2014

Structural Breaks and Portfolio Performance in Global Equity Markets

Harry J. Turtle
West Virginia University

Chengping Zhang
George Fox University, czhang@georgefox.edu

Structural breaks and portfolio performance in global equity markets

HARRY J. TURTLE† and CHENGPING ZHANG‡*

†College of Business and Economics, West Virginia University, Morgantown, WV, USA
‡College of Business, George Fox University, Newberg, OR, USA

1. Introduction

Traditional asset pricing models such as the Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966); Merton's (1973) Intertemporal Capital Asset Pricing Model (ICAPM) and Fama and French's three-factor model (1992, 1993, 1996) often require the assumption of stationary return distributions. Unfortunately, the extant literature provides substantial evidence that financial time series might display structural changes.§ For example, using a long history of aggregate stock returns in a Bayesian framework that incorporates uncertainty about structural break timing, Pastor and Stambaugh (2001) find that the equity premium exhibits a sharp decline during the 1990s. With regard to most emerging markets, Garcia and Ghysels (1998) find that the parameter estimates of a CAPM with world factors show significant structural instability. Bekaert *et al*. (2002) also provide strong evidence of structural breaks in emerging equity markets due to increased world market integration. More recently, structural break analysis by Berger *et al*. (2011) find that pre-emerging or frontier, equity markets remained segregated from other markets over time.

In general, structural breaks produce erroneous inferences and portfolio decisions due to model misspecification. As a simple example, a structural shift in second moments will produce a change in asset betas that might result in a spuriously significant Jensen's alpha, even when the true model displays no marginal performance benefit within regimes. Hillebrand (2005) documents this issue and shows that the sum of estimated autoregressive parameters will be heavily biased towards one in a GARCH model if parameter shifts are not considered. Andreou and Ghysels (2008) point out that ignoring the presence of structural breaks can have costly effects on financial risk management and may produce faulty inferences regarding credit risk. Recently, Pettenuzzo and Timmermann (2011) find structural breaks cause model instability that impacts prediction, optimal asset allocations and resultant investor utility.

*Corresponding author. Email: czhang@georgefox.edu
§See Andreou and Ghysels (2009) for an excellent review of structural breaks in financial time series.

Structural breaks can often be used to describe general financial market conditions. For example, Meligkotsidou and Vrontos (2008) find structural breaks in the hedge fund return series around the internet boom and the Long-Term Capital Management crisis. In addition, Gerlach *et al*. (2006) document a structural shift in the integration of Asia-Pacific real estate markets due to the 1997 Asian financial crisis. The examination of structural breaks in international financial markets remains important because studies by Hardouvelis *et al*. (2006), Carrieri *et al*. (2007), and Pukthuanthong and Roll (2009) find evidence that integration of global stock markets has increased over time.

The potential benefits of adding international investments to a well-diversified US-based portfolio have been debated since the seminal work of Solnik (1974b). For example, Harvey (1995) finds that including emerging market assets in a mean-variance efficient portfolio significantly increases expected returns and reduces volatility from March 1986 to June 1992. Bekaert and Urias (1996) show that, over a similar timeframe, emerging market funds in the UK provide diversification benefits whereas US emerging market funds do not. De Santis and Gerard (1997) estimate that the expected gain from international diversification is more than 2% per year from 1970 to 1994. De Roon *et al*. (2001) find strong evidence of diversification benefits from investing in global emerging markets when market frictions are not taken into consideration, but the benefits largely diminish with short-sale constraints and transaction costs. In contrast, Li *et al*. (2003) show that international diversification benefits remain substantial for US investors even with short selling constraints. Lewis (2007) finds that diversification benefits, either from investing in foreign equities directly or in American Depository Receipts (ADRs) traded in the US, are diminishing due to increased world equity market integration.

In this paper, our goal is to specify an empirical model that admits linkages in the structural relationships between global equity markets. Using a general multivariate regression model with unknown structural breaks, we examine the diversification benefits in international market portfolios when global markets are subject to common structural changes. We apply the spanning tests of Kan and Zhou (2012) to disentangle the portfolio benefits available from various asset classes.

We find two structural breaks (with three resultant regimes) in global equity markets from January 1988 through January 2010. These estimated breaks and their 95% confidence intervals are comparable across multiple beta risk models. We show that emerging markets provide significant marginal portfolio performance benefits to a well-diversified global investor prior to 1997 and post 2003 (post 2006 in our smallest model).

We also find that emerging markets significantly improve the mean-variance frontier relative to a broad set of benchmark assets that include the Morgan Stanley Capital International (MSCI) US equity index, the MSCI developed market index and US small, big, value and growth portfolios. In particular, Latin American emerging market investments improve the location of the tangency portfolio from January 1988 to March 1997, the minimum variance portfolio from April 1997 to March 2003 and both the minimum variance portfolio and the tangency portfolio from April 2003 to January 2010.

Our analysis of structural breaks in global markets suggests that disparate results observed in previous studies might be due to differences in sample periods, the exclusion of particular markets, or both. We document the importance of including even a small number of strongly performing markets within a given sample and show that emerging markets have not fully harmonized with other global markets. Unrecognized inclusion of multiple market regimes in previous studies might be the cause of recent findings of diminishing benefits to diversification.

Overall, we find a continued role for global diversification that does not uniformly diminish over time. Our results suggest that multivariate return moments change over regimes, and that harmonization of global markets is not consistently demonstrated within regimes. For example, although US and emerging market sample means display similar patterns as we move from the first to second regime, the third regime shows a dramatic increase in emerging market index returns as US market returns fall. Intuitively, the positive relationship between means over the initial two regimes is consistent with integration of global markets, whereas the final period may be more indicative of a structural break in the relationship between the US and emerging markets. The difference is important for portfolio managers because integration suggests investment policy might be better structured solely within domestic markets, whereas a structural break interpretation suggests global market diversification remains important for asset allocation decisions.

The remainder of the paper is organized as follows. In Section 2, we review the estimation and testing procedure for unknown breaks in a multivariate regression model and describe our mean-variance spanning test procedure. The data are described in Section 3. Section 4 presents our empirical results. Conclusions are stated in Section 5.

2. Econometric methodology

We consider potential structural breaks in global financial markets comprised of US, developed and emerging markets. We examine the ability of regional emerging market returns to improve the performance of a well-diversified portfolio of US domestic holdings and developed market equities. We use the spanning tests of Kan and Zhou (2012) to analyse the potential of emerging markets to improve portfolio performance.

2.1. Structural break testing and estimation

The literature on testing, estimation and application of structural break models is voluminous. Initial tests of structural changes were restricted to the case of a single break in a univariate regression model (see e.g. Andrews 1993). Later, Bai and Perron (1998, 2003, 2006) studied multiple break points in a univariate regression, while Bai *et al.* (1998) considered a single break in a multivariate context. Lien *et al.* (2003) identify structural changes in the Nikkei spot index and futures price using a nonparametric method. Berger *et al.* (2011) examine various frontier markets using univariate Quandt–Andrews breakpoint tests. Furno (2012) proposes a robust quantile regression test for structural breaks.

We adopt the generalized framework of Qu and Perron (2007) who present the estimation, inference and computation of multiple potential structural breaks using linear multivariate regression models. Our initial multivariate linear specification includes the US market excess return, developed market excess return and emerging market excess return given two risk factors (global market risk and foreign exchange risk). With no structural breaks, our system is written as

$$\begin{bmatrix} r_{USt} \\ r_{DMt} \\ r_{EMt} \end{bmatrix} = \begin{bmatrix} \alpha_{USi} \\ \alpha_{DMi} \\ \alpha_{EMi} \end{bmatrix} + \begin{bmatrix} \beta_{USi} \\ \beta_{DMi} \\ \beta_{EMi} \end{bmatrix} r_{mt} + \begin{bmatrix} \gamma_{USi} \\ \gamma_{DMi} \\ \gamma_{EMi} \end{bmatrix} r_{FXt} + \begin{bmatrix} \varepsilon_{USit} \\ \varepsilon_{DMit} \\ \varepsilon_{EMit} \end{bmatrix}; \qquad \boldsymbol{\varepsilon}_{it} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_i) \tag{1}$$

for $\tau_{i-1} \leq t \leq \tau_i$ for $i = 1, \ldots, M+1$, $\tau_0 = 1, \tau_{M+1} = T$; and where r_{USt}, r_{DMt} and r_{EMt} are monthly excess returns to the MSCI US market, developed market and emerging market indices at time t, respectively; r_{mt} is the monthly excess return of the world benchmark portfolio, proxied by the excess return on the MSCI All Country World Index; and r_{FXt} is the foreign exchange risk factor, proxied by the excess return on a trade-weighted foreign exchange portfolio.

For succinctness, we define $\boldsymbol{r}_t = [r_{USt}, r_{DMt}, r_{EMt}]'$, $\boldsymbol{X}_t = [1, \ r_{mt}, r_{FXt}]'$, $\boldsymbol{\Gamma}_i = [\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \boldsymbol{\gamma}_i]$, $\boldsymbol{\alpha}_i = [\alpha_{USi}, \alpha_{DMi}, \alpha_{EMi}]'$, $\boldsymbol{\beta}_i = [\beta_{USi}, \beta_{DMi}, \beta_{EMi}]'$, and $\boldsymbol{\gamma}_i = [\gamma_{USi}, \gamma_{DMi}, \gamma_{EMi}]'$. Similarly, the vector $\boldsymbol{\varepsilon}_{it} = [\varepsilon_{USit}, \varepsilon_{DMit}, \varepsilon_{EMit}]'$ is a 3×1 vector of disturbances that varies by date within regimes, with a regime-specific variance–covariance matrix, $\boldsymbol{\Sigma}_i$. We consider M structural changes in the system and denote the break dates by the vector $\tau = (\tau_i, \ldots, \tau_M)$. For a given number of M structural breaks, we have $M+1$ regimes.† Our general model admits potential shifts in intercepts, slope coefficients and variance–covariance elements.

Conditional upon a specific number of structural breaks, M, and the definitions above, equation (1) may be written in matrix form as

$$\boldsymbol{r}_t = \boldsymbol{\Gamma}_i \boldsymbol{X}_t + \boldsymbol{\varepsilon}_{it}. \tag{2}$$

Qu and Perron (2007) provide a testing procedure to examine the null hypothesis of no structural breaks against an alternative hypothesis of an unknown number of structural breaks, given an upper bound of M structural breaks. The Qu and Perron test statistic may be written as

$$supLR_T = 2[\log \hat{L}_T(\hat{\tau}_1, \ldots, \hat{\tau}_M) - \log \tilde{L}_T] \tag{3}$$

where $\log \tilde{L}_T$ is the log-likelihood assuming no structural breaks, $\log \hat{L}_T(\hat{\tau}_1, \ldots, \hat{\tau}_M)$ is the log-likelihood assuming M structural breaks and $supLR_T$ represents the supremum of the likelihood ratios over all admissible partitions, $(\hat{\tau}_1, \ldots, \hat{\tau}_M)$, that determine the estimated breakpoints corresponding to the global maximization of the log-likelihood function.

Bai and Perron (1998) present a sequential procedure to test the null hypothesis of l structural changes vs. the alternative of $l+1$ breaks in a single regression model. Qu and Perron (2007) extend this sequential test to the multivariate case. The test procedure considers a model with l breaks under the null hypothesis vs. a less restrictive model with $l+1$ structural breaks. Bai and Perron (2003) provide critical values for these sequential test statistics. For a given system with $(\hat{\tau}_1, \ldots, \hat{\tau}_M)$ estimated break points, the model admits regression coefficient estimates $\hat{\boldsymbol{\Gamma}}_i$ and variance–covariance matrix estimates $\hat{\boldsymbol{\Sigma}}_i$ for each regime $i (i = 1, \ldots, M+1)$ that may be estimated by a quasi-maximum likelihood procedure.‡

Our estimation method allows multiple regimes with differing regression coefficients and covariance structures across regimes. We recognize the importance of conditional means and covariances in portfolio choice, and, therefore, admit variability in our inputs.

2.2. *Mean-variance spanning tests*

A set of K risky assets spans a larger set of N risky assets plus the original K assets if the minimum variance frontier of the K assets coincides with the frontier of the $K+N$ assets. The K risky assets are referred to as benchmark assets and the N assets are called test assets. Huberman and Kandel (1987) first formalized the notion of mean-variance spanning.§ They propose a regression-based likelihood-ratio test of the spanning hypothesis. If we define $\boldsymbol{R}_{1t}$ as the return on the vector of K benchmark assets and $\boldsymbol{R}_{2t}$ as the returns on the N test assets, we can examine spanning in the multivariate regression of $\boldsymbol{R}_{2t}$ on $\boldsymbol{R}_{1t}$ using the equation

$$\boldsymbol{R}_{2t} = \boldsymbol{\alpha} + \boldsymbol{\beta} \boldsymbol{R}_{1t} + \boldsymbol{\varepsilon}_t \tag{4}$$

†We follow the accepted usage in the structural break literature and define the resultant periods with constant population parameters as regimes.

‡The estimation procedure also requires a trimming parameter. Following Bai and Perron (2006), we set the trimming parameter to 0.15, implying a minimal length between breaks of at least 40 months. The use of alternative trimming parameters has little qualitative effect on our findings.

§Spanning regressions and restrictions are closely related to the style regressions of Sharpe (1992) subject to the typical additional restrictions of positive betas for each of the K style portfolios. This approach has been especially helpful in studies of returns-based style analysis, common in the field of portfolio management. Approaches of this sort are valuable when underlying portfolio holdings are unobservable or only observed at discreet intervals.

where $E(\boldsymbol{\varepsilon}_t) = \mathbf{0}_N$ and $E(\boldsymbol{\varepsilon}_t \boldsymbol{R}'_{1t}) = 0_{N\times K}$.† Following Kan and Zhou (2012), we note that when $\boldsymbol{\alpha} = \mathbf{0}_N$ and $\boldsymbol{\beta}\mathbf{1}_K = \mathbf{1}_N$, then $E(\boldsymbol{R}_{2t}) = \boldsymbol{\beta}E(\boldsymbol{R}_{1t})$. Therefore, for each row of $E(\boldsymbol{R}_{2t})$, there exists a combination of the K benchmark assets with the same expected return. Given standard regression assumptions, the covariance matrix of the test assets may be written as $Var(\boldsymbol{R}_{2t}) = Var(\boldsymbol{\beta}\boldsymbol{R}_{1t}) + Var(\boldsymbol{\varepsilon}_t)$. Because $E(\boldsymbol{\varepsilon}_t \boldsymbol{R}'_{1t}) = 0_{N\times K}$ and $Var(\boldsymbol{\varepsilon}_t)$ are positive definite, for any conformable fixed weight vector, $\boldsymbol{\omega}$, the portfolio variance from the N test assets will exceed the portfolio variance from the K benchmark assets. That is, when $\boldsymbol{\alpha} = \mathbf{0}_N$ and $\boldsymbol{\beta}\mathbf{1}_K = \mathbf{1}_N$, the K benchmark assets dominate the N test assets. Huberman and Kandel (1987) first demonstrated that the hypothesis that the minimum variance frontier of K benchmark assets is coincident with the frontier of the universe of $N + K$ assets is equivalent to testing

$$H_0 : \boldsymbol{\alpha} = \mathbf{0}_N, \text{ and } \boldsymbol{\delta} = \mathbf{0}_N \tag{5}$$

where $\boldsymbol{\delta} = \mathbf{1}_N - \boldsymbol{\beta}\mathbf{1}_K$.

Now, let $\hat{\boldsymbol{\Sigma}}$ be the unconstrained maximum likelihood estimator of the covariance matrix of $\boldsymbol{\varepsilon}$ in regression (4) and $\tilde{\boldsymbol{\Sigma}}$ be the maximum likelihood estimator of the covariance matrix of $\boldsymbol{\varepsilon}$ under the additional constraints in (5). The likelihood ratio test of the null hypothesis of spanning is then‡

$$LR = -T \ln\left(\frac{|\hat{\boldsymbol{\Sigma}}|}{|\tilde{\boldsymbol{\Sigma}}|}\right) \sim \chi^2_{2N}. \tag{6}$$

Jobson and Korkie (1989) provide an exact distribution of the likelihood ratio test under H_0. When there is only one test asset, the test statistic can be written as

$$\left(\frac{|\tilde{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}|} - 1\right)\left(\frac{T-K-1}{2}\right) \sim F_{2,T-K-1} \tag{7}$$

where T is the number of observations. For more than one test asset, the test statistic is

$$\left(\left(\frac{|\tilde{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}|}\right)^{1/2} - 1\right)\left(\frac{T-K-N}{N}\right) \sim F_{2N,2(T-K-N)}. \tag{8}$$

Kan and Zhou (2012) provide a step-down testing procedure to disentangle the effects of $\boldsymbol{\alpha} = \mathbf{0}_N$ and $\boldsymbol{\delta} = \mathbf{0}_N$ on the resultant spanning test statistics. The former restriction, $\boldsymbol{\alpha} = \mathbf{0}_N$, can be viewed as a test of whether the tangency portfolio (from the origin) has zero weight in the N test assets and the latter restriction may be viewed as a test of whether the global minimum variance portfolio has zero weight in the N test assets. Intuitively, these tests allow analysis of spanning rejections as in if they are due to the test assets improving the optimal portfolio Sharpe ratio (i.e. the slope of the investment opportunity set [IOS]) or the location of the global minimum variance portfolio (i.e. the vertex of the IOS). The distinction is important because rejections due to improvements in the optimal Sharpe ratio directly impact portfolio policy decisions, resultant portfolio performance and investor utility, whereas rejections due to the location of the global minimum variance portfolio are likely of little interest to most well-diversified investors.

The two component spanning tests may be applied sequentially with the following F-tests

$$\text{Tangency test} : F_1 = \left(\frac{|\bar{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}|} - 1\right)\left(\frac{T-K-N}{N}\right) \sim F_{N,T-K-N} \tag{9}$$

Global Min Variance Test:

$$F_2 = \left(\frac{|\tilde{\boldsymbol{\Sigma}}|}{|\bar{\boldsymbol{\Sigma}}|} - 1\right)\left(\frac{T-K-N+1}{N}\right) \sim F_{N,T-K-N+1} \tag{10}$$

where $\bar{\boldsymbol{\Sigma}}_t$ is the constrained maximum likelihood estimator of the covariance matrix of $\boldsymbol{\varepsilon}_t$ after imposing the constraint, $\boldsymbol{\alpha} = \mathbf{0}_N$, and other terms are as previously defined.

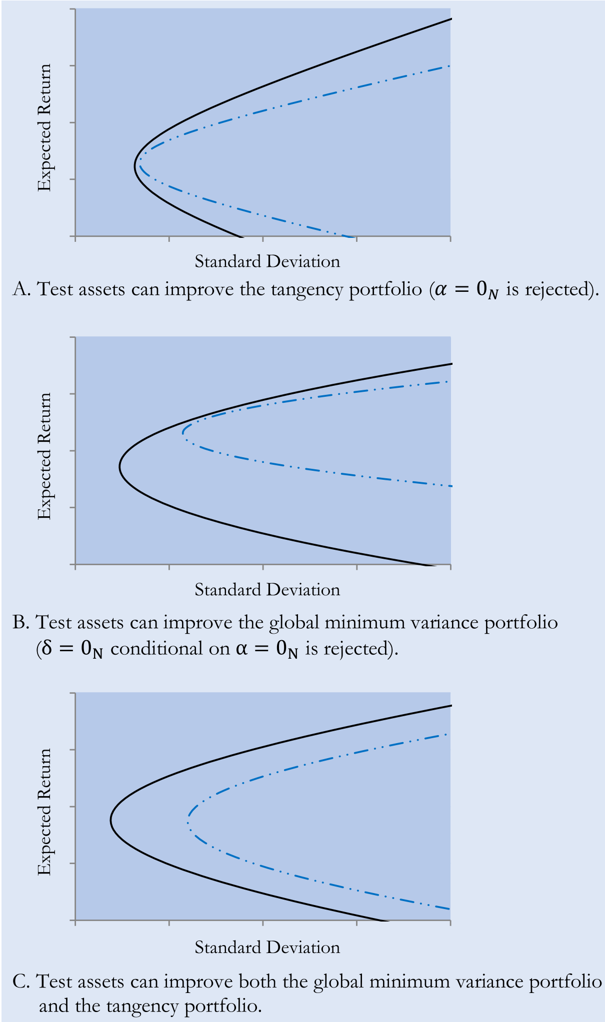
The F_1 test examines the total return regression zero intercept restriction, $\boldsymbol{\alpha} = \mathbf{0}_N$, which is equivalent to a test of whether the K asset set and the full $N + K$ asset set produce the same mean-variance tangency portfolio. The F_2 test of $\boldsymbol{\delta} = \mathbf{0}_N$ conditional on $\boldsymbol{\alpha} = \mathbf{0}_N$ examines if the global minimum variance portfolios for the two asset sets are statistically different. This step-down test procedure provides prescriptive economic guidance regarding the cause of spanning rejections. In general, spanning is rejected if either hypothesis is rejected.

We provide an illustrative description of these spanning tests in figure 1 with minimum variance frontiers based on population parameters. In all panels, the solid line represents the minimum variance frontier for the asset universe and the dashed–dotted line shows the frontier for the subset of K risky benchmark assets. Panel A presents the minimum variance frontier for the benchmark assets when spanning fails due to violation of $\boldsymbol{\alpha} = \mathbf{0}_N$. In this case, the benchmark assets replicate the low risk investment opportunities available from the asset universe; however, at greater risk exposures, the benchmark assets underperform. Panel B provides an example in which the benchmark assets violate $\boldsymbol{\delta} = \mathbf{0}_N$ conditional on $\boldsymbol{\alpha} = \mathbf{0}_N$. Intuitively, the benchmark assets do not provide the available low risk opportunities from the asset universe. In the final panel, spanning fails due to both the vertex location and the resultant IOS shape.

Because investors are primarily concerned with the tangency portfolio, it is important to distinguish between the various cases in figure 1. Arguably, from an investment policy perspective, replicating the tangency portfolio with available benchmark assets is of greater economic importance

†We consider various spanning tests within each regime with fixed regression coefficients. For notational simplicity we do not label coefficients with regime-specific subscripts in this section.

‡In our empirical work, we use the finite sample statistic where we replace T with $T-K-(N+1)/2$ as suggested by Kan and Zhou (2012). Inferences from the modified LR statistic and the asymptotic LR statistic of Huberman and Kandel (1987) are identical. We thank an anonymous referee for this suggestion.



A. Test assets can improve the tangency portfolio ($\alpha = 0_N$ is rejected).

B. Test assets can improve the global minimum variance portfolio ($\delta = 0_N$ conditional on $\alpha = 0_N$ is rejected).

C. Test assets can improve both the global minimum variance portfolio and the tangency portfolio.

Figure 1. Decomposing spanning failures. This figure illustrates the Kan and Zhou (2012) spanning test decomposition. In all graphs, the dash–dot line is the frontier hyperbola of benchmark assets, and the solid line is the frontier hyperbola of the combination of benchmark and all test assets. In Panel A, spanning fails due to different tangency portfolios ($\boldsymbol{\alpha} \neq \mathbf{0}_N$). In Panel B, spanning fails due to different minimum variance portfolios ($\boldsymbol{\delta} \neq \mathbf{0}_N$ conditional on $\boldsymbol{\alpha} = \mathbf{0}_N$). In Panel C, spanning fails due to both components.

than replicating a low risk vertex portfolio. For example, given the existence of a riskless asset, two fund separation suggests our primary interest should be regarding inferences based on tests of F_1 that relate to the IOS shape. Test size is an important consideration in this context. As discussed in Kan and Zhou (2012) if $size_1$ is the size of the F_1 test and $size_2$ is the size of the F_2 test, then the size of the joint test of F_1 and F_2 is $size_1 + size_2 + size_1 size_2$. In reporting later tests, we do not address the issue of choosing between appropriate sizes for each test. Rather, we report p-values (or test sizes) for all joint and individual tests. From a policy perspective, investment managers might be most interested in tests of F_1 and, therefore, may wish to focus solely on these related columns of results. Nonetheless, traditional usage would be to interpret joint tests with a given size, of say 5%, and then to choose reasonable partitions into sizes for F_1 and F_2 as described above.

3. Data

Our sample period includes observations from January 1988 to January 2010. The US market, Developed Markets (DM) and Emerging Markets (EM) monthly indices are from MSCI Barra.† The MSCI developed and emerging markets indices are float-adjusted market capitalization indices designed to measure equity market performance in US dollars. Total monthly returns for portfolios formed on size and book-to-market ratios are collected from Ken French's data library.‡ The US small and big portfolios consist of the smallest 30% and largest 30% of stocks in terms of size as measured by the median market equity value for NYSE firms, respectively. The US growth and value portfolios consist of the bottom 30 and top 30% stocks in terms of book-to-market ratios, respectively.

Much of the international literature uses the MSCI World Index as the global market benchmark. For example, Cumby and Glen (1990) use the indices of 13 developed markets as the dependent variables and the MSCI World Index as the independent variable in their regression analysis. They find that the intercepts are jointly zero, providing support for this index as a mean-variance efficient portfolio. Ferson and Harvey (1993), Fama and French (1998), and Harvey *et al*. (2002) also use the MSCI World Index as their global market proxy. Given our interest in both developed and emerging financial markets, we use the market capitalization-weighted MSCI All Country World Index (ACWI) as our world market benchmark (global market risk factor). As of April 2010, The ACWI consists of 45 country indices including 23 developed markets as in the MSCI Developed Markets Index and 22 emerging markets as in the MSCI Emerging Markets Index. To examine the behaviour of different emerging market regions, we also consider the total returns from the MSCI Barra geographic regions—EM Asia, EM Europe and EM Latin America.§ This breakdown facilitates a differentiation among emerging markets from around the world.

†As of April 2010, the MSCI Developed Markets Index covers 23 countries, including Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom and the United States. The MSCI Emerging Market Index covers 22 countries, including Brazil, Chile, China, Colombia, the Czech Republic, Egypt, Hungary, India, Indonesia, Israel, Korea, Malaysia, Mexico, Morocco, Peru, Philippines, Poland, Russia, South Africa, Taiwan, Thailand and Turkey.

‡We gratefully acknowledge the provision of this data by Ken French.

§The EM Asia includes China, India, Indonesia, Korea, Malaysia, Philippines, Taiwan and Thailand. The EM Europe includes the Czech Republic, Hungary, Poland, Russia, and Turkey. The EM Latin America includes Brazil, Chile, Colombia, Mexico and Peru. Egypt, Israel, Morocco and South Africa are included in MSCI EM index calculation but are not classified within any of the EM regional indices.

Table 1. Descriptive statistics for portfolios and risk factors.

Variable	Median	Mean (p-value)	Standard deviation
r_{US}	0.895	0.477 (0.070)	4.266
r_{DE}	0.490	0.218 (0.479)	4.998
r_{EM}	1.247	0.942 (0.030)	7.032
r_{As}	0.676	0.582 (0.205)	7.461
r_{Eu}	1.680	1.190 (0.031)	8.902
r_{LA}	2.567	1.716 (0.003)	9.249
r_{Small}	1.229	0.670 (0.069)	5.971
r_{Big}	0.980	0.493 (0.061)	4.254
r_{Value}	1.233	0.665 (0.019)	4.581
r_{Growth}	0.810	0.510 (0.067)	4.511
r_m	0.780	0.309 (0.262)	4.466
r_{FX}	0.226	0.221 (0.006)	1.286
R_f	0.413	0.371 (0.000)	0.182

Notes: This table reports summary statistics for monthly excess returns on the MSCI US market (r_{US}), developed market (r_{DE}), emerging market (r_{EM}), Asian emerging market (r_{As}), European emerging market (r_{Eu}), Latin American emerging market (r_{LA}), US small (r_{Small}), US big (r_{Big}), US value (r_{Value}), US growth (r_{Growth}), world benchmark (r_m) and foreign exchange (r_{FX}) portfolios from January 1988 to January 2010. We also report summary statistics for the one-month Eurodollar deposit rate (R_f), which we use as the international risk-free rate. All values are stated in percentage form. Reported p-values test if the mean is significantly different from zero (one sample t-test).

Following Harvey *et al*. (2002), our analysis includes a foreign exchange risk factor defined as the return on the trade-weighted exchange index from the Federal Reserve Bank of St. Louis less the riskless rate.† We use the 30-day Eurodollar deposit rate from the Federal Reserve Bank of St. Louis as the international risk-free rate. Existing international asset pricing literature finds that foreign exchange risk affects expected equity returns.‡ For example, Dumas and Solnik (1995) show that departures from purchasing power parity induce foreign exchange risk premium for the world's largest equity markets. Ferson and Harvey (1993, 1994), and Harvey *et al*. (2002) find that the aggregated exchange risk is another significant factor in unconditional and conditional asset pricing tests. Therefore, similar to Harvey (1995) and Harvey *et al*. (2002), we consider changes in the US dollar against the currencies of a broad group of major US trading partners as our second global risk factor (foreign exchange risk). A positive (negative) change indicates appreciation (depreciation) of the US dollar.

Table 1 provides summary statistics for all variables, including the monthly excess returns on the MSCI US market (r_{US}), developed markets (r_{DM}), emerging markets (r_{EM}), Asian emerging market (r_{As}), European emerging market (r_{Eu}), Latin American emerging market (r_{LA}), US small (r_{Small}), US big (r_{Big}), US value (r_{Value}), US growth (r_{Growth}), world benchmark/global market risk factor (r_m) and foreign exchange risk factor (r_{FX}). We also report summary statistics for the one-month Eurodollar deposit rate (R_f), used as the global risk-free rate. All values are expressed in percentage per month format.

From the three primary equity indices, we observe that the US market portfolio (r_{US}) and the emerging market portfolio (r_{EM}) provide significant positive excess returns for the entire sample period from January 1988 to January 2010 (at 10% and 5% levels, respectively). Reported p-values indicate that excess returns to developed markets and the world market portfolio are not significantly different from the risk-free rate over this sample period. The significant mean excess return of the emerging markets may be heavily influenced by the large Latin American (r_{LA}) mean excess return of 1.7% per month, and also by strong mean excess returns of 1.2% per month in European emerging markets (r_{Eu}). Estimated standard deviations for all emerging markets and breakdowns into regional portfolios show much more volatility than our other portfolios. Estimated sample means for the US small portfolio (r_{Small}) and the US value portfolio (r_{Value}) are greater than the US big (r_{Big}) and US growth (r_{Growth}) portfolios, respectively, consistent with the often documented size and value effects.§, ¶

Our empirical research begins with a multivariate structural break investigation into the number of regimes detected in US, developed and emerging markets, with no accompanying tests of portfolio restrictions. We then extend our analysis to consider one- and two-factor models for the same three primary aggregate markets. These models naturally give rise to traditional tests of Jensen's alpha, which allow performance to vary across regimes. Our final analyses consider an expanded system that partitions the

†The trade-weighted exchange index (broad) is a trade-weighted average of the foreign exchange value of the US dollar against the currencies of a broad group of major US trading partners, including the Euro Area, Canada, Japan, Mexico, China, United Kingdom, Taiwan, Korea, Singapore, Hong Kong, Malaysia, Brazil, Switzerland, Thailand, Philippines, Australia, Indonesia, India, Israel, Saudi Arabia, Russia, Sweden, Argentina, Venezuela, Chile and Colombia.

‡See Solnik (1974a, 1994b), Karolyi and Stulz (2003) for a detailed discussion of foreign exchange risk pricing.

§Banz (1981) first documented the size effect by showing small firm stocks had higher returns than large firm stocks.

¶Stattman (1980) and Rosenberg *et al*. (1985) show that stocks with high book-to-market ratios outperform stocks with low book-to-market ratios.

Table 2. Multivariate structural break model with no risk factors.

A: Structural break estimates

Structural break	Point estimate (95% confidence interval)	$supLR_T$	$seq(2\|1)$
1	January 1994 (March 1993–February 1995)	168.52***	48.87***
2	June 2002 (November 2000–August 2002)		

B: Mean excess return estimates ($\boldsymbol{\mu}_i$)

	No break	Regime 1	Regime 2	Regime 3
	$\boldsymbol{\mu}_i$(*p*-value)	$\boldsymbol{\mu}_i$(*p*-value)	$\boldsymbol{\mu}_i$(*p*-value)	$\boldsymbol{\mu}_i$(*p*-value)
US	0.4771	0.7659	0.5601	0.1529
	(0.069)	(0.061)	(0.198)	(0.776)
DM	0.2177	0.2655	−0.0965	0.5276
	(0.470)	(0.674)	(0.800)	(0.409)
EM	0.9418	2.3140	−0.5135	1.4557
	(0.030)	(0.002)	(0.442)	(0.085)

C: Correlation/volatility estimates

	Regime 1			Regime 2			Regime 3		
	US	DM	EM	US	DM	EM	US	DM	EM
US	3.5292			4.4736			4.5171		
	(0.000)			(0.000)			(0.000)		
DM	0.4218	5.5571		0.7345	4.0972		0.9120	5.3711	
	(0.000)	(0.000)		(0.000)	(0.000)		(0.000)	(0.000)	
EM	0.3785	0.3925	6.4009	0.6571	0.6926	7.0018	0.8278	0.9245	7.2154
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: This table reports break point and parameter estimates for the multivariate structural break model for US, developed and emerging markets from January 1988 to January 2010. Panel A reports break point estimates with associated 95% confidence intervals. The second column reports the $supLR_T$ test statistic to examine the null hypothesis of no break against a maximum of two breaks. The final column reports the sequential test statistic $seq(l+1|l)$ to examine the null hypothesis of l breaks vs. $l+1$ breaks, for $l=1$. Panel B reports mean excess return estimates in each regime with *p*-values in parentheses, along with parameter estimates assuming no structural breaks. Panel C reports portfolio volatilities in percent per month format on the diagonals with correlations in lower triangular elements (associated *p*-values in parentheses).
*Significance levels for the $supLR_T$ and $seq(l+1|l)$ statistics at the 10% levels.
**Significance levels for the $supLR_T$ and $seq(l+1|l)$ statistics at the 5% levels.
***Significance levels for the $supLR_T$ and $seq(l+1|l)$ statistics at the 1% levels.

US market into small, big, high value, and low value firms, as well as a partition of emerging markets into Asian, European and Latin American submarkets. We provide an exhaustive examination of structural breaks, portfolio alphas and spanning tests within regimes in this multivariate context.

4. Empirical analysis

4.1. Characterizing structural breaks in the US, developed and emerging markets

We begin with a general structural break analysis of the US, developed and emerging equity markets as measured by monthly returns in excess of the one-month Eurodollar deposit rate to the MSCI US (r_{US}), developed (r_{DM}) and emerging market (r_{EM}) portfolios. The sample period includes 265 observations from January 1988 to January 2010. Ignoring all risk factors, our simplified model can be written directly from equation (2) with $\boldsymbol{X}_t = [1]$, $\boldsymbol{\Gamma}_i = [\boldsymbol{\mu}_i]$, $\boldsymbol{\mu}_i = [\mu_{USi}, \mu_{DMi}, \mu_{EMi}]'$, and all other terms unchanged.

Break point and parameter estimates for this model are given in table 2. Panel A reports that the $supLR_T$ statistic is 168.52, which strongly rejects the null hypothesis of no structural breaks at conventional levels of significance.† Further, the sequential test statistic $seq(2|1)$ of 48.87 exceeds the 1% critical value of 32.87, and the unreported sequential test statistic $seq(3|2)$ is not significant at any conventional level. These values suggest two break points in the series. Panel A also shows the estimated breakpoints to be January 1994 and June 2002, with 95% confidence intervals that range from March 1993 to February 1995 and from November 2000 to August 2002, respectively. The three regime periods identified by this model are January 1988 to January 1994, February 1994 to June 2002 and July 2002 to January 2010.

Panel B displays mean excess return parameter estimates with *p*-values in parentheses. Assuming no structural breaks over the entire sample period, the US market has an excess return of 0.48% per month (approximately 6% per year). The emerging market provides a significant excess return of 0.94% per month (approximately 12% per year). The developed market return is not significantly different from the risk-free rate during the entire sample period.

With the inclusion of structural breaks, we find that the US equity excess return is 0.77% per month in the first regime (significant at the 10% level), while the emerging

†The critical values are 49.07 (1%), 42.86 (5%) and 39.77 (10%).

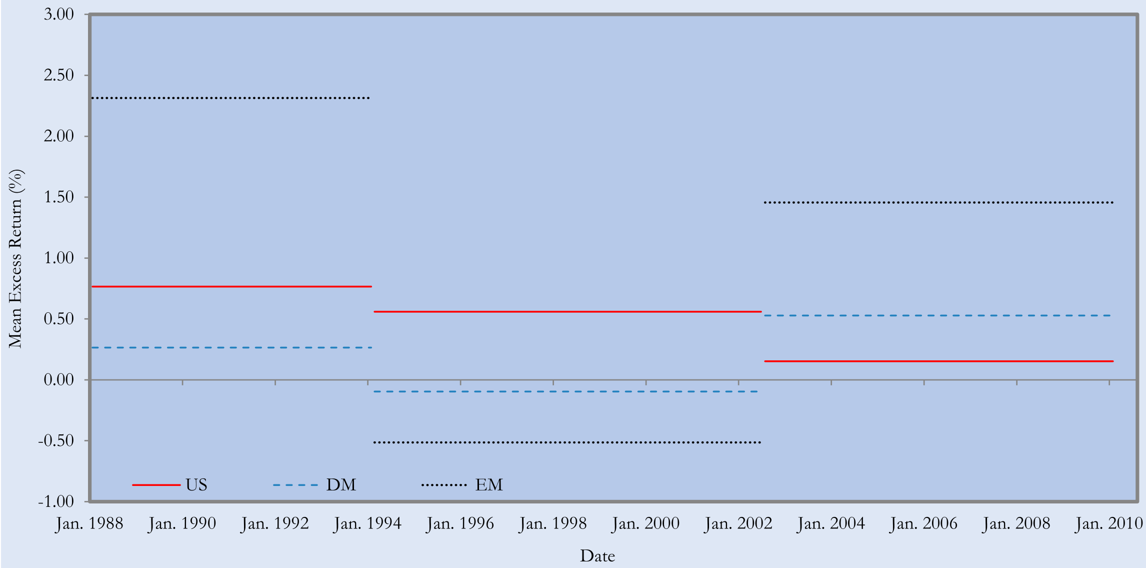


Figure 2. Mean excess returns by regimes for the US, developed and emerging markets. This figure plots the mean excess returns to the US, Developed Market (DM) and Emerging Market (EM) portfolios in three regimes identified using the multivariate structural break model given by equation (1) where r_{USt}, r_{DMt} and r_{EMt} are monthly percent excess returns to the MSCI US market, developed market and emerging market indices, respectively.

market has a very large excess return of 2.31% per month (significant at the 1% level) and the developed market excess return is not significantly different from zero. In the second regime, none of the indices provide significant excess returns. In the third regime, only the emerging market has a significant monthly excess return of 1.46% (at the 10% level).

Panel C reports the correlation and volatility estimates for model disturbances with associated p-values in parentheses. For brevity, we report volatilities in percent per month format along the diagonal with correlations in the lower triangular elements. We find that monthly volatilities and pairwise correlations for all three equity portfolios are greatest in regime three. This interesting result is consistent with previous research that finds integration of global equity markets increased in recent years. Unfortunately, this also suggests that world financial markets are most correlated when volatilities are greatest, diminishing the potential benefits of global diversification.

Figure 2 plots the mean excess returns to the US, developed and emerging market portfolios in each of the three regimes identified by this model. We find that the US and developed markets have similar estimated mean returns across regimes; however, the emerging market portfolio has greater excess return variability across regimes, with greatest mean levels in the first and third regimes.†

4.2. *Structural breaks in the US, developed and emerging markets with risk factors*

In this section, we investigate the performance of US, developed and emerging market portfolios in different regimes after controlling for regime-specific risk factor sensitivities. Our initial model admits varying sensitivities across regimes for all portfolios in relation to global equity market risk. Then, we consider a model including global equity market risk and foreign exchange risk.

Structural break point estimates for our one-factor model (global market risk) and two-factor model (global market and foreign exchange risks) are presented in Panel A of table 3. We find strong evidence of two distinct break points in the joint excess returns of the three market portfolios. The two breaks in our one-factor model occur in October 1997 and May 2006. The $supLR_T$ statistic of 560.77 is significant at the 1% level, and the sequential test statistic $seq(2|1)$ of 174.49 far exceeds the 1% critical value of 37.65. The unreported sequential test statistic $seq(3|2)$ is not significant at conventional levels. The results provide strong evidence of two breaks. The optimal partition of the data suggested by our one-factor model is January 1988 to October 1997 for the first regime, November 1997 to May 2006 for the second regime and June 2006 to January 2010 for the third regime.

In our two-factor model, we again find two break points. The $supLR_T$ statistic of 580.99 and sequential test statistic $seq(2|1)$ of 169.62 are both significant at the 1% level. The sequential test $seq(3|2)$ is not significant. The first break point now occurs in April 1997, six months earlier than the initial break point identified by the one-factor model. The second break point is the same for both models. The three regimes identified by our two-factor model are from January 1988 to April 1997, May 1997 to May 2006 and June 2006 to January 2010.

Panel B presents alpha estimates for our one-factor model with p-values in parentheses. We first report findings with the assumption of no breaks. The no structural break results suggest that the US and emerging market portfolios offer marginal performance benefits to investors. The US market

†Later we provide specific tests of portfolio performance within regimes.

Table 3. Multivariate one- and two-factor structural break models.

A: Structural break estimates				
Structural break	Point estimate (95% confidence interval)		$supLR_T$	$seq(2\|1)$
one-Factor model				
1	October 1997 (August 1997–November 1997)		560.77***	174.49***
2	May 2006 (April 2006–June 2006)			
two-Factor model				
1	April 1997 (December 1996–May 1997)		580.99***	169.62***
2	May 2006 (March 2006–June 2006)			
B: one-Factor model alpha estimates				
	No Break	Regime 1	Regime 2	Regime 3
	$\boldsymbol{\alpha}_i$(p-value)	$\boldsymbol{\alpha}_i$(p-value)	$\boldsymbol{\alpha}_i$(p-value)	$\boldsymbol{\alpha}_i$(p-value)
US	0.2204	0.6803	−0.0827	−0.1744
	(0.089)	(0.003)	(0.533)	(0.372)
DM	−0.1133	−0.3672	0.1069	−0.0277
	(0.206)	(0.008)	(0.426)	(0.832)
EM	0.5886	0.7637	0.3050	1.0502
	(0.045)	(0.120)	(0.454)	(0.000)
C: two-Factor model alpha estimates				
	No Break	Regime 1	Regime 2	Regime 3
	$\boldsymbol{\alpha}_i$(p-value)	$\boldsymbol{\alpha}_i$(p-value)	$\boldsymbol{\alpha}_i$(p-value)	$\boldsymbol{\alpha}_i$(p-value)
US	0.2674	0.6400	0.0848	−0.0269
	(0.037)	(0.006)	(0.410)	(0.892)
DM	−0.1496	−0.3480	−0.0305	−0.1114
	(0.089)	(0.013)	(0.778)	(0.412)
EM	0.6190	1.0076	0.0130	0.8182
	(0.034)	(0.033)	(0.938)	(0.067)

Notes: This table reports break point and parameter estimates for the multivariate structural break model for monthly percentage excess returns to the MSCI US market (r_{USt}), developed market (r_{DMt}) and emerging market (r_{EMt}) indices, respectively, from January 1988 to January 2010. We consider two risk factors given by the monthly excess return of the world benchmark portfolio (r_{mt}) and the monthly excess return from the trade-weighted foreign exchange index (r_{FXt}). Panel A reports break point estimates with 95% confidence intervals. The test statistic $supLR_T$ examines the null hypothesis of no break against a maximum of two breaks. The sequential test statistic $seq(l+1|l)$ examines the null hypothesis of l vs. $l+1$ breaks, for $l=1$. Panels B and C report alpha estimates by regime for the one-factor and two-factor models, with p-values in parentheses. We also report unconditional alpha estimates and p-values assuming no structural breaks.
*Significance levels for the $supLR_T$ and $seq(l+1|l)$ statistics at the 10% levels.
**Significance levels for the $supLR_T$ and $seq(l+1|l)$ statistics at the 5% levels.
***Significance levels for the $supLR_T$ and $seq(l+1|l)$ statistics at the 1% levels.

portfolio yields an alpha estimate of 0.22% per month (significant at the 10% level) and the emerging market portfolio provides an alpha of 0.59% per month (or approximately 7% per year) over the last 20 years.

With structural breaks in the one-factor model, we find the US market in the first regime has an alpha of 0.68% (significant at the 1% level), with insignificant alphas in the second and third regimes. The developed market portfolio has a significant negative alpha of −0.37% in the first regime, but insignificant alphas in the latter two regimes. For the emerging market portfolio, the estimated alpha of 1.05% in the third regime is highly significant, consistent with strong emerging market performance after June 2006.

As shown in Panel C, when we assume no structural breaks in our two-factor model we find alphas that are significant at the 10% level for all three portfolios. The emerging market portfolio in this case provides a positive alpha of 0.62% per month. Comparing our one- and two-factor models, we find similar significance in alpha values across the three regimes. One important exception occurs with the finding of a significantly positive alpha for the emerging market portfolio in the first and third regimes of our two-factor model, rather than only the first regime for the one-factor model.

Unreported results show that sensitivities to the global market risk factor are highly significant in all regimes in the one- and two-factor models. Also, in the two-factor model, the foreign exchange risk factor is typically significant in all three regimes and in the no break model.†

4.3. *Partitions of the US and emerging markets*

We now examine a finer partition of US and emerging markets. In particular, we consider nine variates—the US Market, Developed Market, US small, big, value and growth

†We find significant positive coefficients in all three regimes for the foreign exchange risk factor for the US market portfolio, consistent with the notion that an increase in the value of the US dollar will increase the value of investments in US equities. Concomitantly, an appreciation of the US dollar reduces the value of developed foreign markets. The impact of the foreign exchange risk factor has a mixed impact on the general emerging market index return in different regimes. In general, the direct linkage between US and developed market returns in relation to the FX risk factor is not surprising given it measures the return from a broad trade-weighted foreign exchange portfolio that is dominated by developed markets and their international trade with the US (see also Turtle and Zhang (2012)).

Table 4. Multivariate one- and two-factor structural break models with US and emerging market breakdowns.

A: Structural break estimates			
Structural break	Point estimate (95% confidence interval)	$supLR_T$	$seq(2\|1)$
one-Factor model			
1	March 1997 (February 1997–April 1997)	1127.00***	503.81***
2	March 2003 (January 2003–April 2003)		
two-Factor model			
1	March 1997 (February 1997–April 1997)	1186.37***	518.28***
2	March 2003 (January 2003–April 2003)		

B: Alpha Estimates	B.1: one-Factor alpha estimates			B.2: two-Factor alpha estimates		
	Regime 1 α_i(p-value)	Regime 2 α_i(p-value)	Regime 3 α_i(p-value)	Regime 1 α_i(p-value)	Regime 2 α_i(p-value)	Regime 3 α_i(p-value)
US market	0.6190	0.3123	−0.2247	0.6017	0.3641	−0.1310
	(0.011)	(0.097)	(0.072)	(0.012)	(0.050)	(0.304)
Developed market	−0.3317	−0.2177	0.1613	−0.3204	−0.2723	0.0872
	(0.022)	(0.235)	(0.115)	(0.024)	(0.130)	(0.390)
US small	0.4222	0.7988	0.0230	0.3757	0.7927	0.3193
	(0.248)	(0.257)	(0.948)	(0.272)	(0.268)	(0.361)
US big	0.5934	0.3619	−0.2160	0.5747	0.4231	−0.1136
	(0.015)	(0.054)	(0.088)	(0.016)	(0.021)	(0.358)
US value	0.6599	0.4347	−0.0297	0.6289	0.5132	0.2080
	(0.011)	(0.247)	(0.916)	(0.011)	(0.171)	(0.454)
US growth	0.5726	0.4277	−0.1528	0.5500	0.4609	−0.0506
	(0.049)	(0.056)	(0.344)	(0.054)	(0.041)	(0.755)
EM Asia	0.6315	−0.8538	0.7305	0.5638	−1.0312	0.7502
	(0.220)	(0.295)	(0.068)	(0.239)	(0.204)	(0.073)
EM Europe	1.0692	0.4070	0.7342	1.0420	0.6569	0.3548
	(0.125)	(0.635)	(0.221)	(0.135)	(0.434)	(0.555)
EM Latin America	2.0643	0.1338	1.6323	1.9995	0.1776	1.4148
	(0.020)	(0.857)	(0.001)	(0.022)	(0.813)	(0.004)

Notes: This table reports break point and parameter estimates for the multivariate structural break model for monthly percentage excess returns to the MSCI US market (r_{USt}), developed market (r_{DMt}) and emerging markets in Asia (r_{ASt}), Europe (r_{Eut}) and Latin America (r_{LAt}), respectively, from January 1988 to January 2010. We also include monthly percentage excess returns to the small (r_{Smallt}), big (r_{Bigt}), value (r_{Valuet}) and growth ($r_{Growtht}$) portfolios as additional test assets to provide a richer domestic set of investment opportunities. We consider two risk factors given by the monthly excess return of the world benchmark portfolio (r_{mt}) and the monthly excess return from the trade-weighted foreign exchange index (r_{FXt}). Panel A reports break point estimates with 95% confidence intervals. The test statistic $supLR_T$ examines the null hypothesis of no break against a maximum of two breaks. The sequential test statistic $seq(l+1|l)$ examines the null hypothesis of l vs. $l+1$ breaks, for $l=1$. Panel B reports alpha estimates by regime for the one-factor and two-factor models, with p-values in parentheses.
*Significance levels for the $supLR_T$ and $seq(l+1|l)$ statistics at the 10% levels.
**Significance levels for the $supLR_T$ and $seq(l+1|l)$ statistics at the 5% levels.
***Significance levels for the $supLR_T$ and $seq(l+1|l)$ statistics at the 1% levels.

stocks, and MSCI Barra emerging market (EM) indices for Asia, Europe and Latin America. We use two-factor alpha tests and spanning tests to examine the ability of regional emerging markets, or subsets of these markets, to improve the performance of a well-diversified global portfolio.

Our nine-variate regression model is given by equation (2) with $\boldsymbol{r}_t = [r_{USt}, r_{DMt}, r_{Smallt}, r_{Bigt}, r_{Valuet}, r_{Growtht}, r_{Ast}, r_{Eut}, r_{LAt}]'$, $\boldsymbol{X}_t = [1, r_{mt}, r_{FXt}]'$, $\boldsymbol{\Gamma}_i = [\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \boldsymbol{\gamma}_i]$, $\boldsymbol{\alpha}_i = [\alpha_{USi}, \alpha_{DMi}, \ldots, \alpha_{LAt}]'$, $\boldsymbol{\beta}_i = [\beta_{USi}, \beta_{DMi}, \ldots, \beta_{LAt}]'$, and $\boldsymbol{\gamma}_i = [\gamma_{USi}, \gamma_{DMi}, \ldots, \gamma_{LAt}]'$. The vector $\boldsymbol{\varepsilon}_{it}$ is then a 9×1 vector of disturbances that varies by date within regimes, with a regime specific 9×9 variance-covariance matrix, $\boldsymbol{\Sigma}_i$.

We estimate the system using our one- and two-factor models described earlier. Table 4 reports break point tests and model parameter estimates.

Panel A presents structural break point estimates and their 95% confidence intervals. Both the one- and two-factor models find the same breaks, in March of 1997 and March of 2003. Confidence intervals are also coincident for both models. The $supLR_T$ statistic is 1127 for the one-factor model and 1186 for the two-factor model (both reject the null hypothesis of no structural break at the 1% level). Values for $seq(2|1)$ are significant at all conventional levels, with insignificant values for $seq(3|2)$, providing further evidence of two break points. In these models, the three identified regime periods are January 1998 to March 1997, April 1997 to March 2003 and April 2003 to January 2010.

As shown in Panel B, the one- and two-factor models provide similar alpha parameter estimates by regime. In particular, both models find significant positive alpha estimates for EM Latin America in regimes one and three. Considering our one-factor model, EM Latin America provides a positive and significant (5% level) alpha of 2.06 and 1.63% in the first and third regimes, respectively. In the two-factor model, EM Latin America displays significant alphas of 2.00 and 1.41% per month in the first and third regimes, respectively. In regime three of both models, EM Asia offers a positive alpha at the 10% level of significance. EM Europe does not provide significant alphas in any of the regimes for either model.

Table 5 presents likelihood ratio test results for joint tests alpha equality for our nine-variate model. Panels A and C relate to our one-factor model, with two-factor results given in Panels B and D. The upper left side of table 5 examines the restriction that independently estimated alphas for each of the nine variates are equal across regimes, where the upper right side adds the restriction to each case that alphas are equal to zero. The bottom left section examines whether alphas for all nine variates are equal within each regime (or in all regimes), where the bottom right section jointly tests that all nine variate alphas are equal to zero within each regime (or in all regimes).

We find little qualitative difference in inferences between our one- and two-factor models, even though the addition of a second factor often mitigates the significance of results. The test statistics provide strong evidence to reject the equality of alphas across regimes for individual variates and across variates within regimes. We also find strong evidence to reject that alphas equal zero either across regimes for individual variates or across all nine variates in specific regimes. Our results strongly reject that four of the nine variates have alphas that are equal across the three regimes, that variate alphas are equal to each other or equal to zero within most regimes, and that almost all alphas equal zero either across or within regimes.

In table 6, we report results of the modified likelihood ratio spanning test following Huberman and Kandel (1987), the Jobson and Korkie (1989) F-test, and the Kan and Zhou (2012) step-down tests. We use these mean-variance spanning tests to investigate if IOSs are coincident over our entire sample, and within our estimated regimes. This analysis allows us to examine whether certain emerging market regions, individually or collectively, might improve portfolio performance for an investor who allocates wealth among the US market index, developed market index and US small, big, value and growth indices (our benchmark assets). We consider EM Asia, EM Europe and EM Latin America as individual test assets, as well as all three emerging markets jointly using the various test statistics. We exclude risk factors to avoid the difficulty of foreign exchange risk interpretation. The three regimes in this analysis are from January 1988 to March 1997, April 1997 to March 2003 and April 2003 to January 2010.

For the entire sample period, the Huberman and Kandel (1987) modified likelihood ratio test and the Jobson and Korkie (1989) F-test reject spanning for all emerging market (EM) indices jointly. These tests also reject spanning for EM Europe and EM Latin America individually, at the 5% level. Further examination of the Kan and Zhou (2012) F_1 and F_2 tests provide important information about the source of these rejections. Both F_1 and F_2 tests reject spanning for EM Latin America; however, only the F_2 test appears significant for the other reported rejections. Because F_1 tests relate to the shape of the IOS, whereas F_2 tests relate to the location of the global minimum variance portfolio, an investor concerned primarily with improvements in their optimal tangency portfolio would stress the role of EM Latin America in portfolio improvement.

For the first regime, the Huberman and Kandel (1987) modified likelihood ratio test and the Jobson and Korkie (1989) F-test fail to reject spanning for all individual EM regional indices. However, both procedures marginally reject spanning for all emerging markets considered jointly (at the 10% level). The Kan and Zhou (2012) procedure demonstrates that the joint EM spanning rejection is primarily due to the F_2 test, indicating that emerging markets as a whole can improve the performance of the minimum-variance portfolio when we consider our broad set of benchmark assets. Interestingly, we again find marginal evidence that EM Latin America improves the location of the tangency portfolio according to the F_1 test.

In the second regime, both the modified likelihood ratio and the F-test reject spanning for EM Latin America and for all emerging markets jointly. Decomposing these

Table 5. Test of alpha equality.

Hypothesis	A: 1-Factor χ^2	A: 1-Factor p-value	B: 2-Factor χ^2	B: 2-Factor p-value	Hypothesis	C: 1-Factor χ^2	C: 1-Factor p-value	D: 2-Factor χ^2	D: 2-Factor p-value
$\alpha_{US,1} = \alpha_{US,2} = \alpha_{US,3}$	12.68	0.002	10.70	0.005	$\alpha_{US,1} = \alpha_{US,2} = \alpha_{US,3} = 0$	20.40	0.000	17.60	0.001
$\alpha_{DM,1} = \alpha_{DM,2} = \alpha_{DM,3}$	8.70	0.013	6.28	0.043	$\alpha_{DM,1} = \alpha_{DM,2} = \alpha_{DM,3} = 0$	34.12	0.000	9.94	0.019
$\alpha_{Small,1} = \alpha_{Small,2} = \alpha_{Small,3}$	1.18	0.554	0.38	0.827	$\alpha_{Small,1} = \alpha_{Small,2} = \alpha_{Small,3} = 0$	3.32	0.345	4.56	0.207
$\alpha_{Big,1} = \alpha_{Big,2} = \alpha_{Big,3}$	14.24	0.001	11.76	0.003	$\alpha_{Big,1} = \alpha_{Big,2} = \alpha_{Big,3} = 0$	18.34	0.000	14.66	0.002
$\alpha_{Value,1} = \alpha_{Value,2} = \alpha_{Value,3}$	3.36	0.186	1.34	0.512	$\alpha_{Value,1} = \alpha_{Value,2} = \alpha_{Value,3} = 0$	17.04	0.001	17.48	0.001
$\alpha_{Growth,1} = \alpha_{Growth,2} = \alpha_{Growth,3}$	6.76	0.034	5.42	0.067	$\alpha_{Growth,1} = \alpha_{Growth,2} = \alpha_{Growth,3} = 0$	13.02	0.005	10.88	0.012
$\alpha_{As,1} = \alpha_{As,2} = \alpha_{As,3}$	3.16	0.206	4.20	0.122	$\alpha_{As,1} = \alpha_{As,2} = \alpha_{As,3} = 0$	6.40	0.094	7.16	0.067
$\alpha_{Eu,1} = \alpha_{Eu,2} = \alpha_{Eu,3}$	0.44	0.803	0.86	0.651	$\alpha_{Eu,1} = \alpha_{Eu,2} = \alpha_{Eu,3} = 0$	7.98	0.046	6.78	0.079
$\alpha_{LA,1} = \alpha_{LA,2} = \alpha_{LA,3}$	4.30	0.116	4.40	0.111	$\alpha_{LA,1} = \alpha_{LA,2} = \alpha_{LA,3} = 0$	35.00	0.000	28.80	0.000
Alphas are equal in regime 1	13.60	0.093	13.90	0.084	Alphas are all zero in regime 1	19.60	0.021	20.06	0.018
Alphas are equal in regime 2	6.44	0.598	9.74	0.284	Alphas are all zero in regime 2	17.52	0.041	18.38	0.031
Alphas are equal in regime 3	30.48	0.000	31.30	0.000	Alphas are all zero in regime 3	50.06	0.000	55.82	0.000
Alphas are equal in all regimes	51.44	0.002	57.68	0.000	Alphas are all zero in all regimes	87.30	0.000	94.42	0.000

Note: This table reports the likelihood ratio tests of alpha equity for the nine-variate model. Panel A reports equality restrictions for the one-factor model, Panel B reports equality restrictions for the two-factor model, and Panels C and D impose the additional equality restriction to zero for the one- and two-factor models, respectively.

Table 6. Mean-variance spanning tests.

EM Regions	HK (1987) LR test		JK (1989) test		KZ (2012) step-down test			
	LR test	p-value	F-test	p-value	F_1 test	p-value	F_2 test	p-value
Entire sample: January 1988–January 2010								
EM Asia	2.47	0.290	1.24	0.291	0.13	0.718	2.36	0.126
EM Europe	9.16	0.010	4.66	0.010	1.823	0.178	7.46	0.007
EM Latin America	15.33	0.000	7.90	0.000	5.47	0.020	10.11	0.002
All	16.61	0.011	2.80	0.011	1.83	0.141	3.79	0.011
First regime: January 1988–March 1997								
EM Asia	0.60	0.740	0.30	0.739	0.58	0.447	0.02	0.895
EM Europe	3.65	0.161	1.86	0.161	2.42	0.122	1.26	0.263
EM Latin America	3.99	0.136	2.03	0.136	3.86	0.052	0.17	0.683
All	10.81	0.094	1.83	0.094	1.43	0.240	2.25	0.087
Second regime: April 1997–March 2003								
EM Asia	1.65	0.439	0.83	0.439	1.26	0.265	0.39	0.536
EM Europe	2.08	0.354	1.05	0.354	0.08	0.779	2.06	0.156
EM Latin America	10.03	0.007	5.42	0.007	0.13	0.716	10.85	0.002
All	11.67	0.070	2.00	0.070	0.48	0.697	3.68	0.016
Third regime: April 2003–January 2010								
EM Asia	10.85	0.004	5.85	0.004	0.96	0.331	10.72	0.002
EM Europe	5.90	0.052	3.07	0.052	0.00	0.990	6.22	0.015
EM Latin America	15.47	0.000	8.60	0.000	8.89	0.004	7.41	0.008
All	25.74	0.000	4.63	0.000	4.46	0.006	4.81	0.004

Note: This table presents three mean-variance spanning tests on the MSCI emerging market monthly indices, EM Asia, EM Europe, and EM Latin America and their associated p-values. Other available (benchmark) assets include the MSCI US and DM monthly indices, and the monthly US small, big, value and growth value-weighted portfolio returns. Reported tests include the modified likelihood ratio test of Huberman and Kandel (1987) with a small sample adjustment that we denote as the HK LR test, the JK test of Jobson and Korkie (1989) and the KZ step-down tests of Kan and Zhou (2012) where F_1 is an F-test of $\alpha = \mathbf{0}_N$ and F_2 is an F-test of $\delta = \mathbf{0}_N$ conditional on $\alpha = \mathbf{0}_N$. The tests are performed on each EM regional equity index as well as jointly on the three EM regional indices. The results are presented for the entire sample period and three subperiods separated by two break points defined by the nine-variate structural break model.

rejections, we find the source to be mainly the EM Latin America F_2 test. Thus, there is evidence that the Latin American emerging markets index improves the location of the global minimum-variance portfolio when we consider our benchmark assets.

For the final regime, the modified likelihood ratio test and F-test reject spanning for all EM regional market indices, both individually and jointly (at the 1% level for all tests except EM Europe). The step-down tests show that Asian and European emerging markets improve the global minimum variance portfolio (F_2 test). The Latin American EM index individually, and all emerging market indices jointly can improve both the global minimum variance portfolio and the tangency portfolio during this regime.

Figure 3 provides a graphical representation of our empirical results for the overall sample period and during each regime period identified by the nine-variate model. In the figure, we plot the *ex post* efficient frontier (EF) of benchmark assets (US market, developed market, US small, big, value and growth portfolios); the EF of benchmark assets combined with the test assets (EM Asia, EM Europe and EM Latin America); and mean monthly returns/standard deviations for the three test assets. Improvement in the EF is apparent in the upper left panel (entire sample period) with the addition of emerging market regional indices to the benchmark assets. In conjunction with results from table 6, we conclude that adding emerging markets to a US or global IOS provides significant diversification benefits.

Figure 3 suggests a potentially large increase in the asymptote slope for the entire sample period as well as in the first and third regimes when emerging markets are included. The F_1 tests reported in table 6 for EM Latin America are consistent with this result. Interestingly, when the other emerging markets are included in the F_1 tests in table 6, the tangency portfolio improvement due to Latin America is only detected in the third regime. For example, the p-value for the entire sample period for all emerging markets is 0.141 and the third regime F_1 test statistics is highly significant for all emerging markets with a reported p-value of 0.006. Unfortunately, the most consistent rejections in table 6 seem to relate to the location of the global minimum variance portfolio as seen in F_2 test p-values rather than the F_1 slope test. As Kan and Zhou (2012) clarify, the source of rejections might be critical because an improvement in the location of the global minimum variance portfolio could be of lesser economic importance than an increase in the slope of the efficient frontier.† The distinction is also statistically important because the location of the global minimum variance portfolio can be estimated much more precisely than the efficient frontier slope.

In summary, we find that emerging market indices, especially EM Latin America, are not spanned by a broad set of

†We also developed a figure analogous to figure 3 where we plot the IOSs after imposing the tangency restriction given by F_1. The resultant plots demonstrate the improvement in the global minimum variance portfolio due to the emerging market portfolios in the F_2 test statistics of table 6.

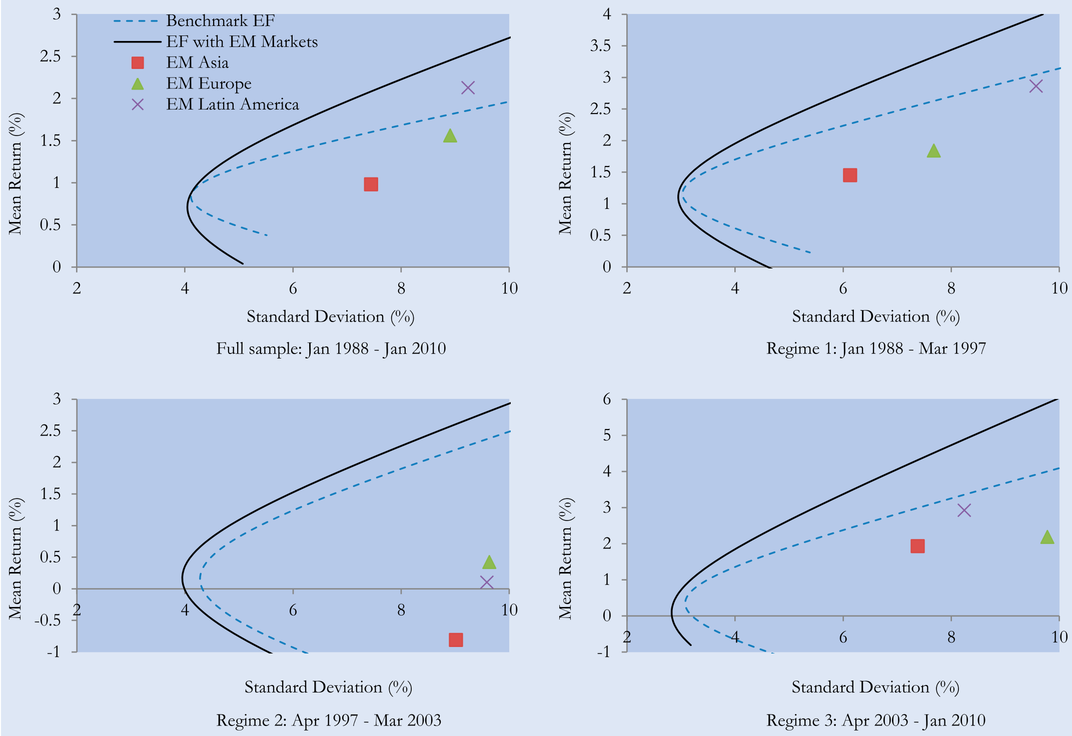


Figure 3. Efficient frontier hyperbolae of benchmark assets with and without test assets. This figure plots the ex-post efficient frontier hyperbolae of the benchmark assets (including the MSCI US and developed market indices, US small, big, value and growth portfolios) and that of the universe of all assets including benchmark and test assets (three MSCI emerging market regional indices), for the entire sample period from January 1988 to January 2010, as well as the three regimes identified by the nine-variate structural break model.

global assets including the US market index, developed market index and US small, big, value and growth portfolios. Thus, we suggest emerging markets provide significant diversification benefits to well-diversified global investors in various time periods. However, we recognize that these benefits are primarily due to an improvement in the location of the global minimum variance portfolio.

5. Conclusion

Literature examining the diversification benefits of investment in international equity markets is mixed. We suggest that variability in findings across studies of emerging market portfolio performance may be due to differences in sample periods, exclusion of specific markets within periods, or both. We find that equity returns from US, developed and emerging markets are subject to two common structural breaks over the period January 1988 to January 2010. We show that emerging market investments might provide significant performance benefits to a well-diversified global investor in various time periods. Specifically, the marginal performance benefit associated with Latin American emerging markets may be as large as 2% per month in some regimes. Our findings are consistent when we apply a one-factor market risk model and two-factor model that includes market and foreign exchange risk.

Our spanning tests show that emerging markets can improve the IOS for an investor considering US and other developed markets. The step-down tests of Kan and Zhou (2012) indicate that the inclusion of Latin American emerging markets improves the efficient frontier relative to that provided by a broad set of benchmark assets. In contrast to existing literature, our results show continued benefits to international diversification in recent periods. Observed structural changes suggest that contradictory findings in the literature may be due to shifts in asset moments over time and not necessarily due to monotonic increases in market integration. Our results suggest that a well-diversified global portfolio continues to benefit from the inclusion of emerging markets, especially Latin American equity investments.

References

Andreou, E. and Ghysels, E., Quality control for structural credit risk models. *J. Econom.*, 2008, **146**, 364–375.

Andreou, E. and Ghysels, E., Structural breaks in financial time series. In *Handbook of Financial Time Series* edited by T.G. Mikosch, J.-P. Kreiß, R.A. Davis and T. Andersen, pp. 839–870., 2009 (Springer: New York).

Andrews, D.W.K., Tests for parameter instability and structural change with unknown change point. *Econometrica*, 1993, **61**, 821–856.

Bai, J., Lumsdaine, R.L. and Stock, J.H., Testing for and dating common breaks in multivariate time series. *Rev. Econ. Stud.*, 1998, **65**, 395–432.

Bai, J. and Perron, P., Estimating and testing linear models with multiple structural changes. *Econometrica*, 1998, **66**, 47–78.

Bai, J. and Perron, P., Computation and analysis of multiple structural change models. *J. Appl. Econom.*, 2003, **18**, 1–22.

Bai, J. and Perron, P., Multiple structural change models: A simulation analysis. In *Econometric Theory and Practice: Frontiers of*

Analysis and Applied Research edited by D. Corbae, S. Durlauf and B.E. Hansen, pp. 212–238., 2006 (Cambridge University Press: Cambridge).
Banz, R.W., The relationship between return and market value of common stocks. *J. Financ. Econom.*, 1981, **9**, 3–18.
Bekaert, G., Harvey, C.R. and Lumsdaine, R.L., Dating the integration of world equity markets. *J. Financ. Econom.*, 2002, **65**, 203–247.
Bekaert, G. and Urias, M.S., Diversification, integration and emerging market closed-end funds. *J. Financ.*, 1996, **51**, 835–869.
Berger, D., Pukthuanthong, K. and Yang, J., International diversification with frontier markets. *J. Financ. Econom.*, 2011, **101**, 227–242.
Carrieri, F., Errunza, V. and Hogan, K., Characterizing world market integration through time. *J. Financ. Quant. Anal.*, 2007, **42**, 915–940.
Cumby, R.E. and Glen, J.D., Evaluating the performance of international mutual funds. *J. Financ.*, 1990, **45**, 497–521.
De Roon, F.A., Nijman, T.E. and Werker, B.J.M., Testing for mean-variance spanning with short sales constraints and transaction costs: The case of emerging markets. *J. Financ.*, 2001, **56**, 721–742.
De Santis, G. and Gerard, B., International asset pricing and portfolio diversification with time-varying risk. *J. Financ.*, 1997, **52**, 1881–1912.
Dumas, B. and Solnik, B., The world price of foreign exchange risk. *J. Financ.*, 1995, **50**, 445–479.
Fama, E.F. and French, K., The cross-section of expected stock returns. *J. Financ.*, 1992, **47**, 427–465.
Fama, E.F. and French, K., Common risk factors in the returns on stocks and bonds. *J. Financ. Econom.*, 1993, **33**, 3–56.
Fama, E.F. and French, K., Multifactor explanations of asset pricing anomalies. *J. Financ.*, 1996, **51**, 55–84.
Fama, E.F. and French, K., Value versus growth: The international evidence. *J. Financ.*, 1998, **53**, 1975–1999.
Ferson, W.E. and Harvey, C.R., The risk and predictability of international equity returns. *Rev. Financ. Stud.*, 1993, **6**, 527–566.
Ferson, W.E. and Harvey, C.R., Sources of risk and expected returns in global equity markets. *J. Bank. Financ.*, 1994, **18**, 775–803.
Furno, M., Quantile regression estimates and the analysis of structural breaks. *Quant. Financ.*, 2012, 1–8. doi:10.1080/14697688.2011.653387.
Garcia, R. and Ghysels, E., Structural change and asset pricing in emerging markets. *J. Int. Money Financ.*, 1998, **17**, 455–473.
Gerlach, R., Wilson, P. and Zurbruegg, R., Structural breaks and diversification: The impact of the 1997 Asian financial crisis on the integration of Asia-Pacific real estate markets. *J. Int. Money Financ.*, 2006, **25**, 974–991.
Hardouvelis, G.A., Malliaropoulos, D. and Priestley, R., EMU and European stock market integration. *J. Bus.*, 2006, **79**, 369–392.
Harvey, C.R., Predictable risk and returns in emerging markets. *Rev. Financ. Stud.*, 1995, **8**, 773–816.
Harvey, C.R., Solnik, B. and Zhou, G., What determines expected international asset returns? *Ann. Econ. Financ.*, 2002, **3**, 249–298.
Hillebrand, E., Neglecting parameter changes in GARCH models. *J. Econom.*, 2005, **129**, 121–138.
Huberman, G. and Kandel, S., Mean-variance spanning. *J. Financ.*, 1987, **42**, 873–888.
Jobson, J.D. and Korkie, B., A performance interpretation of multivariate tests of asset set intersection, spanning, and mean-variance efficiency. *J. Financ. Quant. Anal.*, 1989, **24**, 185–204.
Kan, R. and Zhou, G., Tests of mean-variance spanning. *Ann. Econ. Financ.*, 2012, **13**(1), 139–187.
Karolyi, G.A. and Stulz, R.M., Are financial assets priced locally or globally? In *The Handbook of the Economics of Finance*, Vol. 1, Chapter 16. edited by G.M. Constantinides, M. Harris and R.M. Stulz, pp. 975–1020., 2003 (Elsevier: Amsterdam).
Lewis, K.K., Is the international diversification potential diminishing for foreign equity inside the US? Wharton School. Working paper, University of Pennsylvania, 2007.
Li, K., Sarkar, A. and Wang, Z., Diversification benefits of emerging markets subject to portfolio constraints. *J. Empir. Financ.*, 2003, **10**, 57–80.
Lien, D., Tse, Y.K. and Zhang, X., Structural change and lead-lag relationship between the Nikkei spot index and futures price: A genetic programming approach. *Quant. Financ.*, 2003, **3**, 136–144.
Lintner, J., Security prices, risk, and maximal gains from diversification. *J. Financ.*, 1965, **20**, 587–615.
Meligkotsidou, L. and Vrontos, I.D., Detecting structural breaks and identifying risk factors in hedge fund returns: A Bayesian approach. *J. Bank. Financ.*, 2008, **32**, 2471–2481.
Merton, R.C., An intertemporal capital asset pricing model. *Econometrica*, 1973, **41**, 867–887.
Mossin, J., Equilibrium in a capital asset market. *Econometrica*, 1966, **34**, 768–783.
Pástor, Ľ. and Stambaugh, R.F., The equity premium and structural breaks. *J. Financ.*, 2001, **56**, 1207–1239.
Pettenuzzo, D. and Timmermann, A., Predictability of stock returns and asset allocation under structural breaks. *J. Bank. Financ.*, 2011, **164**, 60–78.
Pukthuanthong, K. and Roll, R., Global market integration: An alternative measure and its application. *J. Financ. Econ.*, 2009, **94**, 214–232.
Qu, Z. and Perron, P., Estimating and testing structural changes in multivariate regressions. *Econometrica*, 2007, **75**, 459–502.
Rosenberg, B., Reid, K. and Lanstein, R., Persuasive evidence of market inefficiency. *J. Portfolio Manage.*, 1985, **11**, 9–16.
Sharpe, W.F., Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Financ.*, 1964, **19**, 425–442.
Sharpe, W.F., Asset allocation. *J. Portfolio Manage.*, 1992, **18**(2), 7–19.
Solnik, B., An equilibrium model of the international capital market. *J. Econ. Theory*, 1974a, **8**, 500–524.
Solnik, B., The international pricing of risk: An empirical investigation of the world capital market structure. *J. Financ.*, 1974b, **29**(2), 365–378.
Stattman, D., Book values and stock returns. *Chicago MBA: A J. Select Papers*, 1980, **4**, 25–45.
Turtle, H.J. and Zhang, C., Time-varying performance of international mutual funds. *J. Empir. Financ.*, 2012, **19**, 334–348.