

IMF WORKING PAPER

© 1996 International Monetary Fund

This is a *Working Paper* and the author(s) would welcome any comments on the present text. Citations should refer to a *Working Paper* of the International Monetary Fund, mentioning the author(s), and the date of issuance. The views expressed are those of the author(s) and do not necessarily represent those of the Fund.

WP/96/28

INTERNATIONAL MONETARY FUND

European II Department

Comovements in National Stock Market Returns:
Evidence of Predictability but not Cointegration

Prepared by Anthony J. Richards 1/

Authorized for distribution by Julian Berengaut

April 1996

Abstract

This paper is a response to the literature that tests for cointegration between national stock market indices. It argues that apparent findings of cointegration in other studies may often be due to the use of asymptotic, rather than small-sample, critical values. In fact, economic theory suggests that cointegration is unlikely to be observed in efficient markets. However, this paper finds some evidence for the long-horizon predictability of relative returns, and the existence of "winner-loser" reversals across 16 national equity markets. A conclusion is that national stock market indices include a common world component and two country-specific components, one permanent and one transitory.

JEL CLASSIFICATION NUMBERS: G15; G12; C32

1/ This paper is forthcoming in the Journal of Monetary Economics. I am grateful to an anonymous referee for very helpful comments on earlier drafts and to Morgan Stanley and Co., Kenneth Kasa, and David Cutler for providing some of the data used in this study. Thanks also to Peter Cornelius, Graham Elliott, Albert Jaeger, Mark Klock, Karen Parker and Heidi Willmann Richards for comments, and Shahbaz Khan for computing advice. The views expressed in this paper are those of the author, and do not necessarily reflect those of the International Monetary Fund.

| <u>Table of Contents</u> | <u>Page</u> |
|---|-------------|
| Summary | iii |
| I. Introduction | 1 |
| II. Cointegration and its Relation to Asset Pricing | 2 |
| a. Background and cointegration | 2 |
| b. Relevance of cointegration for asset pricing | 3 |
| III. Previous Empirical Work Testing for Cointegration between Stock Market Indices | 7 |
| a. Literature review | 7 |
| b. Simulation evidence on small sample properties of multivariate cointegration tests | 8 |
| IV. Tests for Cointegration between National Return Indices and Rest-of-World Return Indices | 10 |
| a. Data | 10 |
| b. Results | 11 |
| V. Testing for Long-Horizon Predictability in National Stock Market Returns | 12 |
| a. Regression tests of predictability of relative returns | 13 |
| b. Testing for winner-loser effects across national stock markets | 15 |
| VI. Conclusion | 17 |
| <u>Text Tables</u> | |
| 1. Analysis of Small-Sample Properties of Johansen Cointegration Tests | 19 |
| 2. Tests for Cointegration between Country and Rest-of-World Equity Return Indices | 20 |
| 3. Tests for Predictability of Relative Returns | 21 |
| 4. Testing for Winner-Loser Effects among National Stock Market Return Indices | 22 |
| <u>References</u> | 23 |

Summary

This paper presents tests for the predictability of long-horizon returns in the national stock markets of 16 industrial countries. It is motivated by a series of papers that have tested for cointegration across national stock market indices, many of which have claimed to find some evidence of cointegration and of "long-run linkages" in national indices.

The literature on cointegration is reviewed in light of standard asset pricing theory, which implies the existence of a random-walk component in asset prices. This suggests that cointegration, and the strong form of predictability that it implies, should not exist between stock market indices in integrated markets. The paper illustrates the fragility of some tests for cointegration by showing that small-sample bias explains one recent finding of four cointegrating vectors (and a single common stochastic trend) in the stock prices of five major industrial countries.

Alternative tests for cointegration are presented, and these confirm that the null hypothesis of no cointegration is typically not rejected. This is further evidence against the notion of cointegration around a single common trend and implies the need for caution in the burgeoning practice of regressing large numbers of country pairs or groups against each other in search of "cointegrating relationships." "Significant" findings of cointegration should therefore be assessed in light of the poor small-sample properties of some of the tests and the economic model, if any, that is offered to explain the estimated relationship.

Two other methodologies are used to investigate whether national stock market returns demonstrate some weaker form of long-horizon relative predictability. First, regression tests are used to examine autocorrelations in relative returns. These suggest the presence of positive autocorrelations at horizons of up to about one year, and negative autocorrelations at horizons of two years or more. Second, tests for "winner-loser" reversals are conducted, confirming this pattern of autocorrelations. These results could be consistent either with an equilibrium model of time-varying risk or with market inefficiency resulting from the existence of fads, bubbles, sentiment, or the like. Regardless of the causes, the predictability in relative returns appears weaker than would be implied by cointegration around a single common trend. The paper concludes, therefore, that national return indices include a common world component and two country-specific components, one permanent and one transitory. The existence of a permanent component implies that there will typically be long-run risk-reduction benefits from investing abroad.

I. Introduction

Recent research into international equity price behavior might be characterized as falling into four groups. First, there is a long literature (e.g., French and Poterba, 1991) that investigates the correlations between national stock markets and the implied benefits of international diversification, typically in a mean-variance framework. Recent increases in cross-border investment (see Frankel, 1994) suggest that these benefits have finally found greater recognition among practitioners and investors. Second, there is a growing amount of research (e.g., Harvey, 1991; Engel, 1994) that investigates the extent to which equity returns can be explained by theories of international asset pricing. A third area of the literature (e.g., King and Wadhvani, 1990), stimulated by the stock market crash of 1987, has been concerned with the transmission of information and shocks between national markets. A fourth branch (e.g., Cutler, Poterba and Summers, 1991) has examined the extent to which equity returns in different countries appear to demonstrate predictability in the same way that returns in the United States have been shown to be somewhat predictable. Such predictability may be due to fads, overreaction or other types of irrationality, or it may be attributable to time-varying risk within an equilibrium framework.

This paper follows the spirit of the fourth branch of research. In contrast to work by Harvey (1991) and others showing that a small number of common variables has some degree of explanatory power for monthly returns in different countries, the focus in this paper is on long-horizon returns. It is motivated by a series of papers that test for cointegration across national stock market indices, including a paper by Kasa (1992) which argues that the price (and dividend) indices for the equity markets of five major industrial countries are all cointegrated around a single common stochastic trend. This finding would imply that returns in these markets may follow different patterns in the short term, but that in the long run, the levels of total return indices (i.e., prices plus reinvested dividends) are very closely linked. The existence of such a cointegrating relationship might provide an explanation for the hitherto low level of international diversification by investors, as the benefits to diversification that are implied by the relatively low short-run correlations would disappear in the long run as correlations become stronger. However, this long-run relationship would also imply that returns will demonstrate significant predictability, and--unless explained by time-varying risk factors or market segmentation--would be contrary to the weak form of the Efficient Markets Hypothesis, which requires that risk-adjusted returns not be predictable based on easily available information such as their own history.

To further investigate this issue, this paper examines the statistical basis for the rejection by Kasa (1992) of the null hypothesis of no cointegration between different markets. Simulation evidence suggests that the finding of such a strong cointegrating relationship is due to a failure to adjust asymptotic critical values to take account of the small number of degrees of freedom that remain in the Johansen (1988) multivariate

estimation procedure. Alternative tests for cointegration are presented, and these confirm that the null of no cointegration is typically not rejected. This is further evidence against the notion of cointegration around a single common trend. The evidence that cointegration does not generally exist between return series is hardly surprising, given that basic models of asset pricing would preclude cointegration.

Two other methodologies are used to investigate whether national stock market returns demonstrate some weaker form of long-horizon relative predictability. These are inspired by the regression approach of Fama and French (1988) and the "winner-loser" effect investigated by DeBondt and Thaler (1985). There is evidence for the predictability of long-horizon relative returns, whereby markets that have performed strongly subsequently underperform, and vice versa. The paper concludes, therefore, that national return indices include a common world component and two country-specific components, one permanent and one transitory.

II. Cointegration and its Relation to Asset Pricing

a. Background on cointegration

Consider two time series, y_t and x_t , which are both integrated of order d (henceforth denoted as $I(d)$). In general, any linear combination of these series will also be $I(d)$. However, if there exists a linear combination of the variables which is $I(d-b)$ where $b > 0$, then y_t and x_t are said to be cointegrated of order (d,b) (see Engle and Granger, 1987; Campbell and Perron, 1991). Engle and Granger suggest a procedure whereby y_t is regressed against x_t using ordinary least squares, and the residuals from this regression are then tested for non-stationarity using a variant of the Augmented Dickey-Fuller test. Engle and Granger also show that the existence of cointegration implies an "error correction representation", by which the change in at least one of the variables will be a function of the "disequilibrium" in a previous period. That is, cointegration implies a particular type of predictability between the variables in the system.

Johansen (1988) and Johansen and Juselius (1990) introduce a maximum likelihood test procedure that allows for the possibility of multiple cointegrating vectors in a multivariate framework. This approach starts with a j -th order vector autoregression (VAR) model for X_t (a vector of n $I(1)$ variables) with independent Gaussian errors,

$$X_t = A_1 X_{t-1} + \dots + A_j X_{t-j} + \mu + v_t, \quad (1)$$

which can be rewritten in an error-correction, or differenced, form

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_{j-1} \Delta X_{t-j+1} + \Pi X_{t-j} + \mu + v_t. \quad (2)$$

Cointegration implies that levels terms (i.e., X_{t-j} or, if reparameterized, X_{t-1}) will still be present in the differenced equation, so one can test for

cointegration by examining the rank of the estimated Π matrix. If the rank of Π is r , where $r \leq n-1$, then there exist r linear independent cointegrating vectors. The rank of Π is tested using two likelihood-ratio tests based on the eigenvalues of a matrix from auxiliary regressions; full details are given in Johansen (1988) and Johansen and Juselius (1990). From the number of cointegrating vectors (r) and the number of variables in the system (n), one can infer the number of "common stochastic trends" driving the system (equal to $n-r$).

In practice, the tests are implemented by first testing the null hypothesis of $r = 0$, and then if it is rejected, testing $r \leq 1$, and so on. Like the Dickey-Fuller test statistics, the critical values do not follow conventional distributions, but simulated asymptotic distributions are provided by Johansen and Juselius (1990). Recent work has suggested, however, that asymptotic critical values may be misleading in small samples. Gregory (1994) shows that the size (the rejection frequency when the null hypothesis of no cointegration is true) of the Johansen tests is significantly higher (i.e., worse) in cases of small samples and a high number of explanatory variables than the size of other tests of cointegration. In particular, as the VAR sample size (T) falls, or the number of variables (n) or lags (j) in the system increases, the critical values should be adjusted upwards (or the test statistics downwards). 1/ Reinsel and Ahn (1988) suggest that the critical values be adjusted upwards by a multiplicative scaling factor or "degrees-of-freedom correction term" given by $T/(T-nj)$. The work of Cheung and Lai (1993) suggests that the implied critical values are a very significant improvement over the asymptotic critical values. 2/

b. Relevance of cointegration for asset pricing

The idea that cointegration has implications for financial market efficiency was introduced by Granger (1986), who argued that two price series that are determined in efficient markets--e.g., the prices of gold and silver--cannot be cointegrated. The implication follows from the correspondence between cointegration and error correction, and the argument that prices should not be predictable in an efficient market.

1/ To be specific, there is an upward bias to the test statistics as the number of degrees of freedom approaches zero, because the canonical correlations between the X_{t-j} and ΔX_t (corrected for the lagged differences) will approach unity even if the series are not cointegrated (in which case they should approach zero); see also Hall (1991).

2/ Cheung and Lai (1993) suggest equations for small-sample critical values that are approximately given by $CR_{\infty} * [0.1 + 0.9 * T/(T-nj)]$, implying that the Reinsel-Ahn scaling factor overcorrects slightly. This conclusion is based on their exclusion of cases of very few degrees of freedom. As is shown in Section 3b, the Reinsel-Ahn correction may actually undercorrect in such cases.

Two qualifications apply, however, to Granger's statement. First, the assertion that predictability of prices is contrary to market efficiency applies strictly only to non-interest- or non-dividend-bearing assets, such as the precious metals in his example. Stock prices, for example, will be highly predictable on days when stocks go ex-dividend, though total returns (including dividends) should not be predictable. The depreciation of the exchange rate of a country experiencing hyperinflation may also be highly predictable, but the total return (including the interest differential) from a strategy that tries to exploit this should not be. This implies that the countless studies that have used tests for cointegration of stock prices or exchange rates as evidence for or against market efficiency may be invalid unless dividends and interest differentials are unimportant (see also Engel, 1995). ^{1/} Second, as is well known, predictability of returns will have no implications for market efficiency unless returns are measured in risk-adjusted terms. While Section 5 does present evidence for the predictability of returns, the issue of risk is not addressed in this paper, so no conclusions about market efficiency can be drawn; see Richards (1996a) for further analysis of this issue.

Under standard assumptions, it is easily shown that a total return index should contain a random-walk component. ^{2/} In the following discussion, excess returns (returns in excess of the risk-free rate) are used for convenience, and the expected (excess) rate of return in continuously compounded terms is denoted by $E_{t-1}(r_{i,t})$. The end-period (excess) return index (in logarithmic form) for asset i , $a_{i,t}$, will be given by:

$$a_{i,t} = a_{i,t-1} + E_{t-1}(r_{i,t}) + \epsilon_{i,t}, \quad (3)$$

where $\epsilon_{i,t}$ is the unexpected return in period t , which should be white noise if expectations are rational. Note that $\epsilon_{i,t}$ should be white-noise even

^{1/} A simple example illustrates how cointegration of stock prices (without dividends) might be theoretically feasible in an efficient market. Assume that managers believed there is an optimal and changing trading price range for their stock where the stock will attract the greatest investor interest. Managers might use dividend policy to keep the stock price within this range, paying very high dividends when earnings are high, and omitting the dividend if the price fell below the desired range. The prices of unrelated stocks might thus appear to be cointegrated around this range, though it is most unlikely that their total returns indices would also be cointegrated. Of course, dividends do not show the level of volatility that would be required in this example (or a related example in Dwyer and Wallace, 1992), so cointegration of stock prices may not be especially likely in practice.

^{2/} The following analysis is implicitly carried out in a common currency. For assets denominated in different currencies, the unexpected return could be disaggregated into its exchange-rate and domestic-currency components; this would simply add a third cumulated error term to equation 6.

with factors such as autocorrelation in the dividend process: if the time-series model for dividends is known, the jump in the return index at the time of any initial dividend shock should fully reflect the present discounted value of all associated innovations. It follows from equation 3 that

$$a_{i,t} = a_{i,0} + \sum_{j=1}^t E_{j-1}(r_{i,j}) + \sum_{j=1}^t \epsilon_{i,j}. \quad (4)$$

That is, the return index will be equal to its starting value, plus the cumulated expected return, plus a random-walk term that represents the cumulated expectational errors.

Now assume that excess returns are generated by the CAPM:

$$E_{t-1}(r_{i,t}) = \beta_{i,t} E_{t-1}(rp_t), \quad (5)$$

where rp_t denotes the continuously compounded market risk premium, and the risk factor $\beta_{i,t}$ is not necessarily constant. Cointegration of the return indices of two assets ($i=a,b$) with a cointegrating vector of $(1, -\alpha)$ would require that:

$$(\beta_{a,t} - \alpha\beta_{b,t}) \sum_{j=1}^t E_{j-1}(rp_j) + \left(\sum_{j=1}^t \epsilon_{a,j} - \alpha \sum_{j=1}^t \epsilon_{b,j} \right) = I(0). \quad (6)$$

Under the assumptions that the cumulated risk premium term is non-stationary, and the cumulated risk premium and $\epsilon_{i,t}$ terms are independent, cointegration would have two implications. First, cointegration would require that the term $(\beta_{a,t} - \alpha\beta_{b,t})$ is stationary with a mean of zero, implying a relationship between the long-run expected returns of two cointegrated assets. Second, the cumulated error terms would have to be cointegrated with coefficient α (or both be $I(0)$), implying a precise relationship between the cumulated unexpected returns of the two assets. ^{1/}

A precise long-run relationship between the expected returns on two assets is not precluded by standard asset pricing models, though it may be unlikely based on evidence that CAPM stock market β 's show significant variation over time, as would be expected given that changes in stock prices cause changes in leverage and thus in risk. However, the second implication regarding the cumulated unexpected components implies that one of the $\epsilon_{i,t}$ terms will predict the other, which is directly at odds with the assumption of rational expectations and with market efficiency. It is clear that $\epsilon_{a,t}$ and $\epsilon_{b,t}$ may be correlated within-period: when the overall national or world

^{1/} A corresponding ARMA representation for the total return is derived by Dwyer and Wallace (1992).

market is stronger than expected, most stocks or national indices will rise. However, cointegration of the cumulated error terms would imply that all company- or country-specific returns shocks are always followed by unexpected but exactly offsetting shocks. At the firm level, for example, this would rule out the possibility that any decisions by corporate management could ever have a permanent effect on a company's return index relative to its competitors. At the national level, a comment by Solnik (1991, p.46) may be relevant: "Some investigators have attempted to find leads or lags between markets. However, no evidence of a systematic delayed reaction of one national market to another has ever been found. The existence of such simple market inefficiencies is, indeed, unlikely, since it would be so easy to exploit them to make an abnormal profit." One should conclude, therefore, that cointegration of return indices is directly at odds with market efficiency.

The above analysis can shed light on attempts by a number of researchers (starting with Taylor and Tonks, 1989) to link the statistical concept of cointegration with the capital markets concept of integration. It has frequently been suggested, for example, that cointegration between national markets will be an indication of a high degree of capital market integration, and that globalization and liberalization of financial markets may cause cointegration of market indices. However, as is well known, integration of financial markets implies simply that the same model of required returns will apply to all assets, regardless of where they are traded. If the CAPM held for all assets, for example, a cointegrating relationship would exist between the cumulative expected returns on any two assets with stationary betas. However, as was shown above, the random-walk component implies that this will not translate into a cointegrating relationship between the actual return indices for the two (or more) assets.

Globalization and increasing integration of economies may well lead to higher short-run correlations between markets as national business cycles and markets become more synchronized (see Bekaert, 1995), but high short-run correlations imply nothing about the nature of the long-run correlations. It is indeed paradoxical that these trends should actually lead to stronger evidence against cointegration. The higher short-run correlations and reduced country-specific noise in price and return indices will mean, for example, that the estimated error term in Engle-Granger "cointegrating regressions" should become smaller on average, and a purer estimate of its time series properties should become possible. And if cointegrating relationships do exist, these trends imply that arbitrageurs will be able to estimate these relationships more accurately, and invest with greater confidence that there will be less noise in the data and thus less risk in their arbitrage positions. In summary, cointegration would be more an indication of market inefficiency than of market integration.

III. Previous Empirical Work Testing for Cointegration between Stock Market Indices

a. Literature review

Since stock price series are almost invariably found to be $I(1)$, tests for cointegration have frequently been used to address questions about the nature of the long-run relationship between the stock prices of different companies or the overall market indices of different countries. An example of tests using individual stock prices is the work of Stengos and Panas (1992), who test for bivariate cointegration between the drachma stock prices of banks listed on the Athens stock exchange. Using daily price data over 1985-1988, they find no evidence of cointegration using the Engle-Granger methodology and interpret their results as being consistent with market efficiency.

Among the numerous papers testing the null hypothesis of no cointegration between national stock price indices, the strongest rejection is provided by Kasa (1992). Using indices in U.S. dollar terms, deflated by the U.S. consumer price index, for the United States, Canada, Germany, Japan and the United Kingdom, Kasa initially presents results using monthly data for 1974:2-1990:8, which suggest the presence of a single cointegrating vector in a Johansen system. Using quarterly data for 1974:1-1990:3 and two lags in the VAR, the results continue to imply the presence of a single cointegrating vector, though for a VAR of ten lags, the results imply the presence of three or four cointegrating vectors. Kasa concludes that there are four cointegrating vectors, implying that stock prices in these countries are all driven by a single common stochastic trend.

This conclusion is somewhat at odds with other work in the area which provides more mixed evidence for cointegration between national indices, and suggests at most a small number of cointegrating vectors in VAR systems (e.g., Arshanapalli and Doukas, 1993; Byers and Peel, 1993; Corhay et al., 1993). 1/ If the finding of a single common trend were robust, it would

1/ A cursory search located nearly 20 recent journal articles testing for cointegration between national market indices, with a substantial majority claiming to find some evidence for cointegration. Three caveats might be noted. First, many papers conduct a large number of tests, and the highlighted rejections of the null are often only slightly more frequent than implied by the size of the tests, which is often set higher than five percent; this suggests a publication bias towards papers that find, rather than fail to find, cointegration. Second, rejections are more common in multi-country tests using the Johansen methodology without a small-sample correction. Third, an economic interpretation of the estimated stationary vectors is rarely provided. How, for example, should one interpret the following near-stationary vector in the five-country dataset studied in this section: $(1.20 \times \text{United States} + 2.00 \times \text{United Kingdom} - 0.21 \times \text{Canada} - 0.65 \times \text{Germany} - 0.98 \times \text{Japan})$?

imply that all of the indices should be cointegrated with the (unobservable) common trend, and therefore that any bivariate system of these variables should also be cointegrated, a proposition for which there is little support. This implication follows from the fact that bivariate cointegration is "transitive" (Taylor and Tonks, 1989). In addition, as noted above, the rejection of the null of no cointegration implies the predictability of some of the series, and may be contrary to the Efficient Markets Hypothesis.

One possible reason for Kasa's apparently anomalous results is that the VARs in his work include up to ten quarters of lagged data, while other researchers using equity price data typically include lags of a few months at most. Kasa argues for long lag structures both to capture the possible effect of mean reversion in equity prices, and to make the error terms from the VARs more consistent with the Gaussian/i.i.d. assumption under which the Johansen methodology is derived. Kasa's tests imply, however, that serial correlation is not present in the low-order VARs, so mean reversion would not appear to be a major factor. And the addition of extra lags to "remove" the non-normality of residuals would be inappropriate if changes in stock prices are fundamentally fat-tailed or otherwise non-normal. Asymptotic critical values for the Jarque-Bera test are also unlikely to be appropriate for Kasa's preferred VAR equations which contain 51 explanatory variables and only 57 observations.

An alternative approach to the choice of lag structure would have been to start from a high-order VAR model and then test the restriction implied by a shorter lag structure. Hall (1991) notes that this may be preferable given that it can be regarded as a general test of whether all important information has been included in the VAR that is finally selected. Tests for the appropriate lag structure for Kasa's data suggest, indeed, that the data might be best represented by a single lag in the VAR. ^{1/} This implies that the VARs used to generate Kasa's results may have been significantly over-parameterized.

b. Simulation evidence on small sample properties of multivariate cointegration tests

To illustrate the possible sensitivity of cointegration tests to the number of lags included in the VAR, a simple simulation experiment was designed to reflect the basic properties of the data used in Kasa's (1992) study. The technique of randomization is used (see Kim et al., 1991, for another example) to obviate the need for the Monte Carlo assumption of

^{1/} The Akaike and Schwarz-Bayes information criteria both suggest only one lag in the VAR. Using the Sims likelihood-ratio test with a five percent significance level, one cannot reject successive reductions from $j=10$ until the restriction of $j=3$ is rejected against $j=4$, though this rejection may be spurious; the restriction of $j=1$ is not rejected against $j=4$ at conventional significance levels.

normally distributed innovations. Systems of five randomized, non-cointegrated series were generated by cumulating the innovations drawn, without replacement, from the innovations (i.e., first differences) in the log real stock price series used by Kasa. The innovations were "reshuffled" in parallel so as to maintain the within-period correlation between the series, though all time dependence is removed. ^{1/} Cointegration tests were conducted for 1000 replications, and following Kasa, the system was tested using a base sample of 67 observations, which for lag lengths (j) of two and ten, results in VAR sample sizes (T) of 65 and 57, respectively. The test statistics from each replication are compared sequentially with the critical values to determine the number of cointegrating vectors in the system.

The simulation results are shown in Table 1. The first four columns show the test statistics obtained by Kasa and the mean test statistics from the 1000 simulations with randomized data. These test statistics can be compared with the data in the next four columns for the asymptotic (Johansen and Juselius, 1990, Table A2) and small-sample (Reinsel and Ahn, 1988) five percent critical values. Finally, the last four columns show the rejection frequencies from 1000 simulations of the randomized data, for the different tests and critical values.

There is a large difference between the results for the different lag structures. For a VAR with two lags, the simulations indicate that the null of no cointegration ($r=0$) is incorrectly rejected in around 20 percent of simulations when using the asymptotic critical values, though the small-sample critical values imply rejection in around only five percent of replications. When ten lags are included in the VAR, the use of the asymptotic critical values implies without exception that the system is cointegrated. In around 96 percent of the simulations one would reject $r \leq 3$, suggesting that there might be four cointegrating vectors. However, in around one-third of the replications, the hypothesis of $r \leq 4$ would also be rejected, implying that the five variables in the system are trend-stationary, and do not belong in a cointegrating system. When the small-sample critical values are used, the hypothesis of no cointegration ($r=0$) is rejected in only around 9 or 18 percent of percent of replications, depending on which likelihood-ratio test is used. That is, the Reinsel-Ahn small-sample critical values are a far better guide to the true critical

^{1/} By simply reshuffling the innovations here and in the remainder of the paper, all simulations impose zero coefficients on all but the first lag in the VAR. This is based on the absence of ARCH effects in these quarterly data, the weak evidence for a higher-order VAR (see footnote 7), and the fact that the data are financial market data which should be close to random walks under an efficient markets null. Alternative simulations suggest that estimates of the small-sample bias are little affected by this assumption. Furthermore, Monte Carlo (with normal innovations) and bootstrapped simulations (sampling with replacement) yield very similar results to the results shown.

values than the asymptotic values, though in this extreme case they are still subject to a modest tendency to over-reject.

If the Reinsel-Ahn critical values are used with the test statistics obtained by Kasa, the hypothesis of $r=0$ is not rejected for either two or ten lags in the VAR. This implies that there are zero cointegrating vectors in the system, rather than the four cointegrating vectors suggested by Kasa. The reason for the large discrepancy in the case of the ten-lag VAR is that the appropriate critical values appear to be at least 8 times higher than the asymptotic values. ^{1/} In cases of such extremely small samples, asymptotic results are unlikely to hold, requiring researchers to examine the robustness of their results through Monte Carlo analysis or some similar technique. In the current case, it would seem that there are no firm grounds for the rejection of the null hypothesis of no cointegration.

IV. Tests for Cointegration between National Return Indices and Rest-of-World Return Indices

The evidence of the previous section suggests that, with relatively short sample periods, there will frequently be small-sample problems in the estimation of multivariate cointegrating systems. If one nonetheless wishes to test for cointegration, it may be necessary to use methods that "collapse" a large number of variables into a smaller number. This section will test for cointegration between the total return indices for a group of countries with a capitalization-weighted "rest-of-world" series calculated for each country. This test offers an indication of the tendency (if any) for the return indices of "domestic" and "foreign" assets not to drift too far apart, and may provide the simplest answer to the question of whether a typical investor will gain long-run diversification benefits from investing abroad.

a. Data

This section describes the data that will be used for the tests of cointegration, as well as for the subsequent tests. As in many other studies, the data used are the Morgan Stanley Capital International indices in U.S. dollars. Unlike most previous work which uses price indices, indices for total equity returns (capital gains plus dividends) are used for all tests, in order to capture the total return received by investors. Data for the following sixteen markets are used; Australia, Austria, Canada, Denmark, France, Germany, Hong Kong, Italy, Japan, the Netherlands, Norway,

^{1/} The Reinsel-Ahn scaling factor is equal to 8.14 (i.e. $57/(57-5*10)$); the simulations suggest scaling factors for the null of $r=0$ of 8.8 (trace statistic) and 10.6 (maximal eigenvalue). If a further lag had been added to the VARs, there would have been 56 explanatory variables and 56 observations, rendering tests of cointegration impossible.

Spain, Sweden, Switzerland, the United Kingdom, and the United States. ^{1/} All tests use end-quarter data from end-December 1969 to end-December 1994, though it should be noted that results for the post-1973 floating exchange rate period are similar to those shown in the paper.

End-quarter market capitalization weights were estimated based on returns data, and end-year market capitalization data from Morgan Stanley, Ibbotson et al. (1982), and the Federation Internationale des Bourses de Valeurs Statistics Yearbook (various issues). Data for the cumulative U.S. risk-free rate for 1969-1993 are taken from the Treasury bills total return index in Ibbotson Associates' SBBI 1994 Yearbook, and are updated for 1994 using series 60c from International Financial Statistics.

b. Results

The Johansen and Engle-Granger methodologies were used to test for cointegration between the return index for each country and the rest-of-world index. In each case, the series are U.S. dollar excess return indices in logarithms (the log of the return indices in U.S. dollars less the log of the index for the cumulative U.S. risk-free rate). ^{2/} Augmented Dickey-Fuller (ADF) tests indicated that the null of non-stationarity could not be rejected for any of the series, while for first differences, the null was rejected in every instance. This provides strong evidence that the series are all I(1).

For consistency with Kasa's (1992) study, results for VARs of two and ten lags are presented for the Johansen methodology, though in the majority of cases a single lag in the VAR seemed appropriate based on the Akaike criterion and likelihood-ratio tests. Statistical inference is based on the 95 percent fractile from 1000 randomizations for each country under the null hypothesis of no cointegration. The average of these critical values is provided in the results, along with the asymptotic and small-sample values. ^{3/}

For the Engle-Granger tests, the long-run equation regresses the log of the excess return index for the individual country against the log of the excess rest-of-world return index, a constant, and a time trend; however,

^{1/} Data for two series calculated by Morgan Stanley since 1969--Belgium and Singapore/Malaysia--were not available and were not included in the study.

^{2/} Similar results (available upon request) are obtained for both simple (not excess) return indices and hedged domestic-currency return indices, where the latter were proxied by assuming the currency risk on a foreign equity position is offset by a short position in a foreign short-term interest-bearing security.

^{3/} As in Kasa (1992), asymptotic critical values are from Table A2 in Johansen and Juselius (1990); the values in Table A1 may be more appropriate but are similar and yield similar conclusions.

the results do not appear especially sensitive to the assumptions in regard to the trend. The number of differenced residual terms included in each test equation was determined by starting with ten lags and testing down. Asymptotic ADF statistics are given by Phillips and Ouliaris (1990, Table 2c), though inference is based on the five percent values derived from 1000 randomizations under the null of no cointegration.

The results are shown in Table 2. First, it might be noted that the asymptotic critical values for the Johansen tests are relatively close to the empirically determined critical values in the case of only two lags in the VAR. However, the Reinsel-Ahn small-sample values are more appropriate in the case of ten lags, providing further evidence that this adjustment is a useful short-cut for researchers wishing to avoid the small-sample problems of the Johansen tests. The small-sample bias is, however, smaller than the example in Table 1 because of the use of longer time series and a VAR model with two rather than five variables. With regard to the cointegration tests, there is some conflict between the test results, as is quite typical (see Gregory, 1994). However, based on the simulated critical values, there are only occasional rejections of the null hypothesis of no cointegration. This provides further evidence that there are no firm grounds for concluding that national stock market indices are cointegrated around a single common trend. ^{1/} As a result, one might conclude that foreign and domestic equity markets will generally move in a significantly different manner in the long run, indicating that there will typically be substantial risk-reduction benefits from investing abroad.

V. Testing for Long-Horizon Predictability in National Stock Market Returns

The previous sections have suggested that there is little evidence for the proposition that return indices for major national stock markets are all cointegrated around a single common trend. This implies that the strong form of predictability of return indices that would be implied by the correspondence between cointegration and error correction can be rejected. An error correction model would imply that when the return index for one country deviated from the common trend, it should revert back, and periods of "overperformance" relative to the common trend would be followed by periods of exactly offsetting "underperformance."

While this strong form of predictability of returns is typically rejected, it is possible that there may be some weaker degree of relative predictability among national equity markets; predictability is necessary, but not sufficient for cointegration. Furthermore, while the predictability

^{1/} An alternative approach is to test for pairwise cointegrating relationships between the 16 countries in the sample. Based on the Engle-Granger test, the 120 country pairs yield only 6 rejections of the null of no cointegration at the 5 percent level, and 18 rejections at the 10 percent level.

of unexpected returns implied by cointegration would be contrary to market efficiency, a weaker form of predictability might be related to time-varying risk and might still be consistent with market efficiency. To test for predictability, it may be useful to apply some of those tests for the predictability of equity returns that have been applied to stocks and indices within national markets.

a. Regression tests of predictability of relative returns

Instead of focussing on the predictability of excess returns (i.e., returns above the risk-free rate), this section will present tests of the predictability of relative returns, which are defined as the return in an individual country relative to the return on the rest-of-world index. Denoting the total return index for country i as $X_{i,t}$ and the rest-of-world index as $ROW_{i,t}$, the relative return index ($RRI_{i,t}$) is defined as

$$RRI_{i,t} = \log X_{i,t} - \log ROW_{i,t}, \quad (7)$$

with all variables defined in U.S. dollars. The k -period relative return, $rr_i(t, t+k)$, is defined as

$$rr_i(t, t+k) = RRI_{i,t+k} - RRI_{i,t}. \quad (8)$$

The predictability of the relative return will be assessed via the regression technique of Fama and French (1988), whereby the k -period return is regressed upon the lagged k -period return as follows:

$$rr_i(t, t+k) = \alpha_i(k) + \beta_i(k)rr_i(t-k, t) + \epsilon_i(t, t+k). \quad (9)$$

Following Fama and French, a general model for the relative return index is assumed, comprising a random walk (or permanent) component $q_{i,t}$, and a stationary component $z_{i,t}$:

$$RRI_{i,t} = q_{i,t} + z_{i,t}, \quad (10)$$

$$q_{i,t} = q_{i,t-1} + \mu_i + \eta_{i,t} \quad (11)$$

where μ_i is a drift term and $\eta_{i,t}$ is a white noise innovation. The expected value of the $\beta(k)$ in equation 9 will depend on the relative importance of the permanent and stationary components. If there is no stationary component (i.e., $z_{i,t} = 0$ for all t), the relative return index is a pure random walk, and after correction for the usual negative bias in estimated autocorrelation coefficients, the estimated $\beta(k)$ should be equal to zero for all k . If there is a stationary component but no random-walk component (i.e., $q_{i,t} = 0$ for all t), the return indices will be cointegrated with a unit cointegrating coefficient, and the $\beta(k)$ should be close to zero for low k and approach -0.5 for high k . Finally, if there is both a stationary and a random-walk component, one might expect a U-shaped pattern for the $\beta(k)$,

being close to zero for small k , then moving towards -0.5 at those horizons where the transitory component is relatively more important, before returning to zero for large k where the random-walk component dominates.

Assessing the statistical significance of the estimated regression coefficients is problematic because of the downward bias in the coefficient that corresponds to the bias in estimated autocorrelation coefficients in small samples, and because of the moving-average error term that is introduced by the use of overlapping data. As in Fama and French, simulation methods are used to estimate the extent of the former bias, though randomization will again be used to avoid the Monte-Carlo assumption of normality for the innovations. By calculating the mean regression coefficient over 1000 replications, the expected value and distribution of the coefficient under the null hypothesis of uncorrelated relative returns can be calculated; the bias-corrected estimate is derived as the estimated $\beta(k)$ less the expected value under the null. The simulated distribution will also be appropriate for accounting for the overlapping observations problem, given that standard techniques for such problems (the Hansen-Hodrick and Newey-West corrections) have been shown to have poor small-sample properties in a related example (see Kim et al., 1991).

The results for tests for the predictability of relative returns based on their own lagged values (equation 9) are shown in Table 3. For each country, results are presented for return horizons of 3, 6, 12, 24, 36, 48, and 60 months, using end-quarter return index data from December 1969-December 1994. In each case, the bias-corrected estimate of $\beta(k)$ is provided, with asterisks indicating those estimates that are significantly different from zero in five or ten percent one-sided tests.

Because of the short sample and high volatility of equity returns, confidence intervals around estimated statistics are wide, and there are relatively few rejections of the null of no predictability in relative returns. However, for return horizons up to one year, there is evidence of positive autocorrelation in relative returns, with eleven of the thirteen rejections being for positive $\beta(k)$. At longer horizons, the autocorrelations appears to turn negative, with all 13 rejections occurring for negative $\beta(k)$. At the three-year horizon, $\beta(k)$ is estimated to be negative for 14 out of 16 countries. The autocorrelations are most strongly negative in the case of the United States and Japan, though these results are not wholly independent: if stock returns in one large country tend to revert to a mean that is dominated by another large country, the reverse is also likely to be true. ^{1/} For the other 14 countries, the unweighted average $\beta(k)$ at the 3-year horizon is -0.19 . The pattern of positive and then negative autocorrelation in relative returns would be consistent with a model in which fads or irrational factors caused price movements to be

^{1/} Tests using data ending in December 1986 suggest, however, that these results are not driven by the rise and fall of Japanese stock prices in the late 1980s and early 1990s.

accentuated at short horizons before fundamental factors reassert themselves at longer horizons. Alternatively, it would also be consistent with a model of stationary but time-varying risk in which changes in risk and required rates of return call forth offsetting movements in asset prices (e.g., Fama and French, 1988).

b. Testing for "winner-loser" effects across national stock markets

A related approach for testing predictability in relative returns is to simulate the performance of a trading rule based on the "winner-loser" effect. A number of researchers, beginning with DeBondt and Thaler (1985), have demonstrated that in the United States, "loser" stocks (those which have underperformed relative to the overall market) in a given ranking period are likely to overperform in the subsequent test period. "Winner" stocks also undergo reversals. For example, DeBondt and Thaler suggest that for portfolios comprising the extreme 35 winner and loser stocks on the New York Stock Exchange, the difference in cumulative average risk-adjusted returns is around 25 percent after three years, though other researchers have questioned if this return differential properly adjusts for risk and measurement problems (e.g., Ball, Kothari and Shanken, 1995).

For this test, indices of total returns in 16 markets are used to calculate returns over ranking periods of one to five years. Countries are assigned to four different portfolios depending on their ranking in this period and the returns of these portfolios are then simulated over the following test period, with portfolio weights based on market capitalizations at the end of the ranking period. The choice of four portfolios of four countries each may reflect the diversification process that an actual investor might follow, and will reduce the effect of country-specific noise on the statistical tests. In addition, the performance of a zero-net-investment portfolio is simulated, based on the assumption that an investor could hold a long position in the extreme loser portfolio and a short position in the extreme winner portfolio, with the return on this portfolio expressed as the difference in the returns on the two extreme portfolios. However, it is not asserted that such a strategy would have been feasible over the entire sample period. In particular, in the early part of the sample period, there were restrictions on capital flows in a number of countries, short selling may have been costly or impossible, and stock index futures contracts non-existent. The trading rules that are simulated here should, therefore, mainly be thought of as analytical constructs for examining the mean-reverting tendencies of market indices in different countries.

The strategies are assumed to be replicated at the end of each quarter, yielding a series of overlapping test outcomes. The distribution of the average returns from such trading rules under the null of no predictability in returns is unclear, so simulation methods are again used to address this issue, with simulated paths for returns and market capitalization weights in all 16 countries created by reshuffling the innovations in each series in parallel. Statistical inference is based on the results from 1000

randomizations. 1/ The number of overlapping quarterly tests ranges from 93 for the one-year test to 61 for the five-year test, and the number of non-overlapping tests ranges from 24 to only four. To the extent that there is a small number of non-overlapping tests for the longer horizons, this will be reflected in wide confidence intervals derived from the simulations which use data sets of similar length, and in a correspondingly low likelihood of rejecting the null hypothesis of no temporal dependence in returns.

The results of the test for the existence of a winner-loser effect are shown in Table 4. For each of the four portfolios and the zero-net-investment portfolio, the table shows: (i) the average annual return in the ranking period; (ii) the average annual return in the test period; and (iii) the 90 percent confidence interval for the test period return, derived from 1000 simulations. Average returns are calculated as the geometric average of return relatives in excess of the total world market return. The results are based on total holding-period returns and are not subject to possible problems from DeBondt and Thaler's use of the sum of monthly returns.

At the one-year horizon, the average return on the zero-net-investment portfolio is estimated to be negative, though not statistically significantly so. However, for k of 2, 3 and 4 years, the strategy of buying losers and selling winners appears to yield statistically significantly positive returns, especially at the three- and four-year horizons where the return is significant at the one percent level. The return differentials for these horizons of 5.3, 6.7 and 6.3 percent per annum, respectively, would also appear to be economically significant. Together, the results of this test suggest again that relative returns may be positively autocorrelated at short horizons and negatively autocorrelated at longer horizons. The stronger evidence for mean reversion in this test than in the previous regression test may be due to the use of portfolios to reduce country-specific noise. Similar results are obtained for equally weighted portfolios, so the results are not purely due to the high weights of the United States and Japan which showed the greatest relative predictability in the previous test. 2/

By comparing the average performance of the zero-net-investment portfolio in the ranking and test periods, one can gain a summary measure of the strength of the mean-reverting tendencies of national equity markets. If the series were fully mean reverting, one would expect that for longer

1/ Monte Carlo simulations show that when innovations for all countries have a similar mean, winner-losers tests have a zero expected return, so that there is no bias corresponding to the autocorrelation bias in regression tests. Simulations using bootstrapped series yield slightly wider confidence intervals than the randomized series but yield similar conclusions on statistical significance.

2/ Similar results indicating winner-loser reversals are also obtained using data for a group of emerging markets (see Richards, 1996b).

horizons, 50 percent of all ranking-period differentials would be reversed in the subsequent test periods. The data suggest that mean reversion is strongest at the four-year horizon where around 31 percent of the ranking-period underperformance of the loser portfolio is reversed on average in the subsequent test period. ^{1/} This implies that there is less than full mean-reversion, but it is nonetheless suggestive of an economically significant mean-reverting component.

VI. Conclusion

This paper has suggested that there is little empirical evidence for the proposition that the stock return indices of different countries are cointegrated. This result may not be especially surprising given that standard asset pricing models imply that return indices should contain random-walk components which would preclude cointegration. To illustrate this, consider first the case of the return indices of two firms in the same industry and country. The indices of these firms will contain a common component that reflects macroeconomic and sectoral shocks. However, it seems reasonable to expect that they will also contain a large company-specific component due to different management decisions: the return indices of competitors such as General Motors and Chrysler, or IBM and Compaq may actually move in different directions following such decisions. Cointegration between the return indices of these companies would, however, require that company-specific shocks to returns were always followed by exactly offsetting shocks.

At the national level, countries often differ significantly in their industrial structures, suggesting that national stock market return indices will also respond differently to many economic shocks. And even in the case of neighboring countries--e.g., France and Germany--that might have quite highly interlinked economies and might be subject to similar shocks, short-run correlations between returns are far from perfect. It may therefore not be unreasonable to expect that long-run correlations will also be less than perfect. Furthermore, if cointegration did exist, it would imply that national return indices were significantly predictable, in violation of the efficient markets hypothesis.

In light of these theoretical arguments against cointegration of return indices, there may be a need for caution in the burgeoning practice of regressing large number of country pairs or groups against each other in search of "cointegrating relationships." "Significant" findings of cointegration should be assessed in light of the poor small-sample properties of some of the tests and the economic model, if any, that is

^{1/} The implied autocorrelation coefficients for the five horizons are: +.07 (1 year), -.16 (2 years), -.25 (3 years), -.31 (4 years), and -.18 (5 years).

offered to explain the estimated relationship. 1/ In the case of the Johansen tests, the simulations in this paper suggest that the use of asymptotic critical values may sometimes be highly inappropriate, though a simple adjustment suggested by Reinsel and Ahn (1988) may yield approximately correctly sized tests, and will offset the bias towards rejection that results when lags are added to the VAR or when variables are added to the system.

The results of the empirical tests conducted in this paper would indicate a model in which the stock return indices of different countries are the sum of a common "world" return index and two country-specific components, one permanent and the other transitory. The existence of a common world component is suggested by the well-documented significant short-run correlations between returns in different national markets. The cointegration tests in this paper imply, however, that national return indices are not in fact cointegrated around the common component. This indicates that country-specific factors also influence long-term stock market performance; that is, there must also be a permanent component that reflects the influence of country-specific shocks. In addition, the evidence of relative return predictability from the regression tests and winner-loser tests implies the existence of a transitory or mean-reverting country-specific component. This may not be surprising in the light of evidence (e.g., Poterba and Summers, 1988) that the price or return indices of a range of countries separately display mean-reverting tendencies. Unless these individual mean reversions were due solely to the common world component and were therefore perfectly correlated across countries, they would also result in some degree of mean reversion in relative returns. Such mean reversion may be evidence of market inefficiency due to fads or irrationality. Alternatively, it may be an indication of the existence of time-varying risk in an equilibrium framework, or an indication of market segmentation, perhaps due to regulatory constraints during part of the sample period. 2/ Regardless of the factors behind the transitory component, the existence of a permanent country-specific component in national return indices implies that there will typically be long-run risk-reduction benefits from investing abroad.

1/ Indeed, Loughran and Newbold (1995) conclude from an empirical study that "It is perhaps overly cynical to suggest that applying cointegration tests to financial markets data is a superb mechanism for generating random numbers. Nevertheless, the inconsistencies in results reported in applied studies ... suggest extreme caution in inferring structural conclusions from the test statistics."

2/ These issues are addressed in Richards (1996a).

Table 1. Analysis of Small-Sample Properties of Johansen Cointegration Tests

| Number of lags in VAR | Null hypothesis | Test statistics | | | | 5 percent critical values | | | | Simulated rejection frequencies (percent) | | | |
|-----------------------------|--------------------|---------------------|--------------------|--------------------------------|--------------------|---------------------------|--------------------|----------------|--------------------|---|--------------------|---------------------------------|--------------------|
| | | Actual: Kasa (1992) | | Mean of 1000 randomizations | | Asymptotic | | Small-sample | | Asymptotic critical values | | Small-sample critical values | |
| | | Trace stat. | Maximal eigenv. | Trace stat. | Maximal eigenv. | Trace stat. | Maximal eigenv. | Trace stat. | Maximal eigenv. | Trace stat. | Maximal eigenv. | Trace stat. | Maximal eigenv. |
| j=2 | r = 0 | 76.56 | 37.98 | 62.57 | 27.52 | 69.98 | 33.26 | 82.70 | 39.31 | 23.7 | 17.6 | 5.3 | 4.5 |
| | r ≤ 1 | 38.58 | 19.42 | 35.06 | 17.38 | 48.42 | 27.34 | 57.22 | 32.31 | 3.3 | 1.6 | 0.3 | 0.1 |
| | r ≤ 2 | 19.15 | 12.70 | 17.67 | 10.55 | 31.26 | 21.28 | 36.94 | 25.15 | 0.4 | 0.0 | 0.0 | 0.0 |
| | r ≤ 3 | 6.45 | 6.12 | 7.12 | 5.82 | 17.84 | 14.60 | 21.09 | 17.25 | 0.1 | 0.0 | 0.0 | 0.0 |
| | r ≤ 4 | 0.33 | 0.33 | 1.31 | 1.31 | 8.08 | 8.08 | 9.55 | 9.55 | 0.0 | 0.0 | 0.0 | 0.0 |
| j=10 | r = 0 | 298.19 | 174.05 | 430.99 | 215.02 | 69.98 | 33.26 | 569.81 | 270.85 | 100.0 | 100.0 | 8.8 | 18.4 |
| | r ≤ 1 | 124.14 | 64.85 | 215.97 | 112.43 | 48.42 | 27.34 | 394.27 | 222.63 | 100.0 | 100.0 | 0.2 | 0.4 |
| | r ≤ 2 | 59.29 | 44.75 | 103.54 | 63.07 | 31.26 | 21.28 | 254.51 | 173.27 | 100.0 | 100.0 | 0.0 | 0.0 |
| | r ≤ 3 | 14.54 | 14.33 | 40.47 | 32.94 | 17.84 | 14.60 | 145.30 | 118.85 | 96.0 | 96.6 | 0.0 | 0.0 |
| | r ≤ 4 | 0.20 | 0.20 | 7.53 | 7.53 | 8.08 | 8.08 | 65.82 | 65.82 | 34.6 | 34.4 | 0.0 | 0.0 |

This table analyses the small-sample performance of the Johansen test statistics using randomized data based on the 67 observation stock price dataset used by Kasa (1992). Asymptotic critical values are from Johansen and Juselius (1990, Table A2) for a system of 5 variables and a constant in the VAR, and are multiplied by $T/(T-n)$ to yield small-sample critical values, as suggested by Reinsel and Ahn (1988). The randomized mean test statistics and rejection frequencies are based on 1000 simulations of a non-cointegrated system generated by reshuffling the innovations in Kasa's data.

Table 2. Tests for Cointegration between Country and Rest-of-World Equity Return Indices:
End-Quarter data, December 1969–December 1994

| Country | Johansen Methodology | | | | Engle–Granger Methodology |
|---|----------------------|-----------------------|---------------------|-----------------------|------------------------------|
| | Number of lags = 2 | | Number of lags = 10 | | ADF Statistic |
| | Trace Statistic | Maximal Eigenvalue | Trace Statistic | Maximal Eigenvalue | |
| **, * denote rejection of H0 of no cointegration at 5, 10 percent level | | | | | |
| Australia | 18.83 ** | 17.21 ** | 9.15 | 8.82 | −3.49 * |
| Austria | 8.86 | 8.27 | 9.97 | 9.56 | −2.36 |
| Canada | 12.50 | 12.50 | 12.20 | 11.76 | −3.76 * |
| Denmark | 9.12 | 8.54 | 30.39 ** | 30.38 ** | −3.73 * |
| France | 9.28 | 8.79 | 14.40 | 14.06 | −2.95 |
| Germany | 8.16 | 7.83 | 7.08 | 7.00 | −2.98 |
| Hong Kong | 7.98 | 7.66 | 6.99 | 6.86 | −3.09 |
| Italy | 8.29 | 7.27 | 6.68 | 6.59 | −4.27 ** |
| Japan | 6.62 | 5.93 | 23.01 ** | 22.61 ** | −3.16 |
| Netherlands | 5.17 | 5.14 | 9.53 | 9.32 | −2.92 |
| Norway | 17.14 * | 16.45 ** | 22.59 ** | 22.34 ** | −3.20 |
| Spain | 7.18 | 5.25 | 6.71 | 6.51 | −2.27 |
| Sweden | 7.62 | 7.49 | 17.73 | 17.70 * | −2.57 |
| Switzerland | 4.11 | 3.68 | 6.09 | 5.65 | −1.74 |
| United Kingdom | 8.80 | 8.39 | 5.86 | 5.69 | −3.03 |
| United States | 5.74 | 5.36 | 16.42 | 16.22 | −2.15 |
| Five percent critical values: | | | | | |
| Asymptotic | 17.84 | 14.60 | 17.84 | 14.60 | −3.80 |
| Small–sample | 18.60 | 15.21 | 22.87 | 18.71 | n.a. |
| Empirical distribution (average) | 17.79 | 15.29 | 21.54 | 18.24 | −3.78 |

This table shows the results of tests for cointegration between excess total return indices (dividends plus capital appreciation less the cumulated U.S. risk-free return) for individual countries with a market capitalization-weighted rest-of-world index, both in logarithms and U.S. dollars. Significance tests are based on the empirical distribution of test statistics for each country under the null hypothesis of no cointegration from 1000 simulations using randomized data. Johansen test statistics are for the null hypothesis of $r=0$ with asymptotic critical values from Johansen and Juselius (1990, Table A2) and are multiplied by $T/(T-n_j)$ to yield small-sample critical values, as suggested by Reinsel and Ahn (1988). Engle–Granger equations include a constant and a trend, with asymptotic critical values from Phillips and Ouliaris (1990, Table 2c).

Table 3. Tests for Predictability of Relative Returns:
End-Quarter Data, December 1969–December 1994

| Country | Bias-corrected regression coefficients | | | | | | |
|--|--|----------|-----------|-----------|-----------|-----------|-----------|
| | 3 months | 6 months | 12 months | 24 months | 36 months | 48 months | 60 months |
| ** (*) denotes significantly different to zero based on 5 (10) percent one-tail test | | | | | | | |
| Australia | -0.10 | -0.16 * | -0.09 | -0.21 | -0.11 | -0.18 | -0.40 |
| Austria | 0.22 ** | 0.11 | 0.08 | -0.14 | 0.02 | 0.15 | -0.06 |
| Canada | 0.00 | 0.15 * | 0.14 | -0.10 | -0.09 | 0.05 | 0.08 |
| Denmark | 0.12 | 0.22 ** | 0.10 | 0.03 | -0.20 | -0.29 | -0.08 |
| France | -0.01 | 0.06 | -0.12 | -0.31 * | -0.43 ** | -0.10 | -0.07 |
| Germany | 0.15 * | -0.02 | -0.07 | -0.27 | -0.09 | -0.08 | -0.12 |
| Hong Kong | -0.07 | -0.10 | 0.00 | -0.25 | -0.18 | -0.09 | 0.13 |
| Italy | -0.08 | 0.30 ** | 0.20 * | -0.30 * | -0.28 | -0.09 | 0.23 |
| Japan | 0.09 | 0.16 * | 0.21 * | 0.09 | -0.11 | -0.53 ** | -0.87 ** |
| Netherlands | 0.02 | -0.15 | -0.02 | -0.12 | -0.32 | -0.38 * | -0.21 |
| Norway | 0.06 | 0.10 | -0.08 | -0.36 ** | -0.52 ** | -0.16 | 0.08 |
| Spain | 0.12 | 0.16 * | 0.26 * | 0.21 | -0.01 | 0.07 | 0.18 |
| Sweden | 0.09 | 0.04 | 0.00 | -0.13 | -0.39 * | -0.49 ** | -0.15 |
| Switzerland | 0.12 | -0.04 | 0.12 | 0.02 | 0.04 | 0.04 | -0.04 |
| United Kingdom | 0.02 | -0.03 | -0.37 ** | 0.04 | -0.14 | -0.29 | -0.26 |
| United States | 0.05 | 0.10 | 0.31 ** | 0.10 | -0.37 * | -0.88 ** | -0.91 ** |
| Average 5 percent one-sided critical values: | | | | | | | |
| H0: Beta < 0 | -0.16 | -0.20 | -0.26 | -0.36 | -0.42 | -0.47 | -0.54 |
| H0: Beta > 0 | 0.16 | 0.20 | 0.27 | 0.38 | 0.46 | 0.54 | 0.64 |

This table shows the bias-corrected regression coefficients from equation 9, a regression of the k-month relative return for each country on a constant and the lagged non-overlapping relative return. The relative return is defined as the change in the logarithm of the ratio of the total returns index for a country to the total returns index for a rest-of-world index, both in U.S. dollars. Significance tests are based on the 5, 10, 90 and 95 percent fractiles from 1000 simulations where the innovations are reshuffled to generate indices with no temporal dependence in returns. Regression coefficients are corrected for bias by subtracting the simulated mean value under the null hypothesis from the estimated value.

Table 4. Testing for Winner-Loser Effects among National Stock Market Return Indices:
End-Quarter Data, December 1969-December 1994

| Number of years in ranking and test period | Portfolio | | | | Zero-net- investment |
|---|-------------|-------------|-------------|-------------|-------------------------|
| | 1 | 2 | 3 | 4 | |
| Ranking—period average annual return | | | | | |
| 1 | -19.78 | -5.21 | 5.98 | 27.08 | -46.86 |
| 2 | -15.56 | -3.28 | 5.42 | 18.42 | -33.97 |
| 3 | -11.53 | -2.04 | 4.36 | 15.06 | -26.58 |
| 4 | -9.85 | -2.02 | 3.51 | 10.33 | -20.18 |
| 5 | -8.45 | -2.09 | 3.03 | 9.31 | -17.75 |
| Test—period average annual return | | | | | |
| 1 | -3.05 | 2.38 | 2.68 | -0.11 | -2.95 |
| 2 | 1.08 | 2.89 | 0.17 | -4.27 | 5.35 |
| 3 | 3.36 | 2.16 | -0.18 | -3.34 | 6.70 |
| 4 | 3.13 | 1.96 | 0.76 | -3.18 | 6.31 |
| 5 | 1.72 | 2.51 | -0.94 | -1.56 | 3.27 |
| Empirical 90 percent confidence interval for test—period return | | | | | |
| 1 | -2.79, 2.70 | -1.90, 2.04 | -1.94, 2.29 | -2.62, 2.87 | -4.70, 4.24 |
| 2 | -2.70, 2.51 | -1.74, 1.81 | -1.73, 1.90 | -2.66, 3.03 | -4.51, 4.24 |
| 3 | -2.57, 2.53 | -1.64, 1.74 | -1.78, 1.99 | -2.72, 2.99 | -4.57, 4.16 |
| 4 | -2.74, 2.56 | -1.67, 1.86 | -1.87, 2.12 | -2.62, 3.21 | -4.32, 4.01 |
| 5 | -2.87, 2.68 | -1.60, 1.91 | -1.82, 2.20 | -2.98, 3.21 | -4.23, 4.18 |

This table summarizes the ex-post simulation of a trading rule in which the performance of national stock markets in a k-year ranking period is used to choose portfolios to be held for the subsequent k-year test period. Portfolio 1 (4) comprises the weakest (strongest) markets in the ranking period, and the zero net investment portfolio is long portfolio 1 and short portfolio 4. Market capitalization-weighted portfolios are formed at the end of each quarter and average annual returns are calculated as the geometric average of all test outcomes. Returns for portfolios 1 to 4 are calculated relative to the return on the world market. 90 percent confidence intervals are based on 1000 simulations using randomized data.

References

- Arshanapalli, Bala and John Doukas, 1993, International stock market linkages: Evidence from the pre- and post-October 1987 period, *Journal of Banking and Finance* 17, 193-208.
- Ball, Ray. S., S.P. Kothari and Jay Shanken, 1995, Problems in measuring portfolio performance: An application to contrarian investment strategies, *Journal of Financial Economics* 38, 79-107.
- Bekaert, Geert, 1995, Market integration and investment barriers in emerging equity markets, *World Bank Economic Review* 9, 75-107.
- Byers, J.D. and D.A. Peel, 1993, Some evidence on the interdependence of national stock markets and the gains from international portfolio diversification, *Applied Financial Economics* 3, 239-242.
- Campbell, John Y. and Pierre Perron, 1991, Pitfalls and opportunities: What macroeconomists should know about unit roots, in: O.J. Blanchard and S. Fischer, eds., *NBER Macroeconomics Annual 1991* (MIT Press, Cambridge, Massachusetts), 141-201.
- Cheung, Yin-Wong and Kon S. Lai, 1993, Finite sample sizes of Johansen's likelihood ratio tests for cointegration, *Oxford Bulletin of Economics and Statistics* 55, 313-328.
- Corhay, A., A. Tourani Rad and J.-P. Urbain, 1993, Common stochastic trends in European stock markets, *Economics Letters* 42, 385-390.
- Cutler, David M., James M. Poterba and Lawrence H. Summers, 1991, Speculative dynamics, *Review of Economic Studies* 58, 529-546.
- DeBondt, Werner F. M. and Richard Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793-805.
- Dwyer, Gerald P. and Myles S. Wallace, 1992, Cointegration and market efficiency, *Journal of International Money and Finance* 11, 318-327.
- Engel, Charles, 1994, Tests of CAPM on an international portfolio of bonds and stocks, in: Frankel (1994).
- _____, 1995, The forward discount anomaly and the risk premium: A survey of recent evidence, forthcoming in *Journal of Empirical Finance*.
- Engle, R.F. and C.W.J. Granger, 1987, Cointegration and error correction: Representation, estimation and testing, *Econometrica* 55, 251-276.
- Fama, Eugene F. and Kenneth R. French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246-273.
- Frankel, Jeffrey A., 1994, ed., *The internationalization of equity markets* (University of Chicago Press, Chicago).

- French, Kenneth R. and James M. Poterba, 1991, Investor diversification and international equity markets, *American Economic Review* (Papers and Proceedings) 81, 222-226.
- Granger, C.W.J, 1986, Developments in the study of cointegrated economic variables, *Oxford Bulletin of Economics and Statistics* 48, 213-228.
- Gregory, A.W., 1994, Testing for cointegration in linear quadratic models, *Journal of Business and Economic Statistics* 12, 347-360.
- Hall, S.G., 1991, The effect of varying length VAR models on the maximum likelihood estimates of cointegrating vectors, *Scottish Journal of Political Economy* 38, 317-323.
- Harvey, Campbell R., 1991, The world price of covariance risk, *Journal of Finance* 46, 111-157.
- Ibbotson, Roger G., Richard C. Carr, and Anthony W. Robertson, 1982, International equity and bond returns, *Financial Analysts Journal*, July-August, 61-83.
- Johansen, S., 1988, Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control* 12, 231-254.
- _____, and K. Juselius, 1990, Maximum likelihood estimation and inference on cointegration - with applications to the demand for money, *Oxford Bulletin of Economics and Statistics* 52, 169-210.
- Kasa, Kenneth, 1992, Common stochastic trends in international stock markets, *Journal of Monetary Economics* 29, 95-124.
- Kim, Myung-Jig, Charles R. Nelson, and Richard Startz, 1991, Mean reversion in stock prices: A reappraisal of the empirical evidence, *Review of Economic Studies* 58, 515-528.
- King, Mervyn A. and Sushil Wadhvani, 1990, Transmission of volatility between stock markets, *Review of Financial Studies* 3, 5-33.
- Loughran, Tim and Paul Newbold, 1995, Tests for cointegration in financial markets time series: What answer do you want?, forthcoming in *Advances in Quantitative Analysis of Finance and Accounting*.
- Phillips, P.C.B. and S. Ouliaris, 1990, Asymptotic properties of residual based tests for cointegration, *Econometrica* 58, 165-193.
- Poterba, James M. and Lawrence H. Summers, 1988, Mean reversion in stock prices: Evidence and implications, *Journal of Financial Economics* 22, 27-59.
- Reinsel, G.C. and S.K. Ahn, 1988, Asymptotic properties of the likelihood ratio test for cointegration in the non-stationary vector AR model, Technical Report, Department of Statistics, University of Wisconsin, Madison.

Richards, Anthony J., 1996a, What explains the winner loser effect across national stock markets?, manuscript, International Monetary Fund.

Richards, Anthony J., 1996b, Volatility and predictability in national stock markets: How do emerging and mature markets differ?, International Monetary Fund Working Paper (forthcoming).

Stengos, T. and E. Panas, 1992, Testing the efficiency of the Athens Stock Exchange: Some results from the banking sector, *Empirical Economics* 17, 239-252.

Solnik, Bruno, 1991, *International Investments*, 2nd. ed., (Addison Wesley, Reading, Mass.).

Taylor, Mark P., and Ian Tonks, 1989, The internationalization of stockmarkets and the abolition of U.K. exchange control, *Review of Economics and Statistics* 71, 332-336.

