Spanning Tests for Replicable Small-Cap Indexes as Separate Asset Classes

Evidence for G-7 countries.

Lorne N. Switzer and Haibo Fan

Lorne N. Switzer
is Van Berkom endowed chair of small cap equities in the John Molson School of Business at Concordia University in Montreal, P.Q., Canada.
switz@jmsb.concordia.ca

Haibo Fan
is research associate in the John Molson School of Business at Concordia University in Montreal, P.Q., Canada

There is growing interest among academicians and practitioners in the performance of small-cap indexes and the benefits of small-caps for international diversification. Banz [1981] first reported the size anomaly that small-cap stocks generate higher returns than large-caps, even after controlling for risk over long investment horizons. Horowitz, Loughran, and Savin [2000], however, using data from the New York and American Stock Exchanges and Nasdaq, find that small-cap portfolios underperformed large-cap portfolios over 1980-1996. Dimson and Marsh [1999] find a small-cap discount for the U.S. and the U.K. after the launch of the Dimensional Fund Advisors 1982-1983 through 1997. Similarly, Reilly and Wright [2002] find that large-cap stocks outperformed small-cap stocks over the period 1984-2000.

Although the evidence of higher returns for small-cap firms is inconclusive over a long horizon, from a strategic asset allocation perspective, small-cap stocks could still be an effective vehicle for international diversification, whose benefits are well documented. Mean-variance spanning tests have been used to study such benefits (see, e.g., Harvey [1995], Bekaert and Urias [1996], Errunza, Hogan, and Hung [1999], and Driessen and Laeven [2003]).

Eun, Huang, and Lai [2004] and Petrella [2005] construct large size-based portfolios and conclude that international small-cap funds are generally not spannable by country stock market indexes over 1980-1999 and 1992-2002. This suggests there are benefits to small-cap diversification, in spite of the evidence of the apparent disappearance of the small-cap anomaly. These studies are not easily replicable for two reasons: 1) the average number of securities required to form portfolios is prohibitive, and 2) there is no adjustment for free float (shares of a companies that are freely available to the investing public).[1]

We use both the traditional spanning test and a step-down spanning test of mean-variance frontiers for U.S. and other G-7 international portfolios. The test results are compared using empirical measures for portfolio optimization for U.S. and small-cap portfolios. Our size-based indexes either are replicable using exchange-traded funds (ETFs) or involve positions in a realistic number of securities.

For the U.S. portfolio, we examine the S&P 500 and Russell 2000 Index, which trade as ETFs (S&P 500 iShares and the Russell 2000 ishare) (an ETF involves an investment in just one security). We also use the MSCI small-cap indexes, which are adjusted for free float, and use only 40% of the full market capitalization of the eligible small-cap universe within each industry group in each country. Using representatives instead of the entire sample could reduce replicating costs. The Russell and MSCI indexes are well accepted by industry professionals as benchmarks of the relevant market segments.[2]

Test results on indexes in the U.S. and other members of the G-7 group show that the composition of benchmark portfolios, the time of the sample, and investor holding periods are all factors that determine whether a small-cap portfolio is a separate asset class (i.e., not spanned by a given benchmark portfolio). Short sale constraints and position limits do not necessarily reduce the diversification benefits of adding a small-cap index to a benchmark portfolio. In fact, we find only a few small-cap portfolios of G-7 countries that behave as separate asset classes with corresponding diversification benefits when benchmarked against the U.S. equity market or an international large-cap portfolio.

DESCRIPTION OF THE DATA

This study uses monthly U.S. dollar total return from January 1989 through December 2004 for indexes in the U.S. and the other countries that constitute the G-7, obtained from Ibbotson and Datastream. Sample data are available for the period January 1989-December 2004 for the U.S., European, and Japanese equity markets (the NIKKEI 300 is cap-weighted price index.). For the other indexes, the data begin with their inception in January 1999.

Exhibit 1 shows the indexes and summary statistics for January 1999-December 2004. Small-cap indexes overall have higher return and lower risk than the large-cap indexes. They also generally have higher Sharpe ratios than the S&P 500 (particularly for the U.K., Canada, and Italy).

Exhibit 2 provides the correlations between the indexes, which range between 0.1 and 0.9. Most of the correlations are well below 1.0, so there may be potential diversification benefits for paired asset combinations.

REGRESSION-BASED MEAN-VARIANCE SPANNING TESTS

Huberman and Kandel [1987] first introduced the concept of mean-variance spanning. A set of K risky assets spans a larger set of $N+K$ risky assets if the minimum-variance frontier of the K assets is the same as the minimum-variance frontier of the $K + N$ assets. Benchmark assets are represented by the first set of K risky assets, while the test assets represent the second set of N risky assets. The tests are based on regressions of the test assets on the benchmark assets:

$$R = XB + E \tag{1}$$

where R is a $T \times 1$ vector of the returns on the test asset (e.g., the small-cap portfolio of a particular country); X is a $T \times (K+1)$ matrix of the benchmark asset returns (e.g., S&P 500, EUT); the coefficient vector is denoted by $B = [\alpha, \beta]'$; and E is a $T \times 1$ vector of error terms ε_t. The null hypothesis of spanning—that there is no net benefit of adding a small-cap portfolio to an established benchmark portfolio—is equivalent to the joint hypothesis:

$$\text{H0}: \alpha j = 0, \text{ and } \sum_j \beta_j = 1 \tag{2}$$

The Lagrange multiplier (LM) and F-test for small samples (Kan and Zhou [2001]) are used to test the null hypothesis.

We assume $T \geq K+2$ and $X'X$ is non-singular. To obtain exact distributions of the test statistics, we also assume that, conditional on returns of benchmark assets, the disturbances ε_t are independent and identically distributed as multivariate normal with mean zero and variance $\boldsymbol{\Sigma}$.

Let U be the ratio of unconstrained and constrained maximum-likelihood estimator of variance:

$$LM = T(1-U) \sim \chi_2^2 \tag{3}$$

Exhibit 1
Sample Indexes and Summary Statistics

Country/ Region	Index	Market Segment	Variable Name	Geometric Mean (%)	Arithmetic Mean (%)	Standard Deviation (%)	Sharpe Ratio
U.S.	S&P 500	Large	SP500	1.25	2.53	16.40	0.05
	Russell 2000	Small	R2000	8.92	11.36	23.76	0.15
Canada	S&P/TSX Composite	Large	SPTSX	12.38	14.51	22.29	0.20
	MSCI Canada Small Cap	Small	CaS	18.66	20.53	21.45	0.30
France	CAC 40	Large	CAC40	4.53	6.79	22.26	0.09
	MSCI France Small Cap	Small	FrS	7.05	10.18	26.57	0.12
Germany	DAX 30 Performance	Large	DAX30	-0.26	3.41	28.20	0.04
	MSCI Germany Small Cap	Small Cap	GeS	-2.81	0.19	25.21	0.00
Italy	S&P MIB	Large	SPMIB	2.69	5.18	23.37	0.07
	MSCI Italy Small Cap	Small	ItS	13.55	15.76	22.65	0.22
Japan	BARRA/Nikko Large Cap	Large	NIKL	3.54	5.90	22.91	0.08
	BARRA/Nikko Small Cap	Small	NIKS	13.22	16.13	26.81	0.19
	NIKKEI 300	Large	NIKI300	2.06	4.22	21.77	0.06
	MSCI Japan Small Cap	Small	JpS	9.54	12.84	28.04	0.14
U.K.	FTSE 100	Large	FTSE100	1.90	3.04	15.49	0.06
	MSCI U.K. Small Cap	Small	UKS	13.61	16.06	24.05	0.21
Emerging Countries	S&P/IFCI Emerging Composite	whole	EMG	14.43	17.22	26.01	0.21
Europe	MSCI Europe	whole	EUT	2.90	4.41	18.04	0.07

The F-test for small samples is:

$$\left(\frac{1}{U}-1\right)\left(\frac{T-K-1}{2}\right)\sim F_{2,T-K-1} \qquad (4)$$

We begin with a U.S. investor with a benchmark portfolio that consists of the S&P 500 Total Return Index (SP500) and the Russell 2000 Total Return Index. We then test other G-7 small-cap indexes against this benchmark, in a descending sequence based on the ranking of the home country's 2004 domestic market capitalization as reported by the World Federation of Stock Exchanges.[3]

When a small-cap portfolio in a particular country is significant in the test, this index is included in the

EXHIBIT 2

Correlation Matrix

	Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	SP500	1.00																	
2	R2000	0.69	1.00																
3	SPTSX	0.78	0.79	1.00															
4	CaS	0.61	0.74	0.78	1.00														
5	CAC40	0.79	0.73	0.73	0.65	1.00													
6	FrS	0.58	0.73	0.59	0.63	0.76	1.00												
7	DAX30	0.79	0.76	0.71	0.63	0.93	0.81	1.00											
8	GeS	0.61	0.67	0.62	0.60	0.70	0.78	0.80	1.00										
9	SPMIB	0.60	0.70	0.65	0.55	0.83	0.77	0.81	0.66	1.00									
10	ItS	0.54	0.60	0.57	0.56	0.72	0.79	0.74	0.63	0.81	1.00								
11	NIKL	0.47	0.42	0.51	0.47	0.37	0.29	0.31	0.27	0.29	0.30	1.00							
12	NIKS	0.28	0.20	0.29	0.34	0.16	0.22	0.12	0.21	0.10	0.19	0.82	1.00						
13	NIKKEI 300	0.49	0.34	0.49	0.45	0.35	0.25	0.27	0.24	0.23	0.28	0.97	0.85	1.00					
14	JpS	0.30	0.24	0.34	0.37	0.19	0.24	0.13	0.20	0.14	0.20	0.83	0.97	0.86	1.00				
15	FTSE100	0.81	0.60	0.65	0.60	0.83	0.59	0.80	0.60	0.67	0.65	0.42	0.21	0.44	0.21	1.00			
16	UKS	0.65	0.80	0.69	0.69	0.71	0.79	0.72	0.69	0.68	0.68	0.52	0.39	0.50	0.40	0.73	1.00		
17	EMG	0.76	0.75	0.76	0.73	0.68	0.68	0.73	0.69	0.60	0.59	0.54	0.37	0.52	0.37	0.68	0.80	1.00	
18	EUT	0.82	0.74	0.73	0.65	0.97	0.77	0.94	0.73	0.85	0.77	0.40	0.18	0.38	0.20	0.91	0.78	0.72	1.00

Correlations based on monthly data over Jan. 1999-Dec. 2004 (N=72).

benchmark to form a revised benchmark against which we test the next-largest G-7 country's small-cap index. If a test asset lacks significance, we do not revise the benchmark for testing the subsequent G-7 country small-cap portfolio.

We begin the analysis with the Japanese small-cap index (JpS), since the Japanese market is the largest market after the U.S. As is shown in Panel A of Exhibit 3, JpS is significant in the spanning test; hence we form a new benchmark portfolio of the SP500 and R2000 and JpS against which we test the U.K. small-cap market (UKS), the third-largest market. UKS is not significant, and neither are the small-cap indexes from France (FrS), Germany (GeS), or Italy (ItS). The only small-cap index aside from JpS that is significant is the Canadian small-cap index (CaS).[4]

We also conduct tests where the initial reference portfolio is a G-7 large-cap benchmark portfolio (LB), and the initial test asset is the U.S. small-stock portfolio, with foreign small-cap portfolios included as the subsequent test assets.[5] The analysis begins with a small-cap U.S. index (R2000) as the test portfolio, and then proceeds to the remaining G-7 countries in descending order of market capitalization.

Test results are shown in Panel B of Exhibit 3. A notable result is that the U.S. small-cap portfolio is spanned by an international large-cap portfolio benchmark.[6]

Overall, Panel B of Exhibit 3 demonstrates that traditional spanning tests may be sensitive to the choice of benchmarks. For example, while the Japanese small-cap portfolio is not spanned by the U.S. market (represented by the S&P 500 and the Russell 2000), it is spanned by an international large-cap portfolio, at the 5% significance level. The Canadian small-cap portfolio remains significant, however, when the U.S. market benchmark is replaced by the international large-cap benchmark portfolio.

In Panel C of Exhibit 3 we examine the stability of this testing procedure over time. Our test asset in this analysis is the Japanese small-cap index, NIKS. The results show that over the entire 16-year holding period 1989-2004, the Japanese small-cap index is not spanned by various benchmarks, at the 1% level of significance. This result is not observed for shorter holding periods, except for the more recent subperiod 1999-2004.

Another salient finding in Exhibits 2 and 3 is that high (low) correlations of returns between the small-cap portfolios of different countries do not preclude (guarantee) some beneficial contribution of combining the assets in a portfolio. For example, we note that CaS has a higher correlation coefficient with each of the assets in the benchmark portfolio than FrS (0.61 versus 0.58 with the S&P 500; 0.74 versus 0.73 with R2000; and 0.37 versus 0.24

with JpS). The spanning test for CaS is significant against these benchmarks, but FrS, on the other hand, is found to be spanned by the same benchmark assets.

We also test more than 90 combinations of small-cap indexes and other international benchmarks with similar findings as those reported here. The adjusted R-squares of ordinary least squares estimations show no difference between the portfolios with significant spanning tests and the portfolios with insignificant spanning tests. Our results imply that the significance of diversification

EXHIBIT 3
Spanning Test Results

Case	Country/Time	Test asset	Benchmark assets	LM Test	p-value	F Test	p-value
Panel A. Spanning test (step-by-step) with U.S. indexes as the basic benchmark							
1	Japan	JpS	SP500 R2000	8.60	0.01	4.68	0.01
2	U.K.	UKS	SP500 R2000 JpS	1.36	0.51	0.65	0.52
3	France	FrS	SP500 R2000 JpS	0.13	0.94	0.06	0.94
4	Germany	GeS	SP500 R2000 JpS	0.89	0.64	0.42	0.66
5	Canada	CaS	SP500 R2000 JpS	8.74	0.01	4.69	0.01
6	Italy	ItS	SP500 R2000 JpS CaS	3.96	0.14	1.90	0.16
Panel B. Spanning test (step-by-step) with large cap indexes in G7 as the basic portfolio							
7	U.S.	R2000	LB	3.68	0.16	1.73	0.19
8	Japan	JpS	LB	5.64	0.06	2.72	0.07
9	U.K.	UKS	LB JpS	2.46	0.29	1.11	0.33
10	France	FrS	LB JpS	3.77	0.15	1.74	0.18
11	Germany	GeS	LB JpS	4.80	0.09	2.25	0.11
12	Canada	CaS	LB JpS	6.56	0.04	3.16	0.05
13		CaS	LB JpS GeS	6.13	0.05	2.88	0.06
14	Italy	ItS	LB JpS CaS	4.98	0.08	2.31	0.11
15		ItS	LB JpS CaS GeS	5.03	0.08	2.29	0.11
Panel C. Different benchmarks and time							
16	Jan.89-Dec.93	NIKS	SP500 R2000 EMG NKIL and EUT	0.38	0.83	0.17	0.84
17	Jan.94-Dec.98	NIKS	SP500 R2000 EMG NKIL and EUT	3.18	0.2	1.51	0.23
18	Jan.99-Dec.04	NIKS	SP500 R2000 EMG NKIL and EUT	10.07	0.01	5.37	0.01
19	Jan.89-Dec.04	NIKS	SP500 R2000 EMG NKIL and EUT	8.81	0.01	4.47	0.01

LB comprises the seven large-cap benchmark indexes: S&P500, NIKKEI 300, SPMIB, SPTSX, FTSE100, CAC40, and DAX30.

depends on the interactions among the assets included in a portfolio. An index may be a separate asset class under some combination of benchmarks but not in other benchmark portfolios. Hence, pairwise correlation is an unreliable guide for constructing a well-diversified portfolio.

STEPDOWN SPANNING TEST RESULTS

The stepdown spanning test examines the two components of the spanning hypothesis of Equation (2) individually and jointly. Its purpose is to isolate each component in terms of the impact of an addition of a small-cap portfolio to an established benchmark on 1) the tangency portfolio, and 2) the global minimum-variance portfolio on the efficient frontier (see, e.g., Kan and Zhou [2001]).

The tangency portfolio effect is captured as a test of $\alpha = 0$ in Equation (2):

$$F_1 = (T - k - 1)\left(\frac{1}{U} - 1\right) \tag{5}$$

where U is the ratio of the unconstrained estimate of variance and the constrained estimate of variance by imposing only the constraint of $\alpha = 0$. Under the null hypothesis,

EXHIBIT 4

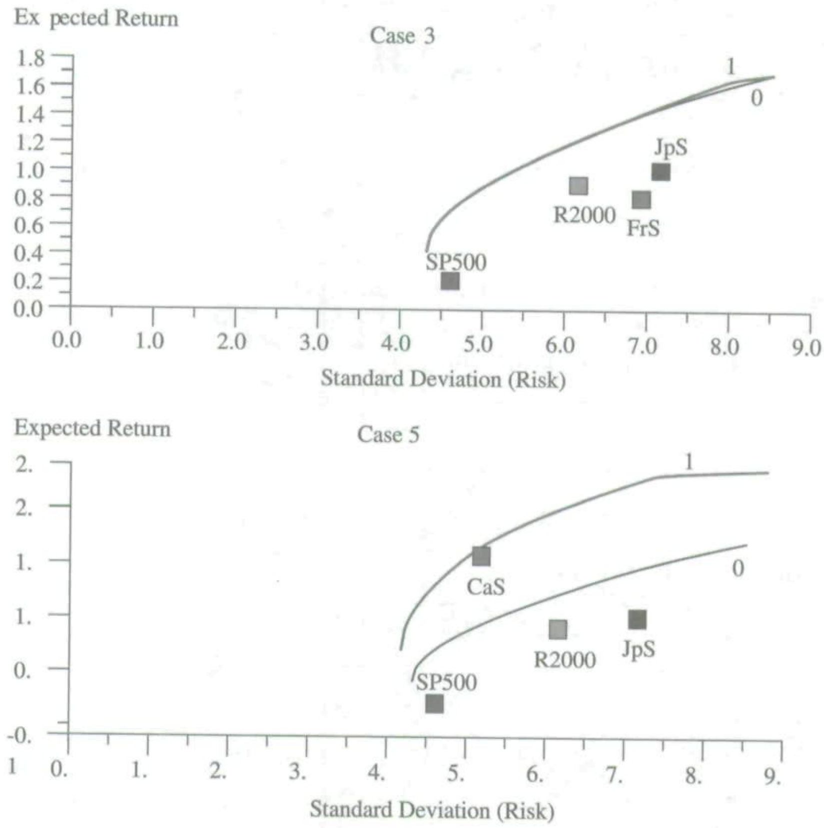
Stepdown Spanning Tests and Associated Performance Measures

Case	Stepdown test				Improvement in performance measures					
					Unconstrained		No short sales		Upper bound	
	F1-test	p-value	F2-test	p-value	SP	GMV	SP	GMV	SP	GMV
1	1.08	0.30	8.27	0.01	0.262	0.058	0.257	0.036	0.531	0.070
2	1.31	0.26	0.00	0.98	0.000	0.041	0.000	0.035	0.007	0.030
3	0.02	0.88	0.10	0.75	0.004	0.003	0.004	0.000	0.014	0.001
4	0.84	0.36	0.01	0.91	0.000	0.021	0.000	0.000	0.011	0.000
5	5.51	0.02	3.64	0.06	0.144	0.122	0.126	0.122	0.206	0.083
6	0.97	0.33	2.93	0.09	0.102	0.040	0.082	0.005	0.094	0.037
7	0.31	0.58	3.17	0.08	0.025	0.087	0.000	0.002	0.022	0.002
8	2.74	0.10	1.05	0.11	0.081	0.083	0.014	0.106	0.044	0.097
9	1.93	0.17	0.29	0.59	0.045	0.004	0.014	0.000	0.025	0.000
10	0.45	0.50	3.05	0.09	0.017	0.086	0.000	0.000	0.004	0.000
11	0.83	0.37	3.68	0.06	0.008	0.085	0.000	0.000	0.000	0.000
12	2.58	0.11	3.64	0.07	0.075	0.116	0.089	0.040	0.073	0.049
13	2.94	0.09	2.74	0.10	0.085	0.088	0.089	0.040	0.073	0.049
14	3.44	0.07	1.13	0.29	0.072	0.050	0.005	0.008	0.024	0.010
15	3.53	0.07	1.01	0.32	0.077	0.043	0.005	0.008	0.024	0.010
16	0.00	0.96	0.35	0.56	0.001	0.256	0.000	0.000	0.000	0.000
17	0.02	0.88	3.06	0.09	0.005	0.316	0.000	0.304	0.000	0.304
18	3.66	0.06	6.81	0.01	0.071	0.640	0.071	0.565	0.074	0.567
19	2.64	0.11	6.24	0.01	0.026	0.082	0.003	0.380	0.003	0.389

The case number reported corresponds to the benchmark portfolio and test asset combinations of Exhibit 3.

EXHIBIT 5

Efficient Frontier—with French and Canadian Portfolios



Frontier 0 is the benchmark portfolio, and Frontier 1 is the new frontier after the test asset is included.

F1 has a central F-distribution with 1 and $T-K-1$ degrees of freedom.

The global minimum-variance portfolio effect is captured in a test that $\Sigma \beta j = 1$, conditional on $\alpha = 0$:

$$F_2 = (T-K)\left(\frac{1}{U}-1\right) \tag{6}$$

where U is the ratio of the constrained estimate of variance by imposing only the constraint of $\alpha = 0$ and the constrained estimate of variance by imposing both the constraints of $\alpha = 0$ and $\Sigma \beta j = 1$. F2 has a central F-distribution with 1 and $T-K$ degrees of freedom, and is independent of F1.

Exhibit 4 provides the results of the stepdown spanning test for all cases (test asset and benchmark combinations) represented in Exhibit 3. Two performance measures that correspond to these test statistics are also reported: 1) the difference of the standard deviation (GMV) of the global minimum-variance portfolios between the benchmark assets and benchmark assets plus the testing asset—to capture risk reduction effects, and 2) changes in the Sharpe ratio of the tangency portfolios after adding the test asset to the benchmark assets. Following Jorion [1985], we assume a zero risk-free rate in the analysis to reduce the bias in the estimates.

When significant (insignificant) spanning tests are reported in many cases in Exhibit 3, the corresponding F1 and F2 stepdown tests are also both found to be significant (insignificant) in Exhibit 4. There are some notable exceptions, however. For example, while the Japanese small-cap portfolio JpS shows significance in the traditional spanning test (case 1), only its F2 test statistic is significant. In addition, while the U.S. small-cap benchmark, R2000, is insignificant in the traditional spanning test (case 7), its F2 test suggests that it may be advantageous in reducing the risk of the global minimum-variance portfolio.

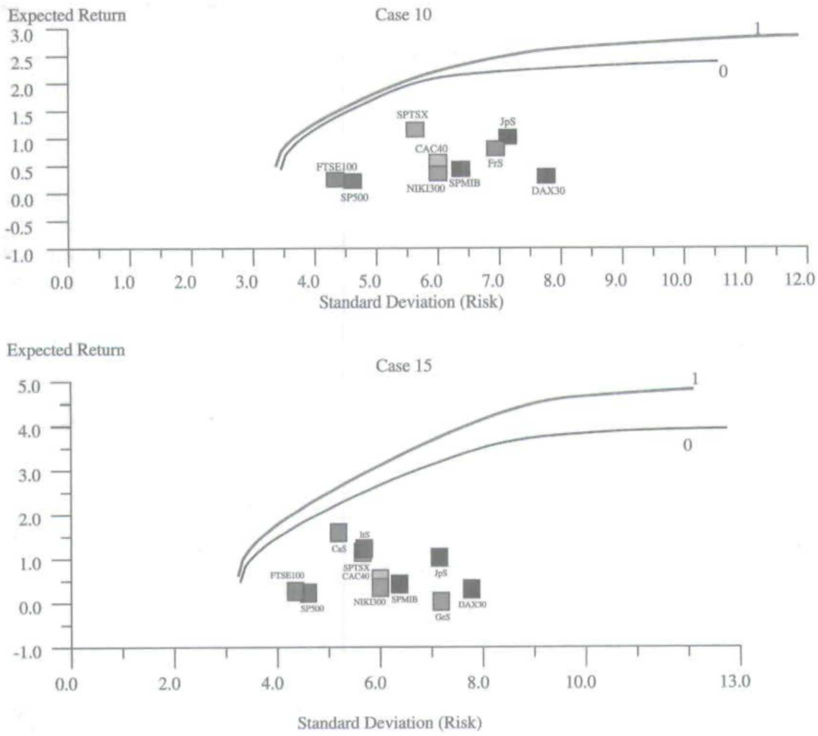
Assets with significant F2 statistics do not necessarily prove benefits of economic significance (Kan and Zhou [2001]), but differences in the tangency portfolios (significant F1 test statistics) will give rise to economic benefits to investors with low risk aversion. Since an asset may have an insignificant spanning test and at the same time a significant F1 test (ItS in cases 14 and 15, for example), the stepdown test is better than the traditional spanning test to find potential assets for diversification.

The stepdown tests also show sensitivity to changes over time and to the composition of the benchmark portfolios. The Japanese small-cap portfolio, NIKS, improves both the tangency portfolio and the global minimum-variance portfolio in the period January 1999 to December 2004 (case 18), although this portfolio improves neither the tangency portfolio nor the global minimum-variance portfolio over January 1989 December 1993 (case 16).

The significance tests of stepdown spanning are also reflected in the performance measures provided in Exhibit 4. Cases with significant F1 (F2) test values show greater improvements in their Sharpe ratios (minimum-variance portfolios).

EXHIBIT 6

Efficient Frontier—with French and Italian Portfolios



Frontier 0 is the benchmark portfolio, and Frontier 1 is the new frontier after the test asset is included.

Like Petrella [2005], we also look at the impact of investment policy constraints on the potential diversification benefits of small-cap portfolios. We report the results of three scenarios: 1) when investors are unconstrained—i.e., they can take long or short positions up to 100% of total capital; 2) when no short sales are allowed; and 3) when investors face upper bound restrictions for investments in a single asset—in our simulations, we set a weight of 0.5 as the upper bound in the optimization process.

As shown in Exhibit 4, in most cases constraints reduce the gains of GMV and SP of adding a small-cap portfolio to a benchmark portfolio. The constraints do not necessarily result in a simultaneous decline in both GMV and SP. For example, the upper bound constraint increases SP in case 5; the no short sale constraint increases GMV in case 8, over the unconstrained condition.

Exhibit 5 shows the efficient frontiers of case 3 and case 5, corresponding to the incorporation of French (FrS) and Canadian (CaS) small-cap portfolios as test assets. Although CaS has a higher correlation than FrS with each of the benchmark assets, CaS greatly increases the portfolio's efficiency. In contrast, FrS provides little diversification benefit for the benchmark portfolio.

The first graph in Exhibit 6 is for case 10, and clearly shows that there is only a slight change in the efficient frontier by adding small-caps from France to the large-cap international benchmark. Recall that for this asset the stepdown test results show an insignificant F1 test statistic and a mildly significant F2 test statistic. Hence, in this instance, only the risk of the minimum-variance portfolio is reduced when the test asset is included.

The second graph in Exhibit 6 corresponds to case 15, where the Italian small-cap market is the test asset against the large-cap international benchmark. Recall that for this case only the efficiency of the tangency portfolio improves after the test asset is included (i.e., only the F1 test statistic is significant), but the favorable shift or rotation of the efficient frontier for risk-averse investors due to addition of the Italian small-cap portfolio is quite evident.

The stepdown spanning test has clear advantages over traditional spanning tests. By separating the sources of diversification benefits rather than combining the test statistics in the traditional spanning test, we are better positioned to identify assets that provide Markowitz mean-variance frontier enhancements.

CONCLUSION

Spanning tests provide a useful and intuitive approach for identifying separate asset classes that enhance portfolio returns and provide diversification benefits. We apply such tests to portfolios of small-caps from G-7 countries. The composition of the benchmark portfolio, the time period examined, and the holding period of the investor are shown to influence the results. Furthermore, investment policy constraints do not necessarily reduce the diversification benefits of small-cap portfolios on benchmark portfolios.

Stepdown mean-variance spanning tests are powerful tools that can be used to search for potential additions to established benchmark portfolios and are consistent with widely used metrics of portfolio performance. Application of such tests to a large battery of replicable small-cap portfolios from G-7 countries indicates that only in a few countries can such portfolios behave as separate asset classes with return and diversification benefits.

ENDNOTES

The authors gratefully acknowledge financial support from the SSHRC (grant # 410-2001-0613).

[1]The U.S. small-cap fund in Eun, Huang, and Lai [2004] consists on average of 1,063 stocks per year. In Petrella's [2005] sample of European stocks, the investor would have to take positions in over 2,232 stocks to replicate the small-cap portfolio.

[2]The MSCI small-cap equity universe in each country incorporates a free float adjustment and has a minimum capitalization of $100 million.

[3]Available at http://www.fibv.com/publications/EQUITY104.pdf.

[4]We also use NIKS and Russell/Nomura Japan Small-Cap Total Return Index as representative of the Japanese small-cap portfolio. The results are very similar to those reported in Panel A of Exhibit 3.

[5]The LB consists of the S&P 500, the NIKKEI 300 (Japan), the FTSE100 (UK), the CAC40 (Germany), the DAX30 (Germany), the SPTSX (Canada), and the SPMIB (Italy).

[6]When we use the alternative Japanese small-cap benchmark indexes, the results are similar to those reported in Panel B of Exhibit 3.

REFERENCES

Banz, R. "The Relationship Between Return and Market Value of Common Stocks." *Journal of Financial Economics*, Vol. 9, No. 1 (1981), pp. 3-18.

Bekaert, G., and M. Urias. "Diversification, Integration and Emerging Market Closed-End Funds." *Journal of Finance*, Vol. 51, No. 3 (1996), pp. 835-869.

Dimson, E., and P. Marsh. "Murphy's Law and Market Anomalies." *The Journal of Portfolio Management*, Vol. 25, No. 2 (1999), pp. 53-69.

Driessen, J., and L. Laeven. "International Portfolio Diversification Benefits: Cross-Country Evidence." Working paper, University of Amsterdam, 2003.

Errunza, V., K. Hogan, and M. Hung. "Can the Gains from International Diversification be Achieved Without Trading Abroad?" *Journal of Finance*, Vol. 54, No. 6 (1999), pp. 2075-2107.

Eun, C., W. Huang, and S. Lai. "International Diversification with Large- and Small-Cap Stocks." Working paper, Georgia Institute of Technology, 2004.

Harvey, C. "Predictable Risk and Returns in Emerging Markets." *Review of Financial Studies*, Vol. 8, No. 3 (1995), pp. 773-816.

Horowitz, J., T. Loughran, and N. Savin. "Three Analysis of the Firm Size Premium." *Journal of Empirical Finance*, Vol. 7, No. 2 (2000), pp. 143-153.

Huberman, G., and S. Kandel. "Mean-Variance Spanning." *Journal of Finance*, Vol. 42, No. 4 (1987), pp. 873-888.

Jorion, P. "International Portfolio Diversification with Estimation Risk." *Journal of Business*, Vol. 58, No. 3 (1985), pp. 259-278.

Kan, R., and G. Zhou. "Tests of Mean-Variance Spanning." Working Paper # 03, University of Toronto, Rotman School of Management, 2001.

Petrella, G. "Are Euro Area Small-cap Stocks an Asset Class? Evidence from Mean-Variance Spanning Tests." *European Financial Management*, Vol. 11, No. 2 (2005), pp. 229-253.

Reilly, F., and D. Wright. "Alternative Small-Cap Stock Benchmarks." *The Journal of Portfolio Management*, Vol. 28, No. 3 (2002), pp. 82-95.