RESEARCH

Mean Reversion in the Dimensions of Expected Stock Returns

November 2014

James L. Davis, PhD
Vice President
RESEARCH

This study looks for evidence of mean reversion in the equity, profitability, size, and value premiums. Regressions test for statistical evidence of mean reversion, and trading simulations examine whether mean reversion in historical premiums was strong enough to permit profitable trading strategies. Evidence of mean reversion is weak, and 780 simulated trading strategies show very limited evidence of reliably positive abnormal returns.[1]

INTRODUCTION

Research in the world's financial markets has documented the existence of equity, size, value, and profitability premiums in the historical sample of stock returns. Since these premiums have been persistent across long time periods and pervasive across various markets, they are sometimes called dimensions of expected stock returns. While the evidence indicates that the averages for these premiums have been reliably positive, their annual realizations have varied substantially. This variation naturally leads some to wonder if the expected values of these premiums are constant over time. In particular, some have speculated that there may be mean reversion in the time series of the premiums, so that high premiums tend to be followed by low premiums and vice versa.

In the case of the equity premium, the discussion of mean reversion has moved beyond mere speculation. Numerous studies have looked for evidence of mean reversion in the equity premium

1. Simulated data has limitations. See "Important Disclosures" in the Appendix.

with mixed results. Other than a study of the value premium by Davis (2008), mean reversion in the other dimensions has received less attention. The objective of the current study is to fill this gap by looking for evidence of mean reversion in all four of these dimensions of expected stock returns.

There are at least two consequences of mean reversion in a return series. First, returns measured over long horizons are not as variable as they would be in the absence of mean reversion. Second, future values of the return are at least partially predictable. If high returns tend to be followed by low returns (and vice versa), investors can learn something about likely future returns simply by looking at past returns. However, it is not clear over what time horizon the mean reversion should be observed, nor is it clear what mean a premium should revert to, nor is it clear how strong the predictability in returns should be. If the predictability is weak, it has little practical relevance for investors.

This study focuses on the second consequence of mean reversion—predictability. Time series regressions look for statistical evidence of predictability, and trading simulations examine whether predictability was strong enough to generate reliable excess returns. The analysis looks at premiums in the US and 14 other markets.

Based on the analysis of Samuelson (1965), it has been a common assumption that mean reversion in returns is a violation of market efficiency. However, later work on consumption-based asset pricing (e.g., Lucas 1978) shows that serial correlation in returns can be present in an efficient market. Expected returns can vary over time in a predictable way. Mean reversion in returns therefore says nothing about the rationality of stock price formation.

The evidence of mean reversion found in this study is quite weak. While the presence of mean reversion in the historical sample cannot be ruled out, there is minimal evidence that it has been strong enough to permit profitable trading strategies. The proportion of simulations generating reliably positive excess returns was similar to what one would expect by random chance. These results indicate that trying to predict future premiums in equity markets using past premiums is not likely to be a fruitful endeavor.

RESEARCH DESIGN

This study includes three different kinds of tests—moving average regressions, moving average simulations, and runs-based simulations. Each test is described below. While the regressions simply look for a statistical relation between each premium and its own lagged moving average, the simulations examine whether mean reversion was strong enough to generate reliably positive excess returns. Thus, the simulations contain additional information not captured by the regressions.

Moving Average Regressions

These regressions look for evidence of a statistical relation between each premium and its own lagged moving average. Since it is not clear over what time horizon mean reversion should be present, different versions of moving average regressions are included in the results. Moving averages of 12 and 60 months are used as explanatory variables, and both one-month and 12-month average premiums are used as dependent variables. In the US, 120-month moving averages are also included as explanatory variables since the longer US sample permits this. The samples outside the US are too short to provide a meaningful test of the 120-month moving averages.

Moving Average Simulations

Since the four premiums in this study are constructed by subtracting one index return from another, each premium has a long side and a short side. For example, the long and short sides of the value premium are a value index and a growth index, respectively. A moving average trading rule tries to trade back and forth between the two indices in a way that generates an average return that is higher than the averages of either of the indices. At each point in time, the trading rule decides which index to invest in by looking at the trailing moving average of the premium. If the trailing average is high, mean reversion suggests that subsequent premiums are likely to be low, so the trading rule buys the index that represents the short side of the premium. For example, if the trailing average value premium is high, the trading rule buys the growth index.

To evaluate the efficacy of a trading rule, the simulated return must be compared to a benchmark. In this case, the benchmark is straightforward. If a trading rule is profitable, it should generate average returns in excess of the simple strategy of remaining invested in the long side of the premium. For example, a rule that switches between value and growth should generate higher average returns than the simple strategy of remaining in value all the time.

The default position for a trading rule is to own the long side of the premium. The trading rule then seeks to find opportune times to switch to the other index. To simulate such a strategy, it is necessary to make several decisions:

1. What is the definition of a "high" moving average premium?
2. Over what horizon should the moving average be calculated?
3. How frequently should the trading decision be made?
4. How long should the trading rule remain in the short index once the switch to that index has been made?

To provide a comprehensive set of tests, different options are included for each of these decisions:

1. High moving averages are defined as either the top 10% or top 20% of the historical distribution, using only premiums that were available as of the trading date (to avoid look-ahead bias). In the discussion below, this percentile that defines what is a high average (either 10% or 20%) is called the *breakpoint*.
2. Similar to the moving average regressions, the simulations use 12-month, 60-month, and 120-month moving averages in the US and 12-month and 60-month moving averages outside the US.
3. Trading rules use both monthly and annual (December 31) rebalance frequencies.
4. One set of simulations switches back to the long side of the premium when the moving average falls below the breakpoint described above (either 10% or 20%). Another set of simulations recognizes that it may take more time for the mean reversion to take effect. These simulations remain in the short side until the trailing moving average falls below its historical median. In the discussion below, this point at which the trading rule switches back to the long side of the premium is called the *switchback*.

Different combinations of these decision rules produce a variety of trading rules. For example, in the US, there are a total of 24 rules for each premium (2 breakpoints × 3 moving average lengths × 2 rebalance frequencies × 2 switchback points). If mean reversion in the premiums is strong, some of these rules should consistently generate reliably positive excess returns.

Runs-Based Simulations

If the long side of a premium has outperformed the short side for several periods in a row, mean reversion suggests that it may be time for the outperformance to end. Runs-based trading rules switch to the short side of the premium (e.g., growth in the case of the value premium) when the trailing premium has been positive for N years in a row. The rules examined in this study include values for N of 3, 4, and 5 years. Similar to the moving average simulations, both monthly and annual (December 31) rebalance rules are included. At each rebalance date the rule calculates the length of the run. If the run length is greater than or equal to the threshold (3, 4, or 5 years), the rule invests in the short side of the premium. Otherwise, it invests in the long side.

Reliability of Excess Returns

As discussed earlier, a trading rule is judged by the difference between its average return and the average return of the long side of the premium. Hereafter, this difference in average returns is called the *excess return* of the trading strategy. There are two important aspects of this excess return. The first is its magnitude. A negative or small positive excess return indicates that the trading rule is not worth the effort. The simulations in this study ignore transactions costs. If the implementation of a trading rule would generate transactions costs (commissions, taxes, etc.), a small positive excess return would likely turn negative after reflecting those costs.

Another equally important aspect of the excess return is its reliability. In assessing the likelihood of a trading rule being successful in the future, it is necessary to determine whether the observed excess return may have occurred by random chance. This study assesses the reliability of excess returns by calculating bootstrapped p-values. For each historical premium sample, I reshuffle the original sample 1,000 times and run the trading rule on these 1,000 reshuffled samples. Since the process of reshuffling removes any mean reversion that may have been present in the original sample, the returns from these 1,000 reshuffled samples give a good idea of the range of returns that could occur by random chance. The p-value for each trading simulation is the proportion of these 1,000 reshuffled samples that generated a higher excess return than the trading rule applied to the original sample.

A low p-value (e.g., less than 5%) for a single simulation indicates that the results are not likely due to random chance. When multiple trials are run, the interpretation of the p-values is different. Some trials will have a low p-value just by chance. If a study conducts a large number of simulations (as this one does), about 5% of the simulations can be expected to have a p-value below 0.05, even if there is no mean reversion in returns. There are 780 simulations in this study, and it is necessary to see a low p-value for substantially more than 5% of these simulations in order to conclude that there is strong evidence of mean reversion.[2] Some methods for adjusting p-values for multiple tests have been developed in the statistics literature, and one of these methods is discussed later.

RESULTS

Exhibit 1 reports summary statistics for annual premiums. (Premium construction is described in the Appendix.) The 14 included countries (in addition to the US) are those that have market, value, and growth returns on Ken French's website back to 1975.[3] This permits tests outside the US for two of the four premiums—equity and value. The premiums with a country designation of "All Non-US Developed" represent a weighted average of the non-US developed markets that were available each year.

In general, Exhibit 1 reports premiums that are positive on average and quite variable. The only premium with a negative average is the Italian value premium, and the standard deviations in Exhibit 1 range from just below 10% to more than 50%. The volatility of these premiums highlights why there is so much interest in mean reversion. If the substantial variation in these premiums is partially predictable, the investment implications are significant. Exhibit 1 also gives an indication of the importance of diversification across markets. For example, the standard deviation of the all non-US developed value premium is less than half of the average of the individual country value premium standard deviations.

2. A coin-flipping analogy may help to convey the logic. If we ask 5,000 people to repeatedly flip a coin, we would expect about five of them to flip 10 heads in a row, just by chance. Upon observing this result, we would not conclude that we have found five really good coin flippers.

3. Ken French is a director and consultant for Dimensional Fund Advisors LP and is co-chair of the firm's Investment Policy Committee.

Exhibit 1 SUMMARY STATISTICS FOR ANNUAL PREMIUMS

Country	Premium	1927–2013 Average	1927–2013 Standard Deviation	1964–2013 Average	1964–2013 Standard Deviation
US	Equity	8.35	20.80	6.52	18.09
US	Size	3.56	14.04	3.80	13.99
US	Value	5.03	13.82	5.06	13.67
US	Profitability	—	—	3.12	9.39

Country	Premium	1975–2013 Average	1975–2013 Standard Deviation
All Non-US Developed	Equity	8.39	21.98
All Non-US Developed	Value	6.42	10.75
Australia	Equity	10.22	27.01
Australia	Value	7.10	19.41
Belgium	Equity	11.57	27.83
Belgium	Value	3.33	16.65
France	Equity	10.63	28.82
France	Value	5.97	20.07
Germany	Equity	9.52	28.89
Germany	Value	5.88	12.80
Hong Kong	Equity	16.73	37.33
Hong Kong	Value	5.07	24.27
Italy	Equity	8.20	36.40
Italy	Value	-3.26	25.67
Japan	Equity	7.04	28.61
Japan	Value	11.46	16.38
Netherlands	Equity	11.35	22.07
Netherlands	Value	3.05	21.98
Norway	Equity	12.44	40.40
Norway	Value	6.35	54.90
Singapore	Equity	11.36	31.39
Singapore	Value	10.30	20.24
Spain	Equity	7.99	31.64
Spain	Value	1.64	29.08
Sweden	Equity	13.46	29.72
Sweden	Value	6.88	29.78
Switzerland	Equity	9.65	24.21
Switzerland	Value	2.53	18.52
UK	Equity	11.97	25.74
UK	Value	2.93	16.43

Exhibit 1 shows averages and standard deviations of annual premiums in percentage points.
See the Appendix for premium construction, data sources, and Important Disclosures.

Exhibit 2 reports results for the US moving average regressions. The results on the left are for regressions in which the dependent variable is the premium in the following month. The results on the right are for regressions in which the dependent variable is the average premium over the next twelve months. Evidence of strong mean reversion would be indicated by a t-statistic on the slope coefficient well below -2.0, coupled with a high adjusted R^2. On the left side, none of the t-statistics are below -2.0, and the R^2 values are consistently at or near zero, indicating no evidence of mean reversion. On the right side, two of the t-statistics are below -2.0 (equity premium and value premium regressed on a 60-month moving average), but even in these cases the R^2 values are small (0.07). Overall, the evidence of return predictability from the US moving average regressions is quite weak.[4]

Exhibit 2 REGRESSIONS OF US PREMIUMS ON LAGGED MOVING AVERAGES

		Dependent Variable: Premium for Month t+1			Dependent Variable: Average Premium for Months t+1 through t+12		
Premium	# Months in Moving Average	Slope	t(Slope)	Adjusted R^2	Slope	t(Slope)	Adjusted R^2
Equity	12	0.15	1.06	0.00	-0.05	-0.40	0.00
Equity	60	-0.42	-0.97	0.00	-0.66	-2.48	0.07
Equity	120	0.17	0.39	0.00	0.14	0.44	0.00
Size	12	0.27	2.36	0.01	0.21	3.31	0.04
Size	60	0.00	-0.08	0.00	-0.14	-0.85	0.00
Size	120	-0.03	-0.08	0.00	-0.14	-0.56	0.00
Value	12	0.22	1.80	0.00	-0.10	-1.12	0.01
Value	60	-0.40	-0.99	0.00	-0.76	-2.41	0.07
Value	120	-0.25	-0.38	0.00	-0.80	-1.33	0.02
Profitability	12	0.29	1.70	0.01	0.02	0.10	0.00
Profitability	60	-0.21	-0.64	0.00	-0.54	-1.73	0.04
Profitability	120	0.12	0.25	0.00	-0.13	-0.34	0.00

Exhibit 2 shows slope coefficients, t-statistics for the slope coefficients, and adjusted R^2 values for regressions of premiums on their own lagged moving averages. Newey-West (1987) standard errors adjust for overlap in the dependent variables. Monthly returns begin July 1926 for the equity, size, and value premiums, and the monthly profitability returns begin in July 1963. Each time series regression begins as soon as there are enough observations to calculate the first moving average. For example, the equity, size, and value premium regressions with a 60-month moving average begin in June 1931. See the Appendix for premium construction, data sources, and Important Disclosures.

4. Two of the t-statistics for the size premium in Exhibit 2 are above +2.0. This indicates positive serial correlation, which is the opposite of mean reversion. However, the low R^2 values for these two regressions (0.01 and 0.04) indicate minimal explanatory power for subsequent size premiums.

Exhibit 3 summarizes the non-US moving average regressions. Since there are so many regressions outside the US, it is not feasible to list each one as Exhibit 2 did for the US. Instead, the regressions are sorted by various characteristics (premium, moving average length, and dependent variable), and Exhibit 3 reports summary results for categories of regressions. (See Exhibit 9 for the results of various tests by country.) The regressions of the 12-month average equity premium on its lagged 60-month moving average provide the strongest evidence of mean reversion in this study. Thirteen of the 15 slope coefficients are more than two standard errors below zero, and the average adjusted R^2 is 0.12. While this average R^2 is substantially higher than any of the others, this does not imply that explaining 12% of the variation in premiums is sufficient to generate profitable trading opportunities. As will be shown later, the trading rules using a 60-month moving average provide limited evidence of being able to generate positive excess returns.

Exhibits 4A and 4B on the next page show results for US moving average trading simulations. Recall that in the US there are 24 moving average simulations for each premium. Showing 24×4=96 simulations is too cumbersome, so Exhibit 4A only shows a subset of the results. Excess returns were consistently higher for annual rebalance than for monthly rebalance, and they were generally higher when a 10% breakpoint was used to define a high moving average. Consequently, Exhibit 4A shows results for simulations that rebalance annually with a 10% breakpoint. The results on the left side of the exhibit are for simulations in which the switchback point is 10%, and the results on the right side are for a switchback at the median of the distribution.

Exhibit 3 SUMMARY OF NON-US MOVING AVERAGE REGRESSIONS

			Number of Regressions		
Premium	Moving Average Length (months)	Dependent Variable	Total	With t(slope)<-2.0	Average Adjusted R^2
Equity	12	1	15	0	0.00
Value	12	1	15	0	0.00
Equity	12	12	15	0	0.01
Value	12	12	15	3	0.03
Equity	60	1	15	3	0.00
Value	60	1	15	1	0.00
Equity	60	12	15	**13**	**0.12**
Value	60	12	15	2	0.05

Exhibit 3 shows summary statistics for regressions of premiums on their own lagged moving averages. Regressions are grouped by the indicated characteristics, and statistics are reported for each group of regressions. Newey-West (1987) standard errors adjust for overlap in the dependent variables. Dependent variable=1 indicates that the dependent variable is the premium for the next month. Dependent variable=12 indicates that the dependent variable is the average premium over the next 12 months. See the Appendix for premium construction, data sources, and Important Disclosures.

Three of the excess returns in Exhibit 4A are reliably greater than zero (with a p-value of 0.05 or lower), but none of the excess returns are above 40 basis points per year. There is no consistency of outperformance across premiums; the rules that generated reliably positive excess returns for one premium did not do so for the other premiums. Eighteen of the 24 excess returns are negative.

Exhibit 4B summarizes all 96 US moving average simulations. Eighty-seven have a negative excess return, and four have a reliably positive excess return (three of the four are shown in Exhibit 4A). The smallest and largest excess returns are -4.08% and 0.40%, respectively, and the average is -1.30%.

The simulations with annual rebalancing trade on December 31 each year. It is helpful to know if the results are sensitive to this choice of rebalance date. In some cases the results are quite sensitive. For example, consider the simulation in Exhibit 4A that has an excess return of 37bp (the equity premium simulation with a 60-month moving average and a 10% switchback). When I vary the rebalance date across the 12 months, the excess return varies between -48bp and +93bp. This is just one example of how sensitive the results can be to small changes in research design.

Exhibit 5 on the next page summarizes the non-US moving average simulations, categorizing the simulations by five important characteristics. **Panel A** provides results for simulations in which the breakpoint and the switchback are the same, and **Panel B** shows simulations for which the switchback point is the median of the distribution. All of the category average excess returns are negative, and no category has reliably positive excess returns for more than one-fifth of the category simulations. As noted earlier, the reliable regression results from Exhibit 3 for the non-US

Exhibit 4A US MOVING AVERAGE SIMULATIONS

		10% Switchback		50% Switchback	
Premium	Moving Average Length	Excess Return	P-Value	Excess Return	P-Value
Equity	12	-0.45	0.31	-0.53	0.23
Equity	60	**0.37**	**0.05**	**0.35**	**0.03**
Equity	120	-1.29	0.60	-3.56	0.77
Size	12	-0.97	0.89	-2.90	1.00
Size	60	0.20	0.11	0.13	0.11
Size	120	0.11	0.18	**0.40**	**0.05**
Value	12	-0.68	0.60	-1.57	0.78
Value	60	-0.47	0.42	-0.23	0.13
Value	120	-0.31	0.32	-1.51	0.55
Profitability	12	-1.08	0.95	-1.72	0.95
Profitability	60	-0.26	0.40	-1.61	0.81
Profitability	120	-0.87	0.72	-1.64	0.74

Exhibit 4B SUMMARY STATISTICS FOR ALL 96 US MOVING AVERAGE SIMULATIONS

Number of Simulations With:				
Negative Excess Return	Reliably Positive Excess Return	Average Excess Return	Smallest Excess Return	Largest Excess Return
87	4	-1.30	-4.08	0.40

Exhibit 4A reports average annual excess returns (in excess of the long side of each premium) and p-values for a subset of the US moving average trading simulations. The simulations switch to the short side of each premium when the trailing moving average is in the top 10% of its historical distribution. The simulation switches back to the long side of the premium when the trailing moving average falls below the percentile designated as the switchback. Simulations in Exhibit 4A use an annual (December 31) rebalance. P-values are based on the distribution of excess returns from each trading rule applied to 1,000 bootstrapped samples. Exhibit 4B provides summary statistics for all 96 US moving average simulations. See the Appendix for premium construction, data sources, and Important Disclosures.

equity premiums did not translate into reliable excess returns in the moving average simulations. The regressions in Exhibit 3 indicate some ability to predict when the equity premium will be above or below average, but that is not enough for a successful trading rule. Such a rule requires an ability to predict when the premium is likely to be negative. The evidence in Exhibit 5 suggests that this is more difficult to do.

Exhibit 5 SUMMARY OF NON-US MOVING AVERAGE SIMULATIONS

Panel A: Switchback equals Breakpoint

				Number of Simulations		
Premium	Rebalance Frequency	Breakpoint (%)	Moving Average Length (months)	Total	With Reliably Positive Excess Return	Average Excess Return
Equity	Monthly	80	12	15	0	-3.34
Equity	Monthly	80	60	15	0	-2.65
Equity	Monthly	90	12	15	3	-1.86
Equity	Monthly	90	60	15	0	-1.46
Equity	Annual	80	12	15	1	-2.77
Equity	Annual	80	60	15	2	-0.34
Equity	Annual	90	12	15	0	-2.08
Equity	Annual	90	60	15	2	-0.45
Value	Monthly	80	12	15	0	-1.99
Value	Monthly	80	60	15	1	-0.97
Value	Monthly	90	12	15	2	-0.94
Value	Monthly	90	60	15	0	-0.78
Value	Annual	80	12	15	0	-1.67
Value	Annual	80	60	15	2	-0.44
Value	Annual	90	12	15	1	-0.90
Value	Annual	90	60	15	2	-0.36

Panel B: Switchback equals 50%

				Number of Simulations		
Premium	Rebalance Frequency	Breakpoint (%)	Moving Average Length (months)	Total	With Reliably Positive Excess Return	Average Excess Return
Equity	Monthly	80	12	15	0	-5.49
Equity	Monthly	80	60	15	0	-2.12
Equity	Monthly	90	12	15	0	-3.35
Equity	Monthly	90	60	15	0	-1.81
Equity	Annual	80	12	15	1	-3.07
Equity	Annual	80	60	15	1	-0.43
Equity	Annual	90	12	15	0	-2.06
Equity	Annual	90	60	15	1	-1.02
Value	Monthly	80	12	15	1	-2.04
Value	Monthly	80	60	15	1	-1.82
Value	Monthly	90	12	15	1	-1.39
Value	Monthly	90	60	15	1	-1.22
Value	Annual	80	12	15	0	-1.73
Value	Annual	80	60	15	2	-1.26
Value	Annual	90	12	15	0	-0.94
Value	Annual	90	60	15	2	-1.09

Summary statistics are shown for non-US moving average simulations sorted by various characteristics. Trading rules move to the short side of the premium when the trailing moving average is above the breakpoint, and they switch back to the long side of the premium when the moving average is below the switchback percentile. Excess returns are annual and are expressed in percentage points. A reliably positive excess return is defined as an excess return that is greater than zero and that has a p-value of 0.05 or below. P-values are based on the distribution of excess returns from each trading rule applied to 1,000 bootstrapped samples. See the Appendix for premium construction, data sources, and Important Disclosures.

Results for US runs-based trading rules are in **Exhibit 6**. The excess return is reliably positive (55bp) for the equity premium simulation using a five-year run length and annual rebalancing. This rule produced negative excess returns for the other three premiums, indicating a lack of consistency. The prevalence of negative excess returns in Exhibit 6 casts doubt on the efficacy of runs-based rules.

Exhibit 6 RESULTS FOR US RUNS-BASED TRADING RULES

		Rebalance Monthly		Rebalance Annually	
Premium	**Run Length**	**Excess Return**	**P-Value**	**Excess Return**	**P-Value**
Equity	3 years	-3.67	0.83	-1.02	0.25
Equity	4 years	-2.13	0.72	0.13	0.09
Equity	5 years	-1.13	0.64	**0.55**	**0.02**
Size	3 years	-2.47	0.98	-1.43	0.90
Size	4 years	-1.35	0.96	-0.46	0.70
Size	5 years	-0.45	0.79	-0.37	0.80
Value	3 years	-1.73	0.64	-1.70	0.63
Value	4 years	-0.84	0.52	-1.12	0.65
Value	5 years	-0.35	0.44	-0.46	0.54
Profitability	3 years	-1.74	0.78	-0.84	0.43
Profitability	4 years	-0.97	0.68	-0.81	0.60
Profitability	5 years	-0.64	0.66	-0.46	0.57

Trading simulations switch to the short side of each premium when the long side has outperformed for N years in a row. N takes values of 3, 4, or 5 years. Simulations switch back to the long side when the run ceases. The exhibit includes results for both monthly and annual (December 31) rebalancing. P-values are based on the distribution of excess returns from each trading rule applied to 1,000 bootstrapped samples. See the Appendix for premium construction, data sources, and Important Disclosures.

The non-US runs-based results are in **Exhibit 7**, with the simulations categorized by premium, rebalance frequency, and run length. With one exception, the evidence of successful trading rules in Exhibit 7 is minimal. The simulations for the equity premium with annual rebalancing and a five-year run length show the strongest results, with reliably positive excess returns for nearly half of the simulations and an average excess return of 76bp/year. One interesting aspect of this result is that it is due mainly to a single year. When 2008 is excluded from the sample, the average excess return is -28bp/year. While there is no reason to exclude 2008 from the sample, it is important to know how sensitive the results can be to a single year. The US results show a similar sensitivity to this year; the excess return of 55bp shown in Exhibit 6 falls to 8bp per year when 2008 is excluded.

The five-year runs test on the equity premium shows the best performance of any rule included in this study. However, there are several reasons to question the robustness of this rule's results. The first reason is the aforementioned sensitivity of the results to a single year. The second reason is the fact that it did not generate reliable excess returns in half of the equity premium trials. A third reason is that the rule's performance for the other premiums was consistently poor. These observations cast doubt on the ability of a five-year run to predict the following year's equity premium.

Exhibit 7 SUMMARY OF NON-US RUNS-BASED SIMULATIONS

			Number of Simulations		
Premium	Rebalance Frequency	Run Length (years)	Total	With Reliably Positive Excess Return	Average Excess Return
Equity	Monthly	3	15	0	-4.07
Equity	Monthly	4	15	0	-1.93
Equity	Monthly	5	15	1	-0.27
Equity	Annual	3	15	0	-2.33
Equity	Annual	4	15	3	0.03
Equity	Annual	5	15	**7**	**0.76**
Value	Monthly	3	15	0	-1.65
Value	Monthly	4	15	1	-0.72
Value	Monthly	5	15	1	-0.54
Value	Annual	3	15	0	-1.22
Value	Annual	4	15	0	-0.64
Value	Annual	5	15	2	-0.55

Summary statistics are shown for non-US runs-based simulations sorted by various characteristics. Excess returns are annual and are expressed in percentage points. A reliably positive excess return is defined as an excess return that is greater than zero and that has a p-value of 0.05 or below. P-values are based on the distribution of excess returns from each trading rule applied to 1,000 bootstrapped samples. See the Appendix for premium construction, data sources, and Important Disclosures.

Exhibit 8 provides an overall summary of all the 780 trading simulations—both US and non-US. If there is no mean reversion at all in the dimensions of return, about 5% of the simulations would likely show reliably positive excess returns just by random chance. The actual proportion is 5.8%. The runs tests do slightly better (7.8%), but overall the evidence of predictability from these simulations is quite weak. Average excess returns are negative for both the runs and the moving average tests.

Exhibit 8 OVERALL SUMMARY OF TRADING SIMULATIONS

Simulation Category	Total Simulations	Fraction with Reliably Positive Excess Return	Average Excess Return
All	780	5.8%	-1.48%
Moving Average	576	5.0%	-1.62%
Runs	204	7.8%	-1.09%

Exhibit 8 provides an overall summary of the simulations reported in Exhibits 4-7. Refer to those exhibits for details. See the Appendix for premium construction, data sources, and Important Disclosures.

As mentioned previously, procedures for adjusting p-values for multiple tests have been developed in the statistics literature. When I apply one of the more well-known of these procedures to the 780 p-values in this study, the proportion of simulations with reliably positive excess returns falls from 5.8% to 0.2%.[5] Based on this adjustment, 99.8% of the simulations fail to accomplish the stated objective of generating a reliably positive excess return.

This study assumes that the objective of an investor in using a timing rule is to achieve a higher average return than the simple strategy of remaining invested in the long side of a given premium. An alternative objective is to achieve a higher Sharpe ratio than the long side of the premium. The trading rules included in this study do not fare any better in achieving that objective. Only 4% of the p-values for Sharpe ratio differences are 0.05 or less. This result is slightly worse than what would be expected by random chance, and it is worse than the 5.8% success rate shown in Exhibit 8 for the average return.

CONCLUSIONS

The procedure followed in this study—trying hundreds of different trading rules in search of some that work in the historical sample—is commonly called data mining. Dimensional has always cautioned investors not to rely on strategies that were found by data mining because the success of the strategies in historical data could very well be spurious. If one dredges through the data long enough, one will eventually find some strategies that perform well in the historical sample. That is not an interesting result. It becomes more interesting if the same rule generates reliable excess returns in multiple samples while underperforming in few samples. The most interesting result of this study is that, in spite of vigorously mining the historical data, I was still unable to find a trading rule that consistently generated reliable excess returns across markets and premiums. While there are certainly more trading rules that I could try, the 780 attempts documented here show little promise. And in my opinion, 780 are enough.

The highest excess return of all the 780 trading simulations was 7.46% per year for one of the moving average simulations of the Italian value premium. An excess return that high might lead some to conclude that they have found a good trading rule. That would be a mistake. The rule in question had an average excess return of -2.29% per year in its other trials (its lowest excess return was -7.16% for the Japanese value premium). This is a good example of the pitfalls associated with using data mining to search for viable trading strategies.

By focusing on mean reversion, this study examines the ability of lagged premiums to predict future premiums. Other variables besides lagged premiums have been proposed as inputs into premium timing rules. For example, differences in book-to-market or earnings/price ratios between the long and short sides of various premiums have been proposed as good candidates. The evidence in Davis (2007, 2011) suggests that these variables may not have much more explanatory power for future premiums than the variables included in this study.

5. See Benjamini and Yekutieli (2001) for details of the procedure.

Exhibit 9 SUMMARY OF NON-US RESULTS BY COUNTRY

Results by Country:	Moving Average Regressions: Total	Moving Average Regressions: With t(slope)<-2.0	Moving Average Regressions: Average Adjusted R^2	Moving Average Simulations: Total	Moving Average Simulations: With Reliably Positive Excess Return	Moving Average Simulations: Average Excess Return	Runs-Based Simulations: Total	Runs-Based Simulations: With Reliably Positive Excess Return	Runs-Based Simulations: Average Excess Return
All Non-US Developed	8	1	0.03	32	0	-2.35	12	2	-0.23
Australia	8	1	0.04	32	1	-0.36	12	3	0.13
Belgium	8	1	0.01	32	0	-3.46	12	0	-2.81
France	8	2	0.02	32	2	-1.30	12	2	-0.89
Germany	8	1	0.02	32	1	-1.97	12	2	-0.87
Hong Kong	8	2	0.03	32	6	0.30	12	0	-1.42
Italy	8	3	0.04	32	8	0.31	12	0	0.08
Japan	8	0	0.01	32	0	-3.78	12	0	-4.21
Netherlands	8	1	0.02	32	0	-1.75	12	0	-1.04
Norway	8	1	0.03	32	0	-1.18	12	1	0.04
Singapore	8	2	0.04	32	2	-1.34	12	0	-1.74
Spain	8	0	0.02	32	1	-4.09	12	1	-1.41
Sweden	8	2	0.04	32	3	-2.15	12	0	-1.49
Switzerland	8	1	0.02	32	1	-1.40	12	1	-0.42
UK	8	4	0.05	32	2	-0.70	12	3	-0.16

Exhibit 9 shows the regressions and simulations from Exhibits 3, 5, and 7 by country. Refer to those exhibits for details. See the Appendix for premium construction, data sources, and Important Disclosures.

APPENDIX

Premium Details

US equity premium: The return on US stocks in excess of the one-month US Treasury bill return. US stocks are represented by the CRSP value-weighted index. Monthly returns begin in July 1926, and annual returns begin in 1927.

US size premium: The return on small cap stocks minus the return on large cap stocks. The small cap and large cap indices are based on the methodology used by Fama and French (1993) to construct the two parts (S and B) of their SMB factor. Monthly returns begin in July 1926, and annual returns begin in 1927.

US value premium: The return on an index of value stocks minus the return on an index of growth stocks. The value and growth indices are based on the methodology used by Fama and French (1993) to construct the two parts (H and L) of their HML factor. Monthly returns begin in July 1926, and annual returns begin in 1927.

US profitability premium: The return on an index of high profitability stocks minus the return on an index of low profitability stocks. Index construction uses 2x3 sorts on size and operating profitability, as those variables are defined in Fama and French (2014). Monthly returns begin July 1963, and annual returns begin in 1964.

Non-US equity premiums: Treasury bill rates for most countries are not available back to 1975, so the equity premiums are the local market returns (in USD) minus the US Treasury bill return. Returns begin in January 1975. Using the dollar-denominated return in excess of the US T-bill return as a proxy for non-US equity premiums could potentially complicate the interpretation of the results. If the tests show evidence of mean reversion for non-US equity premiums, it could be due to mean reversion in currencies. To address this concern, I examined the currency returns for the markets included in this study, looking for evidence of mean reversion. The results (not shown) indicated no mean reversion in currency returns over the horizons examined in this study.

Non-US value premiums: For each market, the value premium is the return on an index of value stocks minus the return on an index of growth stocks. Value and growth indices are based on annual book-to-market rankings, using the "all 4 data items not required" series from Ken French's website. Returns begin in January 1975.

Data are from CRSP, Compustat, and Ken French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html.

IMPORTANT DISCLOSURES

Strategies presented herein are for illustrative purposes only and do not represent actual investments or strategies available during the periods represented. The data does not reflect advisory fees, trading costs, or other expenses associated with the management of an actual portfolio. The securities held in the simulated model may differ significantly from those held in an actual account. Actual management of these types of simulated strategies may result in lower returns than the back-tested results achieved with the benefit of hindsight. Investing involves risks such as fluctuating value and possible loss of principal investment. Past performance of a simulated strategy is no guarantee of future results.

REFERENCES

Benjamini, Yoav, and Daniel Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* 29:1165-1188.

Campbell, John Y., Andrew W. Lo, and A. Carig Mackinlay. 1997. *The econometrics of financial markets*. Princeton, NJ: Princeton University Press.

Davis, James L. 2007. Does predicting the value premium earn abnormal returns? Purely Academic web column, www.dimensional.com.

Davis, James L. 2008. Is there mean reversion in the value premium? Working paper, Dimensional Fund Advisors.

Davis, James L. 2011. Calculating aggregate earnings/ price ratios: the details are important. Working paper, Dimensional Fund Advisors.

Fama, Eugene F. and Kenneth R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33:3-56.

Fama, Eugene F. and Kenneth R. French. 2014. A five-factor asset pricing model. Fama-Miller working paper.

Goyal, Amit, and Ivo Welch. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21:1455-1508.

Lucas, Robert E. 1978. Asset prices in an exchange economy. *Econometrica* 46:1429-1445.

Newey, Whitney K., and Kenneth D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703-708.

Samuelson, Paul A. 1965. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* 6:41-49.