RAFI Replication: Easier Done than Said?[1]

Paskalis Glabadanidis, Ivan Obaydin and Ralf Zurbruegg[2]

Abstract

We investigate whether adding fundamental indices to a portfolio provides increased diversification benefits. Our results show that equity investors who care only about portfolio mean and variance will benefit from including a fundamental index in their portfolios. This benefit is especially pronounced during periods of average stock market volatility. We also find that investors can construct a do-it-yourself buy-and-hold replicating portfolio that frequently outperforms the RAFI ETF out-of-sample.

Key Words: Fundamental indexes, portfolio diversification, mean-variance spanning.

JEL Classification: G11, G12.

Paskalis Glabadanidis obtained his PhD from the Olin School of Business at Washington University in St. Louis. He has worked at Koc University in Istanbul, Turkey before moving to the University of Adelaide. His research interest are in the fields of financial econometrics and applied asset pricing theory.

Ivan Obaydin is an Associate Lecturer at the University of Adelaide Business School with research interests in portfolio analysis.

Ralf Zurbruegg is a Professor of Finance within the University of Adelaide Business School. He has written widely on financial market issues, is the joint-editor of The International Journal of Managerial Finance, as well as a consultant to a number of fund management companies.

[1]We would like to thank Jason Hsu, Bing-Xuan Lin, Ben Marshall and David Michayluk for their insightful remarks on early drafts. Any remaining errors are our own.

[2]Corresponding Author: University of Adelaide Business School, 10 Pulteney Street, Adelaide, SA 5005, Australia. email: Ralf.Zurbrugg@adelaide.edu.au, tel: +61 8 8313 5535, fax: +61 8 8223 4782

1 Introduction

Portfolio diversification benefits arise when investors choose assets with imperfectly correlated returns. The identification of unique asset classes is therefore of significant importance to both institutional and retail investors in achieving risk reduction (Sharpe, 1978). Gallo and Lockwood (1999) demonstrate that investors who fail to correctly identify autonomous asset classes are exposed to unintentional style risk factors. Many studies over the years have examined extensively the benefits of investing in value versus growth stocks as well as small-cap versus large-cap stocks including, but not limited to, Fama and French (1992), Sharpe (1978), Sharpe (1981), and Teo and Woo (2004). The majority of these studies have largely considered only capitalization-weighted indices when examining the various equity styles. The main reason for using cap-weighted indices is the support for such indices in modern portfolio theory. Few studies to date have examined explicitly whether constructing an index using a different methodology would create a separate asset class with desirable diversification properties.

In the presence of efficient market prices, the capital asset pricing model (CAPM) of Sharpe (1964) identifies the market portfolio as the mean-variance efficient portfolio tangent to the capital market line as a portfolio constructed from market capitalization weights. However, where market price inefficiencies or market frictions exist, studies have shown that capitalization based weighting schemes lead to sub-optimal performance. Markowitz (2005), for example, demonstrates that a capitalization-weighted portfolio will not be mean-variance efficient when investors do not have access to unlimited borrowing. Prior studies such as Fama and French (1988), Jegadeesh (1990), and Lo and MacKinlay (1988), among others, present examples in support of the view that markets may not be completely efficient. In light of this evidence, Arnott et al. (2005) suggest that cap-weighted indices suffer unavoidable performance drags. The argument is that overvalued (undervalued) securities would be over- (under-) represented in the index and would hurt the performance of the index when the pricing error is corrected.

Given these concerns, there has been a substantial interest in alternative indexation methodologies. A recent alternative to market cap-weighted indices has been suggested by Arnott et al. (2005) and Arnott and West (2006). Instead of basing index weights on noisy and possibly inefficient market prices, these studies use the economic and fundamental footprint of index components, like cash flows, sales, book values, and dividend yields. Empirical evidence has shown that fundamentally weighted indices have tended to outperform their cap-weighted counterparts, on a risk-adjusted basis, over significant periods of time. For example, the cumulative five-year return of the US 1000 RAFI fundamentally weighted index ETF (US ticker: PRF) is 23.1% by the end of December 2010. At the same time, the cumulative five-year return of the S&P 500 over the same period is 11.99%. Furthermore, back test results reported by Arnott et al. (2005) show that fundamental indices have outperformed the S&P 500 index on average by 1.97% annually during the period 1962–2004. The superiority of fundamental indices' returns over cap-based indices' returns has achieved much attention in the recent literature. However, the usefulness of fundamental indices as investment vehicles has attracted some criticism as well. Perold (2007) and Kaplan (2008), for example, argue strongly against using fundamental indices based on the lack of theoretical support in their favor, their macro-inconsistency, as well as issues arising with re-balancing frequencies and index turnover.

Our objective in this article is two-fold. First, we investigate the merits of using fundamental indices in improving the risk-return profile of a well diversified portfolio. We focus explicitly on fundamental indices since their use may circumvent the potential downside of capitalization-weighted indices due to market price inefficiencies. We do not intend to reconcile the literature regarding the mean-variance efficiency of capitalization versus fundamentally weighted portfolios. Rather, we aim to answer the question of whether the mean-variance efficient frontier generated by a well-diversified portfolio (using only cap-weighted equity indices) can be improved substantially by the addition of a fundamental index. Our results will provide some mixed evidence on this front.

Second, we examine the ease of constructing a buy-and-hold replicating portfolio which mimics the returns of a fundamental index as closely as possible. We use up to six cap-based ETFs in constructing this replicating portfolio and report how well it performs relative to the fundamental index both in-sample as well as out-of-sample. Amazingly, we find that the buy-and-hold replicating portfolios outperform the fundamental index out-of-sample for a period of up to six months post-formation during four separate sub-periods. Surprisingly, this outperformance is not due to higher exposures to systematic risk as the portfolios' stock market betas are very close to one.

Data

In order to determine if fundamental based indices represent a separate equity class we construct a benchmark portfolio with exposures to several different equity styles. Our choice includes a large cap blend, small cap blend, large cap growth, small cap growth, large cap value, and small cap value equity ETFs, essentially, representing the major equity styles. We also deliberately choose publicly listed ETFs rather than equity indices in order to avoid the tracking error involved with replicating an index. The existence of the size and value effects has long been recognized in the finance literature and that is our motivation for our choice of equity styles. Furthermore, institutional investors often invest in these equity styles in order to satisfy their equity diversification mandates as pointed out by Arnott (1985), Hardy (2003), and Sharpe (1981), among others.

Specifically, we use the following iShares ETFs to represent cap-weighted size and style indices in our benchmark portfolio: large-cap (Ticker: IWB), large-cap growth (IWF), large-cap value (IWD), small-cap (IWM), small-cap growth (IWO), and small-cap value (IWN) In almost every style category, the two largest fund families are Vanguard and iShares, with Vanguard having generally greater net asset values but with lesser daily trading volume than iShares. The correlations between any of the six fund family pairs are always in excess of .98 and having examined a number of alternate fund providers, no discernable difference exists in the results reported in this article. In what follows, we will refer to these exchange traded funds as the benchmark assets unless otherwise indicated. The fundamental based index-tracking ETF we use is the US RAFI 1000 ETF (US Ticker: PRF) as it has an investment universe almost the same as the iShares Russell 1000 large cap value fund (IWD). This is important in order to ensure that any potential benefits that may appear from including the RAFI are not directly related to an expanded equity selection base.

We examine daily closing prices adjusted for cash dividends, stock dividends and stock splits starting from December 21, 2005 until December 31, 2010 for the fundamental based RAFI index ETF and six cap-weighted ETFs. Our start date represents the inception date for the traded RAFI ETF. In addition to performing our analysis during the entire five-year sample period, we also report our findings for three sub-periods. The first sub-period begins on December 21, 2005 and lasts until June 29, 2007. This sub-period coincides with a period of rising equity prices. The second sub-period includes the financial crisis and lasts from July 2, 2007 until February 27, 2009. During this sub-period, world financial markets experienced significant volatility and substantial losses. The erosion of portfolio diversification benefits during market downturns has been well documented in Ang and Chen (2002), Campbell et al. (2002), and You and Daigler (2010), among others. This finding has been dubbed as Murphy's law of diversification and is primarily due to an increase in the correlations between the returns of risky assets when stock market volatility is high. Hence, it is of interest to establish whether fundamental indices provide diversification benefits during a financial crisis. The last sub-period we consider begins on March 2, 2009 and lasts until December 31, 2010. This sub-period coincides with increasing equity prices.

We calculate daily simple returns based on the daily adjusted closing price level for each of the seven ETF price series. Table 1 reports descriptive statistics of the daily ETF returns. Note that the US RAFI ETF offers the highest (second highest) Sharpe ratio in the pre- (post-) financial crisis period. However, the RAFI fund turns out to be the worst performer of the group during the financial crisis bear market sub-period. Nevertheless, the RAFI's poor performance is not much different than the performance of the iShares Russell 1000 large cap value fund.

Insert Table 1 here.

Unconditional Mean-Variance Spanning Tests

We employ the mean-variance spanning framework of Huberman and Kandel (1987) to determine whether fundamental indices can be spanned by the components of a well-diversified benchmark portfolio. This approach tests formally whether the mean-variance frontier improves materially with the addition of a new asset. The tests depend largely on the improvement in the risk of the minimum variance portfolio and the improvement of the risk-return trade-off (Sharpe ratio) of the tangent portfolio of the extended mean-variance frontier. We refer the interested reader to Appendix A for the technical details.

In order to examine risk factors we estimate the four-factor model of Carhart (1997) the replicating portfolio *and* US 1000 RAFI fund:

$$r_{p,t} = \alpha + \beta_{MKT} r_{m,t} + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \beta_{UMD} UMD_t + u_t, \quad (1)$$

where $r_{p,t}$ is the daily return time series of either the replicating portfolio or the US RAFI fund, $r_{m,t}$ is the excess return on the market, SMB_t is the (small minus big) Fama-French size factor, HML_t is the (high minus low) Fama and French valuation factor, and UMD_t is the (up minus down) momentum factor.[3]

The motivation for the inclusion of a momentum factor rests with the possibility that any security mispricing will distort the returns on capitalization-weighted indices. Arnott et al. (2005), Perold (2007), as well as Perold and Sharpe (1988) find that cap-weighted indices tend to outperform fundamental indices during periods of positive serial return correlation and vice versa.

The second regression we perform with the RAFI index return is:

$$r_{p,t} = \alpha + \sum_{i=1}^{i=6} \beta_i r_{i,t} + \epsilon_t, \quad (2)$$

where $r_{i,t}$ is the simple daily return of the benchmark ETF i. Intuitively, the spanning tests check whether the intercept α in the above regression is statistically different from zero *and* whether the factor loadings sum up to 1. If these conditions hold, then we can replicate exactly the returns of the fundamental index by constructing a self-financing portfolio of the six ETFs.

Empirical Findings

Portfolio diversification benefits are an outcome of securities being less than perfectly correlated with one another. The lower the correlation between securities, the greater the achievable diversification benefits in terms of risk reduction that can be achieved. It is therefore worth providing some commentary on how the fundamental index exchange traded fund is correlated with the remaining cap-weighted ETFs in our benchmark portfolio, in order to gain some preliminary insight into the level of diversification benefits that may be attainable.

Examining pairwise correlations we note that the US RAFI 1000 fundamental index ETF return has a high correlation with the Russell 1000 fund return, ranging from a minimum of 0.9596 in the first sub-period to a high of 0.9825 during the financial crises period. Based on this evidence alone, the addition of a fundamental

[3] We are grateful to Ken French for providing the data on the factor returns.

index ETF might seem not to provide substantial diversification benefits, in terms of risk reduction, for a portfolio already having exposure to a broad market cap-weighted fund.

Although the correlations are less in magnitude, a similar result is also obtained when examining correlations between the RAFI and Russell 1000 ETFs against other ETF returns over the different sub-periods. These correlations are always above 0.88, with a peak during the financial crises sub-period. What is interesting to note, however, is that the correlations of the other ETFs are generally smaller with the RAFI returns, when compared to the Russell 1000 returns. Moreover, as would be expected in the period following the financial crisis, correlations between the RAFI ETF return and other sample fund returns decline. However, this is not the case for the pairwise correlations between any of the cap-weighted ETF returns. Only in relation to the correlations with the RAFI do we see a reduction after the crises, potentially hinting at the prospect that the RAFI ETF index may yet provide diversification benefits after we include it in our benchmark portfolio.

We now turn to our mean-variance spanning test results for the three sub-periods which we present in Table 2. Panel A reports the regression results which form the basis of the spanning tests tabulated in Panel B. Intriguingly, the large-cap ETFs have statistically significant coefficients before and during the financial crisis, but not in the final sub-period following the crisis. At the same time, the small-cap ETFs have insignificant coefficients outside of the crisis period, while being highly significant during the crisis sub-period.

Insert Table 2 here.

Before and after the financial crisis we find overwhelming evidence against the null hypothesis of mean-variance spanning at any conventional level of significance. This implies that the addition of a fundamental index would indeed improve the mean-variance frontier of our benchmark portfolio. During the crisis sub-period we find that there is slightly less than a 3% chance that the benchmark ETF returns can span the RAFI ETF return. However, most conventional levels of statistical significance will point towards a rejection of the null hypothesis during the crisis sub-period as well.

Next, we investigate whether these rejections are due to a material improvement in the tangent or the minimum variance portfolio of the expanded mean-variance frontier. Panel C of Table 2 reports the results of the step-down tests[4]. The results are quite intriguing in that there appears to be no statistically significant improvement in the tangent portfolio following the addition of the RAFI ETF across all sub-periods. This particular result would, though, be congruent with the earlier discussion on noting high correlations. However, the minimum variance portfolio shows a highly significant improvement after we add the RAFI ETF. The statistical evidence against the null hypothesis that the minimum variance portfolio (MVP) does not improve is overwhelming. Hence, we conclude that the rejection of the joint hypothesis is largely due to the rejection of the hypothesis that the MVP does not improve. The practical implications of this are that the RAFI

[4]In the notation of Appendix A, the null hypothesis of $\hat{\alpha} = 0_N$ refers to the tangent portfolio while the null hypothesis of $\hat{\delta} = 0_N$ refers to the minimum variance portfolio.

ETF can serve a useful purpose in designing low risk portfolios, but does not seem to be able to assist in the risk/return trade-off for other non-minimum variance portfolios.

If we now focus our attention on the Carhart (1997) four-factor loadings of the RAFI ETF and the six benchmark ETFs to understand the inherent risk structure within the fundamental ETF, we observe in Table 3 that the market beta of the benchmark large-cap blend ETF (IWB) is higher than the market beta of the RAFI ETF. Also, both ETFs tend to have negative loadings on the SMB factor, which are of similar magnitude. The major differences between the IWB and the RAFI ETF have to do with the value HML and momentum UMD factor loadings. The HML factor loading of the PRF ETF is positive, whilst negative for IWB. The UMD factor loadings of both ETFs are negative and statistically significant but the RAFI ETF appears to be more contrarian than the IWB ETF. These findings hold across all the sub-periods except the post-crisis period where the IWB UMD loading becomes positive. These factor loadings we note are also not too dissimilar to Arnott et al (2005), albeit for a different sample period.

Insert Table 3 here.

Next, we report the mean, standard deviation, and Sharpe ratio of the global minimum variance portfolio and tangency portfolio for the three sub-periods in Table 4. In the first sub-period we note that the Sharpe ratio increases both for the minimum variance portfolio and the tangent portfolio after we add the fundamental index ETF. Another interesting finding during the crisis sub-period is that adding the fundamental index ETF actually reduces the Sharpe ratio of both the minimum variance and the tangent portfolio. In the last sub-period following the crisis the results are mixed since adding the RAFI ETF improves the Sharpe ratio of the tangent portfolio but decreases the Sharpe ratio of the minimum variance portfolio. We do emphasize that these are simply in-sample findings and that the out-of-sample performance of these portfolios may be quite different. We address this issue in the next section.

Insert Table 4 here.

Rolling Window Mean-Variance Spanning Test

To determine the robustness of our results, we run spanning tests on a rolling window of data. We use a one-year time window and roll it forward on a daily basis. This approach lets us examine whether we are able to span the returns of the fundamental index going forward as new data becomes available.

Figure 1 plots the probability values of the LM test of mean-variance spanning which are largely in agreement with the results presented in the previous section [5]. Note the spike in the line during the fall of 2008 and the gradual decrease in the spring of 2009. This suggests that during the financial crisis an investor is unable to improve diversification benefits by including the RAFI ETF in her portfolio. In fact, the

[5] The results are not dissimilar if Wald or LR tests are conducted instead.

statistical evidence during this sub-period suggests strongly that the fundamental index does not improve on the mean-variance frontier of the six benchmark ETFs. However, the rolling window test results before and after the financial crisis do point in favor of including the fundamental index in the benchmark portfolio.

Insert Figure 1 here.

Furthermore, there is a considerable amount of persistence in the spanning test probability values. This implies that an investor may be able to forecast any future diversification benefits out-of-sample. Nevertheless, it is also clear that the ability to span an asset can change quickly and dramatically. Given these findings, an investor should be able to replicate the mean-variance structure of fundamental indices out-of-sample. We explore this further in the next section.

Insert Figure 2 here.

The above mean-variance spanning test unfortunately does not specifically identify whether failure to span is due to a significant improvement in the tangent or the minimum variance portfolio, following the inclusion of the RAFI. Separating the test into two parts, along the lines of Kan and Zhou (2012), we report two separate series of probability values of the LM test in Figure 2. The results are very intriguing. First, there is no statistical evidence that the tangent portfolio improves following the addition of the RAFI index. Second, there is very strong statistical evidence that the minimum variance portfolio improves significantly before and after the bear market during the financial crisis sub-period. The inability to improve the MVP during the financial crises is to be expected, given the previously noted increased correlation levels the RAFI ETF returns shared with the other ETFs during this period.

RAFI Replicating Portfolios

In this section we apply one of the replicating portfolio algorithms of Glabadanidis (2011) to test the performance of the benchmark ETF replicating portfolio versus the RAFI ETF. In each sub-period we use approximately six months worth of daily return data to estimate all the variances and covariances that we need. Then we use a linear regression, with up to six of the benchmark ETFs, to calculate the replicating portfolio weights and returns. At every stage the linear regression is as follows:

$$r_{p,t} = \alpha + \sum_{i=1}^{i=N_{ETF}} \beta_i r_{ETF_i,t} + \nu_t \tag{3}$$

where $r_{p,t}$ is the simple excess return of the RAFI ETF, $r_{ETF_i,t}$ is the simple excess return of the i-th ETF, and N_{ETF} is the current number of ETFs in the replicating portfolio. The regression slope coefficients β_i represent the optimal replicating portfolio weights.

Table 5 presents the results for the optimal replicating portfolio weights using four different sub-samples. We use a step-wise regression approach to determine how we select which ETF to add sequentially. The first ETF added to the portfolio is the one that leads to the highest goodness of fit, R^2, for the regression. The next ETF we add is the one that results in the next biggest improvement in fit, and so on until all six ETFs are added. From the resulting portfolios we construct buy and hold returns, both in sample as well as out-of-sample, for a period of six months. Focusing our investigation on a buy and hold strategy allows us to avoid nuances in dealing with transaction costs, plus sets a benchmark for potentially more complex approaches that could also be adopted.

Insert Table 5 here.

Overall, the portfolio weights show remarkable stability after the first two or three ETFs are included and the remaining cash position is fairly close to zero in all sub-samples. Most portfolio weights are positive and less than 100% (with one exception in the last sub-sample with only one ETF in the replicating portfolio). The small-cap ETFs (IWO, IWM, and IWN) have small negative weights which are not too excessive and are not surprising given the results reported in Table 3.

Panel A of Table 5 reports the replicating portfolio weights during the first six months after the RAFI ETF was listed. The weights are quite stable with the lion's share being devoted to a large-cap ETF (IWB) and large-cap value (IWD). This is not an unexpected result given that fundamental indices apply similar logic to picking large-cap value stocks. Small caps (IWM) and, to a lesser extent, small-cap growth receive a negative weight which is consistent with the RAFI ETF loading on large-cap value stocks. Large-cap growth stocks receive a negligibly small weight of less than one half of one percent.

Panel B of Table 5 presents the replicating portfolio weights using data from the first half of 2007. Large-cap and large-cap growth stocks again dominate the replicating portfolio with the large-cap value ETF (IWD) having almost twice the weight of the large-cap blend ETF (IWB) by the time all six ETFs are included in the replicating portfolio. The small-cap value ETF (IWN) has a fairly small negative weight and the small-cap growth ETF (IWO) has a negligibly small weight of one tenth of one percent. Note that the cash balance in this sub-sample is higher than the cash balance in the previous sub-sample.

Panel C of Table 5 presents the replicating portfolio weights using data for the six months immediately preceding the onset of the global financial crisis. The weights on all ETFs are remarkably stable after we have included the first three ETFs. The small-cap growth ETF again retains a negative weight while most of the portfolio is again invested in large-cap, large-cap value and small-cap value stocks. The cash position here turns negative but never exceeds 5% of the portfolio value, reflecting borrowing at the daily risk-free rate and investing more than its own capital in the replicating portfolio.

Finally, in Panel D we report the optimal replicating portfolio weights using six months of data following the 5-year market low achieved on March 3, 2009. In this case, we observe a larger amount of borrowing at the risk-free rate though this never exceeds 17% of the portfolio value. The portfolio is again dominated by the large-cap value and large-cap blend ETF. However, large-cap growth and small-cap value have non-negligible positive weights, while small-cap growth has a significant short position (although never exceeding 30% of the portfolio value).

We now move on to examining the performance of the RAFI replicating portfolios. Table 6 reports these results for where there are one to all six ETFs added to the portfolio, and for all four separate sub-periods. The in-sample periods coincide with the sample periods used to construct the portfolio weights reported in the previous exhibit. The out-of-sample periods contain the subsequent six months of daily data.

Insert Table 6 here.

Panel A reports the results for 2006. Note that the replicating portfolio lags the RAFI index in sample by as much as 229 basis points while its performance is roughly on par with the RAFI out-of-sample with a maximum shortfall of 47 basis points up to a small outperformance of 12 basis points. We should stress that this small outperformance is not due to the replicating portfolio taking on too much systematic risk. This can be verified by comparing the market betas in sample *and* out-of-sample.

In Panel B we present the results for 2007 where, in contrast with 2006, the replicating portfolio beats the RAFI in sample by more than it does out-of-sample. Yet, remarkably, the correlation between the RAFI return and the replicating portfolio return exceeds 99% out-of-sample when we use all six ETFs. Note again, that the market beta of the six different replicating portfolios out-of-sample are less than the respective in-sample market betas.

The most interesting sub-period to us is the one where the out-of-sample period includes the severe bear market conditions during the fall of 2008 up to and including the spring of 2009. Panel C reports the results for that particular sub-period. Here the tracking error of the replicating portfolio out-of-sample exceeds the in-sample tracking error which is to be expected. However, our buy-and-hold replicating portfolio still beats the RAFI by between 50 and 206 basis points during the six months out-of-sample. This outperformance was achieved despite having the same amount of market beta risk as contained in the broad stock market proxy.

Finally, we turn to the bull market period following the low equity prices of early March 2009. The in-sample period ends in early September 2009 while the out-of-sample period continues into early March of 2010. The in-sample performance demonstrates that the RAFI considerably outperforms the replicating portfolio. Nevertheless, the RAFI ETF is beaten again by our buy-and-hold replicating portfolio by between 75 and 126 basis points in the six month out-of-sample period.

Conclusion

Our mean-variance spanning results clearly indicate diversification benefits tangibly exist from including a fundamental index within an investor's portfolio. However, these benefits mostly relate to the improvement in the minimum variance portfolio. We do not detect a statistically significant improvement in the tangent portfolio of the extended mean-variance frontier.

Our results also show clearly the out-of-sample power to replicate and outperform the fundamental index returns. This can be achieved without having to load up on systematic beta risk. Nevertheless, we need to caution the reader that this outperformance can only be detected during the past five years for which the RAFI ETF has been publicly listed. Also, our analysis relies on constructing a buy-and-hold portfolio, avoiding issues relating to re-balancing frequencies and transaction costs. However, it would be of interest to consider constructing replicating portfolios with fixed weights which are re-balanced at some frequency (i.e., daily, monthly or quarterly) to investigate how the results might change in the presence of transaction costs. This we leave for future work.

Appendix A. Mean-Variance Spanning Tests

Let R_1 be a $T \times K$ matrix of realized returns of the K benchmark assets over T periods and R_2 be a $T \times N$ matrix of realized returns of the N test assets. The following linear regression is useful in performing tests of mean-variance spanning:

$$R_{2,t} = \alpha + R_{1,t}\beta + \epsilon_t, \tag{4}$$

where β is a $K \times N$ matrix of factor loadings of the test assets onto the benchmark assets. In matrix form, the above regression takes on the following representation,

$$R_2 = XB + E, \tag{5}$$

where X is a $T \times (K+1)$ matrix with a typical row of $[1, R'_{1,t}]$ and E is a $T \times N$ matrix with ϵ'_t as a typical row. The maximum likelihood estimates of B and $\Sigma = \text{var}(\epsilon)$ are

$$\hat{B} = (X'X)^{-1}(X'R_2), \tag{6}$$

$$\hat{\Sigma} = \frac{1}{T}(R_2 - X\hat{B})'(R_2 - X\hat{B}). \tag{7}$$

Mean variance spanning tests involve testing the following restriction, both jointly and independently, on equation (4) in order to infer the existence of diversification benefits.

$$\alpha = 0_N, \quad \beta 1_K = 1_N. \tag{8}$$

where 1_N is a N-vector of ones.

Under these restrictions there is no improvement in the Sharpe ratio of the tangency portfolio ($\hat{\alpha} = 0$) *and* no improvement in the global minimum variance portfolio ($\hat{\beta}1_K = 1_N$). If there is no improvement in the efficient frontier, then the returns of the test assets can be perfectly replicated with a fully invested portfolio of the benchmark assets. Failure to reject this joint test would imply that the addition of a fundamental index fund would not improve on the efficient frontier.

Further, let $B = [\alpha, \beta]'$, $\delta = 1_N - \beta 1_K$ and define $\Theta = [\alpha, \beta]'$. The joint null hypothesis of exact mean-variance spanning is $H_0 : \Theta = 0_{2\times N}$ where $\Theta = AB - C$ with

$$A = \begin{bmatrix} 1 & 0'_K \\ 0 & -1'_K \end{bmatrix}, \quad C = \begin{bmatrix} 0'_N \\ -1'_N \end{bmatrix}, \tag{9}$$

and the maximum likelihood estimator of Θ is $\hat{\Theta} = [\hat{\alpha}, \hat{\delta}]' = A\hat{B} - C$.

To test if we can span the RAFI ETF, we employ the Lagrange Multiplier (LM), Likelihood Ratio (LR), and Wald (W) asymptotic tests. The LM, LR, and W all have an asymptotic χ^2_{2N} distributions. It is also well known that $W \geq LR \geq LM$ in finite samples. We report all three test statistics in our sub-period results but focus our attention on the LM test statistic only in the rolling window and out-of-sample tests in the interest of space.

Expanding on the work of Kan and Zhou (2012), Glabadanidis (2009) shows that LM, LR, and W are given by:

$$LM = T\left(\frac{tr(\hat{D}) + 2det(\hat{D})}{1 + det(\hat{D}) + tr(\hat{D})}\right), \tag{10}$$

$$LR = T\ln\left(1 + det(\hat{D}) + tr(\hat{D})\right), \tag{11}$$

$$W = T tr(\hat{D}), \tag{12}$$

where $\hat{D} = \hat{H}\hat{G}^{-1}$ and

$$\hat{G} = TA(X'X)^{-1}A', \tag{13}$$

$$\hat{H} = \hat{\Theta}\hat{\Sigma}^{-1}\hat{\Theta}'. \tag{14}$$

Following the recommendation of Kan and Zhou (2012), we take a step-down approach to test separately whether $\hat{\alpha} = 0_N$ and $\hat{\delta} = 0_N$. This lets us identify whether failure to span the RAFI ETF returns is due to a significant change in the tangent or minimum variance portfolio. The above formulae work for the step-down tests as well. Testing whether $\hat{\alpha} = 0_N$ obtains with just the first row of A and C above while testing whether $\hat{\delta} = 0_N$ utilizes the second row of A and C. The three test statistics now follow a χ^2_N asymptotic distribution.

References

Ang, A., Chen, J. (2002) 'Asymmetric Correlations of Equity Portfolios,' *Journal of Financial Economics* 63(3), pp. 443–494.

Arnott, R. D. (1985) 'The Pension Sponsor's View of Asset Allocation,' *Financial Analysts Journal* 41(5), pp. 17–23.

Arnott, R. D., Hsu, J., Moore, P. (2005) 'Fundamental Indexation,' *Financial Analysts Journal* 61(2), pp. 83–99.

Arnott, R. D., West, J. M. (2006) 'Fundamental Indexes: Current and Future Applications,' *Institutional Investor Journals*, Fall.

Campbell, R., Koedijk, K., Kofman, P. (2002) 'Increased Correlation in Bear Markets,' *Financial Analysts Journal* 58(1), p. 87.

Carhart, M. M. (1997) 'On Persistence in Mutual Fund Performance,' *Journal of Finance* 52(1), pp. 57–82.

Fama, E. F., French, K. R. (1988) 'Permanent and Temporary Components of Stock Prices,' *Journal of Political Economy* 96(2), pp. 246–273.

Fama, E. F., French, K. R. (1992) 'The Cross-Section of Expected Stock Returns,' *Journal of Finance* 47(2), pp. 427–465.

Gallo, J. G., Lockwood, L. J. (1999) 'Fund Management Changes and Equity Style Shifts,' *Financial Analysts Journal* 55(5), pp. 44–52.

Glabadanidis, P. (2009) 'Measuring the Economic Significance of Mean-Variance Spanning,' *Quarterly Review of Economics and Finance* 49 (May), pp. 596–616.

Glabadanidis, P. (2011) 'Robust and Efficient Ways to Track and Outperform a Benchmark,' Working Paper.

Hardy, R. S. (2003) 'Style Analysis: A Ten-year Retrospective and Commentary,' The Handbook of Equity Style Management, pp. 109–130.

Huberman, G., Kandel, S. (1987) 'Mean-variance spanning,' *Journal of Finance* 42(4), pp. 873–888.

Jegadeesh, N. (1990) 'Evidence of predictable behavior of security returns,' *Journal of Finance* 45(3), pp. 881–898.

Kan, R., Zhou, G. (2012) 'Tests of mean-variance spanning,' *Annals of Economics and Finance* 13(1), pp. 145–193.

Kaplan, P. D. (2008) 'Why fundamental indexation might or might not work,' *Financial Analysts Journal* 64(1), pp. 32–39.

Lo, A. W., MacKinlay, A. C. (1988) 'Stock market prices do not follow random walks: Evidence from a simple specification test,' *Review of Financial Studies* 1(1), pp. 41.

Markowitz, H. M. (2005) 'Market efficiency: A theoretical distinction and so what?' *Financial Analysts Journal* 61(5), pp. 17–30.

Perold, A. F. (2007) 'Fundamentally flawed indexing,' *Financial Analysts Journal* 63(6), pp. 31–37.

Perold, A. F., Sharpe, W. F. (1988) 'Dynamic strategies for asset allocation,' *Financial Analysts Journal* 44(1), pp. 16–27.

Sharpe, W. F. (1964) 'Capital asset prices: A theory of market equilibrium under conditions of risk,' *Journal of Finance* 19(3), pp. 425–442.

Sharpe, W. F. (1978) 'Major Investment Styles,' *The Journal of Portfolio Management* 4(2), pp. 68–74.

Sharpe, W. F. (1981) 'Decentralized investment management,' *Journal of Finance* 36(2), pp. 217–234.

Teo, M., Woo, S.-J. (2004) 'Style effects in the cross-section of stock returns,' *Journal of Financial Economics* 74(2), pp. 367–398.

You, L., Daigler, R. (2010) 'The strength and source of asymmetric international diversification,' *Journal of Economics and Finance* 34(3), pp. 349–364.

Table 1. Descriptive Statistics

Panel A. Pre-financial crisis sub-period: Dec 21, 2005 until Jun 29, 2007.

ETF	Mean	Std. Dev.	Skewness	Kurtosis	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
PRF	0.063	0.635	-0.377	5.757	0.005	-0.113	0.012	0.046	-0.150
IWB	0.054	0.671	-0.462	5.960	-0.022	-0.090	0.005	0.058	-0.169
IWF	0.041	0.714	-0.450	5.071	-0.014	-0.102	0.016	0.008	-0.125
IWD	0.065	0.666	-0.576	6.324	-0.051	-0.034	0.010	0.057	-0.161
IWM	0.067	1.086	-0.219	4.194	0.005	-0.162	0.014	0.030	-0.090
IWO	0.062	1.147	-0.157	4.002	0.017	-0.124	-0.001	0.012	-0.061
IWN	0.067	1.038	-0.147	4.091	-0.006	-0.159	-0.024	0.051	-0.084

Panel B. Financial crisis sub-period: Jul 2, 2007 until Feb 27, 2009.

ETF	Mean	Std. Dev.	Skewness	Kurtosis	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
PRF	-0.171	2.233	0.043	6.974	-0.104	-0.121	0.081	-0.080	-0.031
IWB	-0.142	2.148	0.104	7.550	-0.138	-0.144	0.137	-0.091	-0.041
IWF	-0.120	2.039	0.261	8.738	-0.124	-0.135	0.143	-0.108	-0.049
IWD	-0.165	2.325	0.252	7.650	-0.141	-0.133	0.121	-0.102	-0.038
IWM	-0.143	2.520	-0.174	5.280	-0.137	-0.035	0.074	-0.117	-0.082
IWO	-0.136	2.428	0.021	5.531	-0.119	-0.028	0.109	-0.106	-0.108
IWN	-0.154	2.679	-0.208	5.348	-0.164	-0.041	0.062	-0.078	-0.087

Panel C. Post financial crisis sub-period: Mar 2, 2009 until Dec 31, 2010.

ETF	Mean	Std. Dev.	Skewness	Kurtosis	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
PRF	0.178	1.689	0.178	6.786	-0.040	-0.021	-0.039	0.061	-0.038
IWB	0.136	1.340	0.226	6.002	-0.065	-0.010	-0.038	0.058	-0.009
IWF	0.136	1.220	0.171	5.757	-0.036	-0.036	-0.031	0.031	0.035
IWD	0.137	1.517	0.257	6.403	-0.089	-0.001	-0.018	0.060	-0.043
IWM	0.171	1.809	0.222	4.847	-0.057	-0.047	-0.018	0.030	0.006
IWO	0.173	1.698	0.141	4.558	-0.025	-0.069	-0.012	0.013	0.025
IWN	0.170	1.925	0.276	5.008	-0.074	-0.044	-0.006	0.034	-0.006

Panel D. Full Sample: Dec 21, 2005 until Dec 31, 2010.

ETF	Mean	Std. Dev.	Skewness	Kurtosis	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
PRF	0.028	1.684	-0.012	9.730	-0.064	-0.072	0.042	-0.017	-0.026
IWB	0.019	1.527	0.001	11.142	-0.100	-0.093	0.085	-0.038	-0.027
IWF	0.023	1.444	0.133	12.594	-0.084	-0.097	0.093	-0.063	-0.017
IWD	0.015	1.667	0.147	11.171	-0.110	-0.078	0.080	-0.042	-0.034
IWM	0.035	1.916	-0.160	6.857	-0.090	-0.043	0.046	-0.052	-0.044
IWO	0.037	1.849	-0.043	6.836	-0.067	-0.043	0.067	-0.054	-0.050
IWN	0.032	2.018	-0.167	7.201	-0.115	-0.045	0.040	-0.028	-0.050

Notes: Daily data is used in the above summary statistics. ρ_i is the i-the order serial autocorrelation.

Table 2. Mean-Variance Spanning Tests

Panel A. Regression Coefficients.

Coefficient	Pre-Crisis	Crisis	Post-Crisis	Full Sample
α	0.013	-0.022	0.028	0.009
	(2.044)	(1.721)	(1.686)	(1.143)
β_{IWB}	0.491	0.307	0.174	0.263
	(10.076)	(5.770)	(1.242)	(5.414)
β_{IWF}	0.097	0.094	0.033	0.050
	(3.218)	(3.034)	(0.400)	(1.767)
β_{IWD}	0.228	0.471	0.844	0.589
	(6.777)	(13.538)	(10.334)	(19.215)
β_{IWM}	0.071	0.166	0.015	0.150
	(1.813)	(4.586)	(0.132)	(4.311)
β_{IWO}	-0.016	-0.122	-0.104	-0.143
	(0.551)	(4.903)	(1.431)	(6.081)
β_{IWN}	0.016	0.063	0.131	0.094
	(0.511)	(2.375)	(1.810)	(3.744)
σ_ϵ	0.169	0.338	0.352	0.329
R^2	0.929	0.977	0.956	0.962

Panel B. Mean-Variance Spanning Test Results.

Test	Pre-Crisis	Crisis	Post-Crisis	Full Sample
LM	45.990	7.061	32.894	1.343
	(0.000)	(0.029)	(0.000)	(0.511)
LR	49.003	7.121	34.115	1.344
	(0.000)	(0.028)	(0.000)	(0.511)
W	52.285	7.182	35.398	1.345
	(0.000)	(0.028)	(0.000)	(0.511)

Panel C. Step-Down Mean-Variance Spanning Test Results.

TGP	Pre-Crisis	Crisis	Post-Crisis	Full Sample
LM	2.084	1.728	2.814	1.042
	(0.149)	(0.189)	(0.093)	(0.307)
LR	2.089	1.731	2.823	1.042
	(0.148)	(0.188)	(0.093)	(0.307)
W	2.095	1.735	2.831	1.043
	(0.148)	(0.188)	(0.092)	(0.307)
MVP	Pre-Crisis	Crisis	Post-Crisis	Full Sample
LM	45.444	5.650	28.306	0.288
	(0.000)	(0.017)	(0.000)	(0.591)
LR	48.382	5.688	29.204	0.288
	(0.000)	(0.017)	(0.000)	(0.591)
W	51.580	5.727	30.140	0.288
	(0.000)	(0.017)	(0.000)	(0.591)

Notes: This exhibit presents the mean-variance spanning test results and regression output for the three sub-periods under investigation. Panel A reports the coefficients from regressing the RAFI ETF simple daily returns on a set of style ETF simple daily returns with absolute values of t-statistics in parentheses. Standard errors are based on the Newey-West procedure with 3 lags. Panel B reports the test statistics with p-values in parentheses. Panel C presents the step-down mean-variance spanning tests for the global minimum variance and tangent portfolios with p-values provided in parentheses.

Table 3. Four-Factor Model Regression Results

Panel A. Pre-financial crisis sub-period: Dec 21, 2005 until Jun 29, 2007.

ETF	α	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}	σ_ϵ	R^2
PRF	-0.001	0.968	-0.063	0.078	-0.126	0.154	0.941
	(0.122)	(85.570)	(3.646)	(2.685)	(7.235)		
IWB	-0.008	1.025	-0.155	-0.045	-0.056	0.137	0.958
	(1.844)	(127.970)	(12.569)	(2.210)	(4.570)		
IWF	-0.008	1.014	-0.130	-0.475	-0.066	0.205	0.918
	(1.135)	(75.540)	(6.284)	(13.821)	(3.184)		
IWD	-0.008	1.046	-0.215	0.353	-0.058	0.161	0.941
	(1.367)	(95.378)	(12.701)	(12.560)	(3.467)		
IWM	-0.011	1.191	0.897	0.145	-0.059	0.278	0.934
	(1.283)	(76.402)	(37.394)	(3.636)	(2.459)		
IWO	-0.009	1.231	0.916	-0.107	0.002	0.265	0.947
	(1.068)	(76.748)	(37.093)	(2.593)	(0.095)		
IWN	-0.016	1.183	0.863	0.349	-0.148	0.283	0.926
	(1.615)	(61.776)	(29.272)	(7.116)	(5.034)		

Panel B. Financial crisis sub-period: Jul 2, 2007 until Feb 27, 2009.

ETF	α	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}	σ_ϵ	R^2
PRF	-0.028	0.927	-0.000	0.146	-0.091	0.330	0.978
	(2.158)	(121.809)	(0.021)	(7.243)	(7.333)		
IWB	-0.011	0.959	-0.051	-0.030	-0.017	0.282	0.983
	(1.211)	(185.480)	(4.479)	(2.231)	(2.030)		
IWF	-0.004	0.949	-0.041	-0.314	-0.015	0.385	0.964
	(0.318)	(127.377)	(2.462)	(15.967)	(1.254)		
IWD	-0.019	0.969	-0.125	0.197	-0.035	0.434	0.965
	(1.165)	(100.786)	(5.857)	(7.770)	(2.210)		
IWM	0.022	1.129	0.911	0.177	0.043	0.431	0.971
	(1.469)	(129.827)	(47.324)	(7.688)	(3.049)		
IWO	0.015	1.120	0.849	-0.202	-0.024	0.478	0.961
	(0.891)	(111.371)	(38.173)	(7.604)	(1.438)		
IWN	0.026	1.127	1.033	0.440	0.033	0.542	0.959
	(1.169)	(87.517)	(36.250)	(12.939)	(1.566)		

Panel C. Post financial crisis sub-period: Mar 2, 2009 until Dec 31, 2010.

ETF	α	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}	σ_ϵ	R^2
PRF	0.010	0.955	0.007	0.405	-0.114	0.283	0.972
	(0.832)	(66.768)	(0.290)	(15.956)	(10.513)		
IWB	0.000	0.979	-0.062	-0.030	0.016	0.122	0.992
	(0.007)	(208.032)	(8.256)	(3.627)	(4.577)		
IWF	0.006	0.983	-0.000	-0.317	0.017	0.206	0.971
	(0.832)	(118.369)	(0.015)	(21.489)	(2.634)		
IWD	-0.009	0.955	-0.100	0.289	-0.015	0.192	0.984
	(1.309)	(115.422)	(7.516)	(19.679)	(2.432)		
IWM	-0.023	1.049	0.885	0.074	-0.022	0.253	0.980
	(2.693)	(104.570)	(55.123)	(4.140)	(2.855)		
IWO	-0.013	1.069	0.873	-0.172	-0.004	0.258	0.977
	(1.360)	(94.105)	(48.036)	(8.512)	(0.486)		
IWN	-0.030	1.024	0.890	0.306	-0.036	0.334	0.970
	(2.385)	(69.829)	(37.928)	(11.752)	(3.215)		

Notes: Table 3 reports the four factor model regression coefficients for the three sub-periods under investigation with absolute values of t-statistics in parentheses. Standard errors are adjusted using the Newey-West correction with 3 lags.

Table 3 Continued.

Panel D. Entire Sample: Dec 21, 2005 until Dec 31, 2010.

ETF	α	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}	σ_ϵ	R^2
PRF	-0.001	0.943	0.022	0.264	-0.089	0.294	0.970
	(0.116)	(166.999)	(1.934)	(19.600)	(12.128)		
IWB	-0.006	0.966	-0.059	-0.023	-0.001	0.196	0.984
	(1.578)	(349.451)	(10.333)	(3.485)	(0.394)		
IWF	-0.001	0.963	-0.022	-0.310	-0.000	0.281	0.962
	(0.255)	(227.325)	(2.497)	(30.701)	(0.012)		
IWD	-0.010	0.973	-0.125	0.243	-0.017	0.292	0.969
	(1.586)	(200.392)	(12.526)	(20.996)	(2.725)		
IWM	-0.006	1.104	0.877	0.086	-0.002	0.337	0.969
	(0.922)	(221.350)	(85.383)	(7.184)	(0.275)		
IWO	-0.003	1.123	0.863	-0.210	-0.005	0.354	0.963
	(0.435)	(196.300)	(73.336)	(15.398)	(0.711)		
IWN	-0.012	1.083	0.929	0.320	-0.032	0.416	0.958
	(1.268)	(153.650)	(64.085)	(19.032)	(3.445)		

Table 4. Tangent and Minimum Variance Portfolio Moments and Sharpe Ratios

Panel A. Pre-Crisis sub-period.

	BM TGP	BM+F TGP	BM MVP	BM+F MVP
Mean Return	24.69%	25.47%	11.78%	13.70%
Standard Deviation	12.60%	11.14%	8.71%	8.17%
Sharpe Ratio	1.96	2.27	1.35	1.68

Panel B. Crisis sub-period.

	BM TGP	BM+F TGP	BM MVP	BM+F MVP
Mean Return	-60.65%	-88.81%	-22.53%	-25.79%
Standard Deviation	49.76%	55.90%	30.33%	30.12%
Sharpe Ratio	-1.22	-1.59	-0.74	-0.86

Panel C. Post-crisis sub-period.

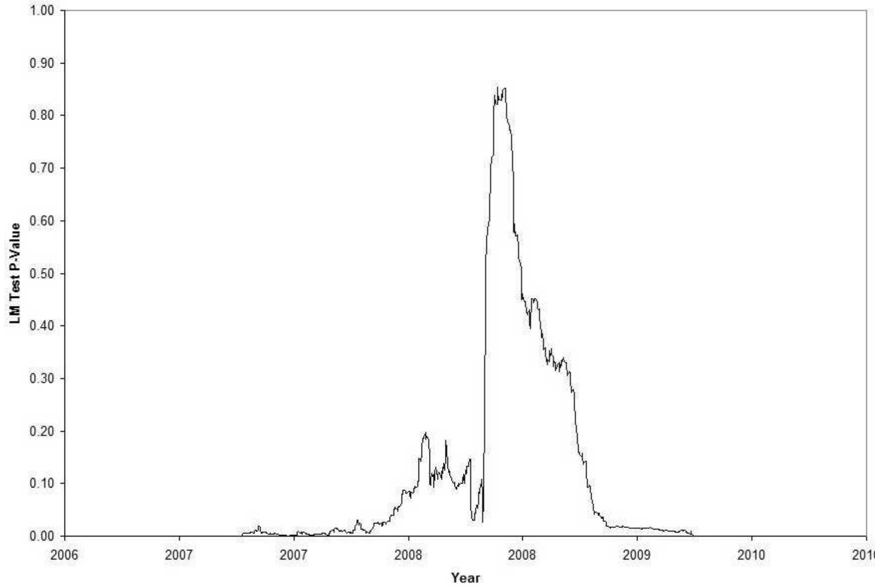
	BM TGP	BM+F TGP	BM MVP	BM+F MVP
Mean Return	33.42%	57.51%	26.34%	20.21%
Standard Deviation	17.15%	24.89%	15.23%	14.75%
Sharpe Ratio	1.95	2.31	1.73	1.37

Panel D. Full sample.

	BM TGP	BM+F TGP	BM MVP	BM+F MVP
Mean Return	24.75%	48.48%	4.01%	3.87%
Standard Deviation	51.42%	73.28%	20.71%	20.70%
Sharpe Ratio	0.48	0.66	0.19	0.19

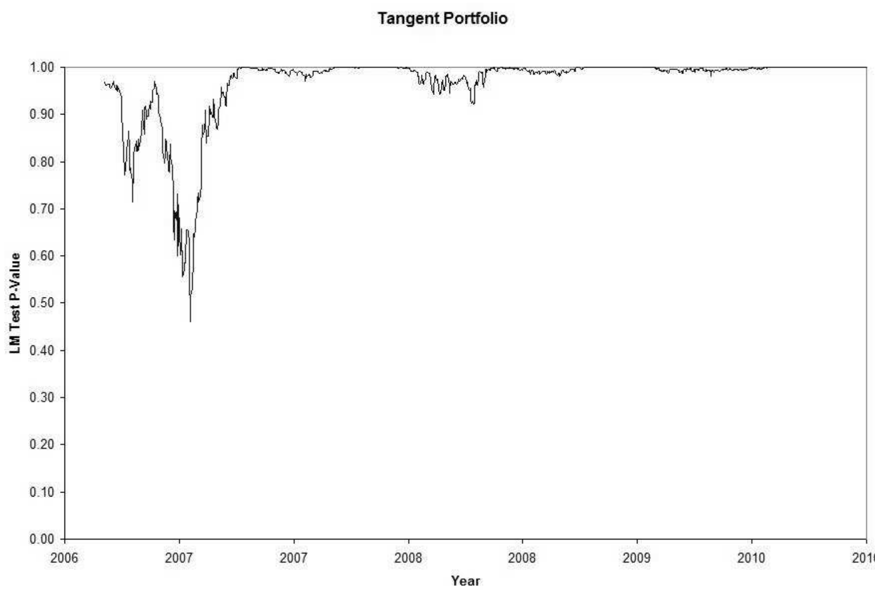
Notes: Table 4 reports the annualized mean, standard deviation and Sharpe ratios of the tangent (TGP) and minimum variance (MVP) portfolios over the three sub-periods. BM is the benchmark replicating portfolio and F is the RAFI ETF.

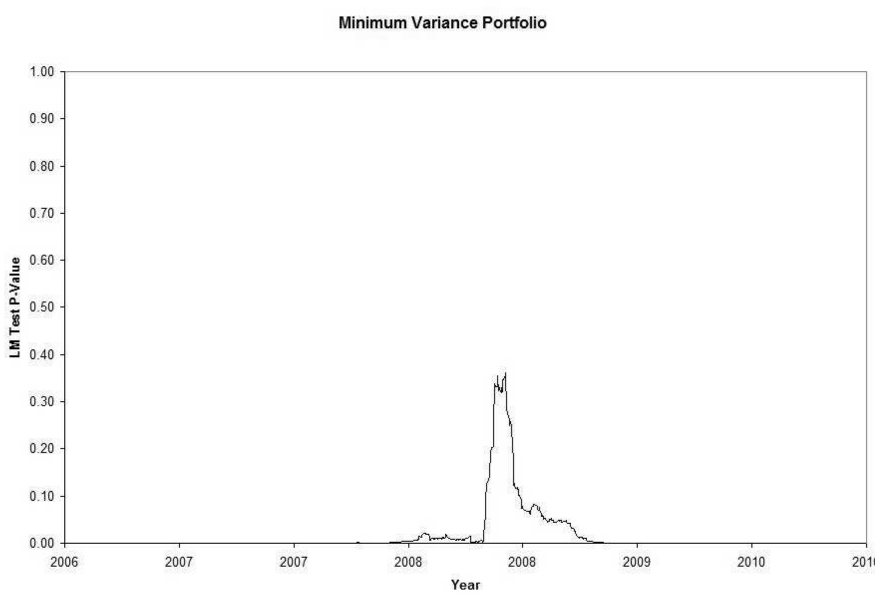
Figure 1. Rolling Window Mean-Variance Spanning LM Test P-Values



Notes: Figure 1 plots the rolling values of the Lagrange Multiplier test (LM) p-values using a year's worth of trading days window rolled over on a daily basis. The sample contains 1016 rolling windows, each containing 250 observations.

Figure 2. Step-Down Rolling Window Mean-Variance Spanning LM Test P-Values





Notes: Figure 2 plots of the rolling values of the Lagrange Multiplier test p-values for the tangent portfolio (top) and the minimum variance portfolios (bottom) using a year's worth of trading days window rolled over on a daily basis. The sample contains 1016 rolling windows, each containing 250 observations.

Table 5. RAFI Replicating Portfolio Weights

Panel A. Sample period is Dec 21, 2005 until Jun 30, 2006.

N_{ETF}							
1	IWB						Cash
	90.689%						9.311%
2	IWB	IWD					
	74.485%	16.879%					8.636%
3	IWB	IWD	IWN				
	70.540%	16.357%	3.124%				9.979%
4	IWB	IWD	IWN	IWM			
	71.030%	17.835%	15.762%	-13.319%			8.691%
5	IWB	IWD	IWN	IWM	IWO		
	71.444%	17.522%	15.688%	-11.728%	-1.535%		8.609%
6	IWB	IWD	IWN	IWM	IWO	IWF	
	71.383%	17.528%	15.693%	-11.735%	-1.537%	0.062%	8.607%

Panel B. Sample period is Jan 3, 2007 until Jul 2, 2007.

N_{ETF}							
1	IWB						Cash
	88.644%						11.356%
2	IWB	IWD					
	47.501%	40.426%					12.073%
3	IWB	IWD	IWM				
	31.939%	36.989%	15.488%				15.584%
4	IWB	IWD	IWM	IWF			
	21.199%	38.031%	13.666%	12.123%			14.981%
5	IWB	IWD	IWM	IWF	IWN		
	20.892%	38.646%	16.406%	12.248%	-3.155%		14.962%
6	IWB	IWD	IWM	IWF	IWN	IWO	
	20.888%	38.650%	16.341%	12.227%	-3.184%	0.107%	14.971%

Panel C. Sample period is Mar 3, 2008 until Sep 2, 2008.

N_{ETF}							
1	IWD						Cash
	98.761%						1.239%
2	IWD	IWB					
	60.554%	44.068%					-4.621%
3	IWD	IWB	IWN				
	57.661%	32.229%	11.985%				-1.875%
4	IWD	IWB	IWN	IWO			
	45.945%	49.464%	21.090%	-14.049%			-2.449%
5	IWD	IWB	IWN	IWO	IWM		
	46.298%	47.954%	16.635%	-18.897%	10.657%		-2.647%
6	IWD	IWB	IWN	IWO	IWM	IWF	
	52.494%	27.602%	19.285%	-23.086%	11.445%	15.872%	-3.612%

Panel D. Sample period is Mar 2, 2009 until Sep 2, 2009.

N_{ETF}							
1	IWD						Cash
	111.391%						-11.391%
2	IWD	IWF					
	99.318%	17.328%					-16.647%
3	IWD	IWF	IWN				
	93.682%	13.547%	7.045%				-14.274%
4	IWD	IWF	IWN	IWO			
	87.229%	29.946%	21.957%	-23.946%			-15.187%
5	IWD	IWF	IWN	IWO	IWB		
	69.620%	15.490%	24.289%	-28.336%	34.948%		-16.012%
6	IWD	IWF	IWN	IWO	IWB	IWM	
	69.289%	15.336%	22.677%	-29.917%	35.218%	3.369%	-15.972%

Table 6. Performance of RAFI Replicating Portfolios

Panel A. In-sample period is Dec 21, 2005 until Dec 29, 2006. Out-of-sample period is Jan 3, 2007 until Jun 29, 2007.

Number of	In-sample					Out-of-sample				
ETFs	ρ_{py}	TE	SF	α_p	β_p	ρ_{py}	TE	SF	α_p	β_p
1	0.9603	0.0017	-0.0333	-0.0000	0.9332	0.9588	0.0021	-0.0085	-0.0001	1.0002
2	0.9609	0.0017	-0.0247	-0.0000	0.9236	0.9621	0.0020	-0.0103	-0.0001	1.0050
3	0.9615	0.0017	-0.0221	-0.0000	0.9522	0.9648	0.0020	-0.0105	-0.0001	1.0187
4	0.9616	0.0017	-0.0256	-0.0000	0.9500	0.9657	0.0020	-0.0099	-0.0001	1.0172
5	0.9617	0.0017	-0.0247	-0.0000	0.9476	0.9657	0.0020	-0.0106	-0.0001	1.0155
6	0.9617	0.0017	-0.0239	-0.0000	0.9475	0.9658	0.0020	-0.0112	-0.0001	1.0159

Panel B. In-sample period is Jul 2, 2007 until Jul 2, 2008. Out-of-sample period is Jul 3, 2008 until Feb 27, 2009.

Number of	In-sample					Out-of-sample				
ETFs	ρ_{py}	TE	SF	α_p	β_p	ρ_{py}	TE	SF	α_p	β_p
1	0.9703	0.0031	0.0760	-0.0001	0.9806	0.9856	0.0054	0.0214	-0.0001	0.9608
2	0.9725	0.0030	0.0566	-0.0002	0.9949	0.9891	0.0047	0.0150	-0.0000	0.9817
3	0.9728	0.0030	0.0551	-0.0002	1.0049	0.9903	0.0044	0.0156	-0.0000	0.9860
4	0.9749	0.0029	0.0521	-0.0002	0.9879	0.9896	0.0046	0.0204	0.0000	0.9885
5	0.9755	0.0028	0.0576	-0.0002	0.9789	0.9895	0.0046	0.0227	0.0000	0.9829
6	0.9757	0.0028	0.0562	-0.0002	0.9781	0.9896	0.0046	0.0243	0.0001	0.9840

Panel C. In-sample period is Mar 2, 2009 until Feb 26, 2010. Out-of-sample period is Mar 2, 2010 until Dec 31, 2010.

Number of	In-sample					Out-of-sample				
ETFs	ρ_{py}	TE	SF	α_p	β_p	ρ_{py}	TE	SF	α_p	β_p
1	0.9729	0.0045	-0.2779	-0.0003	1.0936	0.9924	0.0015	-0.0210	-0.0001	1.0170
2	0.9731	0.0044	-0.2821	-0.0002	1.0631	0.9937	0.0014	-0.0185	-0.0001	1.0060
3	0.9738	0.0044	-0.2742	-0.0002	1.0815	0.9943	0.0013	-0.0149	-0.0001	1.0281
4	0.9743	0.0044	-0.2735	-0.0002	1.0793	0.9936	0.0014	-0.0228	-0.0001	1.0213
5	0.9743	0.0044	-0.2744	-0.0002	1.0756	0.9936	0.0014	-0.0230	-0.0001	1.0189
6	0.9743	0.0044	-0.2745	-0.0002	1.0758	0.9936	0.0014	-0.0229	-0.0001	1.0193

Panel D. In-sample period is Dec 21, 2005 until Mar 3, 2008. Out-of-sample period is Mar 4, 2008 until Dec 31, 2010.

Number of	In-sample					Out-of-sample				
ETFs	ρ_{py}	TE	SF	α_p	β_p	ρ_{py}	TE	SF	α_p	β_p
1	0.9651	0.0023	0.0005	-0.0001	0.9702	0.9739	0.0051	-0.0800	-0.0001	0.9583
2	0.9677	0.0022	0.0010	-0.0001	0.9763	0.9787	0.0045	-0.1027	-0.0001	0.9817
3	0.9686	0.0022	-0.0026	-0.0001	1.0001	0.9801	0.0044	-0.0885	-0.0001	0.9938
4	0.9698	0.0021	-0.0042	-0.0001	0.9879	0.9803	0.0043	-0.0913	-0.0001	0.9964
5	0.9702	0.0021	-0.0046	-0.0001	0.9816	0.9799	0.0044	-0.0882	-0.0001	0.9893
6	0.9703	0.0021	-0.0066	-0.0001	0.9822	0.9801	0.0044	-0.0880	-0.0001	0.9906

Table 6 reports the in-sample and out-of-sample RAFI ETF replicating portfolio results. TE is the tracking error or the daily standard deviation of the difference between the replicating portfolio daily return and the RAFI ETF daily return. SF is the cumulative short-fall of the replicating portfolio relative to the RAFI ETF over the entire six month period under consideration. ρ_{py} is the correlation between the replicating portfolio return and the RAFI ETF return, α_p and β_p are the intercept and slope from the single-factor market model with the value-weighted stock return as the market proxy.