European Finance Review **6:** 189–221, 2002.

The Book-to-Market and Size Effects in a General Asset Pricing Model: Evidence from Seven National Markets

NEAL MARONEY
The University of New Orleans, College of Business Administration, New Orleans, LA. 70148, USA. E-mail: nmaroney@uno.edu

ARIS PROTOPAPADAKIS
University of Southern California, The Marshall School of Business, FBE Los Angeles, CA. 90089-1427, USA. E-mail: aris.protopapadakis@marshall.usc.edu

Abstract. The positive relation of returns with Book-to-Market ratio (*BE/ME*) and their negative relation with Market Value (*MVE*) remains strong under a general stochastic discount function (SDF) that does not depend on a specific asset pricing model and avoids potentially serious simultaneity biases inherent in the Fama and French three-factor model. However, we find that SDFs that include the equivalent of the *HML* portfolio do not span all asset sub-spaces, even with additional conditioning information. Finally, macro and financial variables we introduce to the pricing functions do not offer an alternative explanation of the *BE/ME* effect.

JEL classification codes: G10, G12, G15, G30.

1. Introduction

Several papers show that, in the U.S. as well as internationally, stock returns are correlated with the lagged values of their Book-Value to Market-Value ratio (*BE/ME*) and their market value (*MVE*), and that average returns are greater the higher the *BE/ME* or the smaller the *MVE*.[1] These phenomena are referred to as

[1] See Banz (1981), Basu (1983), Keim (1983) for the *MVE* (Size) effect, and Stattman (1980), Fama and French (1992), and more recently Daniel and Tittman (1997) for the *BE/ME* effect. A related but less pronounced empirical regularity exists between a firm's earnings-price, *E/P* ratio and expected returns. Basu (1977, 1983) and Reinganum (1981) document that firms with high *E/P* ratios have higher average returns than those with low *E/P* ratios. Internationally, Capaul et al. (1993) find a significant "value-growth factor" when they compare the risk-adjusted rates of return from indices composed of value and growth stocks in all five international markets they study (France, Germany, Switzerland, U.K., and Japan). Chan et al. (1991) find a strong relation between *BE/ME* and average returns in Japan, using a SUR procedure. More recent papers by Fama and French (1998), Heston et al. (1999), and Rouwenhorst (1999) extend these early findings. They find cross sectional evidence of the *BE/ME* and *MVE* effects within several international markets.

the "book-to-market effect" and the "size effect".[2] Furthermore, these variables "explain" the average return differences across portfolios that cannot be accounted for by the market portfolio beta (see Fama and French 1992).

Tests reported by Fama and French (1998), Heston et al. (1999), and Rouwenhorst (1999) use a three-factor CAPM variant that includes the *HML* portfolio – an arbitrage portfolio that consists of the difference in returns of the high *BE/ME* and the low *BE/ME* portfolios – as a factor that accounts for the *BE/ME* effect.[3] These papers report that unexplained cross-sectional return differences of portfolios sorted according to *BE/ME* ratios disappear; furthermore, they report R^2s close to one. The authors attribute their results to the significance of the *HML* portfolio.

Tests of the significance of the *BE/ME* and *MVE* effects reported to-date suffer from two potentially serious shortcomings. The first is that they use variants of the CAPM to account for systematic risk.[4] As a result, these tests depend on a specific benchmark model. If the wrong benchmark is chosen, then differences in average returns (or risk premia) may be incorrectly attributed to the existence *BE/ME* or *MVE* effects. Evidence that rejects the CAPM is common in the literature, and there is no agreed-upon empirical model to replace it. Therefore, it is extremely important to construct tests that do not rely on any specific asset pricing model.

The second serious shortcoming is that most of the reported tests use OLS methodology. It is at best doubtful that these tests satisfy the OLS requirement that the explanatory variables are independent of the LHS variable, because the tested portfolios constitute a significant portion of the explanatory variable, the market index! In this paper we report results of tests that are free of these shortcomings. We use a very general asset pricing function to investigate the *BE/ME* and *MVE* effects in seven national equity markets. This pricing model does not rely on the CAPM or ICAPM, and it encompasses all pricing functions that can be constructed from linear combinations of returns. At the same time, our methodology explicitly takes into account the endogeneity of the various portfolios relative to the market

[2] There have been several but as yet not wholly successful attempts to find a theoretical explanation for this empirical regularity. The most obvious explanation is that *BE/ME* or *MVE* proxies for an unobservable risk factor. However, there is no satisfactory theoretical model that makes this connection clear. Ikenberry et ak, (1995) relate this behavior of average returns to share repurchase programs; Loughran and Ritter (1995) and Spiess and Affleck-Graves (1995) connect this behavior to the price behavior of IPOs. Fama and French (1995) find links between returns, earnings, and *Size* and *BE/ME* factors. An intriguing possibility is advanced by Ferguson and Shockley (1999), that the *BE/ME* effect is created by the mismeasurement of the market portfolio due to the omission of corporate bonds.

[3] In a recent paper Ferson, Sergei and Sarkissian (1999) point out plausible conditions under which this approach gives entirely misleading results.

[4] Exceptions are Huberman and Kandel (1987), MacKinlay and Richardson (1991), and Ferson et al. (1993), who use the Latent Variables model to test if a subset of size-sorted portfolios span a larger set of size-sorted portfolios in the U.S. De Santis (1995) derives the spanning test we use based on the Hansen- Jagannathan methodology. He uses spanning tests to compute gains from international diversification and to assess the role of foreign exchange risk, and De Santis (1994) reports tests on diversification that includes emerging markets. No comparable tests have been reported for *BE/ME*.

portfolio. Our tests are by country and, unlike Fama and French (1998), do not rely on the assumption of international capital market integration.

Our tests of the *BE/ME* and *MVE* effects are "spanning" tests that use Stochastic Discount Factors or SDFs. These SDFs are constructed with the methodology proposed by Hansen and Richard (1987) and Hansen and Jagannathan (1991). Since they encompass all rational asset pricing models, they impose the fewest restrictions on the ability of asset returns to price risk; they insure only that the Law of One Price is not violated. The tests consist of constructing candidate SDFs from a subset of portfolios (reference portfolios) and using these to price or "span" all the assets in the market. Successful pricing or spanning implies that the SDFs adequately replicate the average returns of all the portfolios. Unconditional SDFs are constructed using only portfolio returns, while conditional SDFs are augmented with lagged values of other suitable variables, often referred to as managed portfolios (see Cochrane 2001).

Our results confirm many but not all of the results in the recent literature. The *BE/ME* and *MVE* effects survive this very general pricing model for *all* the countries examined. We find that spanning with the unconditional SDFs is overwhelmingly rejected for all the countries. SDFs equivalent to the *HML* portfolio span more frequent but not completely.[5] This finding warns of the danger of concluding that the *HML* portfolio provides a complete explanation of the variation in cross-sectional returns. Formal tests of the significance of the *BE/ME* and *MVE* variables show that the Book-to-Market and Size effects are *not* artifacts of CAPM's shortcomings.

Surprisingly, we find that *conditional* SDFs, which are SDFs augmented by lagged values of *BE/ME*, *MVE*, and each of 15 macroeconomic and financial variables, do not span better than their unconditional counterparts. The patterns of results for the conditional SDFs are similar to those of the unconditional SDFs.

We study stock returns for Australia, Canada, Germany, France, Japan, the U.K., and the U.S. These are the only national stock markets in the COMPUSTAT Global Vantage database with a sufficiently large number of equities traded over the sample period to allow us to form at least seven well-diversified, country-specific portfolios for our tests. This database is shorter than that used by Fama and French (1998) but it includes more stocks for each country.

We show that the statistical features of our data and the results of the standard tests are very similar to those reported by researchers that use the IFC or other datasets. This suggests that the results we present are not dataset-dependent.

Section 2 describes the data used in the study and the construction of the relevant variables and portfolios. Section 3 presents summary statistics for the data and replications of existing results, while section 4 presents the results of our tests.

5 Including the largest and smallest ranked *BE/ME* or *MVE* portfolios is equivalent to including the *HML* portfolio in a regression, except that these portfolios have unrestricted coefficients (rather than equal and opposite).

Section 5 offers concluding comments. The appendices contain some technical information.

2. Data

The firm level equity market data are from COMPUSTAT Global Vantage, compiled by Standard and Poor's Inc.[6] Data for stock indices, other financial variables, and the macro variables, are obtained from DataStream.

The rest of this section outlines in more detail the data used, the selection of the sample, the portfolio formation process, and the construction of the *BE/ME* and *MVE* variables. We discuss the particulars of the macro and other financial variables when they are introduced into our analysis.

2.1. EQUITY MARKET DATA

The COMPUSTAT database lists firms that, (i) are included in the MSCI World Index, (ii) are included in the MSCI country indices, (iii) are included in the major stock exchange indices of the countries, or, (iv) five or more analysts report on the firm. COMPUSTAT also reports that some shares are included because of direct client requests. For the U.S., the database includes all the NYSE and AMEX stocks, and some of the largest companies in the NASDAQ. This indicates that international COMPUSTAT carries only relatively large firms. Since 1991, firms that have stopped trading, for whatever reason, are retained in the database.

There are potential survivorship biases in the COMPUSTAT data that can erroneously lead to finding a *BE/ME* effect. Survivorship has been extensively studied in the U.S. data. Kothari et al. (1995) and Breen and Korajczyk (1993) find that if survivorship bias exists, it is more likely to be found among the smallest firms, most of which are listed on the NASDAQ. The international COMPUSTAT dataset is less likely to have survivorship bias problems, because it contains relatively large firms. Thus, it is unlikely that our results are influenced by survivorship bias.

We construct well-diversified portfolios for each country in order to reduce random noise in the returns. Based on the literature and the number of stocks traded in the various national markets, we form 7 portfolios for tests that involve one-way sorts, and 9 (3×3) portfolios for tests that involve two-way sorts.[7] Therefore, we consider only countries that list a substantial number of stocks. Of the 62 countries represented in COMPUSTAT, only 7 meet this criterion: Australia (AUS), Canada (CAN), Germany (DEU), France (FRA), the U.K. (GBR), Japan (JPN), and the U.S. (USA). Together these countries account for over 80% of the 7,623 firms included in the dataset. Among these, Germany lists the smallest number of stocks,

[6] This is the most complete generally available database that contains the necessary accounting information.

[7] It is desirable to keep the number of portfolios small because the reliability of the test statistics we use deteriorate when larger models are tested on small samples.

167.[8] We use only the common stock of publicly traded non-financial firms that report accounting figures in the same currency as the nationality of their exchange.[9] Our dataset contains monthly data from January 1982 through October 1994 for the U.S. and Canada, and December 1985 through October 1994 for Australia, Canada, France, Germany, Japan, and the U.K. We exclude October 1987 from all estimation because of the unusual and extreme character of the events related to Black Monday.

The MSCI world index is used to proxy for the world market portfolio.

2.2. PORTFOLIO FORMATION

Equally-weighted portfolios for each country are formed at the end of each October of year t by sorting stocks by Market Value of Equity (*MVE*) and Book-to-Market Equity (*BE/ME*) separately (for one-way sorts), and by sorting first by *MVE* and then by *BE/ME* for two-way sorts. For the *MVE* sorting, the Market Value of Equity for each stock is the product of the year t October price and the last reported number of shares outstanding.

For the *BE/ME* sorts, the Book Value (BE) is from the balance sheet reported no later than the end of April of year t. It includes shareholders' equity plus balance sheet deferred taxes (for Japan deferred taxes are zero).[10] Book Value is net of goodwill (for Canada, Japan, and U.S. goodwill is zero). For these sorts, the Market Equity (*ME*) is the end of *April* price (of year t) times the number of shares last reported.[11]

Portfolios are formed once a year, starting in November of year t. To be included in a year t portfolio, firms must have reported an April market value, an April or earlier book value, an October market value, and prices for at least 11 of the 12

---

[8] Of the base set of 7,623 firms from 48 countries, a total of 6,401 firms from 7 countries satisfy our criteria, that each firm reports at least 2 prices, 1 shares outstanding figure, and 1 book value for the entire period. The proportion of the selected firms to total firms by country is: USA-42%, JPN-14%, GBR-13%, CAN-6%, AUS-3%, FRA-3%, DEU-2%.

[9] No preferred stock or identifiable issues of subsidiaries are considered. There are a total of 41 firms that have accounting and market information in two different currencies. This occurs mostly with firms incorporated in New Zealand that trade in Australia and Irish firms that trade in the UK. These firms are excluded from the portfolios. In general, cross-listing of the same firm on two national exchanges is not an issue in this dataset. This is because market information is linked with accounting data, and COMPUSTAT reports consolidated balance sheets in the country of incorporation.

[10] These definitions are consistent with Fama and French (1992).

[11] A consequence of these conventions is that, for sorting purposes, $MVE \neq ME$. Share splits are accounted for in the data.

following months.[12] Firms that report negative Book Value are excluded from that year's *BE/ME* but not the *MVE* portfolios.

The April deadline for accounting data attempts to minimize look-ahead bias by allowing at least six months to elapse between the time accounting data are recorded and when they are known to the market.[13] Firms included in any year's portfolios must also have balance sheet data reported as of April of year t. If that is not available, it is taken from the prior year.[14]

We construct time series for $(BE/ME)_{i,t}$ and $(MVE)_{i,t}$ for each portfolio (i = 1 to 7 [9]). The $(BE/ME)_{i,t}$ variable for the ith portfolio is the average of the *BE/ME*s of the stocks included in each portfolio. We use the latest April Book Value (the same value used for the portfolio selection process, *BE*) but the market value is updated monthly, using the current price and the October number of shares outstanding. The $(MVE)_{i,t}$ variable for each portfolio is the sum of the market values at t of the stocks included in each portfolio, normalized by *last month*'s market value of *all* stocks. We normalize this variable to avoid using a nonstationary variable as an instrument; this normalization preserves the ordering of the *BE/ME* portfolios.

Finally, we follow Fama and French (1998) to construct the international *HML* portfolio. We use 30% of the highest *BE/ME* stocks from each country for the "*Hi*" portfolio and 30% of the lowest *BE/ME* stocks from each country for the "Lo" portfolio. Then we subtract the returns of *Hi* from *Lo*. All returns are first converted to US$s and all portfolios are equally weighted.

3. Summary Statistics and Replication of Existing Results

Average portfolio returns increase as *BE/ME* increases or *MVE* decreases, for all our countries; the *BE/ME* effect seems stronger than the *MVE* effect. Furthermore, the CAPM "beta" does not appear to account for cross sectional differences in average returns. Jensen's alpha (excess return) tests that are based on the CAPM confirm formally the conclusions from the descriptive statistics. Finally we replicate the Fama and French tests with the international *HML* portfolio and confirm that their results hold in our data as well. Adding the international *HML* portfolio

[12] Firms that fail to meet the October portfolio criteria are only excluded for that year's portfolios. If there is a missing price (maximum of 1), the missing return is replaced by the prior year's average return for the firm; and if that is unavailable, the firm is excluded. Firms that are delisted in the post-formation period are not subject to the eleven of the twelve month criterion; instead the final price or the final distribution is used. After such a delisting, the affected portfolio is re-weighted to reflect the increased investment to other firms within the portfolio.

[13] This is consistent with FF (1992). Financial reports of U.S. firms often are not available to the public until four months after fiscal year-end. The fiscal year in Japan is in March for most firms and accounting information is promptly reported (Chan and Lakonishok, 1991). Thus, the six month reporting lag is conservative for the countries that are analyzed.

[14] The maximum reporting lag is 2.5 years for firms with fiscal years ending in May for which data are missing for the prior year but have earlier data.

to an international CAPM model significantly improves the cross-sectional fit of average returns, and makes Jensen's alphas statistically insignificant.

The results in this section are fully consistent with those in the literature, even though we use a somewhat different dataset. Therefore the results of our new tests are unlikely to be dataset-dependent.

3.1. DESCRIPTIVE STATISTICS

Table I, Panel A shows statistics for the portfolios sorted on *BE/ME*, while Panel B shows the same statistics for the portfolios sorted on *MVE*. All the returns are annualized.

For the *BE/ME*-sorted portfolios there is a wide range of *BE/ME* ratios for each country. The range across the countries is similar; Canada, Germany and the U.K. have the widest range and Japan has the narrowest. The associated *MVE*s have a relatively narrow range. Generally, the portfolio *MVE*s and *BE/ME*s are negatively related.

The returns of the *BE/ME*-sorted portfolios generally increase with the *BE/ME* ratio; the highest *BE/ME* ratio portfolios have the highest returns, for every country. For all but the U.K., the returns to the largest *BE/ME* portfolios are much larger than the returns to any of the preceding ones. Furthermore, for all but Germany, the lowest *BE/ME* ratio portfolios have the lowest average returns. Returns are strictly monotonic only for Japan.[15]

For the *MVE*-sorted portfolios, *MVE* has a very wide but dissimilar range across countries. The ratio of largest-to-smallest *MVE* varies from a high of 478 for the U.S. to a low of 45 for Japan.[16] Furthermore, the smallest *MVE* portfolio is in the US; Japan and Germany have much larger *MVE*s both for the smallest and the largest portfolios. For all the countries, the smallest *MVE* portfolios have the highest returns. Again, this relation is strictly monotonic only for Japan.

Table II reports rank correlation coefficients of the portfolios ranked by *BE/ME* or *MVE* with their average returns. These confirm the strong relation between average returns and the *BE/ME* or *MVE* rankings, apparent in Table I. At the same time, it is clear that the *BE/ME* correlations are higher (except for Germany) than the *MVE* correlations. This conclusion is consistent with results in the literature for the U.S. (see Fama and French, 1992).

It is well known that for the U.S. data, traditional measures of systematic risk do not explain the relation of returns to the *BE/ME*- and *MVE*-sorted portfolios. In Table I we also show the relation between average returns and (i) their beta with

[15] Chan et al. (1991) also find a strong relation for Japan.

[16] The next largest ratios are for Canada and the U.K., at approximately 130. These values are in US$s.

Table I. Descriptive statistics for the protfolios

Panel A[1]. Sorted by book-to-market ratio

Country	Book-to Market Ratios (upper) and Number of Firms (*lower*)						
	Low-1	**2**	**3**	**4**	**5**	**6**	**High-7**
AUS	0.317	0.508	0.626	0.751	0.915	1.166	2.077
	20	*20*	*19*	*19*	*19*	*19*	*19*
CAN	0.259	0.486	0.643	0.801	0.983	1.209	2.739
	44	*44*	*44*	*44*	*43*	*43*	*43*
FRA	0.218	0.349	0.453	0.567	0.658	0.799	1.482
	13	*13*	*13*	*12*	*12*	*12*	*12*
DEU	0.190	0.336	0.428	0.521	0.650	0.857	1.923
	20	*20*	*20*	*20*	*20*	*20*	*19*
GBR	0.187	0.343	0.459	0.585	0.782	1.083	1.908
	92	*92*	*91*	*91*	*91*	*91*	*91*
JPN	0.141	0.238	0.305	0.368	0.436	0.522	0.717
	151	*151*	*151*	*151*	*151*	*151*	*150*
USA	0.199	0.385	0.527	0.674	0.841	1.068	1.558
	314	*313*	*313*	*313*	*313*	*313*	*313*

Country	Market Value[2] (upper) and Earnings to Price Ratio (*lower*)						
	Low-1	**2**	**3**	**4**	**5**	**6**	**High-7**
AUS	1,077	957	768	711	629	386	257
	0.028	*0.053*	*0.041*	*0.059*	*-0.046*	*-0.112*	*-0.213*
CAN	343	392	450	520	492	321	230
	-0.057	*0.005*	*0.023*	*0.017*	*0.039*	*0.011*	*1.586*
FRA	225	443	228	579	826	773	683
	-0.002	*0.024*	*0.016*	*0.035*	*0.039*	*0.045*	*0.030*
DEU	5,126	4,822	3,587	2,585	3,172	3,086	2,821
	0.027	*0.046*	*0.059*	*0.065*	*0.048*	*0.018*	*0.142*
GBR	1,305	1,020	1,084	1,266	862	614	567
	0.041	*0.063*	*0.069*	*0.072*	*0.046*	*0.059*	*-0.004*
JPN	2,140	2,095	2,025	1,953	1,581	1,401	2,101
	0.005	*0.013*	*0.015*	*0.020*	*0.020*	*0.023*	*0.021*
USA	1,403	1,326	1,018	960	861	905	639
	-0.022	*0.027*	*0.033*	*0.035*	*0.038*	*0.038*	*0.012*

Country	Average Returns (upper) and Standard Deviation of Returns (*lower*)						
	Low-1	**2**	**3**	**4**	**5**	**6**	**High-7**
AUS	9.6%	13.2%	19.6%	18.8%	14.6%	19.3%	32.1%
	0.70	*0.62*	*0.59*	*0.51*	*0.56*	*0.64*	*0.70*
CAN	12.0%	11.6%	16.4%	16.1%	18.4%	16.2%	30.2%
	0.65	*0.49*	*0.47*	*0.44*	*0.42*	*0.45*	*0.75*
FRA	13.3%	9.7%	13.0%	15.5%	12.1%	17.4%	24.0%
	0.53	*0.65*	*0.65*	*0.64*	*0.61*	*0.60*	*0.74*
DEU	8.5%	13.3%	13.6%	20.3%	21.2%	16.5%	25.7%
	0.68	*0.70*	*0.70*	*0.71*	*0.73*	*0.74*	*0.77*
GBR	10.1%	16.3%	19.4%	21.2%	20.0%	21.1%	23.2%
	0.63	*0.64*	*0.65*	*0.68*	*0.72*	*0.64*	*0.62*
JPN	6.0%	10.4%	11.2%	13.0%	14.2%	15.1%	19.8%
	1.02	*0.97*	*0.95*	*0.94*	*0.90*	*0.90*	*0.91*
USA	14.7%	16.2%	16.6%	18.0%	18.6%	18.3%	24.8%
	0.66	*0.59*	*0.58*	*0.54*	*0.51*	*0.46*	*0.48*

Country	Market Portfolio Betas (upper) and Sharpe Ratios (*lower*)						
	Low-1	**2**	**3**	**4**	**5**	**6**	**High-7**
AUS	0.87	0.96	0.82	0.76	0.67	0.85	0.86
	-0.045	*-0.024*	*0.048*	*0.074*	*0.007*	*0.046*	*0.189*
CAN	1.02	0.90	0.91	0.85	0.79	0.81	0.95
	0.016	*0.016*	*0.102*	*0.104*	*0.151*	*0.114*	*0.244*
FRA	0.67	0.79	0.75	0.83	0.79	0.78	0.79
	0.052	*-0.006*	*0.032*	*0.061*	*0.018*	*0.088*	*0.177*
DEU	0.90	0.94	0.95	0.91	0.92	0.94	0.92
	-0.047	*0.018*	*0.018*	*0.108*	*0.111*	*0.044*	*0.155*
GBR	0.95	0.98	0.98	1.00	1.02	0.96	0.92
	-0.067	*0.023*	*0.059*	*0.088*	*0.067*	*0.094*	*0.123*
JPN	1.06	1.02	1.00	0.98	0.93	0.92	0.89
	0.005	*0.049*	*0.059*	*0.077*	*0.099*	*0.108*	*0.161*
USA	1.19	1.15	1.15	1.03	0.99	0.87	0.84
	0.067	*0.100*	*0.108*	*0.141*	*0.165*	*0.185*	*0.291*

Table I. Continued

Panel B[1]. Sorted by market value

Country	Market Value[2] (upper) and Number of Firms (lower)						
	Low-1	2	3	4	5	6	High-7
AUS	53	107	169	256	446	857	2,958
	20	*20*	*20*	*19*	*19*	*19*	*19*
CAN	14	40	73	127	223	447	1,828
	47	*47*	*47*	*46*	*46*	*46*	*46*
FRA	33	73	126	202	335	602	2,381
	13	*13*	*13*	*13*	*12*	*12*	*12*
DEU	273	674	1,119	1,706	2,532	4,190	14,783
	21	*21*	*21*	*20*	*20*	*20*	*20*
GBR	38	79	133	218	383	844	4,953
	93	*93*	*93*	*93*	*93*	*93*	*92*
JPN	187	321	476	718	1,147	2,088	8,395
	151	*151*	*151*	*151*	*151*	*151*	*151*
USA	12	37	78	151	317	813	5,532
	333	*333*	*332*	*332*	*332*	*332*	*332*

Country	Book-to Market Ratios (upper) and Earnings to Price Ratio (lower)						
	Low-1	2	3	4	5	6	High-7
AUS	1.218	1.034	0.900	0.803	0.813	0.644	0.658
	-0.671	*-0.043*	*0.042*	*0.073*	*0.073*	*0.074*	*0.070*
CAN	0.502	0.972	0.906	0.887	0.828	0.825	0.777
	0.708	*-0.016*	*-0.007*	*0.047*	*0.021*	*0.032*	*0.057*
FRA	0.553	0.654	0.594	0.588	0.641	0.553	0.833
	-0.028	*0.006*	*0.018*	*0.027*	*0.043*	*0.029*	*0.062*
DEU	0.495	0.719	0.571	0.533	0.831	0.522	0.548
	-0.009	*0.020*	*0.061*	*0.049*	*0.094*	*0.077*	*0.076*
GBR	0.809	0.744	0.710	0.699	0.680	0.645	0.586
	-0.158	*0.028*	*0.056*	*0.064*	*0.060*	*0.068*	*0.075*
JPN	0.386	0.416	0.403	0.399	0.384	0.382	0.348
	0.008	*0.014*	*0.015*	*0.018*	*0.018*	*0.020*	*0.021*
USA	0.523	0.733	0.762	0.659	0.662	0.683	0.707
	-0.453	*-0.050*	*-0.003*	*0.036*	*0.056*	*0.061*	*0.077*

Country	Average Returns (upper) and Standard Deviation of Returns (lower)						
	Low-1	2	3	4	5	6	High-7
AUS	30.4%	17.8%	11.1%	15.6%	16.6%	18.9%	18.6%
	1.01	*0.57*	*0.57*	*0.53*	*0.53*	*0.63*	*0.62*
CAN	44.3%	20.3%	12.5%	11.9%	9.3%	11.1%	8.6%
	0.97	*0.52*	*0.49*	*0.45*	*0.46*	*0.46*	*0.45*
FRA	17.5%	13.2%	15.0%	13.9%	17.9%	11.2%	13.8%
	0.66	*0.73*	*0.70*	*0.63*	*0.52*	*0.59*	*0.70*
DEU	26.6%	19.7%	15.5%	12.4%	14.5%	14.8%	12.6%
	0.94	*0.73*	*0.63*	*0.72*	*0.71*	*0.75*	*0.65*
GBR	22.2%	18.8%	18.8%	16.9%	17.5%	17.4%	18.4%
	0.78	*0.67*	*0.69*	*0.67*	*0.67*	*0.67*	*0.61*
JPN	21.5%	16.3%	13.6%	12.9%	10.6%	8.9%	5.7%
	1.12	*1.00*	*1.00*	*0.98*	*0.94*	*0.88*	*0.78*
USA	47.3%	18.0%	16.0%	15.7%	15.7%	16.2%	16.7%
	1.38	*0.60*	*0.56*	*0.58*	*0.58*	*0.52*	*0.48*

Country	Market Portfolio Betas (upper) and Sharpe Ratios (lower)						
	Low-1	2	3	4	5	6	High-7
AUS	0.84	0.73	0.76	0.69	0.79	0.91	1.08
	0.141	*0.045*	*-0.042*	*0.014*	*0.039*	*0.042*	*0.032*
CAN	1.00	0.87	0.87	0.82	0.91	0.95	0.95
	0.312	*0.151*	*0.036*	*0.030*	*-0.017*	*0.009*	*-0.028*
FRA	0.62	0.78	0.82	0.76	0.67	0.78	0.96
	0.104	*0.030*	*0.057*	*0.053*	*0.110*	*0.006*	*0.039*
DEU	1.01	0.91	0.80	0.92	0.94	1.01	0.90
	0.131	*0.101*	*0.050*	*0.002*	*0.032*	*0.028*	*0.008*
GBR	0.91	0.87	0.94	0.99	1.03	1.08	1.02
	0.096	*0.053*	*0.057*	*0.031*	*0.034*	*0.034*	*0.054*
JPN	1.00	0.95	0.97	0.99	1.00	0.98	0.93
	0.153	*0.108*	*0.082*	*0.073*	*0.050*	*0.038*	*0.002*
USA	1.02	0.98	1.04	1.11	1.12	1.08	1.02
	0.265	*0.128*	*0.102*	*0.095*	*0.098*	*0.124*	*0.149*

[1] The table below reports time series averages of important portfolio characteristics by country and by the type of portfolio ranking. Unless otherwise noted, all the statistics are reported in local currency. All balance sheet data come from balance sheets reported no later than April of year t. These statistics are first calculated for the individual firm at year t, and are equally weighted into the portfolio in which it belongs at that date. Time series averages are for each portfolio as a whole. The sample period for USA and CAN is November 1983 to October 1994 and for AUS, FRA, DEU, GBR, and JPN is November 1986 to October 1994.

[2] The Market Values are reported in millions of US$s, translated at the average exchange rate for the period.

Table II. Rank correlations between average returns, market value and book-to-market

Rank	Correlations	
	BE/ME	*MVE*
Australia	0.71	0.18
Canada	0.89	0.96
Germany	0.64	0.46
France	0.89	0.79
U.K.	0.96	0.64
Japan	1.00	1.00
USA	0.96	0.39

The table A shows the rank correlations of average returns with the ranking variables, for each country. A value of 1.00 means that the returns increase monotonically with their *BE/ME* or *MVE* ranking for that country.

respect to the World Market portfolio, and (ii) the Sharpe ratios of excess returns.[17] The table shows a lack of a clear relation between market betas and average returns across the ranked portfolios. The relation is either flat or has the wrong sign. There is no country for which the highest *BE/ME* (or lowest *MVE*) portfolio also has the highest beta! Furthermore, the relation between the ranked portfolios and the Sharpe ratios look remarkably like those for average returns. The largest *BE/ME* and the smallest *MVE* portfolios have the highest Sharpe ratios. These findings imply that the relation of returns to the *BE/ME* and *MVE* rankings is not explained well by standard risk adjustments.

The above descriptive statistics broadly confirm that the empirical regularities observed in the U.S. also exist in the other six countries.[18]

The next section shows the results of CAPM-based regression tests that explore the statistical significance of the relations discussed above. For each country, we estimate Jensen's alphas (excess returns) for each portfolio by regressing adjusted portfolio returns on world and local market indices. Then we examine the excess

[17] The Sharpe ratio is the ratio of the average excess returns to their standard deviations. Since these portfolios are highly diversified in the context of the respective national markets, their standard deviations are likely to be good approximations of traditional measures of systematic risk.

[18] Fama and French (1992) show that the *BE/ME* ratio does better than *MVE* in explaining the cross-section of returns, for the U.S. data. For that reason we also examine two-way sorts. We sort equities first by *MVE* and then by *BE/ME*. If *MVE* contained all the relevant information then the average returns would not be related to *BE/ME*. These two-way sorts suggest that the information contained in *BE/ME* is not duplicated by the *MVE* variable. Plots of average returns reveal that the relation between *BE/ME* and returns remains within each *MVE* portfolio. The relation seems most evident in the smallest *MVE* portfolios, and least evident in the largest *MVE* portfolios. Two-way sorted portfolios are included in all further analysis.

Table III. The relation between Jensen's alphas, market value and book-to-market

Panel A. Zero-beta returns				
	Rank correlations		P-values of wald tests	
	BE/ME	*MVE*	*BE/ME*	*MVE*
Australia	0.68	0.50	0.009	0.575
Canada	0.96	0.96	0.003	0.000
Germany	0.64	0.50	0.103	0.274
France	0.89	0.75	0.002	0.281
U.K.	0.96	0.75	0.000	0.491
Japan	1.00	1.00	0.017	0.162
USA	1.00	0.68	0.000	0.029

Panel A shows the rank correlations of zero-beta returns of portfolios with the ranking variables, for each country. A value of 1.00 means that the returns increase monotonically with their *BE/ME* or *MVE* ranking for that country. This Panel also reports p-values for Wald tests that the zero-beta returns are equal across portfolios ranked by *BE/ME* and *MVE*.

return estimates for apparent regularities. We also replicate recent results reported by Fama and French (1998) using an international *HML* portfolio.

3.2. CAPM-BASED REGRESSION TESTS

We estimate the following SUR regression model for the two sets of 7 portfolios from one-way sorts (by *BE/ME* and *MVE*) and the 9 portfolios from the two-way sorts (by *MVE* and then by *BE/ME*), for each country separately:

$$R^x_{i,t} = \alpha_0 + \alpha_i + \beta_{m,i} R^x_{m,t} + \beta_{w,i} R^x_{w,t} + \varepsilon_{i,t}; \quad i = 1, 7(9), \quad \alpha_1 \equiv 0. \quad \text{(1a)}$$

The variables $R^x_{i,t}$, $R^x_{m,t}$, and $R^x_{w,t}$ are, respectively, local currency returns for the sorted portfolios, the MSCI national market returns, and the World market returns, all adjusted by the local risk-free rate.[19] The sum, $\alpha_0 + \alpha_i$, is Jensen's alpha (or excess return or zero-beta return) for each portfolio.

Table III Panel A shows the rank correlation coefficients between the Jensen's alphas for each set of portfolios and their ranking by *BE/ME* and *MVE*. High correlations imply that the *BE/ME* or the *MVE* effect exists. The table also reports

[19] The world market portfolio is the MSCI World Index. MSCI index returns are translated into local currency returns using the month-end exchange rate. To calculate excess returns we subtract from the returns the following short term interest rates for each country: AUS – average rate on money market; CAN – 1 month T-Bill rate; FRA – call money rate; DEU – call money rate; GBR – 1 month T-Bill rate; JPN – call money rate; USA – 1 month T-Bill rate. All returns are annualized.

Table III. Continued

Panel B. Jensen's alphas with and without the *HML* portfolio

BE/ME-sorted portfolios	Low 1	2	3	4	5	6	Hi 7
Without *HML*	−0.009	0.010	0.036	0.062	0.053	0.061	0.146
	(0.751)	(0.713)	(0.160)	(0.028)	(0.061)	(0.031)	(0.000)
With *HML*	−0.012	−0.001	0.004	0.015	−0.014	−0.025	0.013
	(0.734)	(0.987)	(0.888)	(0.659)	(0.655)	(0.422)	(0.724)
MVE-sorted portfolios	Small 1	2	3	4	5	6	Large 7
Without *HML*	0.198	0.063	0.032	0.023	0.031	0.018	0.016
	(0.000)	(0.063)	(0.278)	(0.397)	(0.245)	(0.502)	(0.487)
With *HML*	0.048	−0.016	−0.020	−0.022	−0.002	−0.015	0.008
	(0.358)	(0.676)	(0.567)	(0.505)	(0.938)	(0.644)	(0.768)

Panel B shows estimates of Jensen's alphas when the pricing model includes only the market portfolio and when it also includes the *BE/ME HML* portfolio. The 7 portfolios are constructed by converting the corresponding country portfolios into U.S.$ returns and weighting them equally. P-values for the hypothesis that the coefficient is not different from zero are in parentheses.

p-values for Wald tests that the zero-beta returns are equal across the portfolios (detailed results are available from the authors on request). Panel A clearly demonstrates that there are strong positive correlations between zero-beta returns and *BE/ME* in all the countries. All the rank correlations are high, and the Wald tests are highly significant, except for Germany. The results for the *MVE*-sorted portfolios are considerably weaker. The hypothesis that the zero-beta returns are the same across portfolios is rejected only for Canada and the U.S. However, the rank correlations are not much lower for the *MVE*-sorted than the *BE/ME*-sorted portfolios. For two-way sorts (not shown) tests for the significance of zero-beta returns confirm those of the one-way sorts. Only Australia and Germany do not reject the equality of zero-beta returns at the 5% level.

These tests quantify and formalize the conclusions from the descriptive statistics. The *BE/ME* "effect" is found in all the countries in our sample in unconditional or CAPM-based tests, while evidence on the *MVE* effect is weaker. These tests do not assume international capital market integration.

3.3. TESTS WITH THE *hml* PORTFOLIO

A somewhat different test has been reported in Fama and French (1998). They construct an international *HML* portfolio as a factor that represents the *BE/ME* effect. They assume that international capital markets are integrated, and they regress international portfolios –sorted on *BE/ME*– on the market and the *HML* portfolio; all the returns are in US dollars. They find that the Jensen's alphas are no longer significantly different from zero, and they do not increase as *BE/ME* increases.

In Table III Panel B we report similar results for our data. The table reports the Jensen's alphas (α_i) estimated from regressions (1b) and (1c), below:

$$R_{i,t}^{\$x} = \alpha_0 + \alpha_i + \beta_{w,i} R_{w,t}^{\$x} + \varepsilon_{i,t}, \tag{1b}$$

$$R_{i,t}^{\$x} = \alpha_0 + \alpha_i + \beta_{w,i} R_{w,t}^{\$x} + \beta_{HML,i} R_{HML,t}^{\$x} + \varepsilon_{i,t};\ i = 1, 7,\ \alpha_1 \equiv 0, \tag{1c}$$

where $R_{HML,t}^{\$x}$ is the *Hi* minus *Lo* portfolio returns, and all other variables are as defined for Equation (1a), except that returns are in US dollars. The results clearly show that the Jensen's alphas estimated from these world portfolios also increase as *BE/ME* increases and as *MVE* decreases. Furthermore, when the *HML* portfolio is added to the regressions, Jensen's alphas are no longer significant and they have no clear pattern.

As we have discussed already, all the above tests use the CAPM as the benchmark model. It is entirely possible that a more general pricing model would account for the apparent pricing anomalies. Indeed, since the *HML* portfolio is constructed from linear combinations of returns, it is reasonable to suspect that such a model could be found.

Furthermore, all these tests share a common difficulty: The LHS portfolios are a significant portion of the market, which is on the RHS; thus the LHS and RHS portfolio returns are determined simultaneously. But since classical regression assumptions require at least that the RHS variables are predetermined, it is not clear to what extent one can rely on the coefficient estimates and the associated test statistics.

4. Spanning Tests

In this section we discuss and implement a set of spanning tests based on the Hansen- Jagannathan – HJ (1991) – Stochastic Discount Factor (SDF) methodology. The type of test that we use was first proposed by De Santis (1994, 1995), and it relies on the idea that one can construct a pricing function (or a stochastic discount factor, SDF) from the returns of portfolios that are traded (or could be traded) in the asset markets. These tests do not depend on the validity of any particular asset-pricing model; all linear combinations of the underlying asset returns are potential pricing models, and all existing linear models are special cases. A short technical description of the construction and properties of a SDF is in Appendix A.[20]

[20] Campbell et al. (1997), and Cochrane (2001) contain excellent detailed expositions of the Stochastic Discount Factor methodology.

If there are a small number of risk factors in the economy, then a subset of portfolios will price themselves and the remaining assets. If such a subset of assets (*reference* assets) price or *span* all the assets (*reference* plus *test* assets), then the reference assets contain all the necessary pricing information.[21] This is generally referred to as the *unconditional spanning* hypothesis (see Huberman and Kandel, 1987); it is labeled "unconditional" because only contemporaneous portfolio returns enter the SDF.[22]

If spanning is not rejected, then no other information or variable is required to fit average returns. If it is rejected, then one may augment the SDF with additional "conditioning" variables, or managed portfolios, to see if spanning can be achieved. The statistic generally used for such tests is the Hansen J-statistic, a goodness of fit measure.

The test evaluates the distance between two bounds in returns-variance space; one bound is constructed from SDFs that include the *reference* plus the *test* assets and the other from SDFs that include only the *reference* assets.[23] If the two bounds are sufficiently close to each other we conclude that the reference assets span or price the test assets.

In Section 4.1 we report on these unconditional spanning tests for 7 portfolios sorted by *MVE*, 7 portfolios sorted by *BE/ME*, and 9 portfolios obtained from two-way sorts (by *MVE* and then by *BE/ME*), for each country. We designate 5 out of 7 (or 7 out of 9) portfolios as *reference* assets and the remaining 2 as *test* assets. We show that these unconditional SDFs generally do *not* span the assets. More specifically, when portfolios with extreme values of *BE/ME* or *MVE* are excluded from the SDFs, spanning is overwhelmingly rejected. But even with the extreme portfolios in the SDFs, spanning is rejected much more frequently than would be expected at random. This raises doubts about the advisability of using the *HML* portfolio as a cure-all factor.

In the first part of Section 4.2 we test the statistical significance *BE/ME* and *MVE*, one at a time, by including their lagged values in the SDFs. They are uniformly significant across all countries. We conclude that the *BE/ME* and *MVE* "effects" exists in all the countries in our sample, even in the context of our very general pricing model. We also test if "conditional" SDFs, augmented with the

[21] For example, in the CAPM framework, any two risky frontier portfolios span all the other risky assets.

[22] DeRoon and Nijman (2001) provide a useful survey mean variance spanning tests.

[23] The bounds are described by an SDF as the risk-free rate varies. Because the bounds are quadratic, only 2 points on each frontier need to be evaluated to assess if the bounds coincide or not. The so-called "intersection" hypothesis relies on only one point on the frontier, and examines possible intersections of the bounds.

lagged values of *BE/ME* or *MVE* span better than the unconditional SDFs.[24] We find that there is no improvement in spanning performance.

We continue the investigation in Section 4.3 by testing if *BE/ME* simply proxies for other macro or financial variables. We examine the marginal improvement of adding each of 15 macro and financial variables to the SDF, one at a time, along with *BE/ME*. The tests of these additional conditioning variables proceed in the same manner as before. None of these additional variables substitutes for *BE/ME* but many of them are significant along with *BE/ME*. However, surprisingly, SDFs augmented by *BE/ME* and these new variables, one at a time, produce only minor improvement in spanning performance.

We use the Generalized Method of Moments (GMM, Hansen, 1982) because it is particularly well suited for our spanning tests. It does not require strong distributional assumptions and it makes it easy to impose orthogonality conditions and overidentifying restrictions.[25] Furthermore it accommodates the endogeneity of the reference assets because it is not necessary to assume orthogonality between the asset returns on the LHS and the test assets on the RHS. The properties of GMM as it applies to our tests are outlined in Appendix B.

4.1. UNCONDITIONAL SPANNING

Consider constructing an SDF, *SDF(ref)*, only from the reference assets. The *unconditional* spanning test consists of testing if the reference portfolios span all the portfolios. The null hypothesis is that the reference portfolios should price or span all the portfolios adequately.

4.1.1. *Test Specification*

We estimate SDFs that maximize the fit across all 7 (9) portfolios while restricting the weights of the test portfolios in the SDF to be zero. For the *BE/ME*- and *MVE*-sorted portfolios, we designate 5 of the 7 portfolios as reference assets and the remaining 2 as test assets. For the two-way sorted portfolios we designate 7 as reference and 2 as test assets. We report results for different combinations of test assets.

---

[24] Such SDFs are formally equivalent to SDFs constructed only from portfolio returns but with nonlinear coefficients, where the nonlinearity is modeled with the conditioning variables. Empirical evidence shows that time-varying betas and risk premia fit U.S. equity market returns significantly better (Campbell, 1987; Harvey, 1989, 1991; Shanken, 1990). Ferson and Harvey (1993) find the lagged MSCI world index to be the most important predictor of international equity market returns. Jagannathan and Wang (1993) find that the importance of *MVE* is greatly reduced when they allow for time variation in market betas over the business cycle, in addition to their human capital variable.

[25] This generality is very useful because the HJ bounds assume little more than finite first and second moments for asset returns.

Let p assets be the reference assets; the $q = n - p$ remaining assets are the test assets. The system that defines the SDF – $m_{c_j,t}$ for the n assets is,

$$\begin{array}{l} R_t - E(R_t \mid m_{c,1}) = \nu_{1,t} \\ R_t - E(R_t \mid m_{c,2}) = \nu_{2,t} \end{array}; \quad \varepsilon_t' = (\nu_{1,t}' \; \nu_{2,t}'),\, E(\varepsilon_t) = 0, \tag{2a}$$

$$m_{c,j,t} = \sum_{p=1}^{5(7)} \beta_{p,j} r_{p,t} + c_j; \quad E(m_{c,j,t}) = c_j,\, j = 1, 2, \tag{2b}$$

where $r_t = R_t - E(R_t)$, and $m_{c,j,t}$ is the SDF as a function of a pre-specified risk-free rate, c_j. *Unconditional* spanning implies that the average pricing error is zero for all asset returns, within the limits of statistical tolerance. The Hansen J-Statistic (Hansen, 1982; Hansen and Singleton, 1982) is used to evaluate the overidentifying conditions implied by spanning. It is distributed χ^2 with degrees of freedom equal to the number of overidentifying conditions.[26]

4.1.2. *Results*

Table IV displays the results of these spanning tests for three pairs of test assets for *BE/ME*, *MVE*, and the 2-way sorted portfolios, by country. We report results for the *extreme* pairs of test assets (1&7 for the *BE/ME* and *MVE* sorting, and 3&7 for the 2-way sorting), for the middle pairs of test assets (3&5 for the *BE/ME* and *MVE* sorting, and 2&6 for the 2-way sorting), and for one set of *in-between* pairs (2&6 for the *BE/ME* and *MVE* sorting, and 1&9 for the 2-way sorting).[27] The *extreme* pair of portfolios rank the smallest and largest in terms of *BE/ME* or *MVE*, the *middle* pair of portfolios rank in the middle, while the *in-between* portfolios are neither in the extremes nor in the middle.[28] Entries are shaded if they statistically reject spanning at the 5% level; we report p-values for the tests.

Overall, the spanning hypothesis is rejected in 27 out of 63 cases. However, there are significant differences in rejection rates across the *test* assets. We define a "case" as 1 portfolio ranking and 1 country. When the *extreme* pairs are the test assets (1&7/3&7), and thus excluded from the SDF, spanning fails in 14 out of 21 cases (67%). Spanning fails in every country for at least one portfolio ranking.

[26] There are $2n$ orthogonality conditions, $2p$ parameters to estimate, and $2q$ overidentifying conditions.

[27] Our numbering convention for the portfolios is shown below.

MVE or *BE/ME*:

Small or Lo						Large or Hi
1	2	3	4	5	6	7

Two-Way Sort:

	BE/ME		
	Lo		Hi
Small	1	2	3
MVE	4	5	6
Large	7	8	9

[28] There are many possible "*in-between*" portfolio pairs. The "middle" portfolio (no. 4 for one-way sorts and no. 5 for two-way sorts) is never used as a test asset.

Table IV. Volatility bounds – unconditional spanning tests (J-Statistic P-Values)

Country	Portfolios sorted by	Test portfolios 1&7	2&6	3&5
AUS	*BE/ME*	0.083	0.519	0.055
CAN		0.000	0.014	0.030
DEU		0.005	0.254	0.090
FRA		0.015	0.739	0.937
GBR		0.000	0.131	0.000
JPN		0.017	0.333	0.045
USA		0.000	0.000	0.684
AUS	*MVE*	0.202	0.671	0.070
CAN		0.058	0.136	0.683
DEU		0.308	0.004	0.008
FRA		0.208	0.390	0.054
GBR		0.001	0.001	0.993
JPN		0.003	0.093	0.514
USA		0.000	0.632	0.147
		Test portfolios 3&7	1&9	2&6
AUS	2 way *MVE & BE/ME*	0.011	0.155	0.058
CAN		0.379	0.011	0.049
DEU		0.180	0.001	0.176
FRA		0.046	0.816	0.934
GBR		0.001	0.047	0.120
JPN		0.001	0.247	0.070
USA		0.000	0.001	0.276

R_t is a n dimensional vector of date t returns on n assets; $r_t = R_t - E(R_t)$ is the vector of demeaned returns. The first p of the n assets are the reference assets and the remaining $q = n - p$ are the test assets. The returns are $R_t = (R'_{p,t}, R'_{q,t})'$. The SDF is m_c and c_j is the predetermined risk-free rate. We use two risk-free rates for the tests. If the n assets are *spanned* by the reference assets, then the volatility bound constructed from the reference assets remains unchanged when the test assets are added. The time-invariant volatility bound is formed by constructing two discount factors that have expected values c_1 and c_2. The system that defines the volatility bound for n assets is:

$$R_t - E(R_t \mid m_{c_j}) = v_{j,t} E(\varepsilon_t) = 0, \quad \text{and } \varepsilon'_t = (v'_{1,t}\ v'_{1,t}),$$

where, $E(R_t \mid m_{c_j}) = [1_n - \mathrm{Cov}(R_t, m_{c_j,t})]/c_j$, and $m_{c_j,t} = \sum_{p=1}^{5(7)} \beta_{p,j} r_{p,t} + \sum_{q=6(8)}^{7(9)} \beta_{q,j} r_{q,t} + c_j$, for $j = 1, 2$.

The hypothesis that p reference assets are sufficient to span n assets implies that $\beta_{q,1} = \beta_{q,2} = 0$; the q test assets are not needed for pricing. *Unconditional* spanning implies the null hypothesis that $E(\varepsilon_t) = 0$, i.e., returns are on average correctly priced. Iterated GMM is used to estimate the discount factor parameters $\beta_{p,1}$, and $\beta_{p,2}$, and the Hansen J-statistic is used to evaluate the overidentifying conditions implied by spanning. This test statistic is distributed χ^2 with degrees of freedom equal to the number of overidentifying conditions. There are $2n$ conditions and $2p$ parameters, which leaves $2q$ overidentifying conditions.

Seven *BE/ME*- and *MVE*-sorted portfolios from each country, and 9 portfolios from 2-way rankings (by *MVE* and then by *BE/ME*) are used for the spanning tests. The construction of *BE/ME*, *MVE*, and the portfolios is described in the Data section. Three tests are reported for each set of sorted portfolios for each country, each for a different pair of test assets.

The table shows the right hand tail probabilities for the unconditional spanning tests. Tests are based on GMM criteria at convergence or the tenth iteration, whichever was first. The GMM estimates are robust to first-order autocorrelation and heteroskedasticity. The sample period is 11/83 – 10/94 (T = 132) in the USA and CAN, and 11/86 – 10/94 (T = 96) in AUS, DEU, FRA, GBR, and JPN. Shaded entries are significant at the 5% level.

But when the *middle* pairs are excluded spanning fails in only 5 out of 21 cases (24%).[29]

This result shows that SDFs that include the *extreme* portfolios span much more successfully than those that exclude them! This is clear evidence that the *BE/ME* and the *MVE* effects exist even under this very general asset pricing model. Including the *extreme* portfolios is the equivalent of including the *HML* portfolio as a pricing "factor" in the context of the CAPM. But unlike the Fama-French results, we find that spanning even when the *extreme* portfolios are included is far from satisfactory. In 5 of the 7 countries spanning fails in this case (CAN, DEU, GBR, JPN, USA) in at least one of the rankings.

The standard approach for improving the ability of the SDF to span the assets is to augment it with objects of the form $\theta_p Z_{t-1} r_p$, where Z_{t-1} is a lagged conditioning variable or "instrument" (for example *BE/ME*), and θ_p is an additional coefficient. These objects are referred to as "managed portfolios" in some applications. The risk premium of each asset now becomes time varying, $\text{cov}(R_i, m_t) = \sum_p (\beta_p + \theta_p Z_{t-1}) \text{cov}(R_{i,t}, r_{p,t})$. In our application, the conditioning variable Z_t (*BE/ME* or *MVE*) is specific to the portfolios, as we discuss below. Therefore, this formulation of the SDF allows both for the conditional risk premium to depend on the instantaneous value of the instrument and for the average risk premium to depend on its average level.

In the next two subsections we investigate if augmenting the SDF by adding *BE/ME*, *MVE*, and other macro and financial variables in conjunction with *BE/ME* improves its pricing and spanning performance.

4.2. CONDITIONAL SPANNING WITH *be/me* AND *mve*

In this section we perform two sets of tests:

(i) The first test examines the statistical contribution of *BE/ME* and *MVE* to the time-series "fit" of the SDFs when they are conditioned by *BE/ME* and *MVE* as instruments; all the assets are included in the SDF.

(ii) The second test investigates how well these conditional SDFs span or price. This is our *conditional spanning* test.

4.2.1. *Test Specifications*

Let Z_s refer to the type of conditioning variable (s = *BE/ME* or *MVE*); a second subscript in $Z_{s,i}$ indicates that the ith Z_s is specific to the ith portfolio. The conditioning information introduces new variables of the form $Z_{t-1} R_{it}$ to the SDF. We require the pricing errors of each portfolio to be orthogonal only to the own lagged *BE/ME* or *MVE*. Thus these new variables are the products of the portfolio

[29] The probability of obtaining 14 or more rejections in 21 tries by chance is 0.00% at the 5% level. The probability of obtaining 5 or more rejections in 21 tries by chance is 0.33% if the tests are uncorrelated.

specific $Z_{s,i,t-1}$ (alternately *BE/ME* and *MVE*) with the corresponding portfolio returns.[30] The set of the new variables can be represented by the Hadamard product, $R_t \oplus Z_{s,t-1}$, rather than the Kroenecker product, $R_t \otimes Z_{s,t-1}$; the latter would imply that the pricing errors for each portfolio are required to be orthogonal to the *BE/ME* or *MVE* of all the portfolios.

The n-dimensional vectors R_t and $Z_{s,i,t-1}$ are, respectively, returns on n assets and n portfolio-specific lagged conditioning variables. The system that defines the volatility bound for the n assets is,

$$\begin{matrix} R_t - E(R_t \mid m_{c,1}) = \nu_{1,t}, \\ R_t - E(R_t \mid m_{c,2}) = \nu_{2,t}, \end{matrix} ;\ \varepsilon_t = (\nu'_{1,t}\ \nu'_{2,t});\ E(\varepsilon_t) = 0,\ E(\varepsilon_{p,t} \oplus Z_{p,s}) = 0, \tag{3a}$$

where $\varepsilon_{p,t}$ and $Z_{p,s}$ refer to the elements of ε_t and $Z_{s,i,t-1}$ that correspond to the reference assets. The conditional SDF is,

$$m_{c,j,t} = \sum_p (\beta_{p,j} + \theta_{p,j} Z_{s,p,t-1}) r_{p,t} + c_j;\ E(m_{c,j,t}) = c_j,\ j = 1, 2. \tag{3b}$$

where s = *BE/ME*, *MVE*, $r_t = R_t - E(R_t)$, c_j is the pre-specified risk free rate, $m_{c,j,t}$ is the SDF. Note that adding conditioning variables to the SDF is equivalent to modeling time variation in the coefficients of the SDF with the conditioning variables.

We test the importance of *BE/ME* or *MVE* for pricing by testing the null hypothesis that the coefficients θ_i of the unconstrained SDFs – p = 7(9), q = 0 – are *not* statistically significant ($\theta_i = 0$). The test statistic is the Wald test; the statistic is χ^2-distributed where the degrees of freedom are the number of restrictions.[31] The null hypothesis of *conditional spanning* implies that the average pricing error with the restricted SDF is zero for all the assets; this is exactly analogous to the unconditional spanning test in equations (2a–b). The test statistic is the Hansen J-statistic, which is χ^2-distributed.[32]

30 In the standard specification every asset return is required to be orthogonal to all the conditioning variables. In our application this approach would result in 7 or 9 orthogonality conditions for each portfolio, which would produce highly unreliable results. Our specification requires only one orthogonality condition for each portfolio.

31 Conditioning on w $\mathbf{Z}_{t-1}$ variables adds $2nw$ orthogonality conditions for a total of $2n(w+1)$ conditions. $2n$ parameters are estimated, which leaves $2nw$ overidentifying conditions. Since $w = 1$, for SDFs conditioned on either *BE/ME* or *MVE* there are $2n$ overidentifying conditions.

32 Since spanning implies that average pricing errors are zero, only the average pricing errors of the test assets are tested against zero. Therefore, the degrees of freedom are the same as in the unconditional spanning case; they are equal to $2q$, which is always four.

4.2.2. *Results*

In Table V, Panel A we report p-values for Wald tests of the null hypothesis, that the SDF coefficients for *BE/ME* and *MVE* ($\theta_{i,j} = 0 \forall i, j$) are insignificant. Entries significant at the 5% level are shaded. *BE/ME* is not significant in only 1 case out of 21 (Germany – DEU – *MVE*-sort) and *MVE* is not significant in 3 cases out of 21. Clearly, both of these variables significantly improve the overall "fit" of the SDFs, separately. This finding implies that the pricing errors of the unconditional SDF are not orthogonal to *BE/ME* and *MVE*. Alternatively, the hypothesis that the SDF coefficients are not time-varying is rejected. We conclude that modeling nonlinearities in the SDF coefficients with our two instruments significantly improves the SDF's time series fit. This finding is in accord with the literature that documents the importance of *BE/ME* and *MVE* in time series regressions.

However, the improved time series fit does not guarantee that the SDFs augmented by *BE/ME* or *MVE* will span significantly better. Table V, panel B reports the spanning performance of these nonlinear (or conditional) SDFs. It is evident that their spanning performance is *not* materially different from the linear (or unconditional) ones. Overall, when *BE/ME* is the instrument, spanning is rejected for 32 out of 63 cases, and once again there are major differences across test assets. When the *extreme* pairs are excluded from the SDF, spanning fails in 17 out of 21 cases (81%), at the 5% level. But when the *middle* pairs are excluded, spanning fails in only 6 out 21 cases (28%). The results when *MVE* is the instrument are very similar. The nonlinearity of the SDFs does not improve their spanning ability, even though it improves the time series fit. Just as in the unconditional case, when the *extreme BE/ME* and *MVE* portfolios are excluded from the SDF spanning is overwhelmingly rejected.

Our results to this point establish the following:

(i) The *BE/ME* and *MVE* effects exist in all the countries we study, even when we use a very general pricing model. The evidence for this is that spanning overwhelmingly fails, when the *extreme BE/ME* and *MVE* portfolios are excluded from the SDFs and that SDFs conditioned by *BE/ME* and *MVE* "fit" the data significantly better.

(ii) The conditional SDFs do not span better than the unconditional ones. Even unconditional and conditional SDFs that include the "extreme" portfolios do not span all the assets.

These findings leave open the possibility that our instruments (*BE/ME* and *MVE*) proxy for other economic or financial variables that may help span the assets.

4.3. ALTERNATIVE MACRO AND FINANCIAL VARIABLES FOR CONDITIONAL SPANNING

In this section we broaden our investigation of the importance of *BE/ME* by testing 15 alternative variables in the SDF, along with *BE/ME*, to determine, if *BE/ME* simply proxies for some other variable, and, if adding any of these variables can

Table V. Time-varying SDFs

Panel A. joint significance of the *BE/ME* and *MVE* coefficients in the SDF Wald test P-values

Country	Portfolios sorted by	Conditional on *BE/ME*	Conditional on *MVE*
AUS	*BE/ME*	0.000	0.000
CAN		0.000	0.013
DEU		0.006	0.001
FRA		0.000	0.000
GBR		0.000	0.002
JPN		0.001	0.000
USA		0.000	0.501
AUS	*MVE*	0.000	0.000
CAN		0.000	0.000
DEU		0.242	0.002
FRA		0.000	0.270
GBR		0.000	0.002
JPN		0.031	0.000
USA		0.000	0.001
AUS	2 way *MVE* and *BE/ME*	0.000	0.000
CAN		0.006	0.068
DEU		0.005	0.000
FRA		0.000	0.000
GBR		0.000	0.000
JPN		0.005	0.000
USA		0.000	0.002

Panel A reports right hand tail probabilities of the null hypothesis that $\theta_{i,j} = 0 \; \forall \; i, \; j$, jointly, i.e., the coefficients associated with *BE/ME* or *MVE*, θ_is, are statistically significant. Let R_t and $Z_{s,t-1}$ be n-dimensional vectors, respectively, of date t returns on n assets and n portfolio specific conditioning variables, s = *BE/ME* or *MVE*, one at a time. The system that defines the volatility bound for the n assets is,

$$\begin{matrix} R_t - E(R_t \mid m_{c,1}) = \nu_{1,t}, \\ R_t - E(R_t \mid m_{c,2}) = \nu_{2,t}, \end{matrix} \; ; \; \varepsilon_t = (\nu'_{1,t} \; \nu'_{2,t}); \; E(\varepsilon_t) = 0, \; E(\varepsilon_{p,t} \oplus \mathbf{Z}_{p,s}) =$$

where $\varepsilon_{p,t}$ and $\mathbf{Z}_{p,s}$ refer to the elements of ε_t t and $\mathbf{Z}_{s,t-1}$ that correspond to the reference assets and $\oplus$ represents the Hadamard product. The conditional SDF is,

$$m_{c,j,t} = \sum_p (\beta_{p,j} + \theta_{p,j} Z_{s,p,t-1}) r_{p,t} + c_j; \; E(m_{c,j,t}) = c_j, \; j = 1, 2.$$

where s = *BE/ME*, *MVE*, and $r_t = R_t - E(R_t)$. To test the null hypothesis that $(\theta_I = 0)$ in the augmented SDF we use the Wald test; the statistic is χ^2-distributed where the degrees of freedom are the number of overidentifying restrictions $(2n)$. The unconditional system is just-identified, so that in this case p = all the portfolios (7 or 9), and q = 0.

Table V. Continued

Panel B. Spanning Performance of Time-Varying SDFs Instruments = *BE/ME*, *MVE* (J-Statistic Test P-Values)

		Instruments					
		BE/ME			*MVE*		
	Portfolios	Test portfolios			Test portfolios		
Country	Sorted by	1&7	2&6	3&5	1&7	2&6	3&5
AUS	*BE/ME*	0.036	0.542	0.043	0.023	0.370	0.072
CAN		0.000	0.005	0.010	0.000	0.002	0.012
DEU		0.010	0.230	0.119	0.025	0.266	0.234
FRA		0.028	0.803	0.567	0.026	0.597	0.614
GBR		0.000	0.057	0.001	0.000	0.175	0.001
JPN		0.007	0.384	0.076	0.010	0.557	0.136
USA		0.001	0.000	0.659	0.000	0.000	0.857
AUS	*MVE*	0.071	0.440	0.054	0.009	0.296	0.074
CAN		0.006	0.033	0.673	0.007	0.082	0.462
DEU		0.443	0.007	0.002	0.840	0.011	0.015
FRA		0.289	0.618	0.199	0.201	0.312	0.084
GBR		0.001	0.008	0.978	0.006	0.018	0.878
JPN		0.000	0.241	0.509	0.009	0.092	0.598
USA		0.000	0.755	0.149	0.000	0.790	0.213
		Test portfolios			Test portfolios		
		3&7	1&9	2&6	3&7	1&9	2&6
AUS	2 way	0.007	0.123	0.223	0.004	0.297	0.177
CAN	*MVE* and	0.266	0.007	0.080	0.174	0.005	0.048
DEU	*BE/ME*	0.207	0.066	0.425	0.242	0.064	0.466
FRA		0.028	0.898	0.821	0.032	0.729	0.983
GBR		0.001	0.073	0.012	0.039	0.074	0.109
JPN		0.003	0.018	0.054	0.044	0.100	0.230
USA		0.000	0.001	0.120	0.000	0.003	0.184

Panel B reports right hand tail probabilities from Hansen J-tests of conditional spanning, using *BE/ME* and *MVE* as conditioning information. Different pairs of assets which are designated test assets q and the remainder are designated reference assets p; $n = q + p$. SDFs formed based only the reference assets (and the conditioning information) are asked to span (price) the test assets, within statistical tolerance. The SDF is given by equations (3a) and (3b). The null hypothesis of *conditional spanning* implies that the average pricing errors of the restricted SDF are zero for all the assets; this is exactly analogous to the unconditional spanning test in Equations (2a–b). We use the Hansen J-statistic, which is χ^2-distributed and where the degrees of freedom are identical to those in the unconditional spanning case, and they are equal to $2q$. Seven *BE/ME*- and *MVE*-sorted portfolios from each country, and 9 portfolios from 2-way sorts (by *MVE* and then by *BE/ME*) are used for the spanning tests. The construction of *BE/ME*, *MVE*, and the portfolios is described in the Data section. Three tests are reported for each set of sorts for each country, each for a different pair of test assets.

The GMM estimates are robust to first-order autocorrelation and heteroskedasticity. The sample period is 11/83 – 10/94 (T = 132) in the USA and CAN, and 11/86 – 10/94 (T = 96) in AUS, DEU, FRA, GBR, and JPN. Shaded entries are significant at the 5% level.

improve spanning performance. We choose to focus only on *BE/ME* both because *BE/ME* performs somewhat better than *MVE*, and because the range of *MVE* for some of the countries is quite narrow.

We use the 1st lags of the following 15 macro and financial variables as additional instruments in the SDF:[33]

r_g	MSCI Global market return (in local currency units),
Gpgold	Gold price growth rate, in local currency units,
i_{euro}	The Eurodollar rate converted to the local currency,
TB1M	The 1-month U.S. T-Bill rate, in local currency,
TPRE	The U.S. "term" premium, $TB30Y - TB3M$, in local currency,
JPRE	The U.S. "junk bond" premium, $BAA - AAA$ rates, in local currency,
WDIVPRE	The global dividend-to-price ratio,
WGINDP	Growth rate of the global industrial production index,
WINFL	The global inflation index,
LSTB	The local short term interest rate,
LTPRE	The local "term" premium, long-term rate – *LSTB*,
LDIVPRE	The local dividend-to-price ratio,
LGINDP	Growth rate of the local industrial production,
GLM2	Growth rate of the local money supply, *M2*,
LINFL	The local inflation index, *CPI*.

4.3.1. *Test Specification*

As before, R_t and $Z_{s,i,t-1}$ are n-dimensional vectors, respectively, of returns on n assets and n portfolio specific *BE/ME*s. Let $V_{k,t-1}$ be the kth instrument (of 15 listed above) to be considered. We require the pricing errors of each portfolio to be orthogonal to its own *BE/ME* and to V_k. The system that defines the volatility bound for the n assets in this case is,

$$\begin{aligned} R_t - E(R_t \mid m_{c,1}) &= \nu_{1,t} \\ R_t - E(R_t \mid m_{c,2}) &= \nu_{2,t} \end{aligned}, \tag{4a}$$

where, s, r_t, c_j, and $m_{c,j,t}$, are as defined for equations (3a, b), $\eta_t = (\nu'_{1,t}\ \nu'_{2,t})'$. $E(\eta_t) = 0$, $E(\eta_{p,t} \oplus Z_{p,s})$, and $E(\eta_{p,t} \otimes V_k) = 0$, where $Z_{p,t}$ and $Z_{p,s}$ refer to the elements of η_t and $Z_{s,i,t-1}$ that correspond to the reference assets. The corresponding conditional SDF is,

$$m_{c,j,t} = \sum_p (\beta_{p,j} + \theta_{p,j} Z_{s,p,t-1} + \lambda_{k,t-1}) r_{p,t} + c_j;\ E(m_{c_j,t}) = c_j,\ j = 1, 2. \tag{4b}$$

[33] This list contains most of the important variables that have been used by various studies to build empirical asset pricing models. In almost all cases there are counterparts to the U.S. and or the global variables. For example, the counterpart of *TB3M* would be *LSTB*, etc. However, there are no available counterparts to the *AAA* and *BAA* rate in foreign countries. Therefore, only the US junk bond premium is used.

We perform three tests:

(i) A test of the null hypothesis that none of these new variables is significant, i.e, $\lambda_{i,j} = 0, \forall\, i,\, j$. If for the kth V_k the null hypothesis ($\lambda_{i,j} = 0\ \forall\, i,\, j$) is rejected in the presence of the *BE/ME*s we may conclude that the kth Vk enhances the time-series "fit" of the SDF. In this case $p = 7(9)$, $q = 0$.

(ii) A test of the null hypothesis that *BE/ME* proxies for one of the new variables, i.e., $\theta_{i,j} = 0, \forall\, i,\, j$. This is a test of the robustness of *BE/ME*. If the null hypothesis ($\theta_{i,j} = 0\ \forall\, i,\, j$) is rejected in the presence of the kth V_k, we may conclude that the kth V_k does not replace or supercede *BE/ME* as an instrument, and that *BE/ME* is robust with respect to the kth V_k in the SDF. The test statistic is the Wald test, where each set of coefficients ($\lambda_{i,j}$ and $\theta_{i,j}$) are tested jointly against the null hypothesis that they are zero. Here again, $p = 7(9)$, $q = 0$.

(iii) A *conditional spanning* test of the reference-asset SDFs conditioned by *BE/ME* and the new variables, one of at a time. This test is of the same form as in section 4.2a above, except that the time-variation of the SDF coefficients is modeled with two variables. The statistic is χ^2-distributed where the degrees of freedom are the number of overidentifying restrictions.[34] *Conditional spanning* implies that average returns for all the assets are correctly priced with this restricted SDF. In this case $p = 5(7)$, $q = 2$. The test statistic is the J-Statistic, which is χ^2-distributed.

4.3.2. *Results*

First we explore whether any of the 15 instruments add significantly to the time-series pricing accuracy of the SDFs when tested one at a time in the presence of *BE/ME*. We do not tabulate the results here because of space considerations; detailed results are available from the authors. For each country, at least 10 of the instruments are significant for each portfolio ranking. On average, 12 of the variables are significant for some portfolio rankings, and in some cases all 15 are significant. We define one "try" as 1 portfolio ranking and 1 instrument. There are 3 portfolio rankings, 3 sets of test assets, and 15 instruments, for a total of 135 tries for each country. The minimum proportion of significant tries (the instrument is significant at the 5% level) is 73% for Canada, France and Great Britain, and the maximum is 91% for the US.

No instrument is significant for all countries and all portfolio rankings, at the 5% level. The local short term interest rate, the growth rate of local Industrial Production, and local Inflation (*LSTB*, *GLINDP*, *LINFL*), are significant in at least 90% of the 21 tries (1 instrument, 7 countries, 3 portfolio rankings each). The least successful instrument is the global Industrial Production index, WGINDP, significant in only 57% of the tries.

[34] The $\mathbf{Z}_{i,t-1}$ and $V_{k,t-1}$ variables each adds $2n$ orthogonality conditions. Thus there are a total of $4n$ overidentifying conditions.

Clearly these instruments, as a group, significantly improve the time series "fit" of the SDFs, for all the countries. Thus, they are legitimate candidates to consider as alternatives to *BE/ME*.

Next we investigate the robustness of *BE/ME* in the presence of these alternative instruments, i.e., whether some of these instruments *displace* it in the SDF. When the 15 instruments are added to the SDFs one at a time, *BE/ME* is not significant only 11% of time, at the 5% significance level. Almost half of the insignificant *BE/ME* coefficients come from the DEU (Germany) *MVE*-sorted portfolios.[35] But this is also the only instance in which *BE/ME* is not significant, even by itself. The *BE/ME* ratio is not replaced in the SDFs by these alternative instruments. Their contributions to pricing are independent and complementary to that of *BE/ME*.

Finally, we investigate the *conditional spanning* hypothesis, whether augmenting the SDFs by *BE/ME* and by these 15 instruments one at a time, improves spanning performance. It is not feasible to present *all* 945 test results. Overall, the SDFs again rarely span when the *extreme* portfolios are the test assets. Spanning is much more common when the *extreme* pair of portfolios are included in the reference assets and the middle pair of portfolios are the test assets.[36] In Table VI we summarize the performance of the "best" instruments along with their marginal contribution to spanning, over *BE/ME* alone.

To get a better understanding of the spanning behavior of the SDFs in the presence of these instruments, we report two best-scenario cases. The left-hand columns report the spanning performance of *WDIVPRE*, the best-performing instrument across all the countries. The right-hand columns report the spanning performance of the best instrument for each country; best means that it spans most frequently across the various tries in that country. To make clear the additional contribution of these instruments we use the following notation in the table. We use "***" where spanning fails both with *BE/ME* and with *BE/ME* plus the instrument; we place "a.s." where spanning was achieved already with *BE/ME* (from Table V, Panel B). Finally, the numbers in the table are p-values for the spanning hypothesis, and they appear only where spanning fails with *BE/ME* alone but it is achieved when the instrument is added to the SDF.

The most striking result is that these instruments make at best a minor contribution to spanning. The best instruments do not help span the *extreme* portfolios. In 14 of 42 instances spanning fails for the *middle* and the *in-between* portfolios with *BE/ME* alone as an instrument; spanning is achieved in only 4 additional cases with *WDIVPRE*.

We conclude that:

(i) Several macroeconomic variables improve the time series fit of the SDFs.
(ii) The macroeconomic variables we tried do not proxy for *BE/ME*, and therefore do not displace it as an explanatory variable.

[35] In this sort, *BE/ME* has a particularly narrow range; see Table 1, Panel B.

[36] The *MVE*-sort in Germany is the exception. In that case, all the instruments help span the *extreme* portfolios and perversely, SDFs that include the *extreme* portfolios fail to span.

Table VI. Spanning performance of the best performing instruments (The SDF includes *BE/ME* and the 15 instruments, one at a time)

			Instruments: *BE/ME* + one variable						
	Portfolios		Test portfolios				Test portfolios		
Country	sorted by	Variable	1&7	2&6	3&5	Variable	1&7	2&6	3&5
AUS	*BE/ME*	*WDIVPRE*	***	a.s.	0.075	*GLINDP*	***	a.s.	a.s.
CAN		*WDIVPRE*	***	***	***	*LINFL*	***	***	***
DEU		*WDIVPRE*	***	a.s.	a.s.	*GLIINDP*	***	a.s.	a.s.
FRA		*WDIVPRE*	***	a.s.	a.s.	*GLINDP*	***	a.s.	a.s.
GBR		*WDIVPRE*	***	a.s.	***	*GLM2*	***	a.s.	***
JPN		*WDIVPRE*	***	a.s.	0.141	*LSTB*	***	a.s.	a.s.
USA		*WDIVPRE*	***	***	a.s.	*GLINDP*	***	***	a.s.
AUS	*MVE*	*Same*	a.s.	a.s.	a.s.	*Same*	0.211	a.s.	a.s.
CAN			***	0.082	a.s.		***	a.s.	a.s.
DEU			a.s.	***	***		a.s.	***	***
FRA			a.s.	a.s.	a.s.		a.s.	a.s.	a.s.
GBR			***	***	a.s.		***	0.126	a.s.
JPN			***	a.s.	a.s.		***	a.s.	a.s.
USA			***	a.s.	a.s.		***	a.s.	a.s.
			Test portfolios				Test portfolios		
			3&7	1&9	2&6		3&7	1&9	2&6
AUS	2 way	*Same*	***	a.s.	a.s.	*Same*	***	a.s.	a.s.
CAN	*MVE* and		a.s.	***	a.s.		a.s.	0.373	0.113
DEU	*BE/ME*		a.s.	a.s.	a.s.		a.s.	a.s.	a.s.
FRA			***	a.s.	a.s.		***	a.s.	a.s.
GBR			***	a.s.	***		***	a.s.	a.s.
JPN			***	0.086	a.s.		***	a.s.	a.s.
USA			***	***	a.s.		***	***	a.s.

The table reports right hand tail probabilities for the J-Statistics of the spanning test detailed below.

The columns on the left report the spanning performance of *WDIVPRE* together with *BE/ME*. *WDIVPRE* is the best-performing instrument across countries and for all the types of SDFs and portfolio sorts. The last 4 columns report the spanning performance of the "best" instrument for each country, when added to *BE/ME*. For these columns, when another instrument performs as well as *WDIVPRE* we report it rather than *WDIVPRE*. We choose the best performing instrument by counting the number of instances in which each instrument enables the associated SDF to span the test assets, only when the test assets are the in-between pair of portfolios (2&6/1&9) or the middle pair of portfolios (3&5/2&6). We do not count the spanning performance for the extreme pair (1&7/3&7) because spanning is rare in those cases.

To enhance exposition, we report p-values *only* for the cells where spanning fails with only *BE/ME*, but is achieved when the instrument is added to the SDF. Where spanning fails both with *BE/ME* and with *BE/ME* and the instrument, we report "***". Where spanning was achieved already with *BE/ME* we report "a.s."

The spanning test is very similar to that in Table V and is specified as follows. Let R_t and Z_{t-1} be n-dimensional vectors, respectively, of date t returns on n assets and n portfolio specific *BE/ME*s. Let $V_{k,t-1}$ be the kth of the 15 instruments to be considered. We require the pricing errors of each portfolio to be orthogonal to their own *BE/ME* and to V_k. Let Z_i = *BE/ME*$_i$ of the various portfolios, and $r_t = R_t - E(R_t)$, $\oplus$ and $\otimes$ represent the Hadamard and Kroenecker products respectively, and $\varepsilon_{p,t}$ refers to the elements of η_t that correspond only to the reference assets. The system that defines the volatility bound for the n assets is,

$$\begin{array}{l} R_t - E(R_t \mid m_{c,1}) = \nu_{1,t} \\ R_t - E(R_t \mid m_{c,2}) = \nu_{2,t} \end{array},$$

where $\eta_{p,t}$ and $\mathbf{Z}_{p,s}$ refer to the elements of η_t and $\mathbf{Z}_{s,t-1}$ that correspond to the reference assets. The corresponding conditional SDF is,

$$m_{c,j,t} = \sum_p (\beta_{p,j} + \theta_{p,j} Z_{s,p,t-1} + \lambda_{k,t-1}) r_{p,t} + c_j;\ E(m_{c_j,t}) = c_j,\ j = 1, 2.$$

The null hypothesis of *conditional spanning* implies that the average pricing error with the restricted SDF is zero for all the assets; this is exactly analogous to the unconditional spanning test in equations (2a–b). The test statistic is the Hansen J-statistic, which is χ^2 – where the degrees of freedom is $2q$, the number of test assets. In this case $p = 5(7)$, $q = 2$.

Seven *BE/ME*- and *MVE*-sorted portfolios from each country, and 9 portfolios from 2-way sortings (by *MVE* and then by *BE/ME*) are used for the spanning tests. The construction of *BE/ME*, *MVE*, and the portfolios is described in the Data section. Three tests are reported for each ranking for each country, each for a different set of test assets.

The GMM estimates are robust to first-order autocorrelation and heteroskedasticity. The sample period is 11/83 – 10/94 (T = 132) in the USA and CAN, and 11/86 – 10/94 (T = 96) in AUS, DEU, FRA, GBR, and JPN.

(iii) These macroeconomic variables make a very marginal contribution to explaining the cross-sectional variation of average returns.

5. Conclusion

A lot of evidence suggests average returns are high when Book Equity to Market Equity ratio (*BE/ME*) is high and when Market Value (*MVE*) is low, internationally. However, most tests in the literature have two serious weaknesses: they are based on a specific asset-pricing model, usually the CAPM, and the estimates suffer from serious simultaneity biases when OLS methods are employed.

We discuss and implement a set of spanning tests based on the Hansen-Jagannathan Stochastic Discount Factor (SDF) approach, which only requires that the Law of One Price is satisfied. These tests suffer neither from the need to specify an asset pricing model nor from simultaneity biases. We perform unconditional tests as well as tests that condition on macroeconomic and financial variables.

Consistent with recent literature, we conclude that the *BE/ME* and *MVE* "effects" are international in character. They remain strong under our general model and against a variety of alternative macroeconomic and financial conditioning variables. These variables do not substitute for *BE/ME* but they do improve the time-series fit of the pricing model significantly. At the same time, none of these additional variables diminish the explanatory power of the lags of *BE/ME* and *MVE* in the SDFs.

Contrary to the findings in the literature, we cannot account completely for the cross-sectional return variation of the *BE/ME*- and *MVE*-sorted portfolios with any of the SDFs we construct, i.e., the assets are *not* universally spanned by our SDFs. Including portfolios with extreme values of *BE/ME* or *MVE* in the SDFs is equivalent to introducing the *HML* portfolio, without the simultaneity problems. Even though such SDFs span much more frequently than the others, they do not span fully and reliably. This continues to be true when we augment the SDFs with macroeconomic and financial variables.

The evidence presented here suggests that the *BE/ME* and *MVE* effects are not artifacts of the inadequacies of the augmented CAPM as an asset pricing model or of omitting potentially important conditioning information from macro and financial variables. Finally, our finding, that we cannot reliably find spanning SDFs suggests that the underlying "factors" that summarize the return-generating process are not present in all the sorted portfolios we have constructed, even though these portfolios are well diversified in other dimensions.

Appendix A. Stochastic Discount Factors

Hansen and Jagannathan (HJ, 1991) derive restrictions from the Law of One Price (LOP) that must apply to any stochastic discount factor (SDF) that can price a set of assets. The LOP articulates a minimal requirement for market efficiency, that assets with the same payoff structure sell for the same price. HJ show how to derive a lower bound on the volatility of SDFs that will price any specified set of assets. These SDFs capture all the information available from linear combinations of asset returns to price average returns exactly.[37] There is a single minimum-variance SDF that will price any set of assets for each possible riskless rate, r_f. Unconditional and conditional versions of these SDFs can be used to construct "spanning" tests that do not depend on any specific asset pricing model or distributional assumptions. Below we discuss the construction and some important properties of SDFs.

The basic requirement of an asset pricing model is that prices correctly reflect expected discounted payoffs. Stated formally,

$$E[\mathbf{X}_{t+1} m_{t+1}] = P_t, \tag{A.1a}$$

or

$$E[\mathbf{R}_{t+1} m_{t+1}] = 1_n. \tag{A.1b}$$

Here X_{t+1} and R_{t+1} are $n \times \omega$ matrices of payoffs and returns of the n assets in each of ω possible states at date $t+1$, P_t is the $n \times 1$ vector of the corresponding prices at date t, m_{t+1} is the $\omega \times 1$ vector of the SDF, and 1_n is the $n \times 1$ unity vector. The Law of One Price (LOP), gives the weakest conditions on m;[38] it merely requires that m exist.[39] This requirement is met if the second moment matrix of payoffs $E[RR']$ is of full rank.

Under the assumption that the asset returns distributions are ergodic, one can treat every observation as a realization of a state ω. This makes it possible to use

[37] The SDF is closely related to state prices; its existence does not depend on market completeness.

[38] Various asset pricing models can be derived by placing alternative restrictions on the discount factor m. For example, the no-arbitrage restriction implies a positive discount factor, $m > 0$ (no negative prices on assets with non-negative payoffs). Expected utility models interpret m as the intertemporal marginal rate of substitution (IMRS), which requires m to be positive, and to be related to consumption through some utility function. HJ (1991) provide a method to guarantee $m > 0$. Stutzer (1993) also provides a functional form that guarantees $m > 0$ consistent with CRRA utility.

[39] This requirement rules out profits from re-packaging assets, and it results in prices that are linear in their payoffs, because payoffs are independent of the initial investment. See Ingersoll (1987) and Ross (1976).

time series data of returns to evaluate the expectations in (1b) and to calculate the SDF (Hansen and Richard, 1987). To compute the SDF, m_c, expand Equation (1b) to get,

$$E[\mathsf{R}m_c] = E[\mathsf{R}]E[m_c] + \mathrm{Cov}[\mathsf{R}m_c] = 1_n,$$

or

$$E[\mathsf{R}] = \frac{1_n - \mathrm{Cov}[\mathsf{R}m_c]}{E[m_c]}. \tag{A.2}$$

where m is now a $T \times 1$ random variable m_c, and R is a $n \times T$ matrix of T observations on n assets. The expected SDF, $E[m_c] = c = 1/r_f$, is the price of a riskless asset.[40] Expanding the covariance and substituting in (2) gives, $E[(\mathsf{R} - E(\mathsf{R}))(m_c - c)] = 1_n - E(\mathsf{R})c$, where the SDF is a linear function of returns, and, $m_c = [\mathsf{R} - E(\mathsf{R})]'\beta_c + c$, where β_c is an $n \times 1$ vector of weights for each asset. The solution of *beta* is given by, $\beta_c = \Sigma^{-1}[1_n - E(\mathsf{R})c]$, where Σ is the covariance matrix of gross returns. The SDF, m_c, replicates *exactly* the sample averages of the included asset returns for any specific risk free rate, r_f. The SDFs over all possible risk-free rates, r_f, form the lower bound on the volatility of all possible discount factors.[41]

Appendix B. Illustration of the GMM Estimation Methodology

The parameters of the HJ model are estimated with Hansen's (1982) Generalized Method of Moments (GMM). The GMM objective is to find a set of parameters that obey a set of sample orthogonality conditions as closely as possible. The model we label unconditional spanning in equations (3a, b) with conditioning information is used here for illustration. First, form the vector of sample moments from the orthogonality conditions in Equation (10a):[42] $E(\varepsilon_t) = E(\varepsilon_t \oplus Z_{i,t-1}) = 0$ give,

$$h_T(b_{p,1}, b_{p,2}, g_{p,1}, g_{p,2}) = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} \varepsilon_t \otimes 1 \\ \varepsilon_t \oplus Z_{s,t-1} \end{bmatrix}, \tag{B.1}$$

[40] The definition of a risk free asset is that it has a unit payoff regardless of the state of the world. Substituting a unit payoff vector for the payoff matrix (**X**) into Equation (1a) and taking expectations $E[1m_C] = p_C = E[m_C] = p_C = c$, and $1/E[m] = 1/p_C = r_f$.

[41] The variance of m_C is given by, $\sigma^2(m_c) = \beta_c'\Sigma\beta_c = \mathrm{Cov}[\mathsf{R}m]'\beta_c$. The minimum variance property of this SDF, and its construction from the first two moments of returns data, gives a duality between the Hansen and Jagannathan bound and the Markowitz mean variance frontier. The duality between the volatility bound and the mean variance frontier can be easily illustrated by noticing that the portfolio that composes m_C will have the maximal Sharpe Ratio (i.e., be tangent to the mean variance frontier with intercept equal to the risk free rate) of any candidate SDF (HJ, 1991; De Santis, 1993; Stutzer, 1995).

[42] The Kronecker product $\otimes$ (i.e., every element in matrix A times each element in matrix B) is usually used in the literature. We use the Hadamard product $\oplus$ (i.e., every element of a row k in matrix A with every element of column k of matrix **B**) because each of our instruments is portfolio specific.

where h_T is a $2(n+w)$ vector of sample moments (i.e., of time-series averages) that dependent on the estimates $b_{p,1}$, $b_{p,2}$, $g_{p,1}$ and $g_{p,2}$ of the SDF population parameters $\beta_{p,1}$, $\beta_{p,2}$, $\theta_{p,1}$, and $\theta_{p,2}$. The objective is to find a set of parameters that minimize a generic least square quadratic form of sample moments,

$$\min_{b_{p,1},b_{p,2},g_{p,1},g_{p,2}} = h_T(b_{p,1}, b_{p,2}, g_{p,1}, g_{p,2})'\mathbf{w}_T h_T(b_{p,1}, b_{p,2}, g_{p,1}, g_{p,2}), \quad \text{(B.2)}$$

where $\mathbf{w}_T$ is a $\{2(n+w) \times 2(n+w)\}$ symmetric weighting matrix that is the inverse of a consistent estimate of the covariance matrix of orthogonality conditions. Hansen describes the form of $\mathbf{w}_T$ that guarantees estimates to be consistent and asymptotically normal. Various choices of this weighting matrix make GMM estimates robust to autocorrelation and heteroskedasticity.[43] A consistent estimate of $\mathbf{w}_T$ in the presence of heteroskedasticity is:

$$\mathbf{w}_T = \left\{ \frac{1}{T} \sum_{t=1}^{T} \left[\begin{pmatrix} \varepsilon_t \otimes 1 \\ \varepsilon_t \oplus Z_{t-1} \end{pmatrix} \begin{pmatrix} \varepsilon_t \otimes 1 \\ \varepsilon_t \oplus Z_{s,t-1} \end{pmatrix}' \right] \right\}^{-1}. \quad \text{(B.3)}$$

The minimized value of the objective function Γ^* provides a test of the $2(q+w)$[44] overidentifying restrictions imposed by spanning in Equations (6). Hansen's J Statistic, $J = T^* \sim \chi^2_{2(q+w)}$, is distributed chi-square with degrees of freedom equal to the number of overidentifying conditions. The J-statistic tests how well the orthogonality conditions are met in the sample.

Numerically, GMM estimation proceeds in stages because the objective is a complicated function of the parameters. In the first step, a set of parameters are estimated $(b_{p,1}, b_{p,2}, g_{p,1}, g_{p,2})$ given an identity matrix for $\mathbf{w}_T$. In each of the successive iterations, a new estimate of $\mathbf{w}_T$ is computed from the parameters of the last iteration, and new parameters are estimated in turn. This process need only continue for two iterations (*two-stage* GMM), to obtain a consistent estimate for $\mathbf{w}_T$. However, additional iterations or iterating until a convergence criterion is met, (*iterative* GMM), may improve the accuracy of small sample estimates.[45]

[43] The form of $\mathbf{w}_T$ without corrections is equivalent to 3SLS if there are overidentifying conditions, and GLS if there are no overidentifying conditions. Refer to Newey and West (1987) for optimal choices of $\mathbf{w}_T$ in the presence of autocorrelation and heteroskedasticity that also ensure this matrix is positive definite. GMM estimates of this paper are made robust to heteroskedasticity and first order autocorrelation using a Bartlett spectral density kernel to maintain positive definiteness [reviewed by Andrews (1991)].

[44] The number of restrictions tested is always 2 times the number of test assets ($2q = 4$). The portfolio-specific instrument w on the test assets is not tested for orthogonality with pricing errors because we seek only to test the average pricing error not the conditional average pricing error.

[45] Successive approximation methods, like the one employed here, do not guarantee convergence to a fixed point. The results reported here are obtained using iterative GMM.

Ferson and Foerster (1994) provide Monte Carlo evidence that iterative GMM is more reliable in small samples, especially when estimating larger systems.[46] They find two-stage GMM tends to reject overidentifying conditions (i.e., the model) too often, with this bias becoming more pronounced in larger systems. Iterated GMM, while more reliable, has a slight tendency to reject less frequently than it should. Surprisingly, the small sample J-statistic from a simple model with few instruments conforms well to its asymptotic distribution with as few as 60 time series observations (using either two-stage or iterative GMM).[47]

Acknowledgements

We want to thank Bruno Gerard, Tom Goodwin, Alex Huemer, Deborah Sherwood, Craig Stubblebine, Vince Warther, Tom Willett, Paul Zak, participants of the Claremont Graduate School economics seminar and the University of Southern California finance seminar for helpful comments and suggestions. We owe special thanks to Giorgio De Santis for his unstinting help and boundless patience.

References

Andrews, W. K. (1991), 'Heteroskedasticity and autocorrelation consistent covariance matrix estimation', *Econometrica* **59**, 817–858.

Arellano, M. and Bond, S. (1991), 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations', *Review of Economic Studies* **58**, 277–297.

Banz, R. W. (1981), 'The relationship between return and market value of common stocks', *Journal of Financial Economics* **9**, 3–18.

Basu, S. (1977), 'Investment performance of common stocks in relation to their price earnings ratios: A test of the efficient market hypothesis', *Journal of Finance* **32**, 663–682.

Basu, S. (1983), 'The relationship between earnings yield, market value and return for NYSE common stocks: Further evidence', *Journal of Financial Economics* **12**, 129–156.

Breen, W. J. and Korajczyk, R. A. (1993), 'On the selection biases in book-to-market based tests of asset pricing models', Working Paper #167, Northwestern University.

Campbell, J. Y. (1987), 'Stock returns and the term structure', *Journal of Financial Economics* **18**, 373–399.

Campbell, J. Y., Lo, A. W. and MacKinlay, C. (1997), *The econometrics of financial markets*, Princeton University Press, Princeton New Jersey.

Capaul, C., Rowley, I. and Sharpe, W. (1993), 'International value and growth stock returns', *Financial Analyst Journal* **41**, 27–36.

Chan, K.C. and Chen, N. (1991), 'Structural and return characteristics of small and large firms', *Journal of Finance* **46**, 1467–1484.

---

[46] The authors use a single-latent-variables model to conduct their simulations. They caution that the results of Monte Carlo studies are usually sensitive to the model and the artificial data used. Iterative GMM is also preferable in the presence of heteroskedasticity (Arellano and Bond, 1991).

[47] The smallest model had three instruments (including a constant) and two test assets (df = 4). Parameter estimates and standard errors also seem unbiased. As model complexity increases, parameter estimates become unreliable and standard errors are biased downward in small samples. However, the J-statistic from iterative GMM is generally well-behaved.

Chan, L. C., Hamao, Y. and Lakonishok, J. (1991), 'Fundamentals and stock returns in Japan', *Journal of Finance* **46**, 1739–1764.

Cochrane, J. H. (2001), *Asset pricing*, Princeton University Press, Princeton New Jersey.

Daniel, K. and Tittman, S. (1997), 'Evidence on the characteristics of cross-sectional variation in stock returns', *Journal of Finance* **52**, 1–34.

De Roon, F. A. and Nijman, T. (2001), 'Testing for mean-variance spanning: A survey', *Journal of Empirical Finance* **8**, 111–115.

De Santis, G. (1994), 'Asset pricing and portfolio diversification: evidence from emerging financial markets', *in* Mike Howell ed., *Investing in Emerging Markets*, Euromoney Books.

De Santis, G. (1995), 'Volatility bounds for stochastic discount factors: Tests and implications from international stock returns', Working paper, University of Southern California.

Fama, E. F. and French, K. (1992), 'The cross-section of expected stock returns', *Journal of Finance* **47**, 427–465.

Fama, E. F. and French, K. (1995), 'Size and book-to-market factors in earnings and returns', *Journal of Finance* **50**, 131–155.

Fama, E. F. and French, K. (1998), 'Value versus growth: The international evidence', *Journal of Finance* **53**, 1975–1999.

Ferguson, M. and Shockley, R. (1999), 'Equilibrium Anomalies', Working Paper, Kelley School of Business Indiana University.

Ferson, W. E., Foerster, S. R. and Keim, D. B. (1993), 'General tests of latent variable models and mean-variance spanning', *Journal of Finance* **48**, 131–156.

Ferson, W. E. and Foerster, S. R. (1994), 'Finite sample properties of the Generalized Method of Moments in tests of conditional asset pricing models', *Journal of Financial Economics* **36**, 29–55.

Ferson, W. E. and Harvey, C. R. (1993), 'The risk and predictability of international equity returns', *Review of Financial Studies* **6**, 527–566.

Ferson, W. E., Sarkissian S., and Simin, T. (1999), 'The Alpha factor asset pricing model: A parable', *Journal of Financial Markets* **2**, 49–68.

Hansen, L. P. (1982), 'Large sample properties of generalized method of moments estimators', *Econometrica* **50**, 1029–1054.

Hansen, L. P. and Jagannathan, R. (1991), 'Implications of security market data for models of dynamic economies', *Journal of Political Economy* **99**, 225–262.

Hansen, L. P. and Singleton, K. J. (1982), 'Generalized instrumental variables estimation of nonlinear rational expectation models', *Econometrica* **50**, 1269–1286.

Hansen, L. P. and Richard, S.F. (1987), 'The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models', *Econometrica* **55**, 587–613.

Harvey, C. R. (1989), 'Time-varying conditional covariances in tests of asset pricing models', *Journal of Financial Economics* **24**, 289–317.

Harvey, C. R. (1991), 'The world price of covariance risk', *Journal of Finance* **46**, 111–158.

Heston, S., Rouwenhorst, G. and Wessels, R. (1999), 'The role of beta and size in the cross-section of European stock returns', *European Financial Management* **4**, 4–28.

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

Ikenberry, D., Lakonishok, J. and Vermaelen, T. (1995), 'Market underreaction to open market repurchases', *Journal of Financial Economics* **39**, 181–208.

Jagannathan, R. and Wang, Z. (1993), *The CAPM is alive and well*, Staff Report #165, Federal Reserve Bank of Minneapolis.

Keim, D. B. (1983), 'Size-related anomalies and stock return seasonality', *Journal of Financial Economics* **12**, 13–32.

Kothari, S. P., Shanken, J. and Sloan, R. G. (1995), 'Another look at the cross-section of expected stock returns', *Journal of Finance* **50**, 1541–1578.

Loughran, T. and Ritter, J. (1995), 'The new issues puzzle', *Journal of Finance* **50**, 23–51.

MacKinlay, C. A., and Richardson, M. (1991), 'Using the generalized method of moments to test mean-variance efficiency', *Journal of Finance* **46**, 511–528.

Newey, W. K. and West, K. D. (1987), 'A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica* **55**, 703–708.

Reiganum, M.R., (1981), 'Misspecification of capital asset pricing: Empirical anomalies based on earnings yields and market values', *Journal of Financial Economics* **9**, 19–46

Rouwenhorst, K. G. (1999), 'Local return factors and turnover in emerging stock markets', *Journal of Finance* **54**, 1439–1464.

Shanken, J. (1990), 'Intertemporal asset pricing', *Journal of Econometrics* **45**, 99–120.

Spiess, K. D. and Affleck-Graves, J. (1995), 'Underperformance in long-run stock returns following seasoned equity offerings', *Journal of Financial Economics* **38**, 243–267.

Stattman, D. (1980), 'Book values and expected stock returns', *The Chicago MBA: A Journal of Selected Papers* **4**, 25–45.

Stutzer, M. (1995), 'A Bayesian approach to diagnosis of asset pricing models', *Journal of Econometrics* **68**, 367–398.