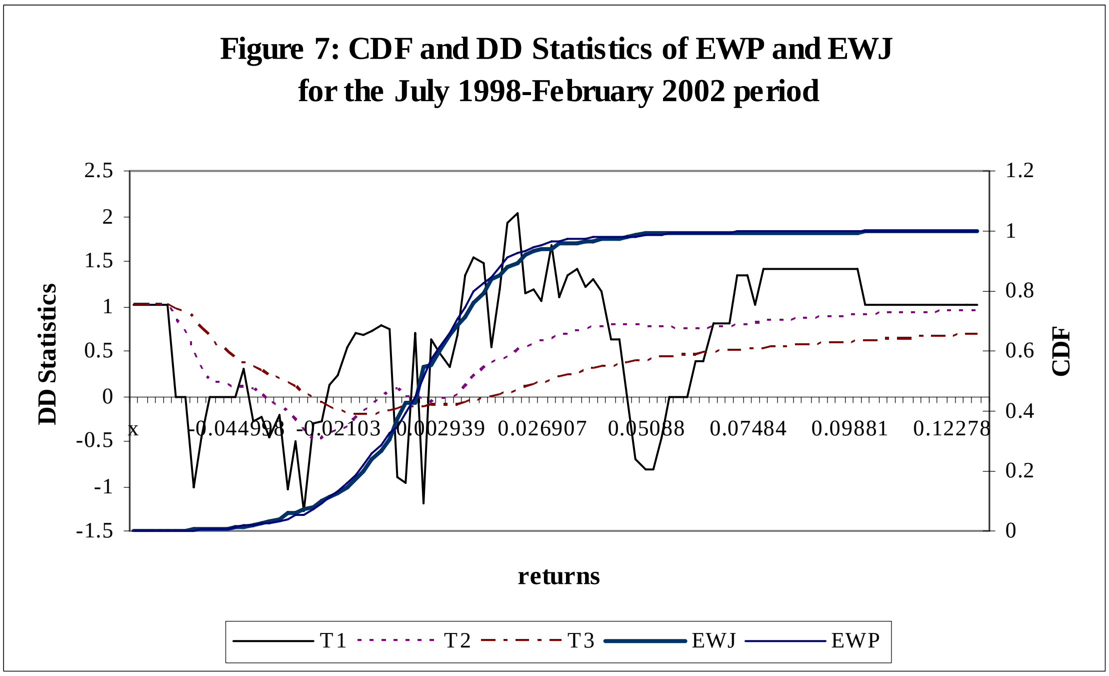


Figure 7: CDF and DD Statistics of EWP and EWJ
for the July 1998-February 2002 period
DD Statistics
CDF
returns
x
-0.044998
-0.02103
0.002939
0.026907
0.05088
0.07484
0.09881
0.12278
T1
T2
T3
EWJ
EWP

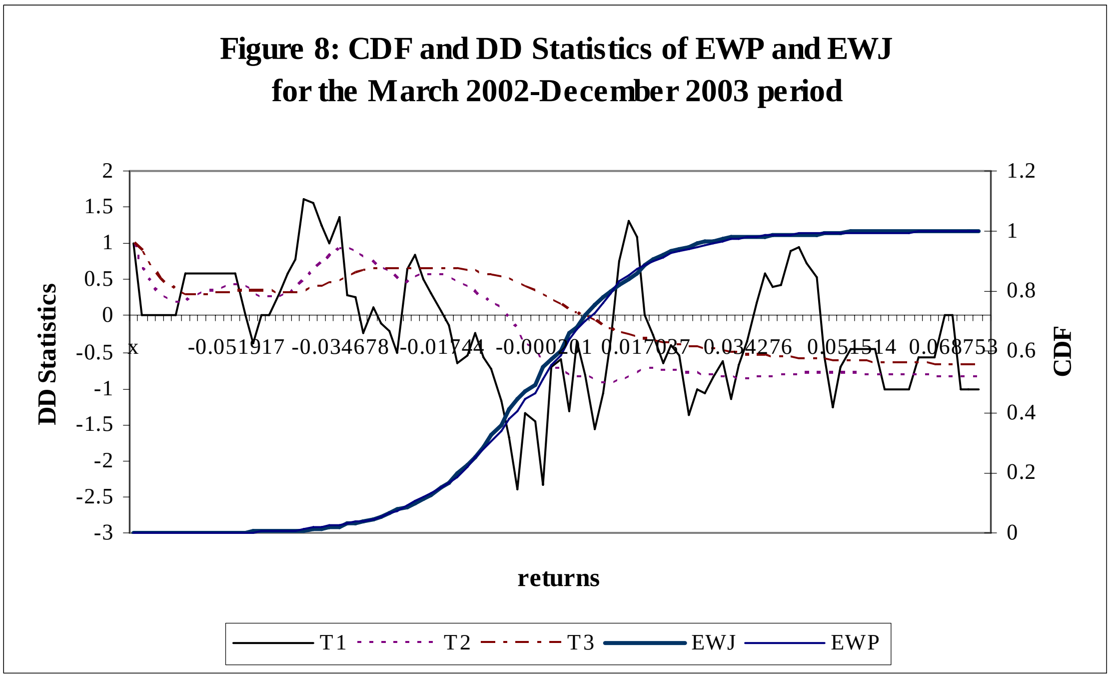


Figure 8: CDF and DD Statistics of EWP and EWJ
for the March 2002-December 2003 period
DD Statistics
CDF
returns
x
-0.051917
-0.034678
-0.01744
-0.000201
0.017037
0.034276
0.051514
0.068753
T1
T2
T3
EWJ
EWP

Figure 9 A: MSCI Index Plot

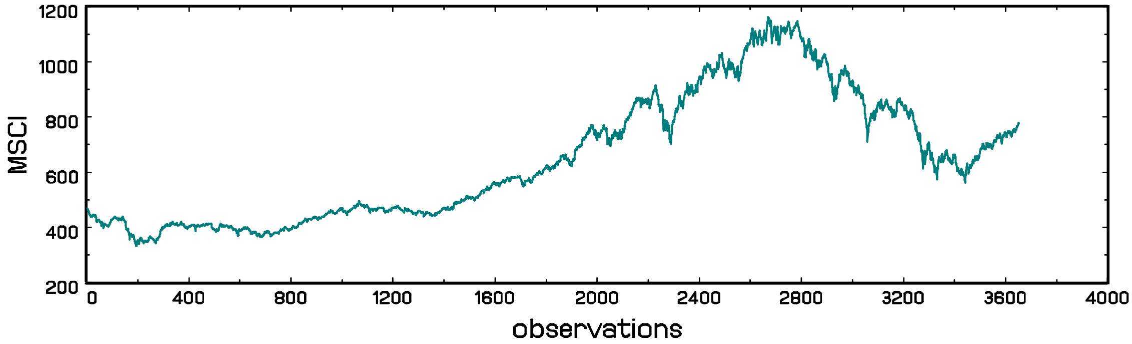
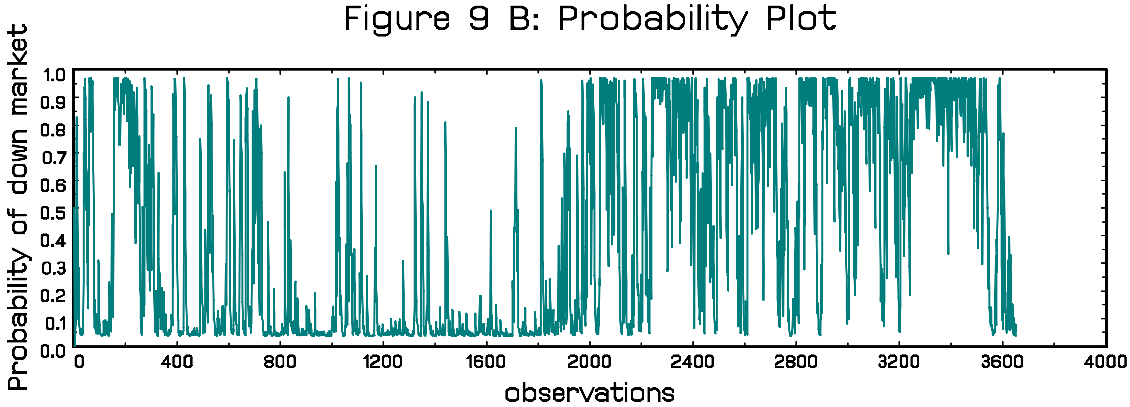


Figure 9 B: Probability Plot



Endnotes

[i] Recently Kaur, et al (1994), Barrett and Donald (2003), and Anderson (1996, 2004) have developed alternative SD tests. However, the DD test developed by Davidson and Duclos (2000) is found by Tse and Zhang (2003) and Lean, et al (2004), to be one of the least conservative but most powerful SD tests.

[ii] Prior to the introduction of the SPDRs, investors could only trade the S&P 500 through the Vanguard 500 Trust mutual fund. Purchases and sales were only made at the net asset value at the closing price once each day.

[iii]. Since we are using daily return data and DD test assumes that each pair of variable is independent and identically distributed (iid) we need to ensure that the DD inference is not problematic. To examine the robustness of our results we use Linton, et al (2002) test, which relaxes the iid assumption and in addition also examine the Barrett and Donald (2003) stochastic dominance test. As the conclusion from these stochastic dominance tests is similar we report only the DD results.

[iv] Davidson and Duclos (2000, pp. 1446) state that the null hypothesis can be rejected if *any* of the *t* statistics is significant with the wrong sign. To minimize type II error of finding dominance when there is none and to avoid from Almost Stochastic Dominance (Leshno and Levy 2002), we use a conservative 10% cutoff point for the proportion of t statistics in our statistical inference.

[v] Refer to Lean, et al (2004) for the reasoning. Critical values are: 3.691, 3.25 and 3.043 for 1%, 5% and 10% level of significance tabulated in Stoline and Ury (1979).

[vi] See Jobson and Korkie (1981) and Memmel (2003) for the development of the ratio test. The results of the test are available from the authors.

[vii] After the SPY, both the Canadian and United Kingdom iShares dominate 7 other funds.

[viii] We note that the comparison of any other pairs of funds (except EWP-EWJ) will either draw to same conclusion as the SPY-EWM comparison or no difference in the funds. Hence we only report the SPY-EWM and EWP-EWJ comparisons. Results of the other comparisons are available upon request.

[ix] Refer to equation (1) for the formula of Tk for k = 1, 2 and 3.

[x] Results of the significance tests are available from the authors.