

# TRIUMPH OF THE OPTIMISTS

## 101 Years of Global Investment Returns

Régie de l'énergie
DOSSIER: R-3495 2003
DÉPOSÉE EN AUDIENCE
Date: 13 mars 2003
Pièces n°: H30-12

**Elroy Dimson**

**Paul Marsh**

**Mike Staunton**

doc. 2.1.6

**Princeton University Press**  
**Princeton, New Jersey and Oxford**

## Chapter 12 The equity risk premium

Investment in equities over the 101 years from 1900–2000 has proved rewarding but, as we have seen, has been accompanied by correspondingly greater risks. In this chapter we examine the historical rewards that investors have enjoyed for bearing this risk.

Clearly, investors do not like volatility—at least on the downside—and will be prepared to invest in equities only if there is some expected compensation for their risk exposure. We can measure the reward for risk that they have received in the past by comparing the return on equities with the return from risk free investments. The difference between these two returns is called the equity risk premium. The risk premium is typically measured relative to either government bills or bonds. In sections 12.1 and 12.2 we therefore analyze the risk premium that investors have enjoyed relative to bills, while in sections 12.3 and 12.4 we examine the risk premium relative to long bonds. In each case, we look first at the United States, and then at worldwide comparisons.

The equity risk premium is an extremely important economic variable. An estimate of the premium is central to projecting future investment returns, calculating the cost of equity capital for companies, valuing companies and shares, appraising investment proposals, and determining fair rates of return for regulated businesses. All of these applications require an estimate of the prospective risk premium, whereas, by definition, the only premium that we can measure is the historical risk premium.

In practice, and perhaps because of its measurability, the historical risk premium is often treated as a proxy for the prospective risk premium. It has certainly been by far the most influential variable in conditioning expert opinion about what the future premium might be. It is frequently assumed that if the measurement interval is long enough, the historical risk premium will provide an unbiased estimate of the future premium. In chapter 13 we will consider the extent to which our findings on the historical risk premium differ from prior estimates, and hence challenge established views about the magnitude of the premium. More fundamentally, we will also examine whether the historical premium really provides a reasonable estimate of the prospective premium. In this chapter, however, the focus is on the historical record of the risk premium.

### 12.1 US risk premia relative to bills

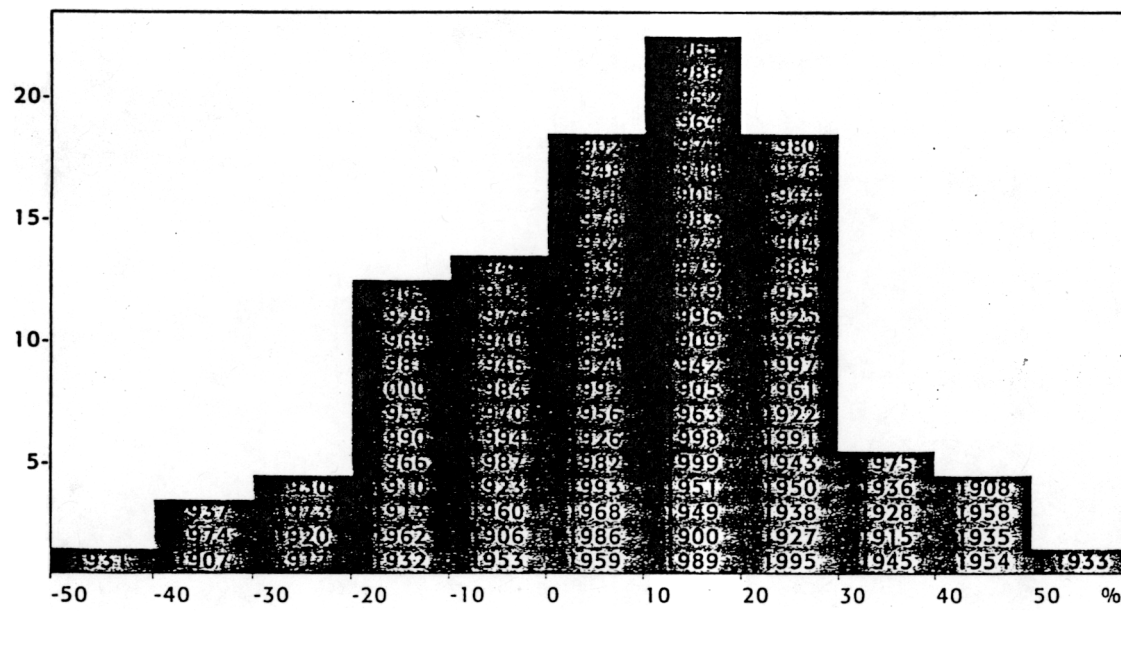
To establish whether equity risk has been adequately rewarded, we need to measure the equity risk premium. This is typically calculated in one of two ways. The first uses treasury bills (very short-term, default-free, fixed-income government securities) as the risk free or “safe” benchmark. The second measures the risk premium relative to long-term default-free government bonds. Of these two, only treasury bills can really be considered risk free, and even here chapter 5 reminded us that hyperinflation can on occasions cause even bill investors to experience large losses in real terms. In this section, we focus on the risk premium relative to bills, while in section 12.3 we contrast this with the premium relative to bonds.

Whether the premium is defined relative to bills or to bonds, we measure it by taking the geometric difference between the equity return and the risk free return. The formula is  $1 + \text{Equity rate of return}$  divided by  $1 + \text{Riskless return}$ , minus 1. For example, if shares with a one-year return of 21 percent were being evaluated relative to treasury bills yielding 10 percent, the equity risk premium would be 10 percent. This is because  $(1 + 21/100)$  divided by  $(1 + 10/100)$  is equal to  $(1 + 10/100)$ . The difference between the equity return of 21 percent and the treasury bill return of 10 percent could also be estimated by subtracting one return from another and the result (11 percent) would be similar.

These measures of the risk premium have no obvious numeraire in terms of currency since the equity risk premium, measured relative to bill or bond returns, is a ratio. It is hence unaffected by whether returns are computed in dollars or (say) pounds, or whether returns are computed in nominal or real (inflation adjusted) terms.

Figure 12-1 shows the year-by-year US risk premium relative to bills. The distribution of outcomes was wide, with the lowest and highest premia occurring, as might be expected, in the worst and best years for stocks. The lowest premium was -45 percent in 1931, when equities returned -44 percent and treasury bills 1.1 percent, while the highest was 57 percent in 1933, when equities gave 57.6 percent and bills 0.3 percent. Figure 12-1 shows that the distribution of annual premia is roughly symmetrical, resembling a normal distribution. The mean is 7.7 percent and the standard deviation is 19.6 percent. On average, therefore, US investors received a positive—and quite large—reward for exposure to equity market risk.

Figure 12-1: Histogram of US equity risk premium relative to treasury bills, 1900–2000



Because the range of premia that are encountered on a year-to-year basis is very broad, it can be misleading to label these as “risk premia.” Investors clearly cannot have expected, let alone required, a negative risk premium from investing in equities, otherwise they would simply have avoided them. All the negative and many of the very low premia shown in Figure 12-1 must therefore reflect nasty surprises. Equally, investors could not have “required” the very high premia, such as the 57 percent in 1933. Such numbers are implausibly high as a required reward for risk, and the high realizations must therefore reflect pleasant surprises. To avoid confusion, many writers choose not to refer to these as “risk premia,” but instead label them as “excess returns,” that is, returns in excess of (or under) the risk free interest rate.

To make sensible inferences about the historical risk premium, it is therefore necessary to look at much longer periods than a single year. Over longer horizons, we might expect good and bad luck to cancel each other out. A common choice of time frame is a decade, and while this is still far too short to produce reliable estimates, we adopt this as our intermediate interval for evaluating risk premia. For long-run comparisons, we use our full 101-year period from 1900–2000.

Over the full 101 years, the annualized (geometric mean) US equity risk premium relative to bills was 5.8 percent, which is 0.9 percentage points lower than the annualized real return on US equities. Figure 12-2 shows the annualized US equity risk premium and US real equity returns, both measured over a sequence of rolling ten-year periods. The two series have tended to track each other. In recent years, the markedly positive real returns on treasury bills have caused the equity risk premium to be lower than the real return on equities.

Figure 12-2: Rolling ten-year US equity real returns and risk premia relative to bills, 1909–2000

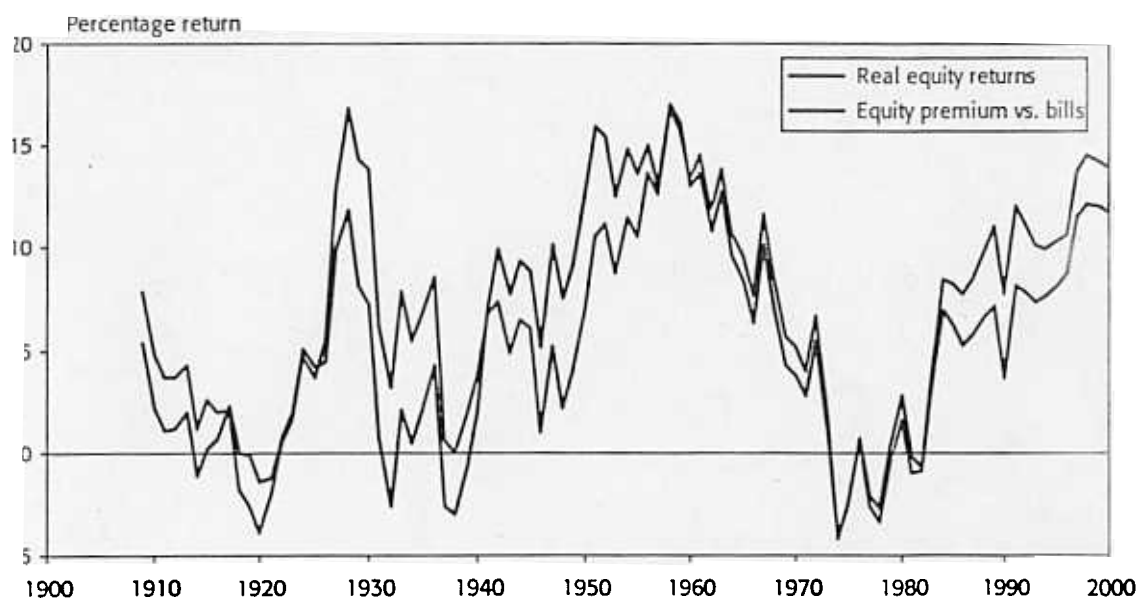


Figure 12-2 shows that even over intervals of a decade, there were periods when the *ex post* US risk premium was negative, most recently in the 1970s and early 1980s. There can be major performance surprises, even over a ten-year interval. It follows that we need very long periods to infer investors' expectations about the reward for exposure to equity market risk. Even then, inference can be problematic, as we will see in chapter 13.

## 12.2 Worldwide risk premia relative to bills

We have seen that US investors enjoyed a large, positive risk premium relative to bills with a 101-year arithmetic mean of 7.7 percent and a geometric mean of 5.8 percent. The US economy has been especially successful, however, and so we need to set the US record in context.

Figure 12-3 shows the annualized, geometric mean risk premia relative to bills for all sixteen countries over the full 101-year period. The bars plot the equity risk premia relative to bills, while the line plot overlay shows the real rate of return on equities. The chart also shows the “world” risk premium, based on our sixteen-country world equity index, calculated in US dollars, with risk premia and real returns computed relative to US bills and US inflation.

Figure 12-3 shows that while the equity risk premium varies across countries, the 101-year averages fall within a fairly narrow range. The US premium of 5.8 percent, while above average, is not the largest, with five other countries enjoying higher premia. France led the field with a premium of 7.4 percent, while the lowest premium was the 1.8 percent for Denmark. The UK risk premium of 4.8 percent was just below the figure of 4.9 percent on the world index, which, in turn, was 0.9 percent below the premium for the United States.

Figure 12-3: Worldwide annualized real equity returns and risk premia relative to bills, 1900–2000

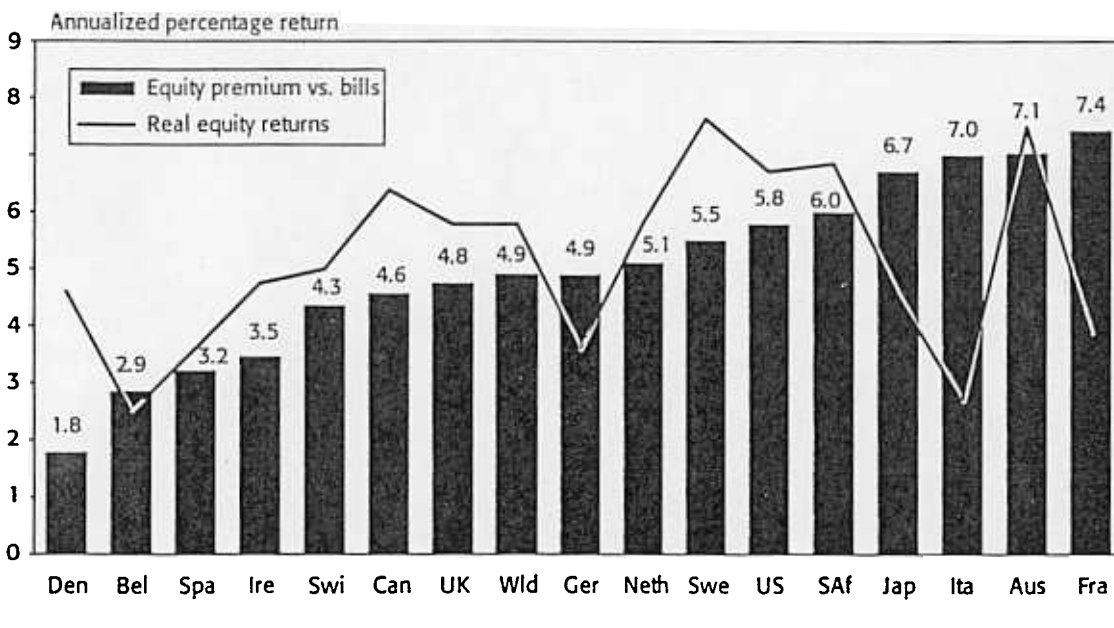


Figure 12-3 shows that the relationship between equity risk premia relative to bills and real equity returns can be complex. In most countries, real equity returns exceeded the risk premium, but in some cases, such as Italy, an above-average risk premium is associated with below-average real returns. The reason for this was identified in chapter 5. In countries like Italy, the returns on cash fell short of investors' *ex ante* expectations due to unexpectedly high inflation. It is therefore possible for an equity market that performs poorly (perhaps also as a consequence of inflation) to have a large *ex post* premium relative to bills.

The equity risk premia relative to bills for the sixteen countries and for the world index are listed in Table 12-1. Note that for Germany the calculations are based on 99, rather than 101 years, since we have excluded 1922–23. While we can calculate a continuous, 101-year real return series for German equities, calculating risk premia is more problematic. The hyperinflation of 1922–23 decimated bill and bond returns, resulting in -100 percent real returns in 1923. We therefore obtain more meaningful statistics by omitting these years.

The first six columns of Table 12-1 provide statistics on the annual risk premia. The geometric means in the first column provided the plotting data for the bars in Figure 12-3. The next five columns show the arithmetic means, standard errors, standard deviations, and maximum and minimum premia. We commented earlier that the distribution of annual risk premia for the United States—displayed in Figure 12-1—was very wide. The standard deviation figures in the fourth column of Table 12-1 show that the dispersion was even greater than in the United States in all but four of the other countries.

Table 12-1: Worldwide equity risk premia relative to bills, 1900–2000

Country	Annual equity risk premium relative to treasury bills						Ten year risk premium		
	Geometric mean	Arithmetic mean	Standard error	Standard deviation	Minimum return	Maximum return	Geometric mean	Arithmetic mean	Standard deviation
Australia	7.1	8.5	1.7	17.2	-30.2	49.4	6.9	7.0	4.2
Belgium	2.9	5.1	2.3	23.5	-38.1	120.6	2.8	3.0	5.4
Canada	4.6	5.9	1.7	16.7	-34.7	49.1	4.6	4.7	4.8
Denmark	1.8	3.4	1.9	19.4	-32.7	87.0	1.4	1.5	3.4
France	7.4	9.8	2.4	23.8	-33.4	78.7	7.5	7.7	7.4
Germany*	4.9	10.3	3.5	35.3	-87.2	165.3	8.1	8.3	8.2
Ireland	3.5	5.4	2.1	20.6	-51.1	73.6	3.2	3.3	4.4
Italy	7.0	11.0	3.2	32.5	-48.6	150.3	6.8	7.2	10.0
Japan	6.7	9.9	2.8	27.9	-48.3	108.6	7.1	7.4	8.7
The Netherlands	5.1	7.1	2.2	22.2	-31.3	126.7	4.7	4.9	6.4
South Africa	6.0	8.1	2.2	22.5	-33.9	106.2	6.6	6.7	5.0
Spain	3.2	5.3	2.1	21.5	-38.6	98.1	2.7	2.9	6.1
Sweden	5.5	7.7	2.2	21.9	-38.3	84.8	4.9	5.1	6.6
Switzerland <sup>†</sup>	4.3	6.1	2.0	19.4	-37.0	54.8	3.5	3.6	5.2
United Kingdom	4.8	6.5	2.0	19.9	-54.6	121.8	5.0	5.0	3.9
United States	5.8	7.7	2.0	19.6	-44.5	57.1	5.6	5.8	5.4
World	4.9	6.2	1.6	16.4	-39.8	70.9	5.0	5.1	4.7

All statistics for Germany exclude 1922–23. <sup>†</sup> Premia for Switzerland are from 1911.

Note that even with 101 years of data, the standard errors around the risk premia estimates are high, ranging from 1.6–1.7 percent (for the world index, Australia, and Canada) to 3.5 percent (for Germany). The standard error for both the United States and the United Kingdom is 2.0 percent. Strictly, these standard errors relate to the arithmetic mean, but as an approximate guide, we can apply them to the geometric mean. This means that while the US figure of 5.8 percent remains our best estimate, we can be only 68 percent confident that the true geometric mean lies within one standard error of this, that is, within the range  $5.8 \pm 2$  percent, or 3.8 to 7.8 percent. Note that the historical geometric mean risk premia of all but four of our sixteen countries fall within this range.

Political and economic conditions have varied over the last 101 years for every country. Some analysts prefer to emphasize recent data since it reflects the current environment. Others recommend using the longest possible data series since this is more likely to encapsulate unusual events that may have current or future relevance. Parallels have been drawn, for example, between the technology and communications revolution of the late twentieth century and the electricity and the automobile innovations of the early twentieth century; or between the nineteenth century railway bubble and the recent internet bubble; or between the 1990 invasion of Kuwait and ensuing Gulf War of 1991 and the September 11, 2001 attacks on the United States and the subsequent “War on Terrorism.” All market analysts agree, however, that repetition of certain historical events is so implausible that the past must be interpreted with care. Extreme hyperinflation is widely regarded as something that will not again afflict major economies, and a world war would be of a different nature if it were to happen in the future.

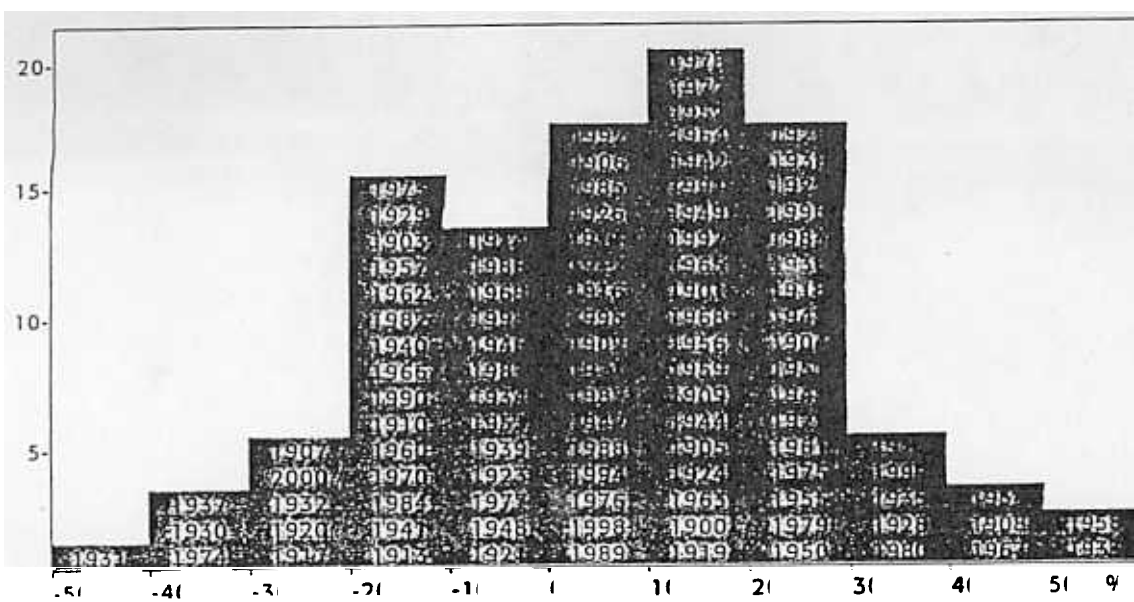
For these, or perhaps for less dramatic reasons, many people wish to review risk premia over shorter intervals than 101 years. We have selected decades as our intermediate time frame, which although arbitrary, is nevertheless a common choice. The final three columns of Table 12-1 present annualized statistics on premia that have been calculated over ten-year intervals. With 101 years of data, there are ninety-two (overlapping) decades for which we can compute premia. Table 12-1 shows the geometric means, arithmetic means, and standard deviations of these ten-year premia. By definition, the arithmetic mean always exceeds the geometric mean, but the difference between them is small when the measurement interval is as long as a decade. Note that, as with measures based on annual data, the United States ranks sixth while the United Kingdom falls in the middle of the distribution of premia for the countries in our sample.

The last column of Table 12-1 shows that over a ten-year horizon, the dispersion of returns is reduced. This is because there is diversification over time: one poor year is not necessarily followed by another. This gives rise to annualized returns that have a tighter distribution (a smaller standard deviation) than the underlying single-year returns. In the United States, for example, the single-year risk premia have a standard deviation of 19.6 percent while the ten-year premia have an annualized standard deviation of 5.4 percent. Assuming that ten-year annualized risk premia are approximately normally distributed, then in about two decades out of three, the annualized ten-year premium lies within  $\pm 5.4$  percent of the mean.

...long-term bonds do have one advantage as a benchmark in that their yields reflect not only today's short-term interest rate but also future expected interest rates. Thus for valuing shares or projects with cash flows extending many years into the future, the promised return on long bonds will encapsulate the expected sequence of returns on short-dated bills over the long term of the bond. The corresponding disadvantage is that long-bond prices are sensitive to changes in real interest rates and to inflationary expectations. Since bonds are riskier than bills, we would expect the equity risk premium relative to bonds to be lower than the premium relative to bills.

With these considerations in mind, we therefore compute both one- and ten-year equity risk premia relative to bonds. In addition, we also compute the average premium relative to bonds, which is 11.2%.

Figure 12.4: Histogram of US equity risk premium relative to bonds, 1900-2000

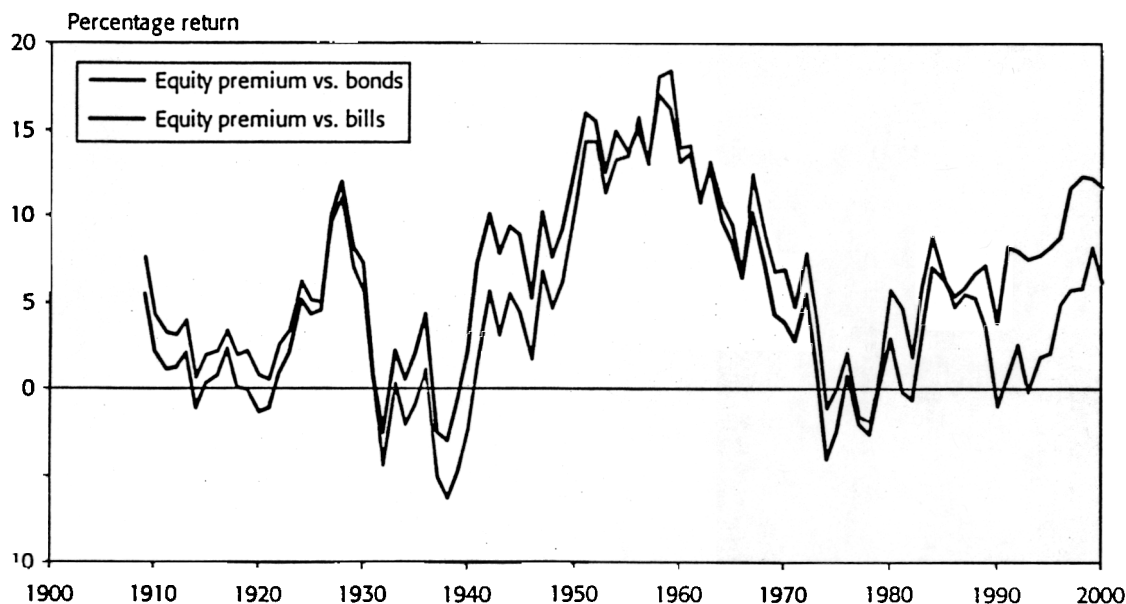


the one-year risk premia relative to bonds for the United States over the 101 years from 1900–2000. The average premium was 5.0 percent, which is below the figure of 5.8 percent relative to bills. The reason for this is that, as discussed in chapter 6, over this 101-year period the return on long-term US government bonds exceeded the return on treasury bills by 0.7 percent. Figure 12-4 shows that there is again a wide dispersion of year-by-year risk premia.

While the one-year equity risk premia in Figure 12-4 summarize historical returns over short periods, it is again helpful to look at longer-term premia. Figure 12-5 plots the rolling ten-year US risk premia relative to bonds over the 101 years from 1900–2000, contrasting this with the equivalent ten-year premia relative to bills. Obviously, over periods when bonds performed poorly, the equity premium relative to bonds has tended to exceed the premium relative to bills, and vice versa. This is especially noticeable over the second half of the twentieth century. During the higher inflation period in the United States from the mid-1950s through until the early 1980s, bond-based premia were consistently higher than bill-based premia. Subsequently, bonds have generally performed well, and from the mid-1980s the ten-year risk premium relative to bonds has been below that on bills.

Note that as with the risk premium relative to bills, there have been periods when the rolling ten-year US risk premium relative to bonds has been negative. Most recently, this occurred in 1990 and 1993. At first sight, this seems surprising since US equities have performed well in recent decades. But over the ten years until the end of 1990, although US equities returned a creditable 12.6 percent per year, US bonds gave an even higher annualized return of 13.7 percent.

Figure 12-5: Rolling ten-year US equity risk premia relative to bonds and to bills, 1909–2000



A similar situation has arisen recently in the United Kingdom. By 2000, the rolling ten-year UK risk premium relative to bonds had fallen to just under 2 percent. This led many UK market professionals to conclude that investors now required a risk premium of just 2 percent, and this was frequently used as a justification for the high stock prices at the turn of the millennium. While it may be true that required premia had become low (we address this issue in the next chapter), we cannot infer this from the recent ten-year historical premium relative to bonds. The main reason the ten-year premium is so low is that, over the decade ending in 2000, UK long bonds performed exceptionally well, almost certainly exceeding investors *ex ante* expectations. Over the same period in the United States, the historical ten-year risk premium relative to bonds was 6.1 percent per year, and yet US stocks were if anything more highly rated than their UK counterparts.

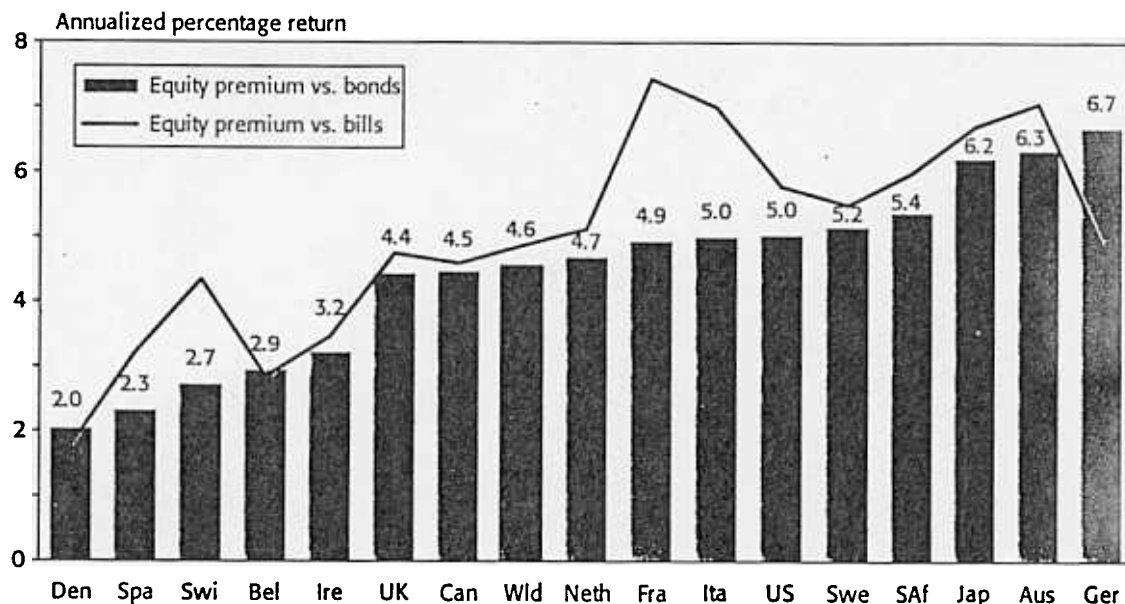
The problem that lies at the source of these confusions is that over many periods, even when we look at ten-year returns, the *ex post* returns on bonds appear to have differed markedly from *ex ante* expectations. One option would be to adjust bond returns for unexpected shocks in the bond market, such as unanticipated decreases or increases in inflation. Approximate adjustments might involve computing equity premia relative to each year's starting yield-to-redemption on bonds, or calculating premia relative to the income yield on bonds (excluding capital gains or losses). Although these computations become possible with the availability of our new data series, it can be misleading to compute premia after adjusting bond returns for shocks such as inflation increases or decreases, without also considering the impact of such shocks or other unexpected events on equity prices. The latter is far more challenging, as we will show in the next chapter.

Nevertheless, the experience of the 1980s and 1990s makes it obvious that we should not draw inferences about the expected future risk premium from periods as short as a decade. This is especially true when we have strong reasons to believe that during the decade in question, investors fared much better or worse than they had expected in advance. During the 1930s or mid-1970s, for example, it would have been extremely foolish to conclude that investors required a negative risk premium for investing in equities. These periods when the historical risk premia turned negative simply reflect the poor performance of equities over the previous decade. While less immediately obvious, interpreting the post-1980 era has been equally problematic, since this was a period that turned out to be particularly rewarding for both bond and equity investors.

## 12.4 Worldwide risk premia relative to bonds

Once again, it is important to set the US findings in their worldwide context. The bars and the line plot in Figure 12-6 show the geometric mean risk premia calculated both relative to bonds (the bars) and to bills (the line plot) over the full 101 years for all sixteen countries plus the world index. This chart shows that for all countries other than Germany, Denmark, and Belgium, the risk premium relative to bonds was lower than the premium relative to bills, consistent with our expectations, given that bonds are higher risk than bills. Figure 12-6 also

Figure 12-6: Worldwide annualized equity risk premia relative to bonds and bills, 1900–2000



shows that for several countries, in contrast to, say, the United States and the United Kingdom, the average risk premium estimated relative to bonds has differed markedly from the premium estimated relative to bills. These include Germany, where the premium relative to bonds is much higher, and France, Italy, and Switzerland, where the premium relative to bonds is much lower. Once again, as was the case with the risk premium relative to bills, the United States ranks in sixth place.

Two of the apparent “winners” in Figure 12-6 are Germany and Japan, which have the highest and third highest premium relative to bonds. France, on the other hand, has the highest premium relative to bills. But earlier, we saw that Sweden (see Figure 4-7) enjoyed the highest real equity return and the highest return in real dollar terms (see Figure 7-8). These at first sight conflicting accounts of which equity market performed best can be better understood when we recognize that Germany and Japan’s high premia relative to bonds arise from their low bond returns rather than high equity returns. Similarly, France’s high premium relative to bills largely reflects its low bill returns. This also helps explain why the premium relative to bonds for our world index seems comparatively low at 4.6 percent. This is because several of the higher-ranking countries earned their positions from the relatively poor showing of their bond markets rather than the strength of their equity markets.

A fuller record documenting equity premia relative to bonds is contained in Table 12-2. Note that the unanticipated losses experienced during the century’s worst inflation episodes afflicted bondholders more than shareholders. The premia shown in this table for those countries that suffered the worst bond returns are thus subject to caveats (see chapter 6).

**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

7.3 percent relative to bills for the period 1926–99. Before the publication of the research for this book, the most widely cited sources for the United Kingdom were Barclays Capital (1999) and CSFB (1999), which both started in 1919. Over the period 1919–98, there had been a UK risk premium relative to bills of 6.2 percent.

In recent years, both practitioners and researchers have grown increasingly uneasy about how to interpret these widely cited estimates, largely because they seemed too high. Apart from biases in index construction—a possibility that had not previously been thought important, but which we saw in section 3.2 is material in relation to the UK figure—the finger of suspicion has pointed mainly at success and survivorship bias among countries. The concern over success bias is that inferences about risk premia worldwide were being heavily influenced by the US experience, yet the United States has been an unusually successful economy. The closely related worry over survivorship bias is that previous attempts to place the experience of other countries like the United Kingdom alongside that of the United States may still have overstated the risk premium since they have focused on just a few selected markets that have survived, typically with continuous trading, over a long period.

To provide better estimates of the equity risk premium we therefore need to focus on the experience of all countries, not just the United States and the United Kingdom. If we look at all markets, then survivorship bias ceases to be an issue. Our sample of sixteen countries is by no means comprehensive. However, it does represent a large proportion by value of the world's stock markets in 1900. Fortunately, we are also able to compute total returns, including reinvested dividends, for this remarkably large sample of countries over a full 101-year period. This has allowed us to estimate long-run risk premia over an extended and uniform research period, thus overcoming another important but overlooked factor in previous studies, namely, easy data bias (see section 3.4). In other words, we mitigate the easy-data tendencies of previous researchers who measure a market's performance after a period of growth or recovery, omitting earlier periods of market turmoil.

We have found that from 1900–2000 the annualized equity risk premium relative to bills was 5.8 percent for the United States and 4.8 percent for the United Kingdom. Across all sixteen countries covered by this study, the cross-sectional (unweighted) average risk premium was 5.0 percent, while the risk premium on our size-weighted world equity index was 4.9 percent. Relative to long bonds, the story is similar. The annualized US equity risk premium relative to bonds was 5.0 percent, and the corresponding figure for the United Kingdom was 4.4 percent. Across all sixteen countries, the cross-sectional average risk premium relative to bonds was 4.5 percent, while for the world index it was 4.6 percent.

These figures are, of course, just long-run averages, and this chapter has also highlighted their high standard errors, the wide dispersion of annual equity risk premia, and how premia fluctuate over time as well as across markets. It has also reminded us that premia relative to bills and bonds must be interpreted with care, taking into account the unusual inflationary conditions earlier in the century and increases in the real interest rate more recently.

Nevertheless, two very important conclusions stand out. First, our evidence does not fully support Jorion and Goetzmann's (1999) claim that "the high equity premium obtained for US

[and, by implication, UK] equities appears to be the exception rather than the rule." While the United States and the United Kingdom have indeed performed well, especially the United States, there is no indication that they are hugely out of line with other countries. The big differences in long-term stock market performances are between the best and worst countries. The United States and the United Kingdom have experienced somewhat similar long-run equity risk premia, both to each other and to the global average. While US equities have performed very well, and the US risk premium was above the global average, it ranked only sixth out of sixteen countries. The United Kingdom is near the middle of the distribution of worldwide equity premia. Concerns about success and survivorship bias, while legitimate, may therefore have been somewhat overstated. In this sense, investors may not have been materially misled by a focus on the United States and, to a lesser extent, UK experience.

The second important conclusion is that the risk premia estimated in this study and reported in this chapter are around 1½ percent lower than those that have been reported in previous studies of long-term US and UK stock market performance. The differences here arise from previous biases in index construction (for the United Kingdom) and (for both countries) from the use of a rather longer time frame, extending back to 1900.

## Chapter 13 The prospective risk premium

In chapter 12, our focus was on the historical risk premium, namely, the premium that investors in different countries have obtained in the past from investing in equities, rather than in less risky assets such as treasury bills or government bonds. In this chapter, we switch from the past to the future, and to the prospective risk premium that investors can reasonably expect to obtain over future years and decades.

In the run-up to the millennial year, there was a series of books with titles like *Dow 36,000*, *Dow 40,000*, and even *Dow 100,000* (see, for example, Elias, 2000, and Kadlec, 1999). Glassman and Hassett (2000), in their book *Dow 36,000*, presented their case for why they believe equities are undervalued. Their view was that investors had come to require an extremely low rate of return on equity investments, and that this required return could get lower still. In short, they agreed that among investors, it was optimists who triumphed—investors in risky securities did well—but Glassman and Hassett believe that, historically, these investors should have done even better. Since stocks therefore remain underpriced in relation to what those authors consider the “perfectly reasonable price,” there are further gains still to come. Accordingly, they regard it as likely that in the future optimists will continue to triumph.

Investors' views on the risk premium can be thought of as a tug-of-war. Glassman and Hassett are at one end of the rope. At the other end is Shiller (2000). Shiller's view is essentially that the equity investors triumphed beyond what might reasonably have been expected: they benefited from the stock market moving to irrationally high levels. Shiller and his colleagues are the pessimists.

If everyone were an optimist, the market would climb new peaks. If everyone were a pessimist, the market would collapse right away. The consensus represents an equilibrium between optimists and pessimists. In estimating the consensus of the market, we therefore have to balance the arguments that are put forward by each side in the tug-of-war. We need a view of what premium the market is offering for risks that may be experienced in the future.

The question of what equity premium we can expect has, for years, been a source of controversy (see Kocherlakota, 1996 and Siegel and Thaler, 1997). We therefore start in section 13.1 by reviewing why the risk premium plays such a crucial role in the theory and practice of investment and corporate finance, and why it is the most important contemporary issue in finance. Section 13.2 takes a more theoretical perspective, asking what the alternative points of view are, and how big the premium should be. In section 13.3 we explain that, when we are concerned with the future expected risk premium, our focus should be on the expected arithmetic risk premium. We illustrate in section 13.4 how to derive plausible estimates of the arithmetic risk premium from the historical record presented in earlier chapters.

In section 13.5, we assess the views of academic experts on the likely magnitude of the future arithmetic risk premium. We conclude that the experts have been heavily influenced by the hitherto accepted wisdom on the historical risk premium. Since our current study finds that

risk premia appear to be lower than previously believed, we argue that experts might now wish to revise downward their long-term future risk premium forecasts.

We then look in section 13.6 at how using our historical analysis may more deeply inform us about likely future premia. We argue both that twentieth century equity returns almost certainly exceeded investors' *ex ante* risk premium requirements, and that the required risk premium itself fell over time. In section 13.7, by decomposing the historical risk premium into the contribution from unanticipated cash flow increases and the impact of a decline in required returns, we are able to infer the likely magnitude of the equity premium. In our concluding observations, in section 13.8, we summarize what we can learn from the historical record as an indicator of the future equity risk premium around the world.

### 13.1 Why the risk premium matters

Investors do not knowingly take on risk unless there is some expected recompense for their risk exposure. For taking on the risks of the equity market, this compensation takes the form of the equity risk premium. Why is it that the size of the equity premium has attracted so much attention? Perhaps the most straightforward answer is that the risk premium is fundamental to valuing financial assets, as will be clear with a simple example.

Consider the Gordon model, described in section 11.4. We will use this approach to value an investment such as buying the portfolio that comprises the Dow Index. Just before the World Trade Center tragedy, the Dow stood at a level of around 10,000 and its prospective dividend yield was 1.4 percent. So for a \$10,000 investment we expected to receive dividends of \$140. Since the dividend yield was much lower than the risk free interest rate, investors were presumably expecting Dow dividends to rise over time.

The Gordon model values this stream of dividends by assuming that dividends grow at a constant rate  $g$  and by discounting them at the expected rate of return  $r$ . This simplifies to the well-known constant-growth formula: *Value* is equal to *Dividend* divided by  $r - g$ . This formula makes a number of strong assumptions, such as that growth will continue indefinitely at the same unvarying rate. However, in the interests of pedagogy, we bypass these assumptions, and explore the implications of the Gordon model (Jagannathan, McGrattan and Scherbina (2001) present the multiperiod Gordon model).

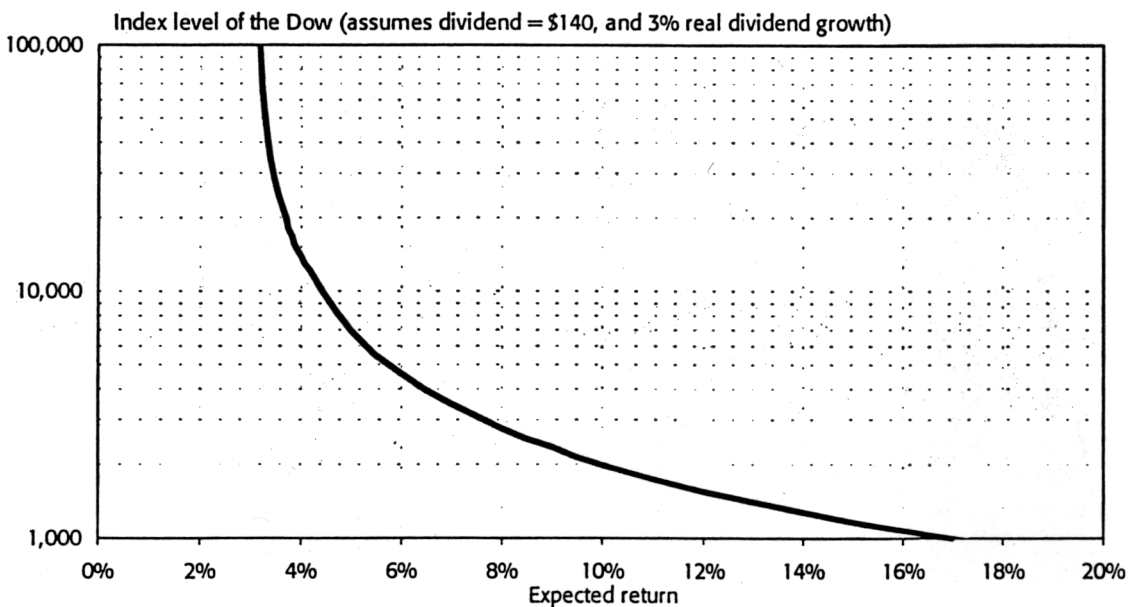
Let us assume real dividends are expected to grow indefinitely at the annual rate of  $g = 3$  percent, and that the real required rate of return is  $r = 4.4$  percent. So  $r - g$  is equal to 1.4 percent or 0.014. With these assumptions the valuation of the Dow is  $140/0.014 = 10,000$ , about its level in early September 2001. But what if dividends had been expected to grow at a real rate of 4 percent per year, so that  $r - g$  is equal to 0.004? That would give us Glassman and Hasset's valuation of the level for the Dow:  $140/0.004 \approx 36,000$ . Of course, this valuation of approximately 36,000 could equally well arise from a view that dividends were set to grow annually at 3 percent, while the required rate of return is not 4.4 but only 3.4 percent.

It is clear that given projected dividend growth, the correct value for the Dow depends on the return that investors expect from common stocks. The relationship is shown in Figure 13-1 which shows how the discount rate, which is the return expected by investors, influences the valuation given by the Gordon model. With dividends growing in real terms at 3 percent, valuations vary from infinity (if the expected return, like the dividend growth rate, is also 3 percent in real terms), through 36,000 (when the expected return is close to 3.4 percent) and 10,000 (if the expected return is 4.4 percent), to just 1,000 (if the expected return is 17 percent in real terms) or less. Values depend on the discount rate.

Using these admittedly crude figures, a tiny move of just half a percent in the return currently demanded by investors can shift the Dow a huge amount. A decline in the required return from 4.4 to 3.9 percent would have thrust the Dow upward by 55 percent. An increase in the required return from 4.4 to 4.9 percent would have been expected to precipitate a 26 percent collapse in the Dow, which is larger than the 14 percent fall in the ten days following the World Trade Center attack. The higher the valuation, the more volatile will be the index. So curiously, while low required rates of return drive up valuations, they also increase volatility. Investors are likely to demand a higher reward when the market is especially volatile. This provides a natural floor to required rates of return.

What required rate of return is appropriate? For providing an indication of the value of a stream of future cash flows, virtually no security is safer than a bond issued by the US Treasury. At the time of writing, the US government has to pay around 3 percent in real terms when it issues inflation-protected government bonds. The real redemption yield had been

Figure 13-1: Gordon model valuation of the Dow



well over 4 percent as recently as late 2000. To the risk free interest rate we may add some additional reward that is expected by investors as compensation for equity market risk. Clearly, the risk premium is the major determinant of asset values. If it is close to zero, the Glassman-Hassett optimists are correct; and if it is large, the Shiller pessimists are correct. This is why most finance professionals and financial economists regard the equity risk premium as the single most important number in finance.

To measure the equity market premium and to establish the reward for risk, we need to look at the markets where equity risk is traded, not only the United States but also other major global stock markets. That was the focus of chapter 12. An unbiased estimate of the *ex ante* risk premium required by investors tells us what equity market returns we can expect in the future relative to bills or bonds. A small risk premium automatically implies higher valuations but lower returns expected in the future from equities; a large risk premium automatically implies lower valuations but higher future returns from equities. If this were not the case, then, as we explained at the start of this chapter, the highly competitive conditions prevailing in the world's leading stock markets would ensure that share prices rapidly rise or fall until promised returns are aligned with required returns.

## 13.2 How big should the risk premium be?

We have explained the central role that the equity premium plays in determining stock prices. Other things being equal, the higher the required risk premium, the lower the value placed on a stream of cash flows. So when the required risk premium goes up, the realized, or *ex post*, risk premium is reduced over the transition period. Conversely, when the required premium goes down, the historical, *ex post* risk premium will be larger over the transition period. Once transition is complete, of course, equity returns will have an expected return that reflects the new required risk premium. Since transition from a high to a lower required risk premium may occupy many years, the historical, *ex post* premium may overstate consensus expectations. It is therefore important not to extrapolate unadjusted, historical equity premia into the future. Rather, one should use history as a means of informing the quest for a sensible estimate of the *ex ante*, forward looking premium.

We might wish to start with a direct estimate of the prospective risk premium. Unfortunately, we cannot *measure* the prospective risk premium since it relates to the unknown future. The prospective premium is the reward investors require for taking on risk. We could in principle simply ask investors. Indeed, we refer below to the results of investor surveys, but mostly they provide a source of amusement rather than useful information. Private investors' opinions appear to reflect what has happened in the recent past, or even sheer fantasy. Professional money managers seem (at best) to provide answers that reproduce, with some noise, the evidence from long-run historical studies. That approach is unlikely to inform us about the market's expectations.

An alternative approach involves analyzing the earnings, dividend, or cash flow projections of security analysts. The premium is inferred from the discount rate that equates these

future cash flows to ruling security prices. This approach, which we discuss in section 13.6, is also fraught with difficulty since there is no evidence that analysts' forecasts are unbiased predictors of the future.

In practice, therefore, we typically measure the historical risk premium, and use this as our starting point for inferences about the future. But to guide us, it is helpful to think about plausible magnitudes for the risk premium. This involves a brief detour into theory.

The starting point is, of course, the capital asset pricing model (CAPM), first formulated by Sharpe (1964). The CAPM was for most of the last three decades the method of choice among finance experts for estimating expected returns. The CAPM asserts that the expected return on a risky asset is the risk free rate of interest plus a risk premium. The latter is equal to the equity market risk premium, scaled up or down by the asset's beta. Beta measures the risk of the asset relative to the market portfolio, and beta estimates are widely available in most stock markets. The average beta for all stocks in a market is, by definition, equal to one. So the equity market risk premium is fundamental in determining CAPM expected returns.

The CAPM links the equity market risk premium to the risk tolerance of investors. The historical reward to equity market investment has, in the United States, looked large. But has it been excessive? The CAPM emphasizes how stock market investments contribute to the level of and uncertainty about an investor's wealth. To say whether the premium has been excessive, we need to look beyond fluctuations in an investor's wealth. What value is wealth unless it is consumed: spent on oneself, spent on others, or given as a bequest? Uncertainty about levels of future wealth leads to uncertainty about levels of future consumption. One variant of the CAPM—Breedon's (1979) consumption CAPM—emphasizes consumption by linking uncertainty about stock returns directly to uncertainty about future consumption.

If equities were to perform well at the same time as consumption peaks, then equities would be a poor hedge against the possibility of low consumption. If equities were to perform well when consumption dips, then equities would be a good hedge against cutbacks in consumption. The equity premium should be large if equities are a poor diversification tool. In other words, the equity premium should be larger when there is a high correlation between stock market returns and consumption.

The equity premium should also be larger when consumers are willing to pay more for insurance against the prospect of reduced consumption. That is, the equity premium should vary positively with the risk aversion of households as well as with the correlation of equity returns and consumption. In reality, however, aggregate consumption changes gradually over time, is not very volatile, and is not obviously correlated with stock market performance. There is a puzzle as to whether the large, historical US equity premium can conceivably be consistent with a focus by households on their potential future consumption. The "equity premium puzzle," first formulated by Mehra and Prescott (1985), thus remains a source of controversy. As Kocherlakota (1996) and Shleifer (2000) point out, traditional finance theory suggests that the equity premium should be much smaller than the (US) historical average.

As one might expect, there are competing theories that suggest differing magnitudes for the risk premium (see Cornell (1999) for a review). Though some writers may give another impression, there is no single figure for the risk premium that theory says is “correct.” Stewart (1991) writes: “Is there any fundamental reason why [the] market risk premium should be 6 percent? Not that I can figure. The question is a little like asking why did God make pi the number 3.14159... Don’t ask. Just memorize it.” If you are not convinced that theory provides a unique constant for you to commit to memory, then you too will wish to read the entrails of capital market history. In the following section, we therefore revert to deeper examination of the historical evidence.

### 13.3 Measuring the premium

We define the risk premium relative to either bills or bonds. For the sake of concreteness, this chapter emphasizes historical and prospective equity risk premia calculated relative to bills. We select bills because they are the more commonly used basis for defining the equity premium—reflecting the fact that only treasury bills can really be considered close to risk free. Nearly all of our discussion, however, applies equally well to premia relative to bonds. Where it is helpful, we supplement our analysis with premia estimated relative to bonds.

The risk premia reported in chapter 12 were computed as geometric means. This has intuitive appeal from an investment perspective. It corresponds to the annualized performance figures you see every day for mutual funds, for indexes, and for pension plans. When past performance is being considered, the geometric mean summarizes the annualized rate of return over a historical period. Computing the geometric mean requires us to calculate the product of  $1 + \text{First return}$  multiplied by  $1 + \text{Second return}$  and so on up to the  $n^{\text{th}}$  return. We then compute the  $n^{\text{th}}$  root of this product and deduct 1. The resulting geometric mean measures the annualized rate of return that equates the initial investment to the final value of a portfolio. The geometric mean risk premium has a similar interpretation. It is the incremental reward from investing in equities in preference to government securities.

We can simplify calculation of the geometric mean return over  $n$  years. In practice, all we require is two numbers, the final and the initial values of the portfolio. The geometric mean involves calculating the  $n^{\text{th}}$  root of *Final value* divided by *Initial value*, and then deducting the value 1. Think of a portfolio that appreciates over a period, such as 101 years, from \$1,000 to (say) \$10,000. It will have precisely the same annualized return regardless of whether the intermediate prices are stable or wildly volatile. By construction, the annualized return is the same over every one of the  $n$  years for which it has been calculated. We cannot meaningfully talk of the volatility of the geometric mean return. Volatility is a characteristic of the intermediate, year-by-year returns within our 101-year interval.

The year-by-year returns clearly experience considerable variability over time. We can capture this by estimating the average around which returns fluctuate, and the range of these fluctuations, that is, the arithmetic mean and the standard deviation of the annual returns. The arithmetic mean over  $n$  periods is the sum of all  $n$  returns, divided by  $n$ . The arithmetic

mean answers the question: If I chose a year at random, what return should I expect? The arithmetic mean may be calculated for premia as well as for raw returns.

The arithmetic mean of a sequence of different returns is always larger than the geometric mean. To see this, consider equally likely returns of +25 and -20 percent. Their arithmetic mean is 2½ percent, since  $(25 - 20)/2 = 2½$ . Their geometric mean is zero, since  $(1 + 25/100) \times (1 - 20/100) - 1 = 0$ . But which mean is the right one for discounting risky expected future cash flows? For forward-looking decisions, the arithmetic mean is the appropriate measure.

To verify that the arithmetic mean is the correct choice, we can use the 2½ percent required return to value the investment we just described. A \$1 stake would offer equal probabilities of receiving back \$1.25 or \$0.80. To value this, we discount the cash flows at the arithmetic mean rate of 2½ percent. The present values are respectively  $\$1.25/1.025 = \$1.22$  and  $\$0.80/1.025 = \$0.78$ , each with equal probability, so the value is  $\$1.22 \times \frac{1}{2} + \$0.80 \times \frac{1}{2} = \$1.00$ . If there were a sequence of equally likely returns of +25 and -20 percent, the geometric mean return will eventually converge on zero. The 2½ percent forward-looking arithmetic mean is required to compensate for the year-to-year volatility of returns.

The more volatile or risky the sequence of returns, the greater will be the difference between the two means. This is confirmed in Table 12-1. The arithmetic and geometric equity premia differ in the relatively low-volatility United States and the United Kingdom by less than 2 percent. The least volatile equity market has been Canada, whose annual risk premium, relative to bills, had a standard deviation since 1900 of 16.7 percent. The most volatile market was Germany with an annual standard deviation of 35.3 percent. For stable assets, the arithmetic mean risk premium is only slightly above its geometric counterpart; for very volatile assets the arithmetic mean premium can be much higher than the geometric mean. Canada has the smallest gap between its arithmetic and geometric means; Germany, the largest.

A reasonable approximation is that the arithmetic mean will exceed the geometric mean by one-half of the variance of annual returns. The variance is the square of the standard deviations presented in Table 12-1 (and also Table 12-2). The difference between arithmetic and geometric real returns therefore varies from Canada's 1.3 percent (which is very close to  $\frac{1}{2} \times .167^2$ ) to Germany's 5.4 percent (fairly close to  $\frac{1}{2} \times .353^2$ ). This approximation is very accurate if returns have a distribution that is close to lognormal. In practice the approximation is surprisingly accurate, not only for the United States but also for our fifteen other countries over the last 101 years. (A more precise formula is given in Levy and Gunthorpe, 1993.)

When we use 101 years to estimate means, the potential inaccuracy of the resulting risk premia remains high. The standard error measures this inaccuracy. It is approximately equal to one-tenth of the annual standard deviation of returns reported in Tables 12-1 and 12-2. In other words, the standard errors for the United States and the United Kingdom are 2 percent, and the range runs from 1.7 percent, for Australia and Canada, to 3.5 percent for Germany. This means that while the US arithmetic mean premium (relative to bills) has a

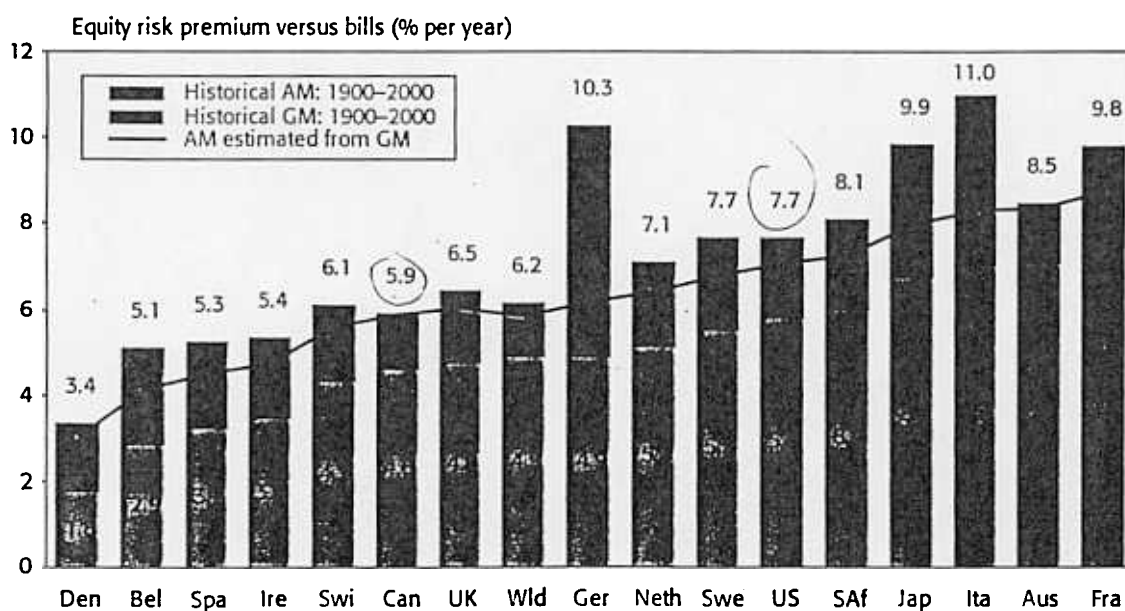
best estimate of 7.7 percent, we can be only two-thirds confident that the true mean lies within one standard error of this, namely, within the range  $7.7 \pm 2$  percent, or 5.7 to 9.7 percent. These high standard errors are the reason why conventional wisdom prescribes that the longest possible series of stock market data should be used to estimate risk premia.

As we explained earlier, one of the key uses of the equity risk premium is to determine investors' required returns, and hence the cost of equity capital. This is used, for instance, as a crucial input when estimating the discount rate for valuing shares, projects and other claims on uncertain future cash flows. In Tables 12-1 and 12-2, we report the historical arithmetic means, and associated standard deviations, for risk premia around the world. To build on that analysis, we now turn to expectations for the future.

### 13.4 Arithmetic and geometric premia

A crude estimate of the expected equity risk premium is the arithmetic mean of the one-year historical premia. In Figure 13-2, the full height of the bars shows the historical arithmetic mean premium relative to bills over the last 101 years for each of the sixteen countries. The US and UK premia are 7.7 and 6.5 percent, respectively, while the world equity risk premium is 6.2 percent. As we would expect, the arithmetic mean premia are noticeably higher than the geometric mean premia. Further, the arithmetic mean is at its largest, both in absolute terms and relative to the geometric mean, for the countries that suffered the greatest turmoil and hence volatility of returns over the last century (see the rightmost bars in Figure 4-14).

Figure 13-2: Arithmetic mean and geometric mean equity risk premia, 1900–2000



The historical arithmetic means in Figure 13-2 are thus clearly influenced by past levels of stock market volatility (among other factors). However, as demonstrated in chapter 8, estimates of the long-term historical standard deviation are not necessarily the best predictors of future volatility. We observed in chapter 12 that repetition of certain historical events is so unlikely that the past must be used with care as a guide to the future. There will be new uncertainties, but we know not what they will be. We therefore need estimates of expected future arithmetic risk premia, conditional on current predictions for market volatility.

It is not clear how investors should ideally adjust historical estimates of the equity risk premium to reflect today's judgments about stock market volatility. The approach we follow is to recalculate the arithmetic means, assuming current projections of early twenty-first century levels of volatility. To do this, we use the result that with lognormally distributed returns, the geometric and arithmetic means are linked by the standard deviation of returns, as described in section 13.3. We bypass some statistical assumptions that underpin our calculations, and which are discussed in Cooper (1996) and Dimson and Marsh (2001). We therefore recalculate the arithmetic mean premium for each country, replacing the historical difference between the arithmetic and geometric means with a difference based on risk estimates that are more contemporary.

For our estimates of arithmetic premia, we illustrate the approach using the same volatility for all national markets. That cannot be correct since markets inevitably expose investors to differing levels of risk. Nevertheless, for simplicity we assume a current volatility level for all sixteen national markets of 16 percent, and for the world index of 14 percent. It follows that our prediction of the difference between the arithmetic and geometric means is  $\frac{1}{2} \times \text{variance} = \frac{1}{2} \times 0.16^2 = .0128 = 1.28$  percent for each country and  $\frac{1}{2} \times 0.14^2 \approx 1$  percent for the world index. At this stage, we adjust the historical record only for volatility changes relative to the past.

Clearly, the volatility of one stock market is not in reality the same as another. Different stock markets have had differing risk levels in the past, and projections for the future should not be uniform. Chapters 4 and 8 show that national stock market risks differ in ways that relate to fundamental attributes, as well as to the impact of political and military history. Some countries are natural resource-based; others are more driven by technology. Some are more highly leveraged; others, less so. Some are concentrated in a narrow range of business activities; others are large, diversified markets which naturally tend to have a lower variability for the risk premium. It is impossible for us to make country-by-country risk predictions that will appear reasonable for the lifetime of this book. We therefore stress that assuming the same projected volatility for all premia is an expositional device, no more.

Given our predicted standard deviation for each country's risk premium, what, then, are our estimates of the arithmetic mean premia? These premia, measured relative to treasury bills, are represented graphically by the line-plot in Figure 13-2. The US equity risk premium is estimated at 7.1 percent, and that for the United Kingdom is 6.0 percent. Arithmetic mean

premia range from 3.1 percent (Denmark) to 8.7 percent (France); the worldwide arithmetic mean premium, relative to bills, is 6.2 percent.

The estimates of the adjusted arithmetic mean equity risk premia, relative to government bonds, are not displayed graphically. They are mostly somewhat lower. The US equity premium relative to bonds is estimated at 6.3 percent, and that for the United Kingdom is 5.7 percent; the range runs from 3.3 percent (Denmark) to 7.6 percent (Australia) or higher (Germany). The adjusted worldwide arithmetic mean premium, relative to bonds, is 5.6 percent.

To obtain an estimate of expected equity market returns, we require a forecast of the future arithmetic risk premium. For those who seek to extrapolate from the long-run series of historical premia, the numbers illustrated by the line plot in Figure 13-2 are the ones to use. For our world index, the adjusted arithmetic mean risk premium is 5.9 percent. Hence, the expected return is the risk free rate of interest, plus a premium of 5.9 percent. However, we can do better than simply extrapolate from the twentieth century to the twenty-first. By better understanding the sources of historical performance, we can extract more about investors' consensus expectations. This is the focus of the remainder of this chapter.

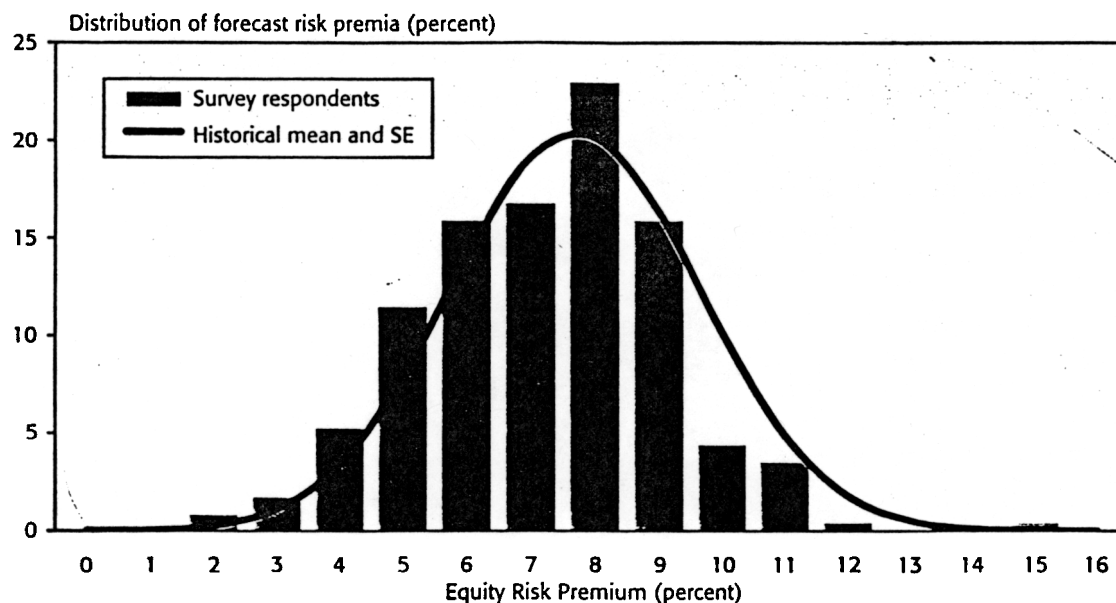
### 13.5 The changing consensus

The key question is whether the worldwide historical risk premium of approximately 6 percent should be our best estimate of the future risk premium. Many leading textbooks advocate the use of the arithmetic mean of historical equity premia, including Brealey and Myers (2000) and Bodie, Kane and Marcus (1999). Until recently, it was widely believed that the best predictor of the equity risk premium was its own past average. Certainly, researchers such as Goyal and Welch (1999) are unable to find variables that robustly predict the equity premium better than simply assuming that the premium will be "like it has been."

Welch (2000) casts light on whether academic finance professionals do, in fact, extrapolate from the historical record into the future. Welch studies the opinions of 226 financial economists who were asked to forecast the arithmetic equity risk premium in the United States. His findings are summarized in Figure 13-3, which shows that the mean forecast of the arithmetic thirty-year equity premium was 7.1 percent. The median was 7.0 percent, and the range ran from 1 to 15 percent. He also finds that the consensus view on a pessimistic outcome (with a 5 percent probability of occurrence) would be an equity premium of 2–3 percent; the consensus regarding an optimistic outcome is for a 12–13 percent premium.

The bars in Figure 13-3 represent the distribution from the Welch survey, while the curved line represents the normal distribution based on the historical mean and the standard error for the US equity risk premium using the full 101 years for which we have data (see chapter 12). An important aspect to grasp is of the inherent spread in both distributions. The uncertainty across financial experts about the risk premium is as large as the uncertainty that arises from statistical analysis of historical returns.

Figure 13-3: Financial economists' risk premium forecasts compared with US capital market history

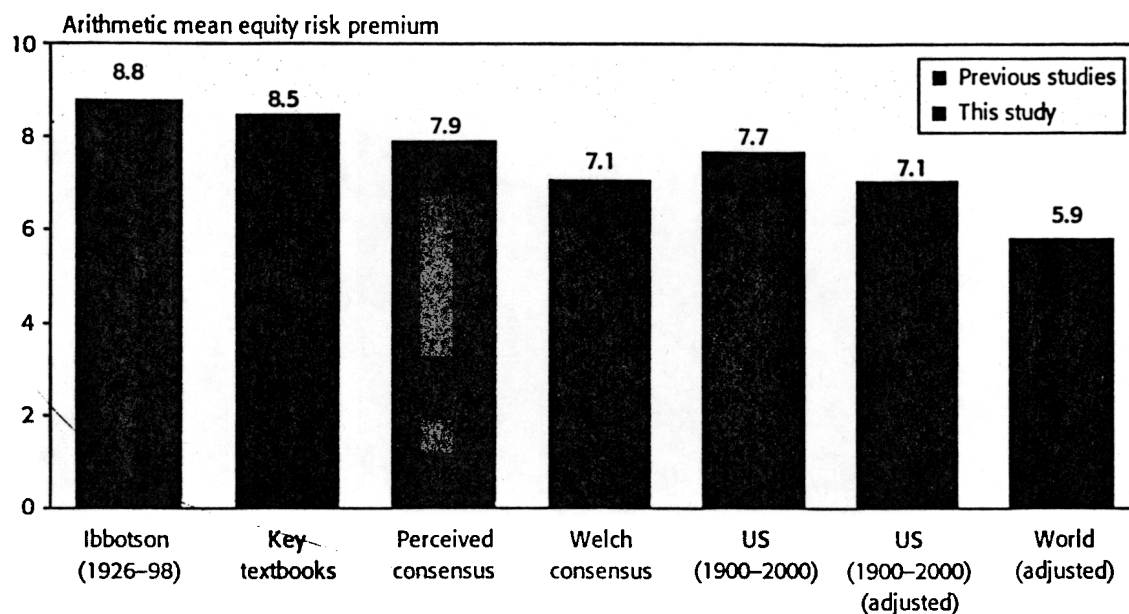


Most respondents to the Welch survey would have regarded the Ibbotson Associates' *Stocks, Bonds, Bills, and Inflation Yearbook* as the definitive study of the historical US risk premium. The first bar of Figure 13-4 shows that, at the time, the most recent long-run arithmetic risk premium computed from Ibbotson data from 1926–98 was 8.8 percent per year. The second bar shows that the key finance textbooks were on average suggesting a premium of 8.5 percent, a little below the Ibbotson figure. The textbook authors may have based their views on earlier, slightly lower, Ibbotson estimates, or else they were shading the Ibbotson estimates downward. The Welch survey mean is in turn lower than the textbook figure, but since respondents claimed to lower their forecasts when the equity market rises, this difference is probably explained by the market's strong performance in the 1990s. Interestingly, the third and fourth bars of the chart show that the survey respondents also perceived the profession's consensus to be higher than it really was. That is, they thought the mean was 0.5 to 1 percent higher than the average revealed in the Welch survey.

These survey and textbook figures represent what is being taught in the world's leading business schools and economics departments. As such, they are also widely used by finance professionals and corporate executives. Similarly, they are cited by regulators and in rate-of-return regulation disputes. Their influence extends from the classroom, to the boardroom, to the dealing room, to the courtroom.

Whether the Welch survey mean of 7.1 percent is appropriate is quite another matter. A large number of the survey respondents were clearly calibrating their forecasts relative to the longest-run historical benchmark available from Ibbotson, and then shading the historical number down based on subjective factors, including their judgment of the impact of recent

Figure 13-4: Alternative estimates of the arithmetic mean risk premium, relative to bills



strong market performance in the late 1990s. The Ibbotson Associates risk premium estimates are based on data from the Center for Research in Securities Prices (CRSP) data that starts in 1926. In the spirit of using the longest possible series of historical data, the fifth bar in Figure 13-4 shows that the arithmetic mean risk premium over the period 1900-2000, as estimated in this current study, is 7.7 percent.

The inclusion of twenty-six years of extra data from the start of the twentieth century (plus adding in 1999 and 2000) lowers the estimated premium by 1.1 percent. To the extent that their forecasts were anchored in long-run historical data, the Welch survey respondents might now wish to make a corresponding further downward revision in their forecasts of the equity risk premium. Welch (2001) is now updating his survey. If his new respondents were to make a full 1.1 percent reduction in their mean, the revised arithmetic mean would fall to 6.0 percent.

Of course, the 101-year mean of 7.7 percent shown in the fifth bar in our chart is the unadjusted historical mean. We argued above that we are likely to obtain more plausible estimates of the future arithmetic mean risk premium if we adjust the historical estimates downward to reflect today's current levels of equity market volatility. The sixth and seventh bars of Figure 13-4 portray the results of making such an adjustment for the United States and the world index. The adjusted US figure is 7.1 percent, while the world index provided a risk premium of 5.9 percent. To the extent that the Welch survey respondents were not already making an adjustment of this kind, this suggests that, in responding to the Welch (2001) survey, they should further lower their forecasts of the risk premium. This assumes, of course, that they still wish to use historical means as the anchor for their future forecasts.

### 13.6 History as a guide to the future

Financial economists may be reluctant to diverge markedly from the historical mean. Decision-makers, on the other hand, need to go beyond using only the past record, and to identify the market's implicit expectation for future performance. This is by no means easy. Recently, a number of researchers have sought to infer expected risk premia from surveys of the financial community, including Claus and Thomas (2001), Fraser (2001), Graham and Harvey (2001a), and Harris and Marston (2001). Unfortunately, such surveys are difficult to interpret. Graham and Harvey find that financial executives underestimate stock market risk substantially, while Brav and Lehavy (2001) and earlier researchers find that investment professionals are over-optimistic about the likely course of stock prices. Despite their wide dispersion, the responses to Welch's (2000) survey are about as usable as any.

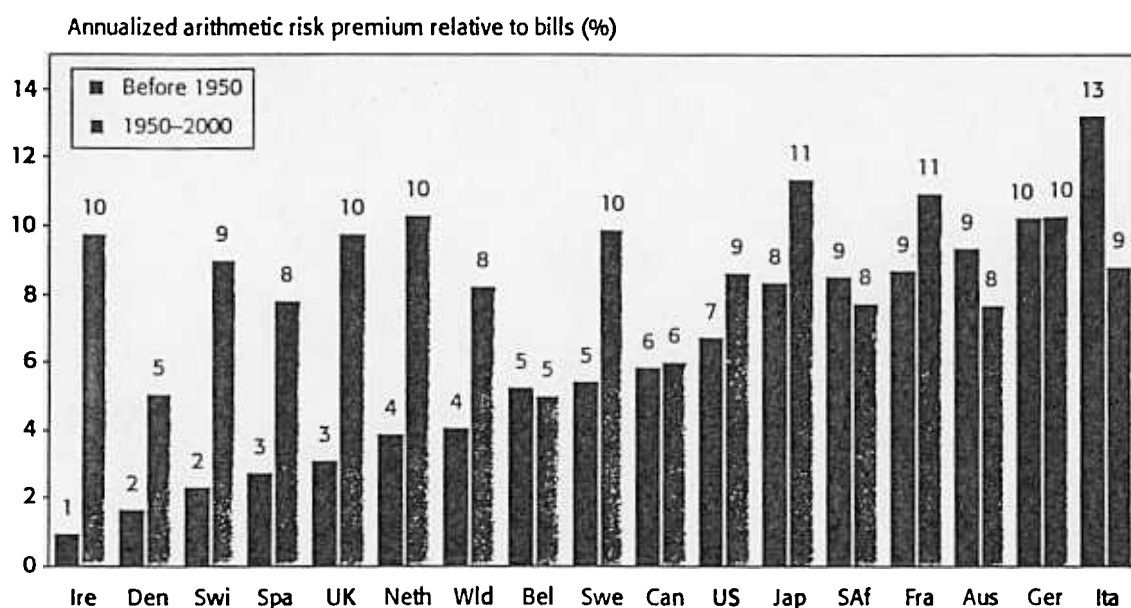
The wide dispersion of views reinforces the case for basing estimates on the historical record. However, since history may have been kind to (or harsh on) stock market investors, there are coherent arguments for going beyond raw historical estimates. First, the whole idea of using the achieved risk premium to forecast the required risk premium depends on having a long enough period to iron out good and bad luck. Yet as we noted in section 13.3, even with 101 years of data, standard errors remain high and our estimates of the average are imprecise.

Second, the expected equity risk premium could change over time. This might be because the underlying risk of equity investment has fluctuated, particularly as the variety of exchange-listed sectors has broadened out over the last century (see chapter 2). The corporate sector as a whole may have become more or less risky. The risks faced by investors have been transformed through reduced dealing costs and the advent of pooled investment vehicles. Enhanced diversification opportunities became available, both domestically and internationally. And as many households became wealthier, there may have been systematic changes in investors' levels of risk aversion.

Third, we must take account of the fact that stock market outcomes are influenced by many factors. Some that were important in the past (removal of barriers to international investment, for example) may be non-repeatable. If so, projections of the future risk premium should deviate from the past. The financial history of our sixteen stock markets has been so variable over time that it is worth exploring this argument further.

A comparison between the first and second halves of our 101-year period makes the point. Over the first half of the twentieth century, the arithmetic average US equity risk premium relative to bills was 6.7 percent, whereas over the period 1950–2000, it was 8.6 percent. As is clear from Figure 13-5, this pattern is common to most of the sixteen countries covered in this book, with Australia, Italy, Belgium, and South Africa being the exceptions. The cross-sectional mean for all countries in the first half of the twentieth century was an arithmetic average risk premium of 6.0 percent, versus 8.6 percent for the sixteen-country mean in the next fifty-one years. The corresponding averages for the equity premium relative to bonds are not shown here, but the pattern is similar: a pre-1950 mean of 5.5 percent as compared to 7.5 percent over the following fifty-one years.

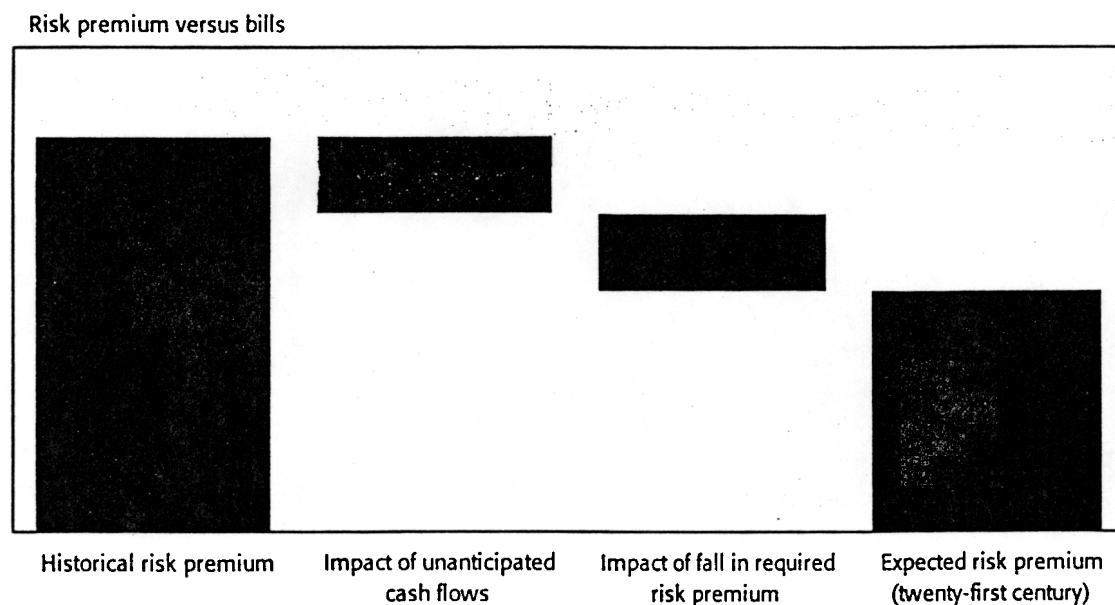
Figure 13-5: The equity premium relative to bills: first half-century versus next fifty-one years



The large risk premia achieved during the second half of the twentieth century are attributable to two factors. First, there was unprecedented growth in productivity and efficiency, acceleration in the pace of technological change, and substantial enhancements to the quality of management and corporate governance. As Europe, North America, and the Asia-Pacific region emerged from the turmoil of the Second World War, expectations for improvement were limited to what could be imagined. Reality almost certainly exceeded investors' expectations. Corporate cash flows grew faster than investors anticipated, and this higher growth is now known to the market and built into higher stock prices.

Second, as noted by Bernstein (1997) and others, stock prices have doubtless also risen because of a fall in the required rate of return due to diminished investment risk. The economic and political lessons of the twentieth century have surely been learned, international trade and investment flows have increased, and the Cold War has ended, leading to a more secure business environment. A further factor that must surely have lowered required returns is that investors now have much more opportunity to diversify, both domestically and internationally, than they had a century ago. Diversification allows investors to lower their risk exposure without detriment to expected return. Transaction costs are also lower now than a century ago. Factors such as these, which led to a reduction in the required risk premium, have contributed further to the upward re-rating of stock prices.

To convert from a pure historical estimate of the risk premium into a forward-looking projection, we need to reverse-engineer the factors that drove up stock markets over the last 101 years. The simplest idea would be to use the Gordon Model (from section 13.1) to infer the impact on returns of the historical changes in dividend yield. But we can go beyond this, as

Figure 13-6: *Ex post* and *ex ante* equity risk premia

illustrated conceptually in Figure 13-6. The left-hand bar in the chart portrays the historical risk premium on the equity market. This includes the contribution from unanticipated growth in cash flows and the gain from falls in the required risk premium. We therefore deduct the impact of these two factors. What remains in the right-hand bar of this chart is an estimate of the prospective risk premium demanded by investors as they look ahead to the remainder of the twenty-first century. The key qualitative point is that this is lower than the raw historical risk premium.

### 13.7 Expectations of the risk premium

To apply this framework, we need some notion of when cash flows (proxied here by equity dividends) have exceeded or fallen short of expectations. What dividends might investors expect each year? There are a variety of possible approaches (see, for example, Ibbotson and Chen, 2001). Perhaps the simplest approach is one that is commonly used today for forecasting the long-run dividend growth rate, namely, to extrapolate from previous long-term dividend growth. The long-term dividend growth rate is then used to project future real growth. That is, we estimate the product of  $1 + \text{Year 1 annual growth}$  multiplied by  $1 + \text{Year 2 annual growth}$  and so on to year  $n$ . We then compute the  $n^{\text{th}}$  root of this product, which is equal to  $1 + \text{Projected growth}$ . To summarize, we calculate the annualized real dividend growth rate to each year-end, over periods that start in 1900. The cumulative real dividend growth rates are illustrated for the United States and United Kingdom in Figure 11-3, and for other countries are summarized in Figure 11-5.

We assume that at every December 31, investors compare the year's real dividend growth to the real growth rate that would have been projected as at January 1 of that year. The difference is defined as  $1 + \text{Annual dividend growth}$  divided by  $1 + \text{Projected growth}$ , minus 1. This error in projecting dividend growth may be thought of as a naive measure of the unanticipated growth rate in dividends. The unanticipated changes in dividend growth are compounded together to produce an estimate of their annualized impact over the last century. This is clearly a rather unsophisticated measure of unanticipated real dividend growth, but it suffices to illustrate the general idea.

Defined this way, the impact in the United States of unanticipated dividend growth amounts to 0.2 percent per year over the period from 1900 to 2000. For the United Kingdom, the impact of unanticipated dividend growth is 1.7 percent per year. This large difference is explained in chapter 11 and reflects two factors. First, in the early decades of the last century, the United States generated more favorable growth expectations than did the United Kingdom; consequently, there was less scope in the United States for a favorable surprise. Second, over the last quarter-century in the United States, dividends became less important than in the United Kingdom as a result of growing rates of repurchases; this further reduced the scope for unanticipated dividend growth.

Since 1900, there has also been a dramatic change in the valuation basis for equity markets. The price/dividend ratio (the reciprocal of the dividend yield) is much higher now than it was in 1900. In the United Kingdom, for example, the price/dividend ratio at the start of 2001 was 42, as compared to 23 a century earlier. Undoubtedly, this change is in part a reflection of expected future growth in real dividends, so we could in principle decompose the impact of this valuation change into both an element that reflects changes in required rates of return, and an element that reflects enhanced growth expectations. To keep things simple, we assume that the increase in the price/dividend ratio is attributable solely to a long-term fall in the required risk premium for equity investment. Given this assumption, the fall in the required risk premium since 1900 is 1.4 and 0.6 percent per year for the United States and the United Kingdom, respectively. These amounts—which arise from the change in the price/dividend ratio—need to be deducted from the historical risk premium to obtain a measure of investors' *ex ante* expectations.

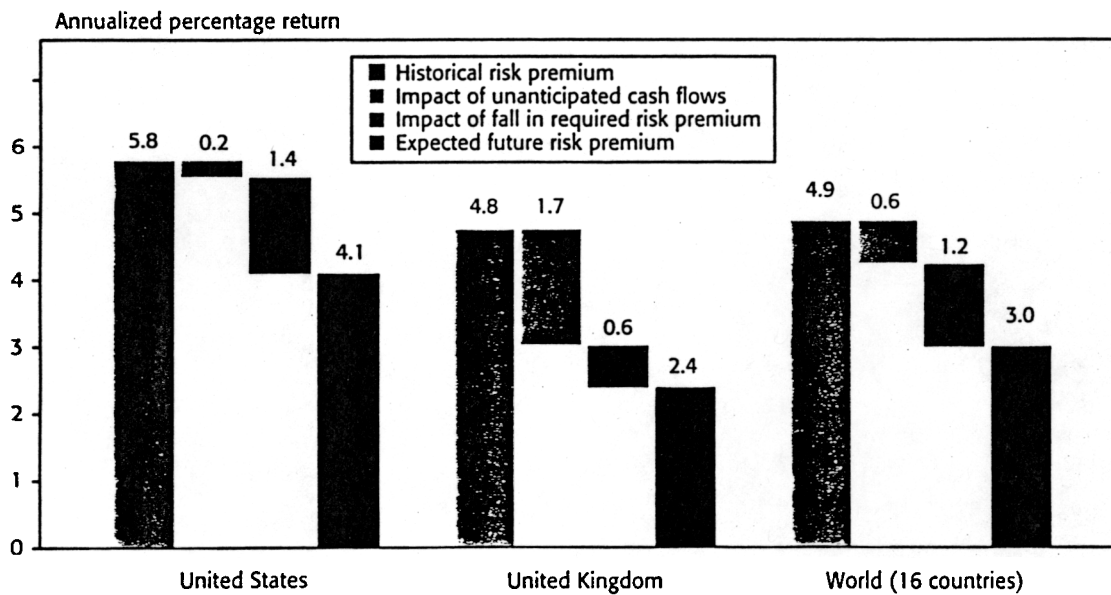
The remainder of the annualized risk premium represents the expected reward for equity investment. Even though we are deviating from some of its underlying assumptions, we can think of the risk premium in terms of the Gordon model for equity valuation. The Gordon model asserts that the expected equity return is equal to the prospective dividend yield plus the projected long-term dividend growth rate. We have already abstracted for the dividend yield, so for a given yield the Gordon model equates the expected return to the long-term dividend growth rate. It follows that the expected risk premium, measured relative to the riskless interest rate, is the projected dividend growth rate, also measured relative to the real interest rate. That is,  $1 + \text{Projected dividend growth}$  divided by  $1 + \text{Riskless rate of interest}$ , minus 1. Linking together these premia for each year after 1900, the expected US and UK equity risk premia are respectively 4.1 and 2.4 percent per year.

More complex decompositions of the historical risk premium are possible; see, for example, Arnott and Bernstein (2001) and Ibbotson and Chen (2001). These analyses, however, all point in the same direction. So using our simplistic decomposition of the backward-looking risk premium, we can now show how the historical premium can be broken down into a series of elements. For the United States, United Kingdom and for our sixteen-country world index, the risk premium estimates are depicted in Figure 13-7.

The differences between the estimates for each index should not be taken too seriously, especially since US corporations pay such low levels of dividends that computations based on dividend growth must be interpreted with care (see section 11.2). We acknowledge that our dividend projections are simplistic, and the reader should not put too much weight on cross-country differences. The key point is that the expected equity risk premium, based on the last century, should be lower than our backward-looking, historical averages suggest.

Our estimates indicate a geometric mean premium for the United States of 4.1 percent, for the United Kingdom of around 2.4 percent, and for our sixteen-country world index of about 3.0 percent. This is similar to the Fama and French (2002) estimate for the United States using a related approach. Also based on dividend yields and dividend growth estimates, Fama and French use the Gordon model to compute the US equity premium from 1872–1999. They find a premium of 3.8 percent before 1949, and a premium of 3.4 percent for the subsequent period. They argue that the difference between these estimates and the larger *ex post* risk premium based on historical realized returns is attributable to a reduction, since 1949, in investors’ required rate of return. Our analysis is consistent with their assertions.

Figure 13-7: *Ex post* and *ex ante* equity risk premia for selected indexes



If they are to be used as a prospective risk premium, our annualized figures need to be converted into arithmetic means, as outlined in section 13.4. Using a projected standard deviation for US and UK equities of 16 percent, the prospective *arithmetic* risk premium is 5.4 percent for the United States and 3.7 percent for the United Kingdom. With a slightly lower standard deviation for the world index of 14 percent, the prospective *arithmetic* risk premium for the world index is 4.0 percent. In fact, whichever country one focuses on, our forward-looking predictions for the equity risk premium are lower than the historically based projections reviewed in section 13.4.

There is scope to finesse our estimates of the expected risk premium. Some of the opportunities to enhance our estimates are the following: First, as explained in chapters 18–33, for many countries we use index series that are spliced together, and this presents difficulties when estimating dividend growth rates. Second, we ignore changes in the tax treatment of dividend income, both cross-sectionally and over time. Third, we estimate annual dividends from annual index returns and capital appreciation, rather than using more frequent (say, monthly) data. Fourth, we use a simplified decomposition of equity risk premia that could easily be enriched. Fifth, and very importantly, we use a naive model of investors' dividend expectations, an issue that merits further research (some of which we present elsewhere; see Dimson and Marsh, 2001). We leave these enhancements as a task for future researchers.

Finally, note that the use of historical averages as indicators of current required returns suggests that France may have a higher equity risk premium, while Denmark's risk premium may be lower (see Figure 13-2). There are obviously differences in risk between markets, but this is unlikely to account for cross-sectional differences in historical premia. Indeed, much of the cross-country variation in historical equity premia is attributable to country-specific historical events that will not recur. When making future projections, there is a strong case, particularly given the increasingly international nature of capital markets, for taking a global rather than a country-by-country approach to determining the cost of capital.

## 13.8 Summary

The equity premium is the difference between the return on risky stocks and the return on safe bonds. The equity market premium is central to investment, financing, and saving decisions. It is often described as the most important number in finance. Yet it is not clear how big the equity premium is today, or how large it has been in the past.

Survey evidence on best practice by corporations and financial advisors, such as Bruner, Eades, Harris, and Higgins (1998), reveals that nearly all respondents estimate the equity risk premium by averaging past data. However, there is considerable variation in their choice of time period and in their method of averaging, whether arithmetic or geometric. Their judgments are usually underpinned by knowledge of one, or at most two, national exchanges: respondents invariably focus on their domestic and/or the US stock market. In this chapter we use evidence from the past to provide a guide to what might be expected in the future.

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.