Time-Variation in Diversification Benefits of Commodity, REITs, and TIPS[1]

Jing-zhi Huang[2] and Zhaodong Zhong[3]

This Draft: July 11, 2006

Abstract

Diversification benefits of three "hot" asset classes, Commodity, Real Estate Investment Trusts (REITs), and Treasury Inflation-Protected Securities (TIPS), are well studied on an individual basis and in a static setting. In this paper, we document that the three asset classes are in general not substitutes for each other and that all ought to be included in investors' portfolios, based on a sample of daily return data from January 1999 through December 2005. We also find that diversification benefits of the three hot assets change substantially over time. For instance, benefits of TIPS were significant before 2001 but have been decreasing gradually since then. On the other hand, diversification benefits from Commodity and REITs fluctuate significantly over the entire sample period. We show that this observed time-variation in diversification benefit can be captured by incorporating time-varying return correlations. To see the implications of this finding for asset allocation in practice, we examine the out-of-sample performance of portfolio strategies constructed based on a variety of correlation structures. We find that Engle's (2002) Dynamic Conditional Correlation model outperforms other correlation structures such as rolling, historical, and constant correlations. Our findings suggest that diversification benefits of the three hot asset classes do vary substantially over time and that investors need to use appropriate correlation estimates in their asset allocation decisions to adjust for such time variation.

[1] We thank Bill Kracaw, Hao Zhou, and seminar participants at the UESTC and the 2006 Journal of Banking and Finance Conference in Beijing for helpful comments and discussions.

[2] Department of Finance, Smeal College of Business, Penn State University, University Park, PA 16802; Tel: (814) 863-3566; jxh56@psu.edu

[3] Department of Finance, Smeal College of Business, Penn State University, University Park, PA 16802; Tel: (814) 865-2446; zxz126@psu.edu

1. Introduction

Diversification plays an essential role in asset allocation. Recently, especially since the burst of internet bubble, there have been many studies on diversification benefits of non-traditional asset classes such as commodity, real estate, and U.S. Treasury Inflation-Protected Securities (TIPS). Rouwenhorst and Gorton (2005) document that commodity future returns are negatively correlated with both equity and bond returns, and argue that "commodity futures returns have been especially effective in providing diversification of both stock and bond portfolios" (p. 28). Roll (2004) analyzes the correlation of TIPS with nominal bond and equity returns and concludes that "an investment portfolio diversified between U.S. equities and nominal bonds would be improved by the addition of TIPS" (p. 31). Chun, Sa-Aadu and Shilling (2004) show that investments in real estate will pay off when consumption growth opportunities are low, and argue that institutional investors should invest more in real estate to eliminate non-market risk. That is, there is clear evidence on diversification benefits of the three asset classes on an individual basis.

As such, one obvious follow up question is if investors should include all three "hot" asset classes in their portfolios given that all of them are considered to be an instrument against inflation. Namely, are these three asset classes substitutes for each other? Another related question is: Do their diversification benefits change over time and, if so, how to adjust for such variation in asset allocation in practice?

In this paper, we try to shed some light on these two questions by doing an integrated analysis of diversification benefits of commodity, real estate investment trusts (REITs), and TIPS. More specifically, we consider seven asset classes including the three hot ones (test asset classes hereafter) and four traditional ones (benchmark asset classes hereafter)—U.S. Equity, U.S. Bond, International Equity, and International

Bond—and examine how the test asset classes benefit an investor's portfolios using daily return data from January 1999 through December 2005.

We test for diversification benefits of the three test asset classes using a variety of methods. Specifically, we use the Gibbons, Ross, and Shanken (1989) test to examine the statistical significance of changes in the Sharpe ratio. To see if a given test asset class can be spanned by other asset classes, we consider two spanning tests. One is Huberman and Kandel's (1987) method in the absence of short-sale constraints. The other is the spanning test developed by De Roon, Nijman and Werker (2001) that incorporates short-sale constraints. Finally, to quantify an asset class's diversification benefits, we use the increase in the tangent portfolio's Sharpe ratio after the addition of the asset (we also look at the weight of the asset in the tangent portfolio).

One finding from our empirical analysis is that the three test asset classes are not substitutes for each other when spanning tests are done using the full sample. However, there is evidence of substitution in certain sub-periods (quarters) when spanning tests are done on a quarterly basis. In most of such cases commodity can be spanned by other assets and thus becomes redundant. Based on the increase in the tangent portfolio's Sharpe ratio after the addition of the test asset class and its weight in the tangent portfolio, TIPS appear to benefit the benchmark portfolio the most, REITs the next, and commodity the least.

We also find that diversification benefits of the three asset classes change substantially over time. For instance, TIPS' improvement in the tangent portfolio is very significant before 2001 and then begins to drop gradually since then. The time series behavior of the asset class's weight in the tangent portfolio tells the similar story. We then show that we can capture the time-variation in diversification benefits of the three test asset classes by using a time-varying return correlation.

If correlations are indeed time varying and diversification benefits of each test asset class do change over time, then how should investors rebalance their portfolio accordingly? To answer this question, we implement portfolio strategies based on a variety of correlation estimates and rebalancing frequencies and do an out-of-sample performance analysis of these strategies. We find that the Dynamic Conditional Correlation (DCC) model (Engle, 2002) outperforms three other popular correlation structures—the unconditional correlation, the rolling correlation, and the constant correlation.

To summarize, our empirical findings suggest that all three test asset classes benefit a portfolio of traditional asset classes but their diversification benefits (and the impact) do vary substantially over time, and that investors need to use appropriate correlation estimates in their asset allocation decisions to adjust for such time variation.

There are several related studies in the literature. Roll (2004) and Kothari and Shanken (2004) both find that substantial weights should be given to TIPS in an efficient portfolio. Simon and Hunter (2005) use a time-varying correlation model (not DCC though) and show that TIPS have superior volatility-adjusted return relative to nominal bonds over time. However, they cannot reject the null that TIPS can be spanning by equities and nominal Treasury securities in their spanning tests. Mamun and Visaltanachoti (2005) find results contrary to Simon and Hunter (2005)'s ones on spanning. Chen, Ho, Lu and Wu (2004) find that REITs also help to improve investors' mean-variance frontier of size/book-to-market stock portfolios in spanning tests. This paper contributes to the literature by providing an integrated analysis of time-varying diversification benefits of Commodity, REITs and TIPS in the literature and demonstrating the importance of DCC (which is statistically a better measure of time-varying correlations than the ad-hoc rolling correlation) asset allocation in the real world.

The organization of the paper is as follows: Section 2 discusses data and empirical methods used in our analysis. Section 3 reports empirical results from spanning tests. Section 4 examines time-variation in diversification benefits of Commodity, REITs, and TIPS. We also run a horse race of four alternative correlation estimates based on an out-of-sample performance analysis. Section 5 addresses the robustness of our empirical results. Finally, we conclude with section 6.

2. Data and Methodology

In this section we describe the return data and the empirical methods used in this study. We first introduce the standard mean-variance spanning test in the absence of short-sale constraints (Huberman and Kandel, 1987). We then describe the spanning test with short-sale constraints proposed by De Roon, Nijman and Werker (2001). Finally, we review the dynamic conditional correlation model (Engle, 2002).

2.1 Data

We use daily return indexes of the following seven asset classes: Commodity, REITs, TIPS, U.S. Equity, U.S. Bond, International Equity, and International Bond. In our analysis, the first three asset classes are used as test asset classes and the last four used as benchmark asset classes. The detailed descriptions of these return indexes are provided in Table 1. As can be seen from the table, the daily return index with the most recent inception date, January 1, 1999, is the International Equity index. As a result, our sample period is from January 1, 1999 to December 31, 2005.

Panel A of Table 2 reports the annualized unconditional means and standard deviations of the seven return series in our sample. As can be seen from the table, commodity has both the highest mean return (21.23%) and highest standard deviation (22.22%) among the seven asset classes considered here. On the other hand, TIPS have a modest mean return (8.24%) and a fairly low standard deviation (4.70%). REITs have both relatively high mean return (17.81%) and relatively high standard deviation (13.38%). The unconditional mean returns of the four conventional asset classes, namely U.S. Equity, U.S. Bond, International Equity, and International Bond, are 3.46%, 5.61%, 6.36% and 4.12%, respectively. The standard deviations of the four are 18.51%, 4.55%, 14.76% and 8.62%, respectively.

Panel B of Table 2 reports the unconditional correlation matrix. Like Rouwenhorst and Gorton (2005), we also find that Commodity is negatively correlated with both U.S. Equity (with a correlation coefficient of -0.0241) and U.S. Bond (with a correlation of -0.0182). It is worth mentioning that Rouwenhorst and Gorton (2005) find negative correlations only in low frequency data (quarterly and annual) but not in monthly data. Our result here is based on daily data and, to some extent, complements to their finding. REITs are positively correlated with U.S. Equity with a correlation of 0.4307 and negatively correlated with U.S. Bond with a correlation of -0.0268. TIPS are negatively correlated with U.S. Equity with a correlation of -0.1715 and positively correlated with U.S. Bond with a correlation of 0.7412. Namely, each of the three test asset classes is negatively correlated with either U.S. Equity or U.S. Bond or both.

Another interesting result shown in Panel B is that the three test assets are only weakly correlated with each other. In fact, the correlation is -0.0668 between Commodity and REITs, is 0.0964 between Commodity and TIPS, and is only -0.0432 between REITs and TIPS. This indicates that although all three assets are considered to be an instrument hedging against inflation, they are not substitutes.

2.2 Spanning Test Specifications without Short-Sale Constraints

Spanning tests are a standard method to study diversification benefits of a new asset. Namely, one examines whether the efficient frontier of the test (or new) asset classes and benchmark asset classes can be "spanned" by the efficient frontier based on the benchmark asset classes only. A related test is the so called intersection test, that examines if the two efficient frontiers (one including test asset classes and the other not) "intersect" at the same tangent point for a given risk-free rate. Figure 1 illustrates what spanning or intersection means. See Kan and Zhou (2001) and De Roon and Nijman (2001) for a survey on spanning tests.

In the absence of short-sale constraints, we use the following regression-based spanning test of Huberman and Kandel (1987). (See also Ferson, Foerster, and Keim (1993), De Santis (1994), Harvey (1995), Bekaert and Urias (1996), and De Roon, Nijman and Werker (2001).)

Assume

$$\vec{r}_t = \vec{a} + \mathrm{B}\vec{R}_t + \vec{e}_t \qquad (1)$$

where the $N \times 1$ vector $\vec{r}_t$ denotes the returns of N test asset classes, the $K \times 1$ vector $\vec{R}_t$ denotes the returns of K benchmark asset classes, B is an $N \times K$ matrix, and $\vec{a}$ and $\vec{e}_t$ are

$N\times 1$ vectors. Under this specification, "intersection" means K benchmark asset classes "intersect" the $K+N$ asset classes and the risk-free asset with return 1/v, and is equivalent to the following conditions

$$v\vec{a} + \mathrm{B}\vec{i}_K - \vec{i}_N = \vec{0}_N \tag{2}$$

where $\vec{0}_N$ is an $N\times 1$ vector of zeros, $\vec{i}_K$ is a $K\times 1$ vector of ones and $\vec{i}_N$ is an $N\times 1$ vector of ones.

On the other hand, "spanning" means K benchmark asset classes "span" the $K+N$ asset classes and is equivalent to the following conditions

$$\vec{a} = \vec{0}_N \text{ and } \mathrm{B}\vec{i}_K = \vec{i}_N \tag{3}$$

As such, both intersection and spanning tests can be conducted by testing the restrictions on the coefficients in a regression-based framework.

Following Ferson, Foerster, and Keim (1993), we use the generalized method of moments (Hansen, 1982) to implement the tests. One advantage of the GMM is that it controls for heteroskedasticity and autocorrelation. We use the Newey-West (1987) correction with a bandwidth of $\frac{1}{2}n^{\frac{1}{3}}$ (Andrews, 1991) for a sample of n observations errors.[1] Since both the regression model and constraints are linear, the GMM versions of the LR and LM tests have exactly the same form as the Wald test (Kan and Zhou, 2001). Therefore, we report only the results of the Wald test of the coefficient restrictions in the paper.

[1] We also try alternative lags (from 4 to 20) and find that our results are robust to different lag choices.

2.3 Spanning Test Specifications with Short-Sale Constraints

In the case where short-sale constraints are imposed, we use the test proposed by De Roon, Nijman and Werker (2001).

Consider first the case where short sale constraints are imposed on the test asset classes only (but not on the benchmark classes). Here, "intersection" for a given value of v implies that

$$\alpha(v) = v\vec{a} + \mathrm{B}\vec{i}_K - \vec{i}_N \leq \vec{0}_N. \tag{4}$$

The Wald test statistic (Kodde and Palm, 1986) is given by

$$\xi(v) = \min_{\{\alpha(v)\leq 0\}} (\hat{\alpha}(v) - \alpha(v))' Var[\hat{\alpha}(v)]^{-1} (\hat{\alpha}(v) - \alpha(v)) \ , \tag{5}$$

where $\hat{\alpha}(v)$ is a consistent estimate of $\alpha(v)$. In this case, $\xi(v)$ follows a mixture of Chi-square distribution and for a given value c,

$$P(\xi(v) \geq c) = \sum_{i=0}^{N} P(\chi_i^2 \geq c)\, w(N, i, Var[\hat{\alpha}(v)]) \tag{6}$$

where $w(N, i, Var[\hat{\alpha}(v)])$ is the probability that i of N elements of a vector with $N(0,, Var[\hat{\alpha}(v)]\ )$ distribution is strictly positive. The *p*-value can be calculated by numerical simulation as suggested by Gourieroux, Holly, and Monfort (1982).

Similarly, "spanning" implies that equation (4) holds for any v = $1/r_{riskfree}$. However, the range of v can be limited before hand if we impose certain economic assumptions on the risk-free rates. Following De Roon, Nijman and Werker (2001), we limit the risk free rate to be between 1 and the intercept of asymptote of the lines tangent to the efficient frontier of returns R_t. Therefore, if we let $v_{\min}$ and $v_{\max}$ be the minimum and maximum of v, then testing for spanning is equivalent to a joint test of equation (4) for both $v_{\min}$ and $v_{\max}$. As before, the Wald test statistics follows a mixture of Chi-square distribution since all intermediate value of v will satisfy equation (4) if both $v_{\min}$ and $v_{\max}$ do so.

Consider next the case where short-sale constraints are imposed on both test and benchmark asset classes. In this case, for any given v, the corresponding portfolio weights of the optimal portfolio on the efficient frontier of the benchmark asset classes are all non-negative. For a given v, denote $\vec{R}_t^{(v)}$ as the vector of those benchmark assets with positive portfolio weight only. Let $\vec{a}^{(v)}$ and $\mathrm{B}^{(v)}$ be the coefficients of

$$\vec{r}_t = \vec{a}^{(v)} + \mathrm{B}^{(v)}\vec{R}_t^{(v)} + \vec{e}_t^{(v)}. \tag{7}$$

De Roon, Nijman and Werker (2001) show that "intersection" for a given value of v implies that

$$\alpha^{(v)}(v) = v\vec{a}^{(v)} + \mathrm{B}^{(v)}\vec{i}_K - \vec{i}_N \leq \vec{0}_N. \tag{8}$$

This equation can be tested in a similar fashion as equation (4).

In terms of spanning, consider a K dimensional return vector $\vec{R}_t = [R_t^1, \cdots, R_t^K]$ and let $\varphi = \{R_t^1, \cdots, R_t^K\}$, a set of the return components. Since the number of the subsets of φ is finite, we can use these subsets to form return vectors, $\vec{R}_t^{[j]}, j = 1, 2, \cdots, J$. Then, all relevant v can be classified into a finite number of disjoint sets $V^{[j]}$, where $V^{[j]} \equiv \{v \mid \vec{R}_t^{(v)} = \vec{R}_t^{[j]}\}$. Let $v_{\min}^{[j]}$ and $v_{\max}^{[j]}$ be the minimum and maximum of $V^{[j]}$, and $\vec{a}^{[j]}$ and $\mathrm{B}^{[j]}$ be the coefficients of

$$\vec{r}_t = \vec{a}^{[j]} + \mathrm{B}^{[j]}\vec{R}_t^{[j]} + \vec{e}_t^{[j]}. \tag{9}$$

De Roon, Nijman and Werker (2001) show that "spanning" implies

$$\begin{matrix} v_{\min}^{[j]}\vec{a}^{[j]} + \mathrm{B}^{[j]}\vec{i}_K{}^{[j]} - \vec{i}_N \leq \vec{0}_N \\ v_{\max}^{[j]}\vec{a}^{[j]} + \mathrm{B}^{[j]}\vec{i}_K{}^{[j]} - \vec{i}_N \leq \vec{0}_N \end{matrix}, \qquad \forall j. \tag{10}$$

The Wald test statistic of the joint test of the above restrictions follows a mixture of Chi-square distribution, similar to equation (6).

2.4 Dynamic Conditional Correlations (DCC)

We use the DCC model introduced by Engle (2002) to model the time-varying correlations among asset classes. The DCC model is based on the idea of estimating volatilities by generalized autoregressive and conditional heteroskedasticity (GARCH) processes (Bollerslev, 1986) and the correlation matrix by a GARCH-like process. Therefore, DCC is computationally more efficient than a multivariate GARCH model due to its two-step estimation of univariate GARCH series and GARCH-like correlation matrix. In addition, DCC has an advantage over other time-varying correlations, such as rolling correlation (which is widely used in the industry) because the asymptotic properties of DCC are fully understood (see Engle and Sheppard (2001), and Cappiello, Engle, and Sheppard (2003)).

We follow Engle and Sheppard (2001) below. The $k \times 1$ demeaned returns series R_t of k assets are assumed to have the following structure:

$$\begin{aligned} & R_t \mid \Omega_{t-1} \sim N(0, H_t) \\ & and \\ & H_t \equiv D_t C_t D_t \end{aligned} \tag{11}$$

where D_t is the $k \times k$ diagonal matrix with the univariate GARCH (p, q) standard deviation $\sqrt{h_{it}}$ on the i^{th} diagonal of H_t, C_t is the $k \times k$ time-varying correlation matrix that follows a GARCH-like process, and Ω_t represents the information set at time t.

The log-likelihood of this estimate is given by:

$$L = -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + 2\log(|D_t|) + \log(|C_t|) + \varepsilon_t^{'}C_t^{-1}\varepsilon_t) \tag{12}$$

where $\varepsilon_t \sim N(0, C_t)$.

The estimation is implemented by two steps. In the first step, k univariate GARCH (p, q) are estimated as:

$$h_{it} = \omega_i + \sum_{p=1}^{P_i}\alpha_{ip}r_{it-p}^2 + \sum_{q=1}^{Q_i}\beta_{iq}h_{it-q} \tag{13}$$

for $i = 1,2,\cdots,k$, and with non-negativity of variance and restrictions $\sum_{p=1}^{P_i}\alpha_{ip} + \sum_{q=1}^{Q_i}\beta_{iq} < 1$ imposed.

The second step is the estimation of the correlation matrix that is assumed to follow a GARCH-like process:

$$Q_t = (1 - \sum_{m=1}^{M}\alpha_m - \sum_{n=1}^{M'}\beta_n)\overline{Q} + \sum_{m=1}^{M}\alpha_m(\varepsilon_{t-m}\varepsilon_{t-m}^{'}) + \sum_{n=1}^{M'}\beta_n Q_{t-n} \tag{14}$$

and

$$C_t = Q_t^{*-1}Q_t Q_t^{*-1} \tag{15}$$

where $\overline{Q}$ is the unconditional covariance of the standardized residuals from the first stage, and Q_t^* is a diagonal matrix composed of the square root of the diagonal elements of Q_t. With appropriate restrictions on the parameters of the k univariate GARCH processes and the correlation's GARCH-like process, C_t is a positive definite correlation matrix (see Engle and Sheppard (2001) for details). This completes the specification of DCC model. In this paper, we implement a simple version of the DCC model with P_i, Q_i, M, and M' all set to 1, for $i = 1, 2, \cdots, k$.

3. Empirical Results from Spanning Tests

In this section, we report the empirical results from spanning tests. We first consider the case when there are no short-sale constraints, and then consider the case where short-sale constraints are imposed.

3.1 Spanning Tests without Short-Sale Constraints

We present first some preliminary evidence on the diversification benefit of the three test asset classes based on the whole sample. Figure 2 plots two efficient frontiers in each panel, one based on the four benchmark assets only (U.S. Equity, U.S. bond, International Equity and International Bond), and the other based on the four benchmark assets plus one test asset class. We can see that addition of each test asset shifts the benchmark efficient frontier noticeably.

To see if the shifts are statistically significant, we test the significance of increases in the Sharpe ratio (of the tangent portfolio) due to such shifts using the Gibbons, Ross, and Shanken (1989) test statistic. Test results, not reported here, indicate that the increase in the Sharpe ratio is statistically significant when adding each of the three test assets.

We then conduct the Huberman and Kandel spanning test using both the full sample and sub-samples and report the test results in Table 3. As can be seen from the table, the results based on the whole sample indicate that none of the three test asset classes can be spanned by the four benchmark asset classes. This is consistent with the result from the Gibbons, Ross, and Shanken test mentioned earlier.

However, observe from the test results based on sub-periods shown in Table 3 that in certain sub-periods especially in more recent quarters, we cannot reject the null that the test assets can be spanned. The asset that displays such a pattern most clearly is TIPS. In this case, the null is rejected soundly with *p*-values less than 0.01% in every quarter (except Q3 2001) in the first half of the entire sample period but cannot be rejected in several quarters in the second half of the period. Overall, the number of quarters with the null rejected at 5% level is 19 out of 28. For commodity and REITs, the number of quarters with the null rejected at 5% level is 10 and 11, respectively. These results provide evidence that diversification benefits of the three test assets are sensitive to time period.

Results reported so far in this subsection are based on portfolios that are constructed using benchmark assets and only one of the test assets each time. In practice, investors may want to include all three test assets in their portfolios. To see if the three test assets are substitutes for each other, we run spanning tests with six benchmark assets, the four original benchmark assets plus two of the three test assets. That is we examine diversification benefits of each test asset when the other two test assets are already included in an investor's portfolio. The test results, shown in Table 4, are similar to what reported in Table 3. In particular, the null that a test asset can be spanned by the benchmark assets (six here) is rejected based on the whole sample but is not rejected in certain sub-periods. Sensitivity of the rejection to time period is also similar to what shown in Table 3. The implication here is that the three test assets are not substitutes for each other and all ought to be included in an investor's portfolio. Nonetheless, the relative importance of the three assets in the portfolio may vary over time. As shown later in the paper, such sensitivity of diversification benefits to time can be explained by time-varying return correlations.

3.2 Spanning Tests with Short-Sale Constraints

Results discussed in Section 3.1 are based on the assumption that short sales are always allowed. In practice, investors sometimes face short-sale constraints. Table 5 reports results from spanning tests with short-sale constraints on both test and benchmark asset classes. Observe first that as before, the null that the benchmark assets span a test asset is still strongly rejected based on the whole sample. Namely, diversification benefits of the three test assets using the whole sample do not disappear when short-sale constraints are incorporated. We also see from the table that results obtained with sub-samples are weaker than those reported in Table 3, in the sense that the number of quarters here where the null is rejected at 5% level is noticeably less than the one shown in Table 3 for each of the test assets. The results still indicate a pattern of time varying diversification benefits of the there test assets.

We also consider the case where short-sale constraints are imposed on test assets only and obtain similar results (not shown in the paper).

3.3 Intersection Tests

Sections 3.1 and 3.2 present results obtained from spanning tests. We also conduct intersection tests with or without short-sale constraints. For brevity, we do not report test results here. Nonetheless, the main implications from the results are the same as those from spanning tests. Namely, diversification benefits of test asset classes are evident based on the whole sample but are sensitive to time within the sample period when sub-samples are used.

4. Time-Variation in Diversification Benefits

In this section we investigate what drives the empirical results presented in the previous section, especially the time sensitivity of diversification benefits of test assets. We first estimate time-varying correlations between asset returns using the DCC model and show that such a correlation structure can explain the time series pattern of diversification benefits documented earlier. We then study asset allocation using correlations based on the DCC model. In particular, we examine how diversification benefits of each test asset class changes over time using two measures of diversification benefits, one based on the increase in the Sharpe ratio of the tangent portfolio and the other based on the change in the tangent portfolio weights. Finally, we examine the out-of-sample performance of portfolio strategies constructed based on a variety of correlation structures in order to shed some light on what correlation estimates to use when doing asset allocation in practice.

4.1 Correlation Estimates Based on the DCC Model

Figure 3 plots correlations between each test asset class with U.S. Equity and U.S. Bond estimated using three alternative methods: the rolling correlation with a-100 day window (in the dotted line on the figure), the unconditional correlation (the solid flat line), and the DCC model (the dark solid line).[2] As we can see from the figure, both DCC and rolling correlations display considerable deviations from the unconditional correlation from time to time during our sample period.

[2] The DCC model is estimated using the UCSD GARCH Toolbox provided by Kelvin Sheppard. See http://www.kevinsheppard.com/research/.

For instance observe from Panel A that although Commodity has a weak negative unconditional correlation with both U.S. Equity and U.S. Bond (-0.0241 and -0.0182, respectively, from Table 1), its DCC and rolling correlations with the two benchmark assets actually fluctuate between -0.4 and 0.4 over the entire sample period.

A closer inspection of the time series behavior displayed in Panel A indicates that time-varying correlations can explain the time variation in diversification benefits of Commodity shown in Table 3. First, the strong rejection of the null in the whole sample reported in the table is consistent with Commodity's weak unconditional correlations with U.S. Equity and U.S. Bond. Secondly, the timing of those sub-periods with a strong rejection of the null (with the p-value less than 0.1%) coincides with the timing of those low DCC estimates for the time-varying correlation between Commodity and U.S. Equity.

Panel B shows the correlations REITs with U.S. Equity and U.S. Bond. Observe that the correlation between REITs and U.S. Equity is always positive but has a positive trend (based on DCC and rolling correlations) over the sample period, increasing from a bit over 0.2 in early 1991 to about 0.7 at the end of 2005. On the other hand, the DCC and rolling correlations between REITs and U.S. Bond fluctuate around zero and more specifically, are initially positive, then stay negative for more than two years, and become positive again since the end of 2003.

Panel C shows the correlations REITs with U.S. Equity and U.S. Bond. Like REITs and U.S. Bond, TIPS and U.S. Equity also have a U-shape correlation structure that fluctuates around zero in terms of both DCC and rolling correlations. The pattern of TIPS's time varying correlations with U.S. Bond is quite striking. They rise gradually from around 0.41 in early 1999 to as high as 0.85 in mid 2003, and stay around that level in the remaining part of the sample period. The implication here is that the market differentiates TIPS from the nominal bonds markedly during the internet bubble period and that after the bubble burst, the two instruments become much more similar as the inflation risk is perceived to be low. This observation also explains the empirical result reported earlier in Table 3 that the null (that TIPS can be spanned) is strongly rejected in the first half of the sample period but is not so in the second half of the period.

In summary, correlations between each of the three test assets and U.S. Equity or U.S. Bond have substantial time variations during the sample period, and such time varying correlations appear to drive the time sensitivity of the test assets' diversification benefits observed earlier.

4.2 Time-Variation of Sharpe Ratios

We now use the increase in the Sharpe ratio of the tangent portfolio due to the addition of a test asset to the (four) benchmark assets as a measure of diversification benefits of the particular test asset. This allows us to quantify the diversification benefit of each test asset and examine its variation over time.

More specifically, we first estimate the mean return for each asset and DCC between each test asset and the benchmark assets using the whole sample. (We do not consider rolling correlations here as they are ad-hoc and their asymptotic properties are not known.) Next, we calculate the Sharpe ratio of the tangent portfolio on a daily basis using daily DCC and volatility series. This is done for portfolios of four benchmark assets first and then for portfolios of the benchmark assets plus each of the three test assets. This gives us four time series of the Sharpe ratio of the tangent portfolio, one for the benchmark portfolio, and three for the benchmark portfolio with the addition of each of the test assets. We then calculate the increase in the Sharpe ratio relative to the series of the benchmark Sharpe ratio for each test asset and obtain three time series of the increase in the Sharpe ratio.

Figure 4 illustrates the three time series of the increase in the Sharpe ratio. We can see that the addition of commodity can increase the Sharpe ratio by the value of 0.1 to 0.5 over the sample period. Compared to commodity, REITs provide a larger diversification benefits as it can raise the Sharpe ratio by as high as 1.5 in some period. While the TIPS can increase the Sharpe ratio by as high as 3 early in the sample period, the diversification benefits decrease gradually after 2001. This pattern is consistent with time variation in TIPS's diversification benefits shown in Table 3. Again, the underlying factor here is the behavior of TIPS's time-varying correlations with U.S. Bond illustrated in Figure 3c. In any case, such dramatic changes of diversification benefits as in the case of TIPS indicate the importance of studying how diversification benefits of non-traditional asset classes vary over time. Finally, notice from Figure 4 that diversification benefits (as measured by the increase in the Sharpe ratio) of the three test asset classes show a tendency of convergence to some extent, a reflection of the fact that correlations among the test assets are getting higher.

Table 6 provides the summary statistics of increases in the Sharpe ratio for each test asset class. In terms of the average increase over the whole sample period, TIPS provide the largest diversification benefit with an average increase of 0.9018. REITs are the second with 0.7643, and Commodity ranks the last with 0.3569. Also, the t-test statistics show that the increase in the Sharpe ratio is significantly greater than zero for all three test asset classes.

4.3 Time-Variation of Tangent Portfolio Weights

Another way to look at the time variation in diversification benefits of test asset classes is to examine how their weights in the tangent portfolio change over time. The implementation here is similar to what is done in Section 4.2. However, here we include all seven assets when constructing portfolios since investors should have access to all three test asset classes in practice. For illustration, we do not calculate the weights on a daily basis but rather take 18 snap-shots of these weights (assuming that investors rebalance their portfolio every 100 days).

Results shown in Figure 5 indicate that TIPS make up the largest portion of the tangent portfolio in the beginning of sample period but the portion become smaller in the later period. In contrast, REITs take up a significant weight across the entire sample period. The weights of Commodity are relatively stable but are much lower relatively to those of REITs and TIPS. This pattern is consistent with what observed in the results from spanning tests, time-varying correlations, and increases in the Sharpe ratio.

4.4 Performance of Portfolio Strategies Using Alternative Correlation Estimates and Rebalancing Frequencies

We have documented that diversification benefits of each test asset class have substantial time variations and so do optimal portfolio weights. However, the results presented so far, in particular, those on the tangent portfolio's Sharpe ratio and weights are based on in-sample estimates. In reality, investors have to estimate parameters using historical data available at the time of estimation. In this section, to see the implications of our analysis for asset allocation in the real world, we construct the optimal portfolio using only the information available at the time of the construction, allow portfolio rebalances to take into account time-varying correlations, and examine the out-of-sample performance of such portfolio strategies. In particular, we run a horserace among four alternative estimates of return correlations. Again, investors are assumed to have access to all seven asset classes considered here.

We now describe how to form our portfolios. First, we use the initial 1,000 observations to construct the first portfolio for a pre-specified (target) expected return by a proper mix of the tangent portfolio and the risk-free asset. After that, the portfolio is rebalanced regularly. Again, only the past information up to the time of rebalancing is used to form a new portfolio.

Given a particular portfolio strategy, we calculate its realized standardized deviation from the targeted return in each rebalancing period and then average these realized standard deviations. We then evaluate the performance of a given portfolio strategy based on its average realized standard deviation. This evaluation criterion is suggested by Engle and Colacito (2003),

We consider four alternative correlation estimates in this exercise. They include DCC, rolling correlation (with a 100-day window), historical correlation, and constant correlation (which is equal to the average of all pair-wise historical correlations).[3] Levels of the target return considered are 5%, 10%, 15%, and 20%. Portfolios are assumed to rebalance every 20, 50, or 100 days.

Table 7 reports the average realized standard deviation of portfolio strategies based on each of the four correlation structures. One observation is that the higher the rebalancing frequency, the better the performance, regardless of the correlation structure used. The intuition here is that more rebalancing leads to more updated adjustment for

[3] The constant correlation is studied in Elton and Gruber (1973), Elton, Gruber, and Ulrich (1978) and Elton, Gruber, and Spitzer (2005).

time-varying correlations. Another observation is that portfolios constructed using DCC have the smallest (average) realized standard deviation from the targeted returns in most cases. Furthermore, the advantage of DCC is larger when the rebalancing frequency is higher. This is not surprising given the nature of DCC.

To get a better sense on the impacts of correlation estimates on asset allocation, we now look at the tangent portfolio's weights. Panels A through D in Figure 6 illustrate the weights of the tangent portfolio under each of the four alternative correlation structures with rebalancing every 100 days. It is easy to see that results obtained from using alternative correlation estimates differ significantly. This finding suggests that investors need to use appropriate correlation estimates to adjust for the time-variation in diversification benefits of the test asset classes.

5. Robustness Checks

Results reported so far are all based on daily return data. To examine the robustness of our test results on the frequency of the return data used, we repeat spanning tests using both weekly and monthly data. The results, shown in Tables 8 and 9, respectively, indicate that the main findings obtained using daily data still hold. For instance, none of the three test assets can be spanned by the benchmark assets in the whole sample and diversification benefits are still time sensitive. In addition, TIPS are shown to be more important in the early period than in the later period.

However, it is worth mentioning that the power of spanning tests decreases noticeably when the frequency of data is lower. Kan and Zhou (2001) show that the power of spanning tests increases with the number of observations used in the test. Therefore, it is possible that our results based on daily data have more power than the results based on weekly and monthly data merely because we have more observations in daily data. A better understanding of this issue is certainly of great interest but is beyond the scope of this paper.

Spanning tests considered in Section 4 are unconditional tests. Here we repeat our empirical studies using conditional spanning tests. There are several methods that can incorporate conditioning information in spanning tests. (See De Roon and Nijman (2001) for a survey on this subject.) For example, Cochrane (1996) and Behaert and Urias (1996) use scaled returns. Alternatively, Shanken (1990) and Ferson and Schadt (1996) assume that the regression coefficients $\bar{a}$ and B in Eq. (1) are a linear function of instruments. We follow the first approach in our conditional spanning tests.

Table 10 reports the results from conditional spanning tests. The instruments used in the tests are also described there. Notice that sub-periods used here are no longer quarters as the number of daily observations in a quarter is not high enough to do a sensible conditional spanning test. We can see that results shown in the table are similar to those from unconditional tests (reported in Table 3). Namely, the main findings are robust under conditional spanning tests.

Finally, we consider an extended sample period. As mentioned earlier, the sample period January 1999 to December 2005 is chosen because the daily return index of International Equity is incepted on January 2, 1999. Since the TIPS were first introduced in the early 1997, most studies that involve TIPS have sample periods starting from 1997. For better comparison with existing studies, we use the MSCI World Equity Price Index (excluding U.S.) as a proxy for International Equity and, as a result, are able to extend our sample period to March 1997. We then repeat the analysis using this extend sample. We find that the unconditional mean and standard deviation of indexes returns are quite sensitive to the length of sample period. However, the main findings of this paper—the patterns of time-varying diversification benefits of test asset classes—still hold. In particular, diversification benefits of TIPS change significantly after 2001. The only exception is that the overall diversification benefit of Commodity is lower using the extended data. To some extent, this result actually reinforces the importance of studying time-variation in diversification benefits of asset classes.

6. Conclusions

In this paper, we examine diversification benefits of three hot asset classes, Commodity, REITs, and TIPS, using daily return data from January 1999 through December 2005. We test for evidence of diversification benefits using Gibbons, Ross, and Shanken (1989) test and the Huberman and Kandel (1987) mean-variance spanning test. We quantify an asset's diversification benefits using the increase in the tangent portfolio's Sharpe ratio after the addition of the asset. We focus on answering two questions: Are the three hot asset classes substitutes for each other, given that they are all considered to be an instrument against inflation? Do diversification benefits of the three asset classes change over time and if so, how to take it into account in asset allocation in practice?

We find that the three asset classes are not substitutes for each other based on the analysis in a full sample. However, there is evidence on the substitution effect in certain sub-periods (quarters) when the analysis is done on a quarterly basis. In most of such cases commodity is redundant. Based on the increase in the tangent portfolio's Sharpe ratio after the addition of an asset and the asset's weight in the tangent portfolio, TIPS dominate REITs which in turn dominate commodity.

We find that diversification benefits of the three asset classes change substantially over time. For instance, TIPS's improvement in the tangent portfolio is very significant before 2001 and then begins to drop gradually since then. The time series behavior of the asset class's weight in the tangent portfolio tells the similar story. We show that the time variation in diversification benefits of the three hot asset classes can be explained by using a time-varying return correlation.

Finally, based on an out-of-sample performance analysis, we find that the Dynamic Conditional Correlation (DCC) model (Engle, 2002) outperforms three other popular correlation structures – the unconditional correlation, the rolling correlation, and the constant correlation. That is, in order to take into account time varying diversification benefits, the best correlation estimate to use in asset allocation is the DCC estimate.

References:

Andrews, D. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858

Bekaert, G., and Urias, M. (1996), "Diversification, Integration, and Emerging Market Closed-End Funds*," Journal of Finance*, 511, 835-870

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327

Cappiello, L., Engle, R., and Sheppard, K. (2003), "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns," ECB Working Paper No. 204

Chen, H., Ho, K., Lu, C., and Wu, C. (2004), "An Asset Allocation Perspective of Real Estate: the Case of Real Estate Investment Trusts," Yuan Ze University working paper

Chun, G., Sa-Aadu, J., and Shilling, D. (2004), "The Role of Real Estate in an Institutional Investor's Portfolio Revisited," *Journal of Real Estate Finance and Economics*, 29, 295-320

Cochrane, J. (1996), "A Cross-sectional Test of An Investment-based Asset Pricing Model," *Journal of Political Economics*, 104, 572-611

De Roon, F., Nijman, T. (2001), "Testing for Mean-Variance Spanning: a Survey", *Journal of Empirical Finance*, 8, 111-155

De Roon, F., Nijman, T., and Werker, B. (2001), "Testing for Mean Variance Spanning with Short Sales Constraints and Transaction Costs: the Case of Emerging Markets," *Journal of Finance*, 56, 721-742

De Santis, G. (1994), "Asset Pricing and Portfolio Diversification: Evidence from Emerging Financial Markets," in Mike Howell, Ed.: *Investing in Emerging Markets*, Euromoney Books, London

Elton, E. and Gruber, M. (1973), "Estimating the Dependence Structure of Share Prices – Implication for Portfolio Selection," *Journal of Finance*, 28, 1203-1232

Elton, E., Gruber, M., and Ulrich T. (1978), "Are Betas Best?" *Journal of Finance*, 23, 1375-1384

Elton, E., Gruber, M., and Blake, C. (2004), "The Adequacy of Investment Choices Offered by 401(k) Plans," *Journal of Public Economics,* Forthcoming

Elton, E., Gruber, M., and Spitzer, J. (2005), "Improved Estimates of Correlation Coefficients and Their Impact on the Optimum Portfolios," New York University working paper

Engle, R. (2002), "Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models," *Journal of Business and Economic Statistics,* 20, 339-350

Engle, R., and Colacito, R. (2003), "Testing and Valuing Dynamic Correlations for Asset Allocation", New York University working paper

Engle, R., and Sheppard, K. (2001), "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH," NBER Working Paper

Ferson, W., Foerster, S., and Keim, D. (1993), "General Tests of Latent Variable Models and Mean-Variance Spanning," *Journal of Finance*, 48, 131-156

Ferson, W., and Schadt, R. (1996), "Measuring Funds Strategy and Performance in Changing Economic Conditions," *Journal of Finance*, 51, 425-462

Gibbons, M., Ross, R., and Shanken, J. (1989), "A Test of the Efficiency of A Given Portfolio", *Econometrica*, 57, 1121-1152

Gorton, G., and Rouwenhorst, K. (2005), "Facts and Fantasies about Commodity Futures," *Financial Analyst Journal*, Forthcoming

Gourieroux, C., Holly, A., and Monfort, A. (1982), "Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on Regressions Parameters," *Econometrica,* 50, 63-80

Harvey, C. (1995), 'Predictable Risk and Returns in Emerging Markets," *Review of Financial Studies*, 8, 773-816

Hansen, L. "Large Sample Properties of Generalized Method of Moments Estimators," *Econometria*, 50, 1029-1054

Huberman, G., and Kandel, S., (1987), "Mean-Variance Spanning," *Journal of Finance*, 42, 873-888

Kan, R., and Zhou, G. (2001), "Tests of Mean-Variance Spanning," Washington University working paper

Kodde, D., and Palm, F. (1986), "Wald Criteria for Jointly Testing Equality and Inequality Restrictions," *Econometria,* 54, 1243-1248

Kothari, S., and Shanken, J. (2004), "Asset Allocation with Inflation-Protected Bonds," *Financial Analysts Journal*, 60, 54-70

Mamun, A., and Visaltanachoti, N. (2004), "Diversification Benefits of Treasury Inflation Protected Securities," Massey University working paper

Newey, W. and West, K. (1987), "A Simple Positive Semi-Definite, Heteroskedasticity, Autocorrelation Consistent Covariance Matrix," *Econometria*, 55, 703-708

Roll, R. (2004), "Empirical TIPS," *Financial Analysts Journal,* 60, 31-53

Shanken, J. (1990), "Intertemporal Asset Pricing: An Empirical Investigation," *Journal of Econometrics*, 45, 99-120

Simon, D., and Hunter, D. (2005), "Are TIPS the 'Real' Deal: A Conditional Assessment of their Role in a Nominal Portfolio," *Journal of Banking and Finance,* 29, 347-368

Table 1: Data Descriptions

This table describes the sources and inception dates of the daily total return indexes used in the paper.

Asset Class	Description	Frequency	Source	Inception Date	End Date
U.S. Equity	S&P500 Index	Daily	Bloomberg	01/04/1988	12/30/2005
U.S. Bond	U.S. Broad Investment Grade (USBIG) Bond Index	Daily	Datastream	12/31/1993	12/30/2005
Int'l Equity	MSCI World Index (Excluding U.S.)	Daily	Bloomberg	01/01/1999	12/30/2005
Int'l Bond	MSCI Sovereign Debt Indices (Excluding U.S.)	Daily	Datastream	12/31/1993	12/30/2005
Commodity	Goldman Sachs Commodity Index (GSCI)	Daily	Datastream	12/31/1969	12/30/2005
REITs	Dow Jones Wilshire REIT Index	Daily	Datastream	01/31/1996	12/30/2005
TIPS	U.S. Inflation-Linked Securities Index (ILSI)	Daily	Datastream	03/03/1997	12/30/2005

Table 2: Summary Statistics on Returns and Correlations

This table reports the summary statistics of the seven daily data series used in this paper. Sample period is from January 1999 to December 2005.

Panel A: Unconditional Mean and Standard Deviation (annualized percentage)

	Mean	Standard Deviation
U.S. Equity	3.46	18.51
U.S. Bond	5.61	4.55
International Equity	6.36	14.76
International Bond	4.12	8.62
Commodity	21.23	22.22
REITs	17.81	13.38
TIPS	8.24	4.70

Panel B: Unconditional Correlation Matrix

	U.S. Equity	U.S. Bond	Int'l Equity	Int'l Bond	Commo-dity	REITs	TIPS
U.S. Equity	1.0000	-0.1758	0.4299	-0.1406	-0.0241	0.4307	-0.1715
U.S. Bond	-0.1758	1.0000	-0.1234	0.2907	-0.0182	-0.0268	0.7412
Int'l Equity	0.4299	-0.1234	1.0000	0.1448	0.0609	0.2603	-0.1399
Int'l Bond	-0.1406	0.2907	0.1448	1.0000	0.0718	-0.0178	0.2931
Commodity	-0.0241	-0.0182	0.0609	0.0718	1.0000	-0.0668	0.0964
REITs	0.4307	-0.0268	0.2603	-0.0178	-0.0668	1.0000	-0.0432
TIPS	-0.1715	0.7412	-0.1399	0.2931	0.0964	-0.0432	1.0000

Table 3: Spanning Tests with Four Benchmark Asset Classes: U.S. Equity, U.S. Bond, International Equity and International Bond

This table reports the Wald-test *p*-values of spanning tests. The asset class being tested is listed in the top of each column. The benchmark asset classes are U.S. Equity, U.S. Bond, International Equity and International Bond. We conduct spanning tests for each quarter in our sample period. The last row reports the spanning test results for the whole sample period. Newey-West standard errors are used to control for heteroskedasticity and autocorrelation. Daily data from January 1999 to December 2005 are used.

Test Period	*p*-values for Each Test Asset Class		
	Commodity	REITs	TIPS
1999 Q1	**0.0340**	**0.0010**	**<.0001**
1999 Q2	**0.0064**	0.0587	**<.0001**
1999 Q3	**<.0001**	**0.0006**	**<.0001**
1999 Q4	**0.0251**	0.3107	**<.0001**
2000 Q1	0.0731	0.0763	**<.0001**
2000 Q2	**0.0001**	**<.0001**	**<.0001**
2000 Q3	0.3099	0.2348	**<.0001**
2000 Q4	0.6239	0.8025	**<.0001**
2001 Q1	0.0631	**<.0001**	**<.0001**
2001 Q2	0.2897	**0.0002**	**<.0001**
2001 Q3	0.3641	0.8443	0.0672
2001 Q4	0.0571	**0.0012**	**0.0002**
2002 Q1	0.1230	**0.0103**	**<.0001**
2002 Q2	**0.0266**	0.1253	**<.0001**
2002 Q3	**0.0260**	0.1236	**0.0033**
2002 Q4	**0.0141**	0.6798	0.2019
2003 Q1	0.8691	0.5288	0.6480
2003 Q2	0.2643	0.1022	**0.0080**
2003 Q3	0.3314	**0.0040**	**0.0019**
2003 Q4	0.8378	0.0959	0.4767
2004 Q1	0.1443	0.2329	0.3124
2004 Q2	0.6520	**0.0130**	0.3414
2004 Q3	**0.0282**	0.3769	0.9149
2004 Q4	0.7064	0.2358	**0.0118**
2005 Q1	**0.0195**	**<.0001**	**<.0001**
2005 Q2	0.7368	**0.0447**	0.3610
2005 Q3	0.3369	0.1011	**0.0058**
2005 Q4	0.2523	0.0901	0.2596
Whole Sample	**<.0001**	**<.0001**	**<.0001**
Number of quarters with $p < 5\%$	10	11	19
Percentage of quarters with $p < 5\%$	36%	39%	68%

Table 4: Spanning Tests with Six Benchmark Asset Classes: U.S. Equity, U.S. Bond, International Equity, International Bond, and Two of the Three Test Asset Classes

This table reports the Wald-test *p*-values of spanning tests. The asset class being tested is listed in the top of each column. The benchmark (spanning) asset classes are U.S. Equity, U.S. Bond, International Equity, International Bond, and two of the three test asset classes (for example, when Commodity is the test asset class, we also include REITs and TIPS in the spanning asset classes). We conduct spanning tests for each quarter in our sample period. The last row reports the spanning test results for the whole sample period. Newey-West standard errors are used to control for heteroskedasticity and autocorrelation. Daily data from January 1999 to December 2005 are used.

Test Period	*p*-values for Each Test Asset Class		
	Commodity	REITs	TIPS
1999 Q1	0.2760	**0.0001**	**<.0001**
1999 Q2	0.6927	0.1375	**<.0001**
1999 Q3	**<.0001**	**0.0006**	**<.0001**
1999 Q4	0.6431	0.3133	**<.0001**
2000 Q1	0.1797	0.0850	**<.0001**
2000 Q2	**0.0109**	**0.0110**	**<.0001**
2000 Q3	0.9749	0.4055	**<.0001**
2000 Q4	0.9378	0.1361	**<.0001**
2001 Q1	0.1518	**0.0006**	**<.0001**
2001 Q2	0.4693	**0.0024**	**<.0001**
2001 Q3	0.5255	0.2833	**0.0412**
2001 Q4	**0.0349**	**0.0008**	**0.0022**
2002 Q1	0.1141	0.0635	**<.0001**
2002 Q2	0.2100	0.0813	**<.0001**
2002 Q3	0.1669	0.1445	0.1286
2002 Q4	**0.0115**	0.9000	0.2103
2003 Q1	0.8000	0.3372	0.6820
2003 Q2	0.7213	**0.0135**	**0.0003**
2003 Q3	0.1947	**0.0474**	**0.0026**
2003 Q4	0.7140	0.0915	0.2624
2004 Q1	0.0985	0.1988	0.3006
2004 Q2	0.3911	**0.0125**	0.0887
2004 Q3	**0.0128**	0.2883	0.7153
2004 Q4	0.4178	0.2787	**0.0074**
2005 Q1	**0.0025**	**0.0008**	**<.0001**
2005 Q2	0.6859	**0.0316**	0.3137
2005 Q3	0.7619	0.2482	**0.0414**
2005 Q4	0.2818	0.1346	0.1874
Whole Sample	**<.0001**	**<.0001**	**<.0001**
Number of quarters with $p < 5\%$	6	11	19
Percentage of quarters with $p < 5\%$	21%	39%	68%

Table 5: Spanning Tests with Short-Sale Constraints

This table reports the p-values of spanning tests with short-sale constraints on both test asset classes and benchmark asset classes. The asset class being tested is listed in the top of each column. The benchmark asset classes are U.S. Equity, U.S. Bond, International Equity, International Bond. We conduct spanning tests for each quarter in our sample period. The last row reports the test results using the whole sample. p-values are calculated using numerical simulation. Daily data from January 1999 to December 2005 are used.

Test Period	p-values for Each Test Asset Class		
	Commodity	REITs	TIPS
1999 Q1	0.2620	0.8756	0.8051
1999 Q2	0.7057	0.0764	0.0540
1999 Q3	**0.0009**	0.8713	0.8472
1999 Q4	0.8099	0.8824	0.8680
2000 Q1	0.2128	0.8259	**<.0001**
2000 Q2	**0.0140**	**0.0077**	0.1239
2000 Q3	0.8553	0.2820	0.0837
2000 Q4	0.7573	0.5948	**0.0013**
2001 Q1	0.8799	0.7375	**0.0226**
2001 Q2	0.8725	**0.0042**	0.4377
2001 Q3	0.8738	0.7782	0.8652
2001 Q4	0.8889	0.7516	0.8663
2002 Q1	0.1850	**0.0085**	0.0724
2002 Q2	0.7944	0.3764	**<.0001**
2002 Q3	**0.0447**	0.8853	**0.0288**
2002 Q4	0.7515	0.8686	0.8685
2003 Q1	0.8610	0.7651	0.8048
2003 Q2	0.8684	0.2800	0.8810
2003 Q3	0.8546	0.0992	0.3185
2003 Q4	0.7823	0.7993	0.3143
2004 Q1	0.5471	**0.0210**	0.1856
2004 Q2	0.8175	0.8558	0.8376
2004 Q3	0.1276	0.4556	0.8694
2004 Q4	0.8639	0.1809	0.4338
2005 Q1	**0.0225**	0.8697	0.6007
2005 Q2	0.8789	**0.0014**	0.8600
2005 Q3	0.1966	0.8543	0.3569
2005 Q4	0.8849	0.8814	0.8698
Whole Sample	**0.0067**	**<.0001**	**<.0001**
Number of quarters with $p < 5\%$	4	5	5
Percentage of quarters with $p < 5\%$	14%	18%	18%

Table 6: Summary Statistics on Difference in Sharpe Ratios

This table reports the summary statistics of the difference in Sharpe ratios with and without each test asset class. For each test asset class, we use the DCC estimates and mean returns to calculate the Sharpe ratio of tangent portfolios of (1) benchmark asset classes only and (2) benchmark asset classes and the test asset class. Difference in Sharpe Ratio is the difference of these two Sharpe ratios (with and without each test asset class). We use 3-month U.S. Treasury yield as the risk-free rate. Daily data from January 1999 to December 2005 are used.

Test Asset	Mean	Standard Deviation	Median	Min	Max
Commodity	0.3568***	0.0948	0.3653	0.0711	0.5466
REITs	0.7643***	0.3098	0.7886	0.0630	1.4873
TIPS	0.9018***	0.7188	0.5722	0.0963	3.1120

*** Significantly larger than zero at one percent level.

Table 7: Empirical Performance of Portfolio Strategies using Alternative Correlation Estimates and Rebalancing Frequencies

This table reports the average realized volatilities (annualized percentage) of portfolio strategies using alternative correlation estimates and rebalancing frequencies. Correlation estimates used are DCC estimates, rolling correlation (100-day), historical correlation, and constant correlation (all correlation is equal to the average pair-wise historical correlation). The first 1000 observations are used to make initial estimation and form the first investment portfolios. The portfolios are then rebalanced every 20, 50, or 100 days. All historical data from the start of sample period to the time of rebalancing are assumed to be available in making rebalancing decisions. All seven asset classes are included in each strategy. Daily data from January 1999 to December 2005 are used.

Panel A: Realized Standard Deviation – Rebalancing every 20 days

	Average Realized Standard Deviation (annualized %)			
Targeted Return (annualized %)	DCC	Rolling Correlation	Historical Correlation	Constant Correlation
5	1.85	1.91	1.94	2.13
10	4.69	4.81	4.89	5.33
15	7.39	7.58	7.70	8.39
20	9.99	10.23	10.40	11.32

Panel B: Realized Standard Deviation- Rebalancing every 50 days

	Average Realized Standard Deviation (annualized %)			
Targeted Return (annualized %)	DCC	Rolling Correlation	Historical Correlation	Constant Correlation
5	1.97	2.06	2.04	2.23
10	4.90	5.08	5.05	5.48
15	7.70	7.96	7.93	8.59
20	10.37	10.72	10.68	11.57

Panel C: Realized Standard Deviation - Rebalancing every 100 days

	Average Realized Standard Deviation (annualized %)			
Targeted Return (annualized %)	DCC	Rolling Correlation	Historical Correlation	Constant Correlation
5	2.16	2.21	2.15	2.33
10	5.24	5.34	5.24	5.64
15	8.19	8.34	8.18	8.80
20	11.01	11.21	11.01	11.83

Table 8: Spanning Tests Using Weekly and Monthly Data —No Short-Sale Constraints

This table reports the Wald-test *p*-values of spanning tests. The asset class being tested is listed in the top of each column. The benchmark asset classes are U.S. Equity, U.S. Bond, International Equity and International Bond. In Panel A, the data frequency is weekly. We conduct spanning tests for each year in our sample period. The last row reports the spanning test results for the whole sample period. In Panel B, the data frequency is monthly. Newey-West standard errors are used to control for heteroskedasticity and autocorrelation. The sample period is from January 1999 to December 2005.

Panel A: Results based on Weekly Data

Test Period	*p*-values for Each Test Asset Class		
	Commodity	REITs	TIPS
1999	**0.0332**	**0.0003**	**<.0001**
2000	**<.0001**	**0.0026**	**<.0001**
2001	0.1135	0.0780	**0.0083**
2002	**0.0019**	0.9611	0.3494
2003	0.4141	**0.0063**	0.1272
2004	0.3878	0.1117	0.4526
2005	0.9468	**0.0081**	0.0544
Whole Sample	**0.0434**	**0.0025**	**<.0001**

Panel B: Results based on Monthly Data

Test Period	*p*-values for Each Test Asset Class		
	Commodity	REITs	TIPS
Whole Sample	0.0836	**0.0333**	0.0508

Table 9: Spanning Tests Using Weekly and Monthly Data —with Short-Sale Constraints

This table reports the *p*-values of spanning tests with short-sale constraints on both test asset classes and benchmark asset classes. The asset class being tested is listed in the top of each column. The benchmark asset classes are U.S. Equity, U.S. Bond, International Equity, International Bond. In Panel A, the data frequency is weekly. We conduct spanning tests for each year in our sample period. The last row reports the spanning test results for the whole sample period. In Panel B, the data frequency is monthly. *p*-values are calculated by numerical simulation. The sample period is from January 1999 to December 2005.

Panel A: Results based on Weekly Data

Test Period	*p*-values for Each Test Asset Class		
	Commodity	REITs	TIPS
1999	0.0688	0.8740	0.2050
2000	**0.0001**	**0.0098**	**<.0001**
2001	0.8890	0.0584	0.8221
2002	**0.0097**	0.7875	**0.0256**
2003	0.6007	0.0804	0.2679
2004	0.6554	0.3164	0.2307
2005	0.7830	0.4724	0.8591
Whole Sample	**0.0164**	**<.0001**	**<.0001**

Panel B: Results based on Monthly Data

Test Period	*p*-values for Each Test Asset Class		
	Commodity	REITs	TIPS
Whole Sample	0.0543	**0.0013**	**0.0036**

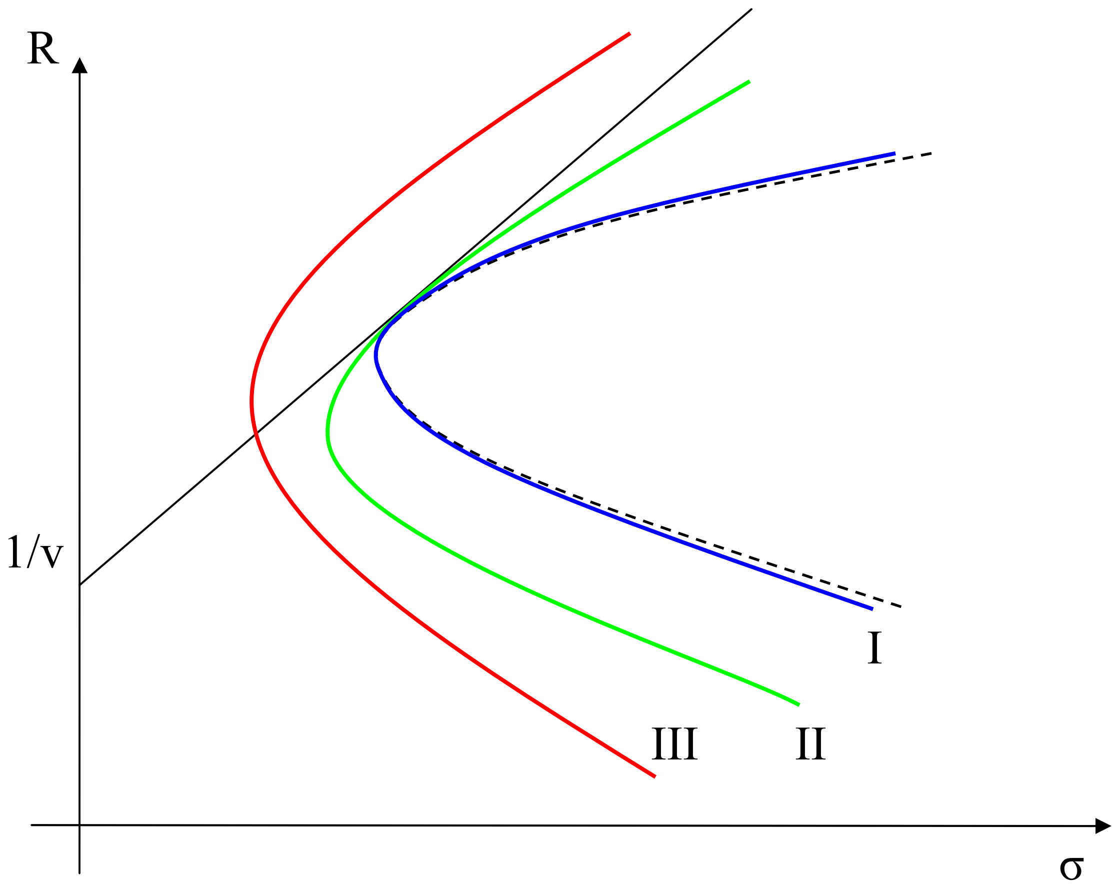
Table 10: Conditional Spanning Tests

This table reports the p-values of conditional spanning tests. The asset class being tested is listed in the top of each column. The benchmark asset classes are U.S. Equity, U.S. Bond, International Equity and International Bond. The model is:

$$\vec{e}_t = \vec{r}_t - \mathrm{B}\vec{R}_t$$
$$\mathrm{B}\vec{i}_K = \vec{i}_N$$

where N×1 vector $\vec{r}_t$ denotes the returns of N test asset classes, the K×1 vector $\vec{R}_t$ denotes the returns of K benchmark asset classes, B is an N×K matrix, and $\vec{e}_t$ is an N×1 vector. The orthogonality condition is $E\left(\vec{e}_t[\vec{R}_t, \vec{Z}_{t-1}]\right) = 0$, where $\vec{Z}_{t-1}$ is a L×1 vector of instrumental variables. The instruments we used are: a constant, the lagged returns of benchmark asset classes, the U.S. short-term risk-free rate (3-month U.S. Treasury constant maturity) and the yield curve slope of U.S. Treasury rates (10-year minus 2-year). We conduct the Hansen's (1982) over-identification test using data for each year in our sample period. The last row reports the spanning test results for the whole sample period. Daily data from January 1999 to December 2005 are used.

Test Period	*p*-values for Each Test Asset Class		
	Commodity	REITs	TIPS
1999	**0.0147**	**0.0099**	**<.0001**
2000	**0.0114**	0.0563	**<.0001**
2001	0.0871	**0.0015**	**<.0001**
2002	**0.0188**	0.3380	**0.0009**
2003	0.3021	**0.0087**	**0.0064**
2004	0.1787	0.0891	0.1597
2005	0.3667	**0.0196**	**0.0102**
Whole Sample	**<.0001**	**<.0001**	**<.0001**



--- Mean-variance efficient frontier of benchmark asset classes

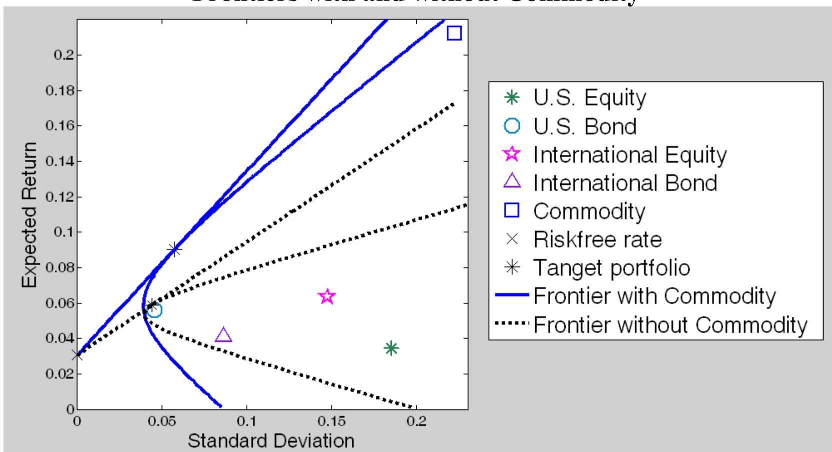
I. Benchmark asset classes span test asset classes and benchmark asset classes

II. Benchmark asset classes "intersect" test asset classes, benchmark asset classes, and risk-free asset, 1/v

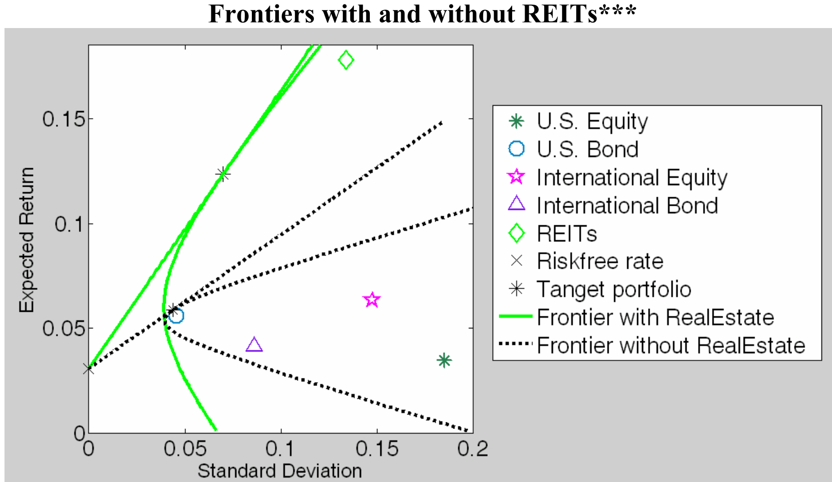
III. Benchmark asset classes do not span test asset classes and benchmark asset classes

Figure 1. Illustrations of Spanning and Intersection Tests. The dotted line represents the frontier of benchmark asset classes only. The solid lines denoted by I, II, III, represent different cases of frontiers generated by the test and benchmark asset classes.

Frontiers with and without Commodity*



Frontiers with and without REITs***



Frontiers with and without TIPS***

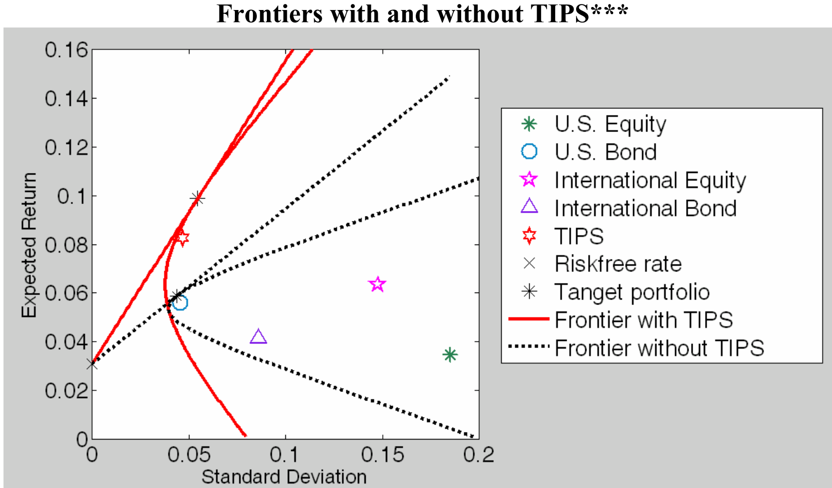


Figure 2. Efficient Frontiers with and without Each Test Asset Class. We use 3-month U.S. Treasury yield as the risk-free rate. Daily data from January 1999 to December 2005 are used.

* GRS test statistic for the increase in the Sharpe ratio due to the addition of a test asset class is significant at 5% level.
*** GRS test statistic for the increase in the Sharpe ratio due to the addition of a test asset class is significant at 1% level.

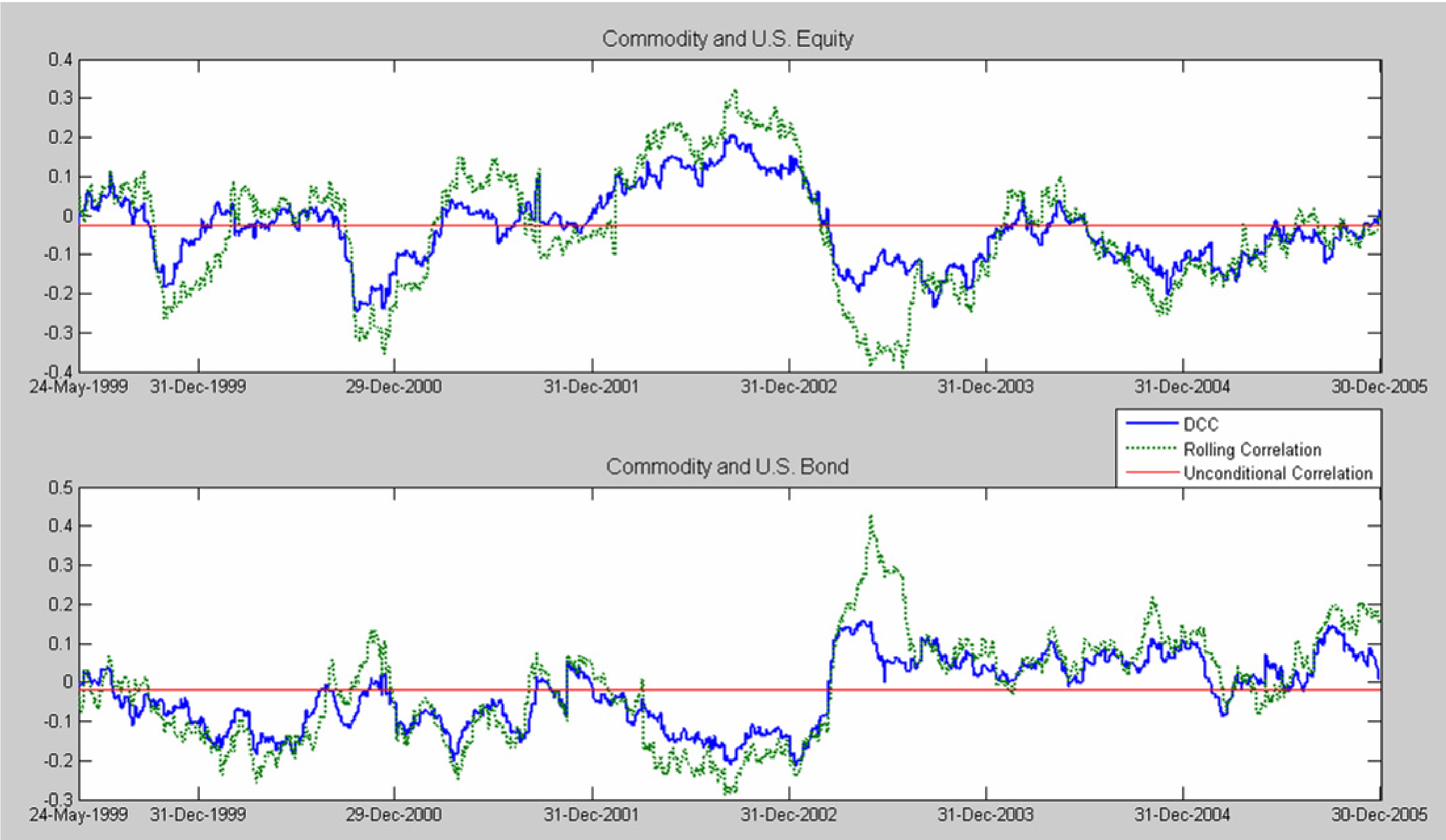


Figure 3a. DCC, Rolling Correlations (with a 100-day window), and Unconditional Correlations of Commodity with both U.S. Equity and U.S. Bond over Time. Daily data from January 1999 to December 2005 are used.

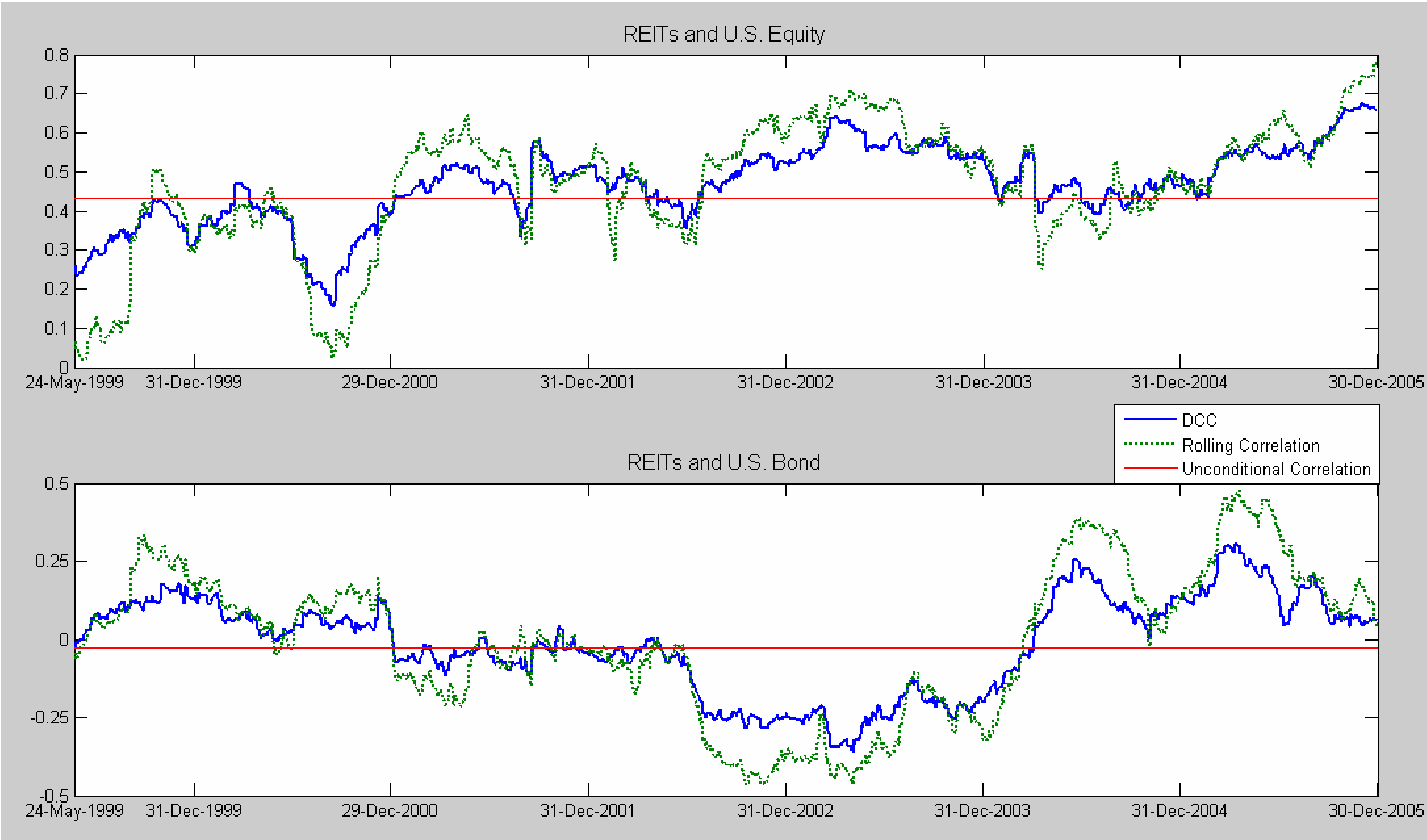


Figure 3b. DCC, Rolling Correlations (with a 100-day window), and Unconditional Correlations of REITs with both U.S. Equity and U.S. Bond over Time. Daily data from January 1999 to December 2005 are used.

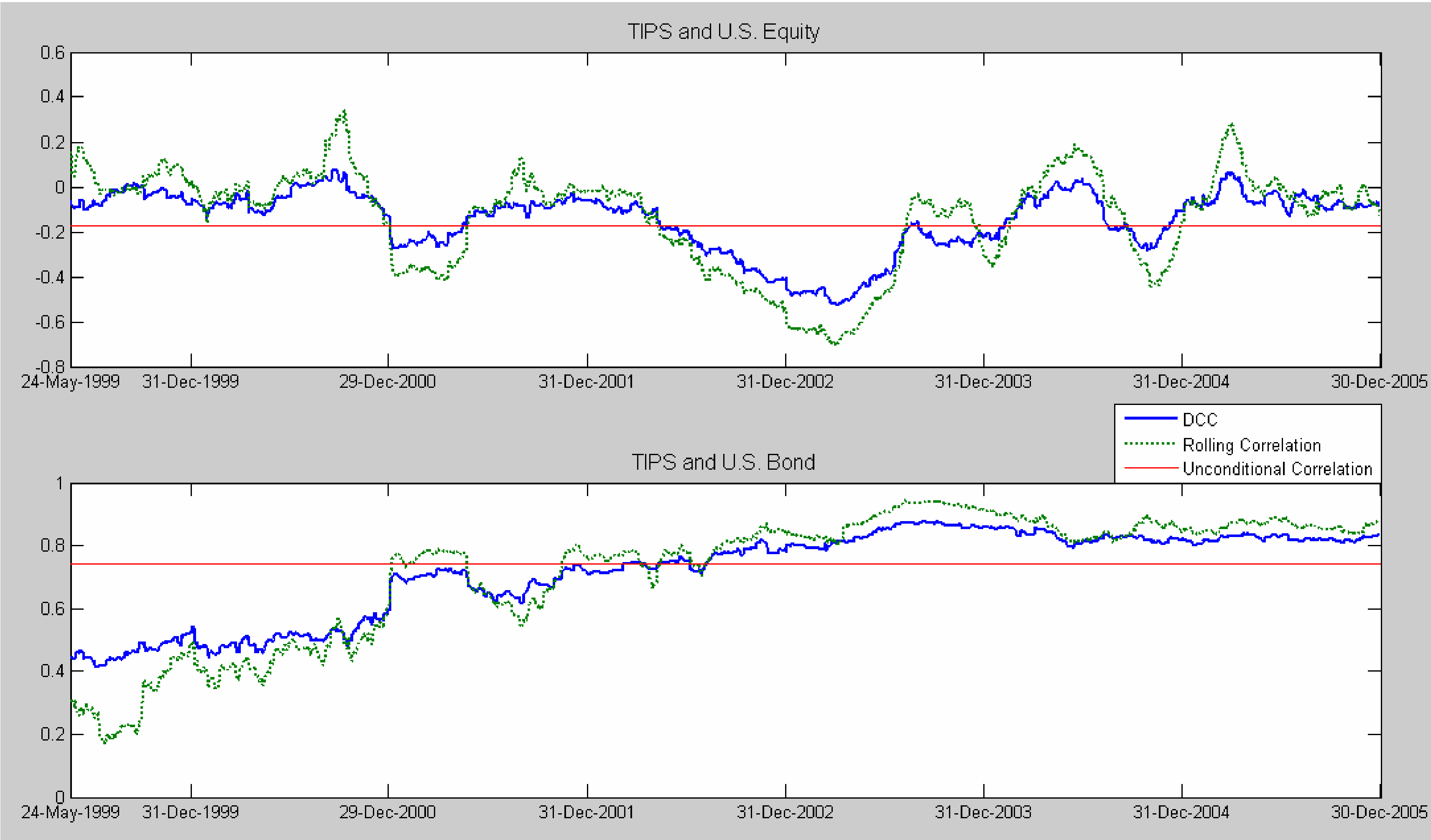


Figure 3c. DCC, Rolling Correlations (with a 100-day window), and Unconditional Correlations of TIPS with both U.S. Equity and U.S. Bond over Time. Daily data from January 1999 to December 2005 are used.

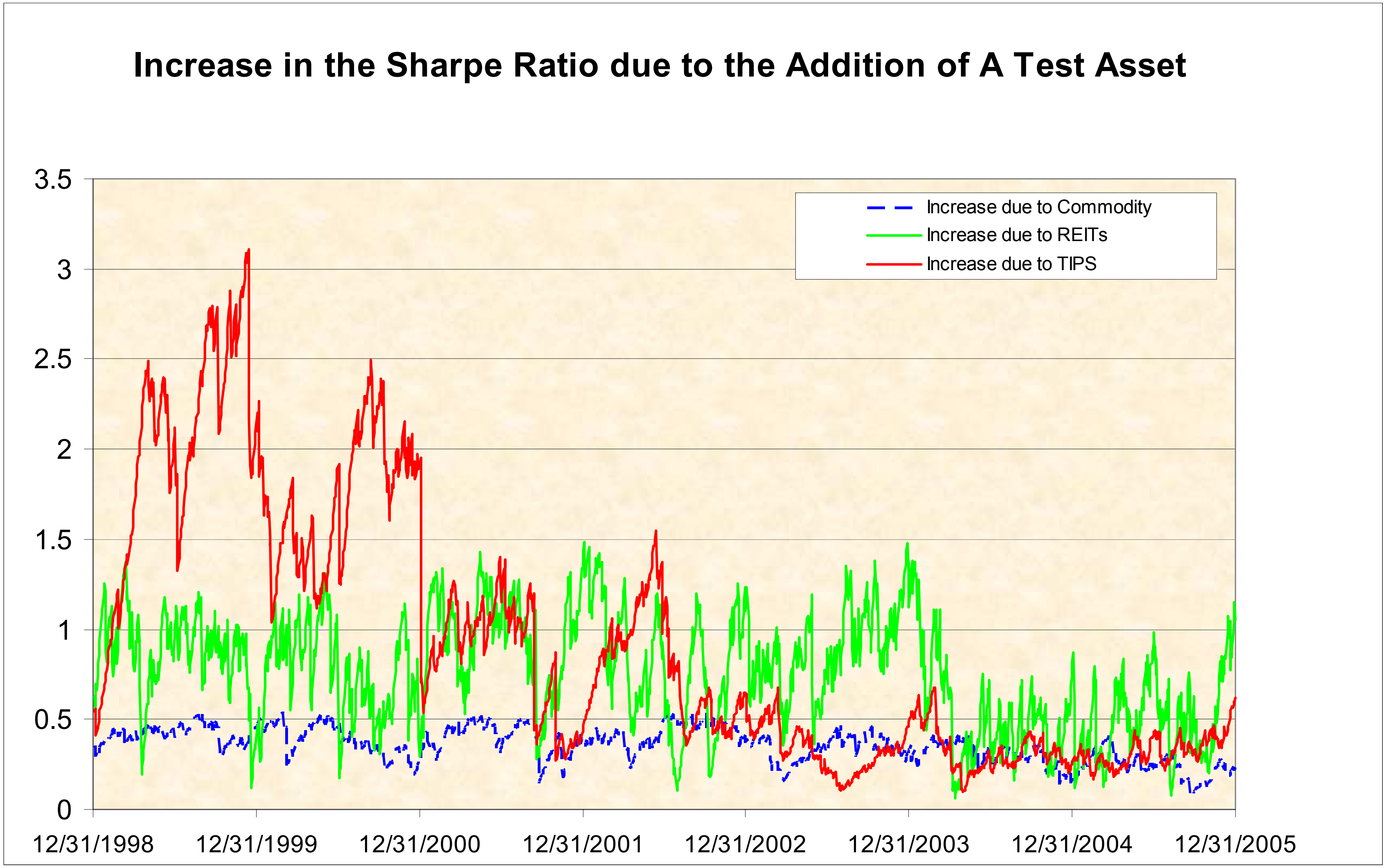


Figure 4. Changes in the Sharpe Ratio of the Tangent Portfolio due to the Addition of A Test Asset. The benchmark portfolio includes the (four) benchmark assets. We then add a test asset and calculate its impact on the Sharpe ratio of the tangent portfolio. Correlation estimates are based on DCC. We use 3-month U.S. Treasury yield as the risk-free rate. Daily data from January 1999 to December 2005 are used.

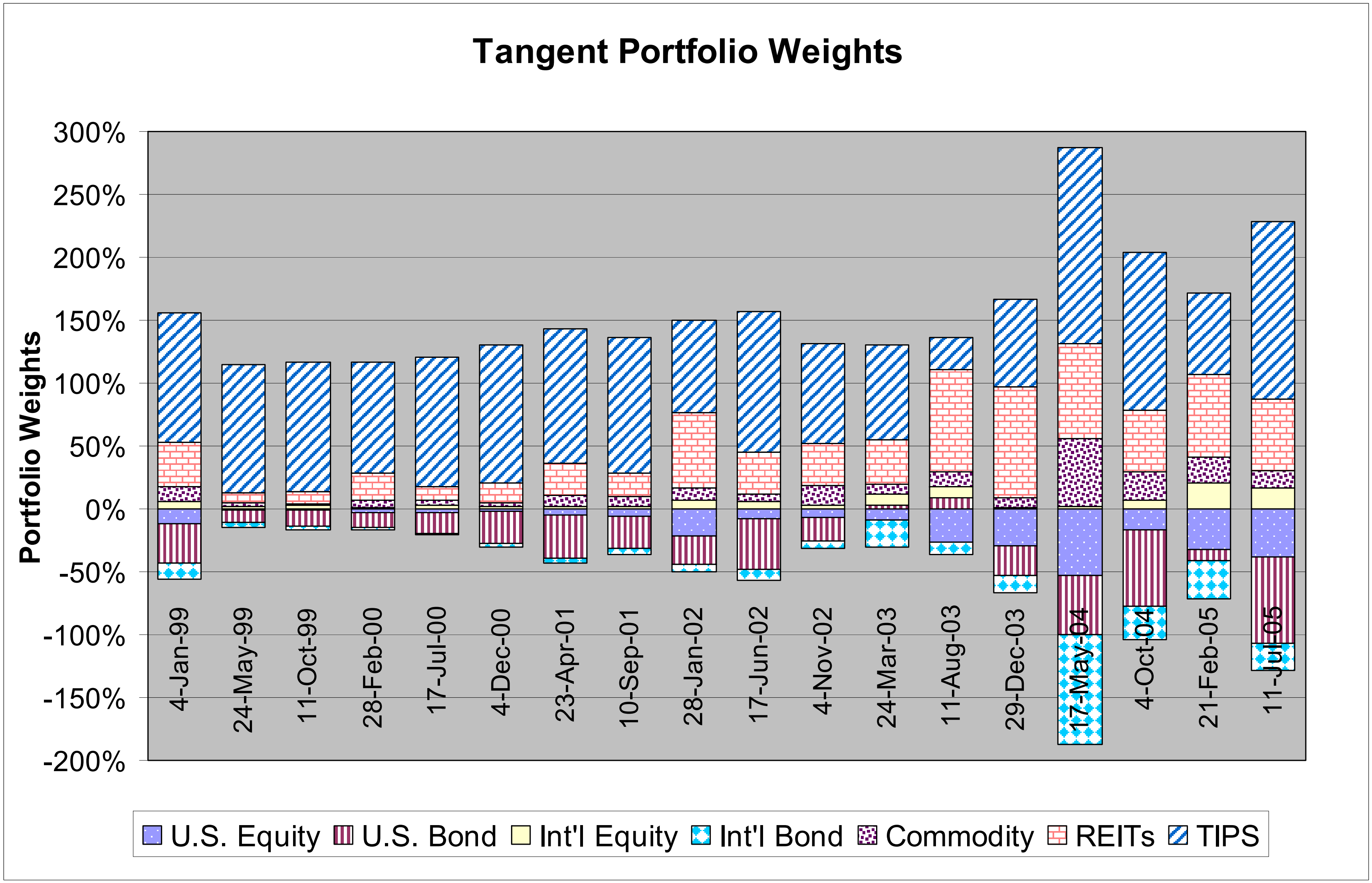


Figure 5. Tangent Portfolios Weights. Shown here are the portfolio weights of the tangent portfolio constructed using all seven asset classes. Correlation estimates used are based on DCC. We use 3-month U.S. Treasury yield as the risk-free rate. Daily data from January 1999 to December 2005 are used.

Figure 6. Tangent Portfolio Weights of Portfolio Strategies using Alternative Correlation Estimates. The weights of tangent portfolios with all seven asset classes are calculated using alternative correlation estimates on each rebalancing day. The first 1,000 observations are used to estimate and form the first investment portfolio. Portfolios are rebalanced every 100 days. Only historical data up to the time of rebalancing are used for estimation when a portfolio is rebalanced. We use 3-month U.S. Treasury yield as the risk-free rate. Daily data from January 1999 to December 2005 are used.