



ELROY DIMSON • PAUL MARSH • MIKE STAUNTON

Triumph of the Optimists

101 YEARS OF GLOBAL INVESTMENT RETURNS



countries, over the whole of the 101-year period from 1900–2000. We also have century-long evidence on the small-firm and value/growth phenomena. We have put significant effort into compiling complete financial market histories, so that we can present consistent and comparable records for different countries. But *Triumph of the Optimists* is about much more than just data, since it has description and analysis at its core.

There is an obvious need for a reliable and truly international dataset for the investment industry as it continues relentlessly toward full market globalization. One of the many changes taking place in the investment business is the increasing demand for locally sourced research placed in a global context. Another innovation is the growing number of truly global mandates being given to fund managers. Globalization may be a cliché, but for portfolio managers it is fast becoming a reality. Access to a properly constituted and rigorously maintained international database is a sine qua non for the start of any investment process.

The period since spring 2000 has come as a shock to those who had become used to the bull market conditions of previous years. The bursting of the technology bubble, the rapid decline in economic growth rates, especially in the United States, and the advent of international terrorism raised questions about what we can expect for the future. We assert in this book that the single most important variable for making investment decisions is the equity risk premium, and we argue that high long-term returns on equities, relative to bonds, are unlikely to persist. Even after the setbacks of 2000–01, it is necessary to justify the relatively high rating of today's stock markets in terms of a historically low forward-looking equity risk premium. For the investment strategist this raises the most fundamental question of all: Do investors realize that returns are likely to revert to more normal levels, or do current valuations embody exaggerated expectations based on an imperfect understanding of history?

Good data is the key to understanding history. With this as our guiding principle, assembling the data for this book was a major task. For the United Kingdom, ABN AMRO supported us in compiling an authoritative record of UK equity market performance over the last 101 years. We did this because we were not satisfied with the data that previously existed, and there was anyway no comprehensive record of equity returns extending back to 1900. To construct our UK indexes, we devoted intensive efforts to financial archaeology. This involved transcribing original source data from dusty newspaper archives and ancient reference books into our database. A resulting benefit is that we have not simply assembled an index, but we also have the underlying stock-by-stock data, so we can now study the performance of segments of the market, such as industry sectors and market-capitalization bands. We also compiled a series of UK government bond indexes especially for this study.

For the other fifteen countries covered in this book, we have linked together the best quality indexes and returns data available from previous studies and other sources, a number of which are previously unpublished, and some of which are still work in progress. In addition to the United Kingdom, we cover two North American markets, the United States and Canada; ten other European markets, namely, Belgium, Denmark, France, Germany, Ireland, Italy, The Netherlands, Spain, Sweden and Switzerland; two Asia-Pacific markets, Australia

Chapter 3 Measuring long-term returns

Good measures of long-run returns should accurately reflect the outcome of an implementable investment strategy. The strategy should be one that could have been set up in advance, and followed in real life, and which is representative of the asset class and country in question. It is only too easy for researchers to fail to meet these criteria.

This chapter begins in section 3.1 by setting out the principles that need to be followed in constructing long-run return indexes. These provide a benchmark for assessing previous studies, and have been the guiding framework for this book. Given that our data goes back to the beginning of the last century and covers sixteen countries, we have not always been able to adhere to every principle, especially in the earliest years. Nevertheless, these standards have guided our choices, and we indicate where compromises have been necessary.

Next, in section 3.2 we take a closer look at equity index construction and at a bias that has afflicted some previous studies. When an index is compiled retrospectively, a crucial issue is how to avoid tilting its composition toward companies that, with hindsight, are known to have survived and/or to have been successful. In section 3.3, we review other issues that arise in index design, such as dividend reinvestment, index coverage, and index weighting.

In section 3.4 we consider how best to assemble a sample of international indexes. We show that reliance on data that is easy to acquire, such as indexes that start after the end of a war, tends to result in overstated performance. Both success bias and easy-data bias arise from a focus on assets that have survived or prospered over a particular period, and both can lead to overestimates of index returns and risk premia.

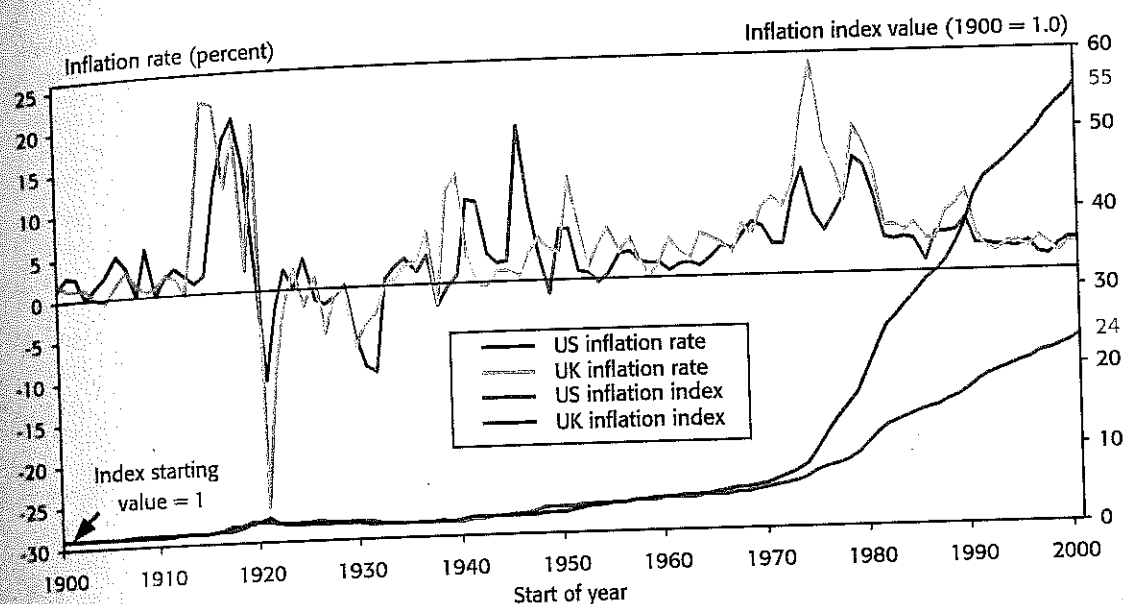
In section 3.5, we focus on the special problems that can arise when measuring inflation rates, as well as long-term returns on bonds, bills and currencies. We conclude in section 3.6 with a summary of the chapter.

3.1 Good indexes and bad

There are five guiding principles that underpin our measures of long-term performance. They are to avoid bias in index construction, to focus on total returns, to ensure the widest possible coverage within each market, to apply appropriate methods of weighting and averaging, and to maximize the extent to which comparisons can be made across national boundaries.

First, equity indexes should avoid bias. Good indexes follow an investment strategy that could be followed in real life. Apart from dealing costs, an investor should in principle have been able to replicate index performance. Indexes, especially when they are constructed retrospectively, must therefore be free of any look-ahead bias. They must be constructed solely from information that would have been available at the time of investment. Serious bias can arise if index constituents are tilted toward companies that subsequently survived or

Figure 5-2: US and UK annual inflation rates and cumulative inflation, 1900–2000



much higher than in the United States, peaking at 25 percent in 1975. The cumulative impact of these higher rates can be seen by comparing the two inflation indexes, which are plotted against the right-hand axis in Figure 5-2. The red line for the United Kingdom and the blue line for the United States are virtually coincident until the mid-1960s. From that point onward, the UK index rises to a value of fifty-five by end-2000, compared with twenty-four for the United States. From 1900–2000, UK consumer prices thus rose 55-fold, a factor of 2.3 times more than in the United States, with the difference almost entirely attributable to the last thirty-five years. Over the full 101-year period, the average annualized UK inflation rate was 4.1 percent per year, versus 3.2 percent for the United States.

5.2 Inflation around the world

While inflation was fairly similar in the United States and the United Kingdom, a number of other countries had quite different experiences. Table 5-1 provides international inflation rate comparisons across all sixteen countries covered in this book, showing the mean inflation rates from 1900–2000. Clearly, there were large differences between countries. At the same time, the standard deviations for each country show that there was also considerable variation in inflation rates over time. Taking the full 101-year period as a whole, there have been four high inflation rate countries, France, Germany, Italy, and Japan; two runner-ups, Belgium and Spain; and one low inflation country, Switzerland. The other countries fall in between, with inflation rates in the region of 3–4 percent per year. Note that the true 101-year means and standard deviations for inflation in Germany are much higher than shown in Table 5-1 since the statistics in the table omit the hyper-inflationary years of 1922–23.

and some of the studies span periods as brief as five years. These differences in research periods, methodologies, and definitions of "smallness" mean that the premia shown in Figure 9-5 are not directly comparable. In particular, it would be unwise to make inferences about the magnitudes or significance of any apparent size premium differences between countries.

In spite of this caveat, Figure 9-5 paints a very clear picture, namely, that the size premium was not restricted to the United States but was present in almost every country studied by the researchers. The sole exception was Korea, where a negative premium was reported, although this study used just five years of data. Furthermore, in most countries, researchers also looked at risk differences. They concluded, like Banz (1981), that the size premium could not be explained away by risk.

The pervasiveness and magnitude of the size effect, and the striking outperformance of smaller companies in most countries around the world, meant that the size effect rapidly became recognized as the premier stock market anomaly.

9.4 The reversal of the size premium

The "discovery" of the size effect in the United States by Banz (1981) and Reinganum (1981), and the publication and dissemination of their research, led to considerable interest in small-caps among investors in the United States. This spurred the launch of significant new small-cap investment vehicles led by Dimensional Fund Advisors, who raised several billion dollars within a couple of years of their 1981 launch. This honeymoon period lasted for approximately two years, until the end of 1983, and during this period, US small-caps continued to outperform. But subsequently, and over much of the period since, US small-caps have underperformed.

The UK experience was remarkably similar. When the HGSC was launched in 1987, its back-history showed that smaller companies had outperformed the UK market by 5.2 percent per year. This dramatic outperformance attracted substantial media attention, and there were over two hundred follow-up articles in the UK press. By the end of 1988, at least thirty open- and closed-end funds had been launched to exploit the perceived outperformance of small-caps, and numerous investment institutions developed a strategy of investing in smaller companies as a distinct asset class. Again, the honeymoon lasted just two years. In the decade that followed, smaller companies were to underperform by a large margin.

This reversal in the fortunes of US and UK small-cap stocks led us to write an article in 1999 entitled "Murphy's Law and Market Anomalies." Murphy's Law is often summarized as "bread always falls with the buttered side down." Figure 9-6 shows the performance record of US and UK small-caps at the time of our article, and shows why this appeared like a classic case of Murphy's Law. The left-hand side of Figure 9-6 shows the historical small- and micro-cap premia for the United States and the United Kingdom from the start date of the original research studies until the end of the post publication honeymoon period (i.e., 1926-

Our subsequent research has shown that the small-cap reversal extended beyond the United Kingdom and United States, and was a worldwide phenomenon. The line of investigation we followed here was to revisit all of the research studies that have been conducted into the size effect in different countries, and to estimate the size premium over the years since the research was published. These studies were discussed earlier in section 9.3 and their findings were summarized in Figure 9-5. We found that they showed evidence of a significant size premium in every country examined, with the sole exception of Korea, where the research covered just a five-year period. Most of these research studies were published in the 1980s.

To update these studies, we estimated the size premium in each country over the period since each study was published. For consistency, we again measured the size premium as the difference between the average monthly returns on the smallest and the largest stocks. For the United States, we use the CRSP NYSE Decile 10 and Decile 1 returns as our respective measures of small and large stock returns, as this most closely approximates Banz's (1981) earlier research, and gives results close to his over his earlier period. Similarly, for the United Kingdom we adopt the same definition as was used in Figure 9-5, namely, the difference between HGSC returns and overall UK equity returns.

For all other countries, we use the size-based indexes published by either Independence International Associates (IIA) or by FTSE International. IIA publish large- and small-cap indexes for a number of countries starting in 1975. They define small as the bottom 30 percent by capitalization of their universe, and large as the top 70 percent. FTSE publish a similar set of large and medium-small-cap indexes for a larger population of countries, but only from 1987, with some countries starting even later. FTSE define medium/small-cap as the bottom 25 percent by capitalization, and large-cap as the balancing 75 percent. For countries where we had a choice between both IIA and FTSE Indexes, we have used the IIA series since they provide a longer time series and generally have somewhat wider coverage.

The results of our research are shown in Figure 9-7. Countries are listed in alphabetical order, and for each country, the size premium reported by the original research studies and plotted earlier in Figure 9-5 is shown in green. Alongside this, the yellow bar shows the size premium calculated over the period since the original research was published, that is, over the period starting at the beginning of the year immediately following publication and ending at New Year 2001. No size-based indexes were available for Korea or Taiwan, so we omitted these countries. We have, however, included the four countries covered in this book, but which did not feature in Figure 9-5 due to the absence of any research study on the size premium. For these countries, we have omitted the "initial research" bars in Figure 9-7, while the "subsequent period" bars show the size premium over the period from 1990–2000.

It is clear from Figure 9-7 that there was a global reversal of the size effect in virtually every country, with the size premium not just disappearing but going into reverse. Researchers around the world universally fell victim to Murphy's Law, with the very effect they were documenting—and inventing explanations for—promptly reversing itself shortly after their studies were published. The only country experiencing a size premium, as opposed to a size discount, in the period subsequent to the original research was Switzerland. However, the Swiss size premium was statistically insignificant, and its magnitude was just 0.05 percent.

Table 12-2: Worldwide equity risk premia relative to long bond returns, 1900–2000

Country	Annual equity risk premium relative to long-term bonds						Ten year risk premium		
	Geometric mean	Arithmetic mean	Standard error	Standard deviation	Minimum premium	Maximum premium	Geometric mean	Arithmetic mean	Standard deviation
Australia	6.3	8.0	1.9	18.9	-30.6	66.3	6.3	6.4	4.6
Belgium	2.9	4.8	2.1	20.7	-35.1	76.6	3.0	3.2	5.1
Canada	4.5	6.0	1.8	17.8	-36.8	54.7	4.6	4.7	5.4
Denmark	2.0	3.3	1.7	16.9	-35.9	74.9	1.8	1.9	4.1
France	4.9	7.0	2.1	21.6	-32.7	83.7	4.9	5.1	6.8
Germany	6.7	9.9	2.9	28.4	-38.6	117.6	8.2	8.5	9.1
Ireland	3.2	4.6	1.7	17.4	-37.0	73.3	3.0	3.2	4.8
Italy	5.0	8.4	3.0	30.0	-39.6	152.2	5.0	5.4	9.2
Japan	6.2	10.3	3.3	33.2	-43.3	193.0	6.7	7.2	11.5
The Netherlands	4.7	6.7	2.1	21.4	-43.9	107.6	4.3	4.5	6.5
South Africa	5.4	7.1	2.0	19.7	-29.2	70.9	6.2	6.3	5.0
Spain	2.3	4.2	2.0	20.3	-34.0	69.1	2.2	2.3	5.5
Sweden	5.2	7.4	2.2	22.1	-38.3	87.8	4.8	5.0	7.7
Switzerland [†]	2.7	4.2	1.9	17.9	-34.4	52.2	2.0	2.1	5.1
United Kingdom	4.4	5.6	1.7	16.7	-38.0	80.8	4.8	4.9	4.5
United States	5.0	7.0	2.0	20.0	-40.8	57.7	4.9	5.0	5.2
World	4.6	5.6	1.4	14.5	-31.2	37.4	4.6	4.7	4.8

* All statistics for Germany exclude 1922–23. † Premia for Switzerland are from 1911.

In this table, the first six columns give summary statistics for the annual premia, while the last three columns relate to rolling ten-year premia. The first column shows the geometric means that were plotted as bars in Figure 12-6. The fourth column shows the standard deviations. The 20.0 percent figure for the United States is close to the 19.6 percent standard deviation for the premia relative to bills shown earlier in Table 12-1. For some countries, however, the distribution of premia relative to bonds is narrower than relative to bills. For the United Kingdom, for example, the standard deviation is 16.7 percent, compared with 19.9 percent relative to bills. This is because, in the United Kingdom, there was a fairly high correlation between annual equity returns and long bond returns (0.56), while the correlation between equities and bills was lower (0.29). This propensity for good bond years to coincide with good equity years, and vice versa, has tended to lower the annual difference between equity and bond returns in the United Kingdom. This was particularly marked in the best and worst years on record for UK equities, namely, 1975 and 1974 respectively.

12.5 Summary

In this chapter, we have used 101 years of stock market history for sixteen different countries and for the world index to take a fresh look at the equity risk premium. In the past, the historical evidence for the US market, and to a lesser extent for the United Kingdom, has heavily influenced views about the magnitude of the risk premium. For the United States, the most widely cited source is Ibbotson Associates (2000), who estimate a geometric risk premium of

The chapter addresses four questions: Which historical equity risk premium should one use as the starting point? Why has it typically been so high? What is a good forward-looking predictor for the future? How can one use variables such as the dividend yield to improve forecasts of the risk premium?

We stress the central role in finance of the equity premium. The historical premium is often summarized in the form of an annualized rate of return. This is a geometric mean. It provides information on past performance. For the future, what is required is the arithmetic mean of the distribution of equity premia, which is larger than the geometric mean. For markets that have been particularly volatile, the arithmetic mean of past equity premia may exceed the geometric mean premium by several percentage points. We adjust the arithmetic mean for (i) the differences between the variability of the stock market over the last 101 years, and the variability that we might anticipate today, and (ii) the impact of unanticipated cash flows and of declines in the required risk premium. The result is a forward-looking, *geometric* mean risk premium for the United States, United Kingdom and world of around 2½ to 4 percent and an *arithmetic* mean risk premium for US, UK, and world equities that falls within a range from a little below 4 to a little above 5 percent.

These equity risk premia are lower than those cited in surveys of finance academics. They are also lower than frequently quoted historical averages, such as those from Ibbotson Associates, which cover a somewhat briefer interval. We show that the historical risk premium, even if it embraces countries that have been less successful than the United States, is supported by two factors. Over the second half of the last century, equity cash flows almost certainly exceeded expectations, and the required rate of return doubtless fell as investment risk declined and the scope for diversification increased. Stock markets rose, in both the United States and other countries, for reasons that are unlikely to be repeated. Even after the setbacks of 2001, the prospective risk premium is markedly lower than the historical risk premium.