Equity Premium: Historical, Expected, Required and Implied

Pablo Fernández*
PricewaterhouseCoopers Professor of Corporate Finance
IESE Business School
Camino del Cerro del Aguila 3. 28023 Madrid, Spain
Telephone 34-91-357 08 09. e-mail: fernandezpa@iese.edu

ABSTRACT

The equity premium designates four different concepts: Historical Equity Premium (HEP); Expected Equity Premium (EEP); Required Equity Premium (REP); and Implied Equity Premium (IEP). We highlight the confusing message in the literature regarding the equity premium and its evolution. The confusion arises from not distinguishing among the four concepts and from not recognizing that although the HEP is equal for all investors, the REP, the EEP and the IEP differ for different investors.

A unique IEP requires assuming homogeneous expectations for the expected growth (g), but we show that there are several pairs (IEP, g) that satisfy current prices. We claim that different investors have different REPs and that it is impossible to determine the REP for the market as a whole, because it does not exist. We also investigate the relationship between (IEP – g) and the risk free rate.

There is a kind of schizophrenic approach to valuation: while all authors admit different expectations of equity cash flows, most authors look for a unique discount rate. It seems as if the expectations of equity cash flows are formed in a democratic regime, while the discount rate is determined in a dictatorship.

JEL Classification: G12, G31, M21

Keywords: equity premium; equity premium puzzle; required market risk premium; historical market risk premium; expected market risk premium; risk premium; market risk premium; market premium.

February 16, 2007

* I would like to thank, without implicating, Michael Brennan, Jose Manuel Campa, George Constantinides, Javier Estrada, Christophe Faugere, Roger Ibbotson, Juan Ignacio Peña, Julio Pindado, Jay Ritter, Jake Thomas and Tuomo Vuolteenaho for feedback and very helpful comments, and to my research assistant, Jose Maria Carabias, for his wonderful assistance.

Electronic copy of this paper is available at: http://ssrn.com/abstract=933070
1. Introduction

The equity premium (also called *market risk premium, equity risk premium, market premium* and *risk premium*) is one of the most important, but elusive parameters in finance. Some confusion arises from the fact that the term equity premium is used to designate four different concepts:

1. **Historical** Equity Premium (HEP): historical differential return of the stock market over treasuries.
2. **Expected** Equity Premium (EEP): expected differential return of the stock market over treasuries.
3. **Required** Equity Premium (REP): incremental return of the market portfolio over the risk-free rate required by an investor in order to hold the market portfolio. It is needed for calculating the required return to equity (cost of equity). The CAPM assumes that REP and EEP are unique and that $\text{REP} = \text{EEP}$.
4. **Implied** Equity Premium (IEP): the required equity premium that arises from a pricing model and from assuming that the market price is correct.

The four concepts are different. The HEP is easy to calculate and is equal for all investors, but the REP, the EEP and the IEP are different for each investor and are not observable magnitudes. We also claim that there is not an IEP for the market as a whole: different investors have different IEPs and use different REPs. A unique IEP requires assuming homogeneous expectations for the expected growth (g), but there are several pairs (IEP, g) that satisfy current prices.

An anecdote from Merton Miller (2000, page 3) about the expected market return in the Nobel context: “I still remember the teasing we financial economists, Harry Markowitz, William Sharpe, and I, had to put up with from the physicists and chemists in Stockholm when we conceded that the basic unit of our research, the expected rate of return, was not actually observable. I tried to tease back by reminding them of their neutrino—a particle with no mass whose presence was inferred only as a missing residual from the interactions of other particles. But that was eight years ago. In the meantime, the neutrino has been detected”.

Different authors claim different relations among the four equity premiums defined above. These relationships vary widely:

- **HEP = EEP = REP** according to Brealey and Myers (1996); Copeland *et al* (1995); Ross *et al* (2005); Stowe *et al* (2002); Pratt (2002); Bruner (2004); Bodie *et al* (2003); Damodaran (2006); Goyal and Welch (2006); Ibbotson Ass. (2006).

---

1 Or the extra return that the overall stock market must provide over the Government Bonds to compensate for the extra risk.
2 We agree with Bostock (2004) when he says that “understanding the equity premium is largely a matter of using clear terms”.
3 Provided they use the same time frame, the same market index, the same risk-free instrument and the same average (arithmetic or geometric).
Bostock (2004, 0.6 to 1.8%); Dimson et al (2006c, 3 to 3.5%); Siegel (2005b, 2 to 3%); Ibbotson (2002, < 4%); Campbell (2002, 1.5 to 2%).

- **EEP is near zero** according to McGrattan and Prescott (2001); Arnott and Ryan (2001); Arnott and Bernstein (2002).

- Authors that try to find the EEP doing surveys, as Welch (2000, 7%); Welch (2001, 5.5%); Graham and Harvey (2000, 4.65%); Graham and Harvey (2005, 2.93%); O'Neill et al (2002, 3.9%).

- There is a unique IEP and REP = IEP, according to Damodaran (2001a); Arzac (2005); Jagannathan et al (2000); Harris and Marston (2001); Claus and Thomas (2001); Fama and French (2002); Goedhart et al (2002); Harris et al (2003); Vivian (2005).

- Authors that “have no official position”, as Brealey and Myers (2000, 2003, 2005).

- Authors that claim “that no one knows what the REP is”, as Penman (2003).

- Authors that claim that “it is impossible to determine the REP for the market as a whole, because it does not exist”, as Fernandez (2002).

- Authors that claim that “different investors have different REPs”, as Fernandez (2004).

The rest of this paper is organized as follows. In section 2 we revise different estimates of the Historical Equity Premium (HEP), note that not all the authors get the same result for the HEP, and analyze the data. We highlight the change in the market around 1960. Before that date, the dividend yield was higher than the risk-free rate, but after that date has been always smaller. In sections 3 and 4 we discuss different estimates of the Expected Equity Premium (EEP) and of the Required Equity Premium (REP). In section 5 we revise the equity premium puzzle. Section 6 is a revision of the prescriptions of the main finance textbooks about the risk premium. We highlight the confusing message of the textbooks regarding the equity premium and its evolution. In section 7, we show that there are several pairs (IEP, g) that explain current market prices and we argue that there is no a REP for the market as a whole, but rather different investors use different REPs. We also show a positive relationship between (IEP – g) and the risk free rate after 1960. Section 8 explains which REP uses the author. Finally, section 9 concludes.

### 2. Historical Equity Premium (HEP)

The HEP is the historical average differential return of the market portfolio over the risk-free debt⁴. The most widely cited source is Ibbotson Associates whose U.S. database starts in 1926. Another frequently used source is the Center for Research in Security Prices (CRSP) at the University of Chicago.

---

⁴ This average differential return may be arithmetic or geometric. Different stock market indexes are used as the market portfolio, and Government bonds of different maturities are used as risk-free debt. A good discussion of the geometric and arithmetic average is Jacquier, Kane, and Marcus (2003).
2.1. First studies of the historical equity return

Smith (1926) made the first empirical estimate of the long run return on stocks (only price changes) for the most actively traded stocks from 1901 to 1922, and showed that an equity investor (even without market timing or stock selection ability) outperformed a bond investor over this period5.

Cowles (1939) published the first empirical study carefully done on the performance of the stock market. Cowles calculated the total return to equity from 1872 to 1937 for the NYSE, documenting a positive long term equity performance.

Fisher and Lorie (1964), using for the first time the database of stock prices completed at the University of Chicago's Center for Research in Security Prices (CRSP), showed that the average return from a random investment in NYSE stocks from 1926 to 1964 was 9.1% a year6.

2.2. Estimates of the historical equity premium of the US

Table 1 contains the 1926-2005 average returns and HEP for the US according to Ibbotson Associates (2006). The HEP in table 1 is the difference between the average return on the S&P 500 and the return of Gov. Bonds or T-Bills. However, Ibbotson Associates (2006, page 73), use the income return (the portion of the total return that results from a periodic bond coupon payment) of the Gov. Bonds (5.2%) and consider that the relevant HEP during the period 1926-2005 is 7.1% (12.3-5.2).

Schwert (1990) and Siegel (1994, 1999, 2002, 2005a) studied the relationship between U.S. equity and bonds before 1926. The data on which they base their studies is less reliable than recent data, but the results are interesting, nevertheless. Table 2 shows their conclusions: the HEP and the inflation in the period 1802-1925 were substantially smaller than in subsequent years7. Note that table 1 provides a higher HEP than table 2 for the period after 1926 because Ibbotson do not consider the income return of the bonds.

Wilson and Jones (2002) provide a monthly stock price index from 1871 through 1999. They note that the S&P Index returns have often been misrepresented8 and reconstruct the weekly S&P Composite for the period 1926-56 containing more than 400 stocks (instead of 90 as the daily S&P Composite). They get some differences versus other used indexes that are summarized on table 3.

Ibbotson and Chen (2003) use 1926-2000 historical equity returns and conclude that the expected long-term equity premium (relative to the long-term government bond yield) is 5.9% arithmetically, and 3.97% geometrically.

---

5 Three years after publication, the market crash happened. Benjamin Graham blamed Smith's book for inspiring an “orgy of uncontrolled speculation”.

6 For a more detailed history see Goetzmann and Ibbotson (2006).

7 Siegel (1999) argues that this is because bond returns were exceptionally low after 1926, while total equity returns were relatively stable over the whole time period.

8 Standard & Poor's first developed stock price indices in 1923 and in 1927 created the Composite Index (90 stocks). On 1 March 1957, the Composite was expanded to 500 stocks and renamed S&P 500 Index (its market value was $173 billion, 85% of the value of all NYSE listed stocks). From 1926 to 1957 there were 2 different S&P Composite indexes: one was weekly and the other was daily. The S&P Composite daily covered 90 stocks until 1957; The S&P Composite weekly covered more than 400.
Goetzmann and Ibbotson (2006) employ a new NYSE database for 1815–1925 to estimate the U.S. equity returns and the HEP since 1792 (but they mention that dividend data is absent pre-1825, and is incomplete in the period 1825–71). Their main results are in table 4.

### Table 1. Returns and HEP according to Ibbotson Associates (2006). 1926-2005

<table>
<thead>
<tr>
<th>Nominal Returns 1926-2005</th>
<th>Average return</th>
<th>Standard deviation</th>
<th>Serial correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arithmetic</td>
<td>Geometric</td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>12.3%</td>
<td>10.4%</td>
<td>20.2%</td>
</tr>
<tr>
<td>Income</td>
<td>4.2%</td>
<td>4.2%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Capital appreciation</td>
<td>7.8%</td>
<td>5.9%</td>
<td>19.5%</td>
</tr>
<tr>
<td>Long-Term Gov. Bonds</td>
<td>5.8%</td>
<td>5.5%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Income</td>
<td>5.2%</td>
<td>5.2%</td>
<td>2.7%</td>
</tr>
<tr>
<td>Capital appreciation</td>
<td>0.5%</td>
<td>0.4%</td>
<td>4.4%</td>
</tr>
<tr>
<td>T-Bills</td>
<td>3.8%</td>
<td>3.7%</td>
<td>3.1%</td>
</tr>
<tr>
<td>Inflation</td>
<td>3.1%</td>
<td>3.0%</td>
<td>4.3%</td>
</tr>
</tbody>
</table>

HEP over Gov. Bonds: 6.5\%  HEP over T-Bills: 8.5\%  

### Table 2 - Real returns and HEP from Siegel (2005a)

<table>
<thead>
<tr>
<th>Average real returns (%)</th>
<th>Stocks</th>
<th>Bonds</th>
<th>HEP (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>arith.</td>
<td>geom.</td>
<td>arith.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1802-1870</td>
<td>8.28</td>
<td>7.02</td>
<td>5.11</td>
</tr>
<tr>
<td>1871-1925</td>
<td>7.92</td>
<td>6.62</td>
<td>3.93</td>
</tr>
<tr>
<td>1926-2004</td>
<td>8.78</td>
<td>6.78</td>
<td>2.77</td>
</tr>
<tr>
<td>1802-2004</td>
<td>8.38</td>
<td>6.82</td>
<td>3.88</td>
</tr>
</tbody>
</table>

### Table 3. Geometric average of the returns of different indexes in selected periods

<table>
<thead>
<tr>
<th>(%)</th>
<th>Cowles</th>
<th>S&amp;P</th>
<th>Wilson and Jones</th>
<th>Ibbotson</th>
<th>CRSP NYSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1871-1925</td>
<td>7.24</td>
<td>7.28</td>
<td>7.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1926-1940</td>
<td>3.27</td>
<td>4.20</td>
<td>3.23</td>
<td>4.04</td>
<td>3.01</td>
</tr>
<tr>
<td>1941-1956</td>
<td>15.60</td>
<td>15.20</td>
<td>16.11</td>
<td>15.36</td>
<td></td>
</tr>
<tr>
<td>1957-1999</td>
<td>12.10</td>
<td>12.28</td>
<td>12.24</td>
<td>11.79</td>
<td></td>
</tr>
<tr>
<td>1926-1999</td>
<td>11.08</td>
<td>11.00</td>
<td>11.35</td>
<td>10.70</td>
<td></td>
</tr>
<tr>
<td>1871-1999</td>
<td>9.51</td>
<td>9.40</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4. Average return of the US according to Goetzmann and Ibbotson (2006)

<table>
<thead>
<tr>
<th>(%)</th>
<th>1972-1925</th>
<th>1926-2004</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arithmetic</td>
<td>Geometric</td>
</tr>
<tr>
<td>Stocks</td>
<td>7.93%</td>
<td>6.99%</td>
</tr>
<tr>
<td>Bonds</td>
<td>4.17%</td>
<td>4.16%</td>
</tr>
<tr>
<td>Comm. Paper</td>
<td>7.62%</td>
<td>7.57%</td>
</tr>
<tr>
<td>Inflation</td>
<td>0.69%</td>
<td>0.61%</td>
</tr>
<tr>
<td>HEP (Bonds)</td>
<td>3.70%</td>
<td>2.83%</td>
</tr>
<tr>
<td>HEP (Bills)</td>
<td>6.63%</td>
<td>6.71%</td>
</tr>
</tbody>
</table>

Total returns from 1871 to 1925 are constructed from the Price-Weighted NYSE and the Cowles Income Return Series.

In a very interesting article, Siegel and Schwartz (2006) calculate the return of the original S&P 500 companies since 1957 until 2003 and find that their return has been higher than the return of the

---

S&P 500\(^{10}\). The average geometric return of the S&P 500 was 10.85% (standard deviation of 17%), while the return of the original 500 companies was 11.31% (standard deviation of 15.7%).

### Table 5. Different Historical Equity Premiums (HEP) in the US according to different authors

<table>
<thead>
<tr>
<th>HEP vs. LT Gov. Bonds</th>
<th>Geometric</th>
<th>Arithmetic</th>
<th>Geometric</th>
<th>Arithmetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ibbotson</td>
<td>1926-2005</td>
<td>4.9%</td>
<td>1926-1957</td>
<td>6.0%</td>
</tr>
<tr>
<td>Shiller</td>
<td>5.5%</td>
<td>1968-2005</td>
<td>4.1%</td>
<td></td>
</tr>
<tr>
<td>WJ</td>
<td>4.4%</td>
<td>1926-1957</td>
<td>6.5%</td>
<td></td>
</tr>
<tr>
<td>Damodaran</td>
<td>5.1%</td>
<td>1968-2005</td>
<td>7.0%</td>
<td></td>
</tr>
<tr>
<td>Siegel</td>
<td>4.6%</td>
<td>1926-1957</td>
<td>5.8%</td>
<td></td>
</tr>
<tr>
<td>Max-min</td>
<td>1.0%</td>
<td>1968-2005</td>
<td>4.0%</td>
<td></td>
</tr>
</tbody>
</table>


Note that not all the authors get the same result, even for the HEP. Table 5 is a comparison of the HEP in the US according to different authors. The differences are substantial, especially for the period 1926-1957. The differences are mainly due to the stock indexes chosen. It is also important to keep in mind that the data from the 19th century and from the first part of the 20th century is quite poor and questionable. Table 6 shows the differences among the different indexes commonly used.

### Table 6. Number of securities in the US indexes commonly used

<table>
<thead>
<tr>
<th>S&amp;P composite weekly</th>
<th>Ibbotson</th>
<th>CRSP NYSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1926-1957</td>
<td>228 stocks in 1927; 410 in 1928, 480 in 1956</td>
<td>Growing number of stocks: 592 in 1927; 1059 in 1957</td>
</tr>
<tr>
<td>1957-2006</td>
<td>abandoned</td>
<td>S&amp;P Composite daily: 500 stocks</td>
</tr>
</tbody>
</table>

#### 2.3. A closer look at the historical data

Figure 1 shows that interest rates were lower than dividend yields until 1958 and than the earnings to price ratio until the 1980s. It suggests that many things have changed in the capital markets and that the last 40 years have been different than the previous ones. It is quite sensible to assume that the portfolio theory, the CAPM, the APT, the VAR analysis, the futures and options markets, the appearance of many mutual and hedge funds, the increase of investors, the legislation to protect investors, financial innovation, electronic trading, portfolio insurance, market participation,… have changed the behaviour and the risk attitudes of today’s investors vs. past investors. In fact, financial markets are so different that the relative magnitude of dividend yields to interest rates has been reversed.

---

\(^{10}\) The market value of the S&P 500 companies that have survived from the original 1957 list was only 31% of the 2003 year-end S&P 500’s market value. Since the S&P 500 was formulated, more than 900 new companies have been added to the index (and an equal number deleted from).
It is interesting to look at historical data to know what happened to our grandparents (or to our great grandparents), but it is not sensible to assume that their markets and their investment behaviour were similar to ours\textsuperscript{11}.

Figure 2 shows the evolution of the 20-year rolling correlation of (dividend yield – R\textsubscript{F}) versus R\textsubscript{F} (the yield on Government long-term bonds). Again, we may see that something has changed in the markets because that correlation after 1960 has been lower than ever before. Figure 3 shows the raw data used to calculate the correlations of Figure 2 and permits to contrast the different behavior of the markets in the periods 1871-1959 and 1960-2005. In section 7 we analyze this data and derive implications.

Figure 4 shows the evolution of the 20-year rolling HEP (arithmetic and geometric) relative to the T-Bills. It may be seen that the periods with equity returns much higher than the T-Bill rates were the 50s and the 90s.

Figure 5 compares the 20-year rolling HEP with the current T-Bond yield. From 1960 to 2000 the HEP increased when the yield decreased and vice versa. It did not happen so clearly in previous years.

\textsuperscript{11} Neither the exam of Ec1010 in 1932 is very useful for a student today.
2.4. Estimates of the Historical Equity Premium (HEP) in other countries

Blanchard (1993) examined the evolution of stock and bonds rates over the period 1978 to 1992 for the US, Japan, Germany, France, Italy and the UK. He constructed ‘world’ rates of return (using relative GDP weights for the countries) and documented a postwar decline in the dividend yield and in various measures of the HEP.

Jorion and Goetzmann (1999) constructed a database of capital gain indexes for 39 markets, with 11 of them starting in 1921 (see table 7). However, they obtained pre-1970 dividend information
only for 6 markets. They concluded that “for 1921 to 1996, US equities had the highest real return for all countries, at 4.3%, versus a median of 0.8% for other countries. The high equity premium obtained for U.S. equities appears to be the exception rather than the rule”. According to the authors, “there are reasons to suspect that [the US] estimates are subject to survivorship”.

However, Dimson and Marsh (2001) do not find survivorship bias for the US. They calculate the geometric HEP for 1955-1999 of US, UK, Germany and Japan and get 6.2%, 6.2%, 6.3% and 7.0%.

Table 7. Equity return of selected countries, according to Jorion and Goetzmann (1999)

<table>
<thead>
<tr>
<th>Country</th>
<th>Period</th>
<th>Nominal Return</th>
<th>Real Return</th>
<th>Dollar Return</th>
<th>Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>21-96</td>
<td>6.95%</td>
<td>4.32%</td>
<td>6.95%</td>
<td>2.52%</td>
</tr>
<tr>
<td>Sweden</td>
<td>21-96</td>
<td>7.42%</td>
<td>4.29%</td>
<td>7.00%</td>
<td>3.00%</td>
</tr>
<tr>
<td>Germany</td>
<td>21-96</td>
<td>4.43%</td>
<td>1.91%</td>
<td>5.81%</td>
<td>2.47%</td>
</tr>
<tr>
<td>Canada</td>
<td>21-96</td>
<td>5.78%</td>
<td>3.19%</td>
<td>5.35%</td>
<td>2.51%</td>
</tr>
<tr>
<td>U.K.</td>
<td>21-96</td>
<td>6.30%</td>
<td>2.35%</td>
<td>5.20%</td>
<td>3.86%</td>
</tr>
<tr>
<td>France</td>
<td>21-96</td>
<td>9.09%</td>
<td>0.75%</td>
<td>4.29%</td>
<td>8.28%</td>
</tr>
<tr>
<td>Belgium</td>
<td>21-96</td>
<td>4.45%</td>
<td>-0.26%</td>
<td>3.51%</td>
<td>4.73%</td>
</tr>
<tr>
<td>Italy</td>
<td>28-96</td>
<td>10.10%</td>
<td>0.15%</td>
<td>3.22%</td>
<td>9.94%</td>
</tr>
<tr>
<td>Japan</td>
<td>21-96</td>
<td>7.33%</td>
<td>-0.81%</td>
<td>1.80%</td>
<td>8.21%</td>
</tr>
<tr>
<td>Spain</td>
<td>21-96</td>
<td>4.66%</td>
<td>-1.82%</td>
<td>1.53%</td>
<td>6.61%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Median 39 countries</th>
<th>Real Return</th>
<th>Dollar Return</th>
<th>Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.75%</td>
<td>4.68%</td>
<td></td>
</tr>
<tr>
<td>11 countries with continuous histories into the 1920s:</td>
<td>Mean</td>
<td>1.88%</td>
<td>5.09%</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>2.35%</td>
<td>5.20%</td>
</tr>
</tbody>
</table>

Table 8. HEP vs. short (30 days) and long term (10 or 30 years) fixed income in 17 countries. 1900-2005. Annualized returns. Source: Table 3 of Dimson, Marsh and Staunton (2006c)

<table>
<thead>
<tr>
<th>% p.a.</th>
<th>Bills</th>
<th>Bonds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Geometric Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>Australia</td>
<td>7.08</td>
<td>8.49</td>
</tr>
<tr>
<td>Japan</td>
<td>6.67</td>
<td>9.84</td>
</tr>
<tr>
<td>South Africa</td>
<td>6.20</td>
<td>8.25</td>
</tr>
<tr>
<td>Germany</td>
<td>3.83</td>
<td>9.07</td>
</tr>
<tr>
<td>Sweden</td>
<td>5.73</td>
<td>7.98</td>
</tr>
<tr>
<td>U.S.</td>
<td>5.51</td>
<td>7.41</td>
</tr>
<tr>
<td>U.K.</td>
<td>4.43</td>
<td>6.14</td>
</tr>
<tr>
<td>Italy</td>
<td>6.55</td>
<td>10.46</td>
</tr>
<tr>
<td>Canada</td>
<td>4.54</td>
<td>5.88</td>
</tr>
<tr>
<td>France</td>
<td>6.79</td>
<td>9.27</td>
</tr>
<tr>
<td>Netherlands</td>
<td>4.55</td>
<td>6.61</td>
</tr>
<tr>
<td>Ireland</td>
<td>4.09</td>
<td>5.98</td>
</tr>
<tr>
<td>Belgium</td>
<td>2.80</td>
<td>4.99</td>
</tr>
<tr>
<td>Norway</td>
<td>3.07</td>
<td>5.70</td>
</tr>
<tr>
<td>Spain</td>
<td>3.40</td>
<td>5.46</td>
</tr>
<tr>
<td>Denmark</td>
<td>2.87</td>
<td>4.51</td>
</tr>
<tr>
<td>Switzerland</td>
<td>3.63</td>
<td>5.29</td>
</tr>
<tr>
<td>Average</td>
<td>4.81</td>
<td>7.14</td>
</tr>
<tr>
<td>World-ex U.S.</td>
<td>4.23</td>
<td>5.93</td>
</tr>
</tbody>
</table>

Dimson et al (2006c) use a unique database to calculate the historical equity premium for 17 countries over 106 years (1900-2005). Their estimates (see Table 8) are lower than frequently quoted
HEPs mainly due to the incorporation of the earlier part of the 20th century as well as the opening years of the 21st century.  

But, apart from the historical interest, how useful and accurate is that data? As Dimson et al. (2006c) point out, “virtually all of the 16 countries experienced trading breaks … often in wartime. The U.K. and European exchanges, and even the NYSE, closed at the start of World War I … Similarly, the Danish, Norwegian, Belgian, Dutch and French markets … when Germany invaded in 1940, and even the Swiss market closed from May to July 1940 for mobilization. … Japan after the Great Tokyo Earthquake of 1923. … Germany and Japan from towards the end of World War II, and Spain during the Civil War”. They claim that “we were able to bridge these gaps”, but this assertion is questionable. They admit that “the end-year index levels recorded for Germany for 1943–47, Japan for 1945, and Spain for 1936–38 cannot be regarded as market-determined values”. Dimson et al. (2006c) explain in their footnote 7 that “In Spain, trading was suspended during the Civil War from July 1936 to April 1939, and the Madrid exchange remained closed through February 1940; over the closure we assume a zero change in nominal stock prices and zero dividends”. It is not clear why this assumption is a reasonable one. They also mention one “unbridgeable discontinuity, namely, bond and bill (but not equity) returns in Germany during the hyperinflation of 1922–23, when German bond and bill investors suffered a total loss of –100%. … bonds and bills can become riskier than equities. When reporting equity premiums for Germany … we thus have no alternative but to exclude the years 1922–23”.

In a previous work Dimson, Marsh and Staunton (2002) show that the HEP was generally higher for the second half century: the World had 4.7% in the first half, compared to 6.2% in the second half.

Table 9 contains some of the HEPs reported by different authors for the US.

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Reference/average</th>
<th>Period for</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siegel (2005a)</td>
<td>T-Bonds, geo.</td>
<td>1926-2004</td>
<td>4.53%</td>
</tr>
<tr>
<td>Dimson &amp; al.(2006c)</td>
<td>T-Bonds, geo. US</td>
<td>1900-2005</td>
<td>4.52%</td>
</tr>
<tr>
<td>Dimson &amp; al.(2006c)</td>
<td>T-Bonds, geo. World</td>
<td>1900-2005</td>
<td>4.04%</td>
</tr>
</tbody>
</table>

This section has revised different estimates of the Historical Equity Premium (HEP) and permits to note that not all the authors get the same result for the HEP. We highlight the change in the

---

12 Their database contains annual returns on stocks, bonds, bills, inflation, and currencies for 17 countries from 1900–2005, and is described in Dimson et al (2006a and 2006b). They construct a World equity index (U.S. dollars index of 17 countries weighted by its starting-year market capitalization or by its GDP, before capitalizations were available) and a World bond index, constructed with each country weighted by its GDP. The series were compiled to avoid the survivorship bias that can arise from backfilling. Their choice of international markets was limited by their requirement to have data for the whole century.
market around 1960. Before that date, the dividend yield was higher than the risk-free rate, but after that date has been always smaller. We question the usefulness of historical data to predict the future.

3. Expected Equity Premium (EEP)

The Expected Equity Premium (EEP) is the answer to a question we would all (especially analysts and fund managers) like to answer accurately in the short term, namely: what incremental return do I expect from the market portfolio over the risk-free rate over the next years?

Estimates of the EEP based on historical analysis presume that the historical record provides an adequate guide for future expected long-term behaviour. However, the HEP changes over time, and it is not clear why capital market data from the 19th century or from the first half of the 20th century may be useful in estimating expected returns in the 21st century.

Numerous papers assert that there must be an EEP common to all investors (to the representative investor). But it is obvious that investors do not share “homogeneous expectations”13 and, also, that many investors do not hold the market portfolio but, rather, a subgroup of stocks and bonds14. Heterogeneous investors do not hold the same portfolio of risky assets; in fact, no investor must hold the market portfolio to clear the market.

We claim in section 7 that without “homogeneous expectations” there is not one EEP (but several), and there is not one REP (but several).

3.1. The Historical Equity Premium (HEP) is not a good estimator of the EEP

Although many authors consider that the equity premium is a stationary process, and then the HEP is an unbiased estimate of the EEP (unconditional mean equity premium), we do not agree with that statement: the HEP is not a good estimator of the EEP. For example, Mehra and Prescott (2003) state that “…over the long horizon the equity premium is likely to be similar to what it has been in the past”.

The magnitude of the error associated with using the HEP as an estimate of the EEP is substantial. Shiller (2000) points out that “the future will not necessarily be like the past”. Booth (1999) concludes that the HEP is not a good estimator of the EEP and estimates the later in 200 basis points smaller than the HEP15. Mayfield (2004) suggest that a structural shift in the process governing the volatility of market returns after the 1930s resulted in a decrease in the expected level of market risk, and concluded that EEP = HEP – 2.4% = 5.9% over the yield on T-bills (4.1% over yields on T-bonds).

---

13 Brennan (2004) also admits that “different classes of investor may have different expectations about the prospective returns on equities which imply different assessments of the risk premium”.

14 But, even with “homogeneous expectations” (all investors have equal EEP), the REP would not be equal for all investors. In that situation, the investors with lower REP would clear the market.

15 He also points out that the nominal equity return did not follow a random walk and that the volatility of the bonds increased significantly over the last 20 years.
Survivorship bias\(^\text{16}\) was identified by Brown, Goetzmann and Ross (1995) as one of the main reasons why the results based on historical analyses can be too optimistic. They pointed out that the observed return, conditioned on survival (HEP), can overstate the unconditional expected return (EEP). However, Li and Xu (2002) show that the survival bias fails to explain the equity premium puzzle: “To have high survival bias, the probability of market survival over the long run has to be extremely small, which seems to be inconsistent with existing historical evidence”. Siegel (1999, p. 13) mentions that “Although stock returns may be lower in foreign countries than in the U.S., the real returns on foreign bonds are substantially lower”.

Pastor and Stambaugh (2001) present a framework allowing for structural breaks in the risk premium over time and estimate that the EEP fluctuated between 4% and 6% over the period from 1834 to 1999, declined steadily since the 1930s (except for a brief period in the mid-1970s) and had the sharpest drop in the last decade of the 20th century. Using extra information from return volatility and prices, they narrow the confidence interval of their estimation (two standard deviations) to plus or minus 280 basis points around 4.8%.

Constantinides (2002) addresses different ways in which we may account for biases in the sample mean premium in order to estimate the expected premium and draws a sharp distinction between conditional, short-term forecasts of the mean equity premium and estimates of the unconditional mean. He says that the conditional EEPs at the end of the 20th century and the beginning of the 21st are substantially lower than the estimates of the unconditional EEP (7%) “by at least three measures”. But he concludes that “the currently low conditional, short-term forecasts of the equity premium do not necessarily imply that the unconditional estimate of the mean premium is lower than the sample average. Therefore, the low conditional forecasts do not necessarily lessen the burden on economic theory to explain the large sample average of the equity return and premium over the past 130 years”.

Dimson et al (2003) highlight the survivorship bias relative to the market, “even if we have been successful in avoiding survivor bias within each index, we still focus on markets that survived” and concluded that the geometric EEP for the world’s major markets should be 3% (5% arithmetic). Dimson et al (2006c) admit that “we cannot know today’s consensus expectation for the equity premium”, but they conclude that “investors expect an equity premium (relative to bills) of around 3-3½% on a geometric mean basis”, substantially lower than the HEP found in their own study.

3.2. Surveys

A direct way to obtain an expectation of the equity premium is to carry out a survey of analysts or investors although Ilmanen (2003) argues that surveys tend to be optimistic: “because of behavioural biases, survey-based expected returns may tell us more about hoped-for returns than about required returns”.

\(^{16}\) “Survivorship” or “survival” bias applies not only to the stocks within the market (the fact that databases contain data on companies listed today, but they tend not to have data on companies that went bankrupt or filed for bankruptcy protection in the past), but also for the markets themselves (“US market’s remarkable success over the last century is typical neither of other countries nor of the future for US stocks” (Dimson et al 2004)).
Welch (2000) performed two surveys with finance professors in 1997 and 1998, asking them what they thought the EEP was over the next 30 years. He obtained 226 replies, ranging from 1% to 15%, with an average arithmetic EEP of 7% above T-Bonds.\textsuperscript{17} Welch (2001) presented the results of a survey of 510 finance and economics professors performed in August 2001 and the consensus for the 30-year arithmetic EEP was 5.5%, much lower just 3 years earlier.

Graham and Harvey (2005) indicate that U.S. CFOs reduced their average EEP from 4.65% in September 2000 to 2.93% by September 2005. Over this period, the HEP had fallen only 0.4%.

Goldman Sachs (O’Neill, Wilson and Masih, 2002) conducted a survey of its global clients in July 2002 and the average long-run EEP was 3.9%, with most responses between 3.5% and 4.5%. The magazine *Pensions and Investments* (12/1/1998) carried out a survey among professionals working for institutional investors and the average EEP was 3%.

### 3.3. Regressions

Attempts to predict the equity premium typically look for some independent lagged predictors (X) on the equity premium:  
\[
\text{Equity Premium}_t = a + b \cdot X_{t-1} + \varepsilon_t
\]

Many predictors have been explored in the literature. Some examples are:

- The inflation rate (money illusion): Fama and Schwert (1977), Fama (1981), and Campbell and Vuolteenaho (2004a,b), and Cohen, Polk and Vuolteenaho (2005).
- Interest rate and dividend related variables: Ang and Bekaert (2003).

Goyal and Welch (2006) used most of the mentioned predictors and could not identify one that would have been robust for forecasting the equity premium and, after all their analysis, they recommended “assuming that the equity premium is ‘like it always has been’”. They also show that

\textsuperscript{17} The interest rate paid by long-term T-bonds in April 1998 was approximately 6%. At that time, the most recent Ibbotson Associates Yearbook was the 1998 edition, with an arithmetic HEP versus T-bills of 8.9% (1926–1997).
most of these models have not performed well for the last thirty years, that are not stable, and that are not useful for market-timing purposes.

However, Campbell and Thompson (2005) claim that some variables (ratios, patterns, levels of sort and long term interest rates) are correlated with subsequent market returns and that “forecasting variables with significant forecasting power in sample generally have a better out-of-sample performance than a forecast based on the historical average return”.

3.4. Other estimates of the expected equity premium

Siegel (2002, page 124) concluded that “the future equity premium is likely to be in the range of 2 to 3%, about one-half the level that has prevailed over the past 20 years”\(^{18}\). Siegel (2005a, page 172) affirms that “over the past 200 years, the equity risk premium has averaged about 3%”. Siegel (2005b) maintains that “although the future equity risk premium is apt to be lower than it has been historically, U.S. equity returns of 2-3% over bonds will still amply reward those who will tolerate the short-term risk of stocks”. However, in a presentation at the SIA annual meeting (November 10, 2005) Siegel maintained that “equity premium is 4% to 5% now”.

In the TIAA-CREF Investment Forum of June 2002, Ibbotson forecasted “less than 4% in excess of long-term bond yields”, and Campbell “1.5% to 2%”.

McGrattan and Prescott (2001) did not find corporate equity overvalued in 2000 and forecasted that the real returns on debt and equity should both be near 4%: “Therefore, barring any institutional changes, we predict a small equity premium in the future”.

Arnott and Ryan (2001) claim that the expected equity premium is near zero. They base their conclusion on the low dividend yield and their low expectation of dividend growth. Arnott and Bernstein (2002) also conclude that “the current risk premium is approximately zero”.

Bostock (2004) concludes that according to historical average data, equities should offer a risk premium over government bonds between 0.6% and 1.8%.

Grabowski (2006) concludes that “after considering the evidence, any reasonable long-term estimate of the normal EEP as of 2006 should be in the range of 3.5% to 6%”.

Maheu and McCurdy (2006) claim that the US Market had “three major structural breaks (1929, 1940 and 1969), and possibly a more recent structural break in the late 1990s”, and suggest an EEP in 2004 between 4.02% and 5.1%.

Table 10. Estimates of the EEP (Expected Equity Premium) according to different authors

<table>
<thead>
<tr>
<th>Authors</th>
<th>Conclusion about EEP</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surveys</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pensions and Investments (1998)</td>
<td>3%</td>
<td>Institutional investors</td>
</tr>
<tr>
<td>Graham and Harvey (2000)</td>
<td>4.65%</td>
<td>CFOs</td>
</tr>
</tbody>
</table>

\(^{18}\) Siegel also affirms that: “Although it may seem that stocks are riskier than long-term government bonds, this is not true. The safest investment in the long run (from the point of view of preserving the investor’s purchasing power) has been stocks, not Treasury bonds”.

14
4. Required and implied equity premium

The Required Equity Premium (REP) of an investor is the incremental return that she requires, over the risk-free rate, for investing in a diversified portfolio of shares. It is a crucial parameter in valuation and capital budgeting because the REP is the key to determining the company’s required return to equity and the required return to any investment project. The HEP is misleading for predicting the REP. If there was a reduction in the REP, this fall in the discount rate led to re-pricing of stocks, thus adding to the magnitude of HEP. The HEP, then, overstates the REP.

The IEP is the implicit REP used in the valuation of a stock (or a market index) that matches the current market value with an estimate of the future cash flows to equity. The IEP is also called the ex ante equity premium. However, the existence of a unique IEP implies to consider that the equity market can be explained with a representative consumer, or to consider that all investors have at any moment the same expectations about future cash flows and use the same discount rate to value each company.

Two models are widely used to calculate the IEP: the Gordon (1962) model (constant dividend growth model) and the residual income (or abnormal return) model.

According to the Gordon (1962) model, the current price per share ($P_0$) is the present value of expected dividends discounted at the required rate of return ($k$). If $d_1$ is the dividend per share expected to be received at time 1, and $g$ the expected long term growth rate in dividends per share\(^{19}\),

\[
P_0 = \frac{d_1}{k - g}, \quad \text{which implies: } \quad k = \frac{d_1}{P_0} + g.
\]

\(1\)

\(^{19}\) Although we say “dividends per share”, we refer to equity cash flow per share: dividends, repurchases and all expected cash for the shareholders.
The abnormal return method is another version of the Gordon (1962) model when the “clean surplus” relation holds (\(d_t = e_t - (BV_t - BV_{t-1})\)), being \(d\) the dividends per share, \(e\) the earnings per share and \(BV\) the book value per share:

\[
P_0 = bv_0 + (e_1 - k bv_0) / (k - g),
\]

which implies:

\[
k = e_1/P_0 + g (1 - bv_0/P_0) \quad (2)
\]

Jagannathan, McGrattan and Scherbina (2000) use the Gordon model, assume that dividends will growth as fast as GNP, and come with an estimate of 3.04%. They mention that “to get the estimate up to Brealey and Myer’s 9.2%, we would need to assume nominal dividend growth of 13.2%. This is an unreasonable assumption”. They also revise Welch (2000) and point out that “apparently, finance professors do not expect the equity premium to shrink”.

Claus and Thomas (2001) calculate the equity premium using the Gordon model and the residual income model, assuming that \(g\) is the consensus of the analysts’ earnings growth forecasts for the next five years and that the dividend payout will be 50%. They also assume that the residual earnings growth after year 5 will be the current 10-year risk-free rate less 3%. With data from 1985 to 1998, they find that the IEP is smaller than the HEP, and they recommend using a REP of about 3% for the US, Canada, France, Germany, Japan and UK.

Harris and Marston (2001), using the dividend discount model and estimations of the financial analysts about long-run growth in earnings, estimate an IEP of 7.14% for the S&P 500 above T-Bonds over the period 1982-1998. They also claim that the IEP move inversely with government interest rates, which is hard to believe.

Easton, Taylor, Shroff and Sougiannis (2002) used the residual income model with IBES data for expected growth\(^{21}\), and estimated an average IEP of 5.3% over the years 1981-1998.

Goedhart, Koller and Wessels (2002) used the dividend discount model (considering also share repurchases), with GDP growth as a proxy for expected earnings growth and with the average inflation rate of the last 5 years as a proxy for expected inflation. Table 11 contains their results that they report. They conclude that “we estimate that the real cost of equity has been remarkably stable at about 7% in the US and 6% in the UK since the 1960s. Given current, real long-term bond yields of 3% in the US and 2.5% in the UK, the implied equity risk premium is around 3.5% to 4% for both markets”.

| Table 11. IEP and real cost of equity in the US and the UK according to Goedhart et al (2002) |
|------------------------------------------|-----------------|-----------------|-----------------|-----------------|
|                                        | US |                |                |
| Market risk premium                    | 5.0% | 3.6%   | 4.3% | 3.0%   |
| Real risk-free rate                    | 2.2% | 3.1%   | 1.4% | 2.8%   |
| Real cost of equity                    | 7.2% | 6.7%   | 5.7% | 5.8%   |

\(^{20}\) Comparing the two models, it is clear than in a growing perpetuity, \(D_1 = E_1 - g \cdot BV_0\). The equivalence of the two models may be seen in Fernandez (2005)

\(^{21}\) Although Chan, Karceski and Lakonishok (2001) report that “IBES forecasts are too optimistic and have low predictive power for long-term growth”.
Fama and French (2002), using the discounted dividend model, estimated the IEP for the period 1951-2000 between 2.55% and 4.32%, far below the HEP (7.43%). For the period 1872-1950, they estimated an IEP (4.17%) similar to the HEP (4.4%). They claimed that in the period 1951-2000 “a decline in the expected stock return is the prime source of the unexpected capital gain”, and that “the unconditional EEP of the last 50 years is probably far below the realized premium”\(^\text{22}\).

Ritter and Warr (2002) claim that in 1979-1997, the IEP declined from +12% to -4%. However, Ritter estimate of the IEP in 2006 is a little over 2% on a geometric basis.

Harris, Marston, Mishra and O’Brien (2003) estimated discount rates for several companies using the dividend discount model and assuming that g was equal to the consensus of the analysts’ growth of dividends per share forecasts. They found an IEP of 7.3% (if betas calculated with a domestic index) and 9.7% (when betas calculated with a world index).

Many authors use an expected growth of dividends per share (g) equal to the consensus of the analysts’ forecasts, but Doukas, Kim and Pantzalis (2006) find that stock returns are positively associated with analyst’s divergence of opinion, and consider the divergence of opinion as risk.

Vivian (2005) replicated Fama and French (2002) to the UK, obtained similar results (see table 12), and concluded that the discount rate (REP) declined in the later part of the 20th Century.

<table>
<thead>
<tr>
<th>Table 12. REP and HEP in the US and in the UK according to Fama and French (2002) and Vivian (2005)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table I of Fama and French (2002)</strong></td>
</tr>
<tr>
<td><strong>US</strong></td>
</tr>
<tr>
<td>1872-2000</td>
</tr>
<tr>
<td>1872-1950</td>
</tr>
<tr>
<td>1951-2000</td>
</tr>
<tr>
<td>1966-2002</td>
</tr>
</tbody>
</table>

O’Hanlon and Steele (2000) proposed calculating the REP using accounting figures and got a variety of estimates between 4 and 6%.

Glassman and Hassett (2000) calculated in their book Dow 36,000 that the REP for the U.S. in 1999 was 3%, arguing that stocks should not carry any risk premium at all, and that stock prices will rise dramatically further once investors come to realize this fact\(^\text{23}\).

Faugere and Erlach (2006) claimed that the equity premium tracks the value of a put option on the S&P 500. However, their conclusion is not very helpful: “using an 8.1% premium in valuation formulas and capital budgeting problems may be appropriate, since the observed level of the long-run equity premium is fully consistent with the observed steady-state GDP growth and consistent with risk explanations as well. However, if one believes that the recent 1990’s trends in dividend yields, interest rates, taxes and inflation represent permanent regime shifts, our model can be parameterized to yield a 3.5% equity premium”.

---

\(^{22}\) Fama and French (1992) report that in the period 1941-1990 an equally weighted index outperformed the value weighted (average monthly returns of 1.12% and 0.93%) in the whole period and in most sub sample periods.

\(^{23}\) Not to be outdone, Kadlec and Acampora (1999) gave their book the title, *Dow 100,000: Fact or Fiction?*
Donaldson, Kamstra and Kramer (2006) simulate the distribution from which interest rates, dividend growth rates, and equity premia are drawn and claim that “the true ex ante equity premium is 3.5% plus or minus 50 basis points”. They say that previous studies “estimate the equity premium with great imprecision: often a 5% to 6% ex post estimate can not be statistically distinguished from an ex ante value as low as 1% or as high as 10%”.

One problem of all these estimates is that they depend on the particular assumption made for the expected growth.

Table 13. Implied Equity Premium (IEP) and Required Equity Premium (REP) according to different authors

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Method</th>
<th>IEP</th>
<th>REP</th>
</tr>
</thead>
<tbody>
<tr>
<td>O'Hanlon and Steele (2000)</td>
<td>accounting</td>
<td>4 to 6%</td>
<td></td>
</tr>
<tr>
<td>Jagannathan &amp; al. (2000)</td>
<td>DDM</td>
<td>3.04%</td>
<td></td>
</tr>
<tr>
<td>Glassman and Hasset (2000)</td>
<td></td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Harris and Marston (2001)</td>
<td>DDM</td>
<td>7.14%</td>
<td></td>
</tr>
<tr>
<td>Claus and Thomas (2001)</td>
<td>RIM</td>
<td></td>
<td>3%</td>
</tr>
<tr>
<td>Fama and French (2002)</td>
<td>DDM</td>
<td>1872-1950</td>
<td>4.17%</td>
</tr>
<tr>
<td>Goedhart, Koller and Wessels (2002)</td>
<td>DDM</td>
<td>1990-2000</td>
<td>3.5 to 4%</td>
</tr>
<tr>
<td>Ritter (2002)</td>
<td>DDM</td>
<td>2001</td>
<td>0.7%</td>
</tr>
<tr>
<td>Ritter and Warr (2002)</td>
<td>RIM</td>
<td>1979-1997</td>
<td>+12% to -4%</td>
</tr>
<tr>
<td>Harris &amp; al. (2003)</td>
<td>DDM</td>
<td></td>
<td>7.3%</td>
</tr>
<tr>
<td>Vivian (2005)</td>
<td>DDM &amp; RIM</td>
<td>1951-2002 UK</td>
<td>4.6%</td>
</tr>
<tr>
<td>Donaldson, Kamstra and Kramer (2006)</td>
<td>DDM</td>
<td>1952-2004</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

DDM = dividend discount model. RIM = residual income model

5. The equity premium puzzle

The equity premium puzzle, a term coined by Mehra and Prescott (1985), is the inability of a standard representative consumer asset pricing model, using aggregate data, to reconcile the HEP. To reconcile the model with the HEP, individuals must have implausibly high risk aversion according to standard economics models. Mehra and Prescott (1985) argued that stocks should provide at most a 0.35% premium over bills. Even by stretching the parameter estimates, Mehra and Prescott (2003) concluded that the premium should be no more than 1%. This contrasted starkly with their HEP estimate of 6.2%.

5.1. Attempts to solve the equity premium puzzle

This puzzle has lead to an extensive research effort in both macroeconomics and finance. Over the last 20 years, researchers have tried to resolve the puzzle by generalizing and adapting (weakening one or more of the assumptions) the Mehra-Prescott (1985) model, but still there is not a solution generally accepted by the economics profession. Some of the adapted assumptions include:

24 Kocherlakota (1996) reduces the models to just 3 assumptions: individuals have preferences associated with the standard utility function, asset markets are complete (individuals can write insurance contracts against any contingency), and asset trading is costless.
• alternative assumptions about preferences (state separability, leisure, precautionary savings) or generalizations to state-dependent utility functions: Abel (1990); Constantinides (1990); Epstein and Zin (1991); Benartzi and Thaler (1995); Bakshi and Chen (1996); Campbell and Cochrane (1999); and Barberis, Huang, and Santos (2001),
• narrow framing\(^{25}\): Barberis and Huang (2006),
• probability distributions that admit disastrous events such as fear of catastrophic consumption drops: Rietz (1988); Mehra and Prescott (1988), Barro (2005),
• survivorship bias: Brown, Goetzmann, and Ross (1995),
• liquidity premium: Bansal and Coleman (1996),
• taxes and regulation: McGrattan and Prescott (2005),
• the presence of uninsurable income shocks or incomplete markets: Mankiw (1986); Constantinides and Duffie (1996); Heaton and Lucas (1996) and (1997); Storesletten, Telmer, and Yaron (1999),
• relative volatility of stocks and bonds: Asness (2000)
• distinguishing between the cash flows to equity and aggregate consumption: Brennan and Xia (2001), who claim to be able to justify an equity premium of 6%.
• borrowing constraints: Constantinides, Donaldson, and Mehra (2002),
• other market imperfections: Aiyagari and Gertler (1991); Alvarez and Jermann (2000),

There are several excellent surveys of this work, including Kocherlakota (1996), Cochrane (1997) and Mehra and Prescott (2003 and 2006). Kocherlakota (1996) says that “while there are several plausible explanations for the low level of Treasury returns, the large equity premium is still largely a mystery to economists”.

Rietz (1988) and Barro (2005) suggest that low-probability disasters, such as a small a large “crash” in consumption, may justify a large equity premium. However, Mehra and Prescott (1988) challenge Rietz to identify such catastrophic events and estimate their probabilities.

McGrattan and Prescott (2005) argue that the 1960-2001 HEP is mainly due to changes in taxes and regulatory policy during this period. They also say that “Allowing for heterogeneous individuals will also help quantify the effects of increased market participation and diversification that has

\(^{25}\) Narrow framing is the phenomenon documented in experimental settings whereby, when people are offered a new gamble, they sometimes evaluate it in isolation, separately from their other risks.
occurred in the past two decades. Until very recently, mutual funds were a very expensive method of creating a diversified equity portfolio”.

Limited stock market participation can increase the REP by concentrating stock market risk on a subset of the population. To understand why limited participation may have quantitative significance for the REP, it is useful to review basic facts about the distribution of wealth, and its dynamics over time. Mishel, Bernstein and Allegretto (2006) document that wealth and stock holdings in the U.S. remain highly concentrated in dollar terms: in 2004, the wealthiest 10% held 78.8% of the stocks (84% in 1989 and 76.9% in 2001), and the wealthiest 20% held over 90% of all stocks. Only 48.6% of U.S. households held stocks in 2004 (51.9% in 2001 and 31.7% in 1989) and only 34.9% (40.1% in 2001 and 22.6% in 1989) held stock worth more than $5,000. Of this 34.9%, only 13.5% had direct holdings. Mankiw and Zeldes (1991) reported that 72.4% of the 2998 families in their survey held no stocks at all. Among families that held more than $100,000 in other liquid assets, only 48% held stock. The covariance of stock returns and consumption of the families that hold stocks is triple that of no stockholders and it may explain part of the puzzle.

Brennan (2004) highlights the “democratization of Equity Investment”: “The increase in the number of participants in equity markets was accompanied by a massive increase in the scale of the equity mutual fund industry: the assets under management rose from $870 per capita in 1989 to over $14,000 per capita in 1999, before declining to a little over $12,000 per capita in 2001. On the other hand, holdings of bond mutual funds grew only from $966 per capita in 1989 to $2887 in 1989. In other words, while bond funds roughly tripled, equity funds went up by a factor of over 14!” and “the share of corporate equity held by mutual funds rose from 6.6% in 1990 to 18.3% in 2000”.

Heaton and Lucas (2000) introduced Limited Participation and Limited Diversification in an overlapping generations model and concluded that the increases in participation of the past two decades are unlikely to cause a significant reduction in the EEP, but that improved portfolio diversification might explain a fall in the EEP of several percentage points.

There is some promising research on heterogeneity. Abel (1991) hoped that “incorporating differences among investors or more general attitudes toward risk can explain the various statistical properties of asset returns”. Levy and Levy (1996) mentioned that the introduction of a small degree of diversity in expectations changed the dynamics of their model and produced more realistic results. Constantinides and Duffie (1996) introduced heterogeneity in the form of uninsurable, persistent and heteroscedastic labor income shocks. Bonaparte (2006) used micro data on households' consumption and provides a new method on estimating asset pricing models, considering each household as living on an island and taking into account its lifetime consumption path. Due to the great deal of heterogeneity across households, he replaced the representative agent with an average agent.

Bakshi and Chen (2006) claim that “disentangling the equity premium into its cash flow and discounting components produces an economic meaningful equity premium of 7.31%”.

Shalit and Yitzhaki (2006) show that at equilibrium, heterogeneous investors hold different risky assets in portfolios, and no one must hold the market portfolio.
It is interesting the quotation in Siegel and Thaler (1997): “no economic theorist has been completely successful in resolving the [equity premium] puzzle” ... but ... “most economists we know have a very high proportion of their retirement wealth invested in equities (as we do)”.

6. The equity premium in the textbooks

This section contains the main messages about the equity premium conveyed in the finance textbooks and valuation books. More details may be found in Fernandez (2006). Figure 6 collects the evolution of the Required Equity Premium (REP) used or recommended by the textbooks and by the academic papers mentioned on previous sections. Table 14 contains the equity premium recommended and used in different editions of several textbooks. Ritter (2002) mentions the use of the historical equity risk premium in textbooks as an estimate of the future as one of the "The Biggest Mistakes We Teach". Looking at Figure 6 and at Table 14, it is quite obvious that there is not much consensus, creating a lot of confusion among students and practitioners (and finance authors, also) about the Equity Premium.

Brealey and Myers considered REP = EEP = HEP in the 2nd, 3rd, 4th and 5th editions (1984, 1988, 1991 and 1996), using Ibbotson data that ranged from 8.2 to 8.5% (arithmetic HEPs over T-Bills in periods starting in 1926). In the 6th, 7th and 8th editions (2000, 2003 and 2005 with Allen), they said that “Brealey, Myers and Allen have no official position on the exact market risk premium, but we believe that a range of 5 to 8.5 percent is reasonable for the risk premium in the United States.” (In the previous editions the ranges was 6 to 8.5%).

Copeland, Koller and Murrin (McKinsey) used a REP = geometric HEP versus Government T-Bonds in the two first editions (1990 and 1995). However, they changed criteria in the 3rd and 4th editions: they advised to use the arithmetic HEP of 2-year returns versus Government T-Bonds reduced by a survivorship bias. In the 1st edition (1990), they recommended 5-6%, in the 2nd edition (1995) they recommended 5-6%, in the 3rd edition (2000) they recommended 4.5-5% (“we subtract a 1.5 to 2% survivorship bias from the long-term arithmetic average of 6.5%”) and in the 4th edition (Koller, Goedhart and Wessels, 2005) they recommended 3.5-4.5% (“we subtract a 1% to 2% survivorship bias from the long-term arithmetic average of 5.5%”).


Bodie, Kane and Marcus (1993, 2nd edition) used a REP = EEP = 6.5% to value Hewlett-Packard. In the 3rd edition (1996, page 535), they used a REP = EEP = HEP − 1% = 7.75% to value Motorola. In the 5th edition (2002, page 575), they valued Motorola using a REP = 6.5%. In the 6th edition (2003), they used in the examples different REPs: 8% (pages 426, 431) and 5% (page 415).
Damodaran (1994, 2002) recommended REP = EEP = geometric HEP versus T-bonds. In 1997 he used a REP = arithmetic HEP versus T-Bills. In 2001a and 2006 he recommended REP = EEP = IEP. Damodaran on Valuation (1994), recommended an EEP of 5.5%, the geometric HEP using T-bonds for the period 1926-1990. Damodaran (2001a, 2006, 2nd edition) used a REP = IEP of 4% for the US. because “the implied premium for the US and the average implied equity risk premium has been about 4% over the past 40 years”. Damodaran (1996, 1997, 2001b, 2001c and 2002), however, used a REP of 5.5%. In (1996, page 48) he shows that 5.5% is the geometric HEP versus T-bonds in the period 1926-90.

Copeland and Weston (1979, 1988) used a REP = 10%. However, Weston and Copeland (1992), used a REP = 5%.

Van Horne (1968, 1st ed.) still did not mention the CAPM or the equity premium. In (1983, 6th ed.), he used a REP = 6% He justified it: “Suppose, for easy illustration, that the expected risk-free rate is an average of the risk-free rates that prevailed over the ten-year period and that the expected market return is average of market returns over that period”. In (1992, Fundamentals, 8th ed.), he used a REP = 5% and justified it: “Assume that a rate of return of about 13% on stocks in general is expected to prevail and that a risk-free rate of 8% is expected”.

Penman (2001, 1st ed.) said that “the market risk premium is a big guess. Research papers and textbooks estimate it in the range of 4.5% to 9.2%. … No one knows what the market risk premium is”. In (2003, 2nd ed.), he admitted that “we really do not have a sound method to estimate the cost of capital… Estimates [of the equity premium] range, in texts and academic research, from 3.0% to 9.2%”, and he used 6%.

Weston and Brigham (1968) still did not defined equity premium. In (1982, 6th edition) they said that “the market risk premium can be considered relatively stable at 5 to 6% for practical application”. Weston, Chung and Siu (1997) recommended 7.5%. Bodie and Merton (2000) used 8% for USA.

Fernandez (2002) is the only finance textbook claiming that “it is impossible to determine the premium for the market as a whole, because it does not exist”. He also mentions that we “could only talk of a market risk premium if all investors had the same cash flow expectations... However, expectations are not homogeneous”. Fernandez (2004, 2001) also mentioned that “the HEP, the EEP and the REP are different concepts” and that “different investors have different REPs”. In the examples he uses REP = 4%.

Palepu, Healy and Bernard (2004, page 8-3) mention that the HEP “constitutes an estimate of the REP” and use REP = 7% in the examples (page 8-5).

Weston, Mitchel and Mulherin (2004) mention that the arithmetic HEP over T-bonds in the period 1926-2000 according to Ibbotson was 7.3% and (page 260) they use REP = EEP = 7%.

Bruner (2004) used a REP of 6% because “from 1926 to 2000, the risk premium for common stocks has averaged about 6% when measured geometrically”.

Arzac (2005) uses a REP = IEP = 5.08% for a valuation done in December 2002 (the IEP equity premium as of that date calculated using the Gordon equation).

In the following section we claim that the confusion comes from the fact that there is not a REP for the market as a whole: different investors use different REPs. Last sentence may me rewritten as: there is not an IEP for the market as a whole: different investors use different IEPs. A unique IEP requires assuming homogeneous expectations for the expected growth (g), but there are several pairs (IEP, g) that satisfy current prices.
Table 14. Equity premiums recommended and used in textbooks

<table>
<thead>
<tr>
<th>Author(s) of the Textbook</th>
<th>Assumption</th>
<th>Period for HEP</th>
<th>REP recommended</th>
<th>REP used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brealey and Myers</td>
<td>REP=EEP= arith HEP vs. T-Bills</td>
<td>1926-81 1926-85 1926-88 1926-88</td>
<td>8.3% 8.4% 8.4% 8.2 - 8.5%</td>
<td>8.3% 8.4% 8.4% 8.0%</td>
</tr>
<tr>
<td></td>
<td>REP=EEP= arith HEP vs. T-Bills</td>
<td>No official position No official position</td>
<td>6.0 - 8.5%</td>
<td>5.0 - 8.5%</td>
</tr>
<tr>
<td>Copeland, Koller and Murrin (McKinsey)</td>
<td>REP=EEP= arith HEP vs. T-Bills</td>
<td>1903-2002</td>
<td>3.5 - 4.5%</td>
<td>4.8%</td>
</tr>
<tr>
<td>Ross, Westerfield and Jaffe</td>
<td>REP=EEP= arith HEP vs. T-Bills</td>
<td>1926-88</td>
<td>8.5%</td>
<td>8.5%</td>
</tr>
<tr>
<td>Van Horne, 6th edition. 1983</td>
<td>REP=EEP= geo HEP vs. T-Bonds</td>
<td>1926-88</td>
<td>5 - 6%</td>
<td>6%</td>
</tr>
<tr>
<td>Weston and Copeland (1992)</td>
<td>REP=EEP= arith HEP vs. T-Bills</td>
<td>1926-93</td>
<td>8.5%</td>
<td>8.5%</td>
</tr>
<tr>
<td>Bodie, Kane and Marcus</td>
<td>REP=EEP= arith HEP vs. T-Bills</td>
<td>1926-94</td>
<td>8.5%</td>
<td>8.5%</td>
</tr>
<tr>
<td>Damodaran</td>
<td>REP=EEP= geo HEP vs. T-Bonds</td>
<td>1926-90</td>
<td>5.5%</td>
<td>5.5%</td>
</tr>
<tr>
<td>Weston &amp; Brigham (1982)</td>
<td>REP=EEP= geo HEP vs. T-Bonds</td>
<td>1926-99</td>
<td>5.5%; 8%</td>
<td>4.8%</td>
</tr>
<tr>
<td>Bodie and Merton (2000)</td>
<td>REP=EEP= HEP</td>
<td>1926-2001</td>
<td>5%</td>
<td>8%</td>
</tr>
<tr>
<td>Stowe et al (2002)</td>
<td>REP=EEP= geo HEP vs. T-Bonds</td>
<td>1928-00</td>
<td>5.51%</td>
<td>5.51%</td>
</tr>
<tr>
<td>Hawaiinini and Viallet (2002)</td>
<td>REP=EEP= geo HEP vs. T-Bonds</td>
<td>1926-99</td>
<td>5.7%</td>
<td>5.7%</td>
</tr>
<tr>
<td>Pratt (2002)</td>
<td>REP=EEP= HEP</td>
<td>1926-2002</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>Fernandez (2002)</td>
<td>REP=EEP= HEP</td>
<td>1926-2000</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>Weston, Mitchel &amp; Mulherin (2004)</td>
<td>REP=EEP= arith HEP vs. T-Bonds</td>
<td>1926-2000</td>
<td>7.3%</td>
<td>7%</td>
</tr>
<tr>
<td>Arzac (2005)</td>
<td>REP=IEP</td>
<td>1926-2000</td>
<td>5.08%</td>
<td>5.08%</td>
</tr>
</tbody>
</table>
7. There is not an IEP, but many pairs (IEP, g) which are consistent with market prices

Even if market prices are correct for all investors, there is not a unique REP common for all investors. In a simple Gordon model, there are many pairs (Ke, g) that satisfy equation (1). As Ke is the sum of the Implied Equity Premium (IEP) plus the risk-free rate (Rf), there are many pairs (IEP, g) that satisfy equation (1). A unique IEP requires assuming homogeneous expectations for the expected growth (g). If equation (1) holds, the expected return for the shareholders is equal to the required return for the shareholders (Ke), but there are many required returns (as many as expected growths, g) in the market. On top of that, IEP and g change over time.

If investors’ expectations were homogeneous, it would make sense to calculate a unique IEP, as all investors would have the market portfolio and the same expectations regarding the portfolio\[^{26}\]. However, as expectations are not homogenous\[^{27}\], different investors use different REPs: investors who expect higher growth will have a higher REP. Heterogeneous investors do not hold the same portfolio of risky assets; in fact, no investor must hold the market portfolio to clear the market: it does not make sense to search for a common REP because it does not exists.

We can find out an investor’s REP by asking him, although for many investors the REP is not an explicit parameter but, rather, an implicit one that manifests in the price they are prepared to pay for shares\[^{28}\]. However, it is impossible to determine the REP for the market as a whole, because it does not exist. Even if we knew the market premiums of all the investors who operated on the market, it would be meaningless to talk of a premium for the market as a whole.

A rationale for this may be found in the aggregation theorems of microeconomics, which in actual fact are non-aggregation theorems. One model that works well individually for a number of people may not work for all of the people together\[^{29}\]. For the CAPM, this means that although the CAPM may be a valid model for each investor, it is not valid for the market as a whole, because investors do not have the same return and risk expectations for all shares. Prices are a statement of expected cash flows discounted at a rate that includes the risk premium. Different investors have different cash flow expectations and different future risk expectations. One could only talk of an equity premium if all investors had the same cash flow expectations.

Reallocating terms in equation (1), we get:

26 Even then, this method requires knowing the expected growth of dividends. A higher growth estimate implies a higher premium.
28 An example: An investor is prepared to pay 80 euros for a perpetual annual cash flow of 6 euros in year 1 and growing at an annual rate of 3%, which he expects to obtain from a diversified equity portfolio. This means that his required market return is 10.5% ([6/80] + 0.03).
29 As Mas-Colell et al. (1995, page 120) say, “it is not true that whenever aggregate demand can be generated by a representative consumer, this representative consumer’s preferences have normative contents. It may even be the case that a positive representative consumer exists but that there is no social welfare function that leads to a normative representative consumer.”
IEP – g = d₁/P₀ - R_F \tag{3}

There are many pairs (IEP, g) that satisfy the Gordon equation at any moment. All the papers that we revised on section 5 assume that there is an “expected growth rate for the market” and get an “IEP for the market”. But without homogeneous expectations, there is not an “expected growth rate for the market”.

Similarly, for having an EEP common for all investors we need to assume homogeneous expectations (or a representative investor) and, with our knowledge of financial markets, this assumption is not reasonable. A theory with a representative investor cannot explain either why the annual trading volume of most exchanges more than double the market capitalization.

We also find that the difference (IEP – g),\(^\text{30}\) is related to the risk free rate in the period after 1960. Figure 7 shows the relationship for the period after 1980 for the US, Spain and the UK. It may be seen the high negative correlation between (IEP – g) and the risk free rate in the three markets. Table 15 presents the regressions for more countries.

\[ (d₁/P₀ - R_F) = \text{IEP} - g. \]

Source of the data: Datastream

\textbf{Figure 7. Correlations (d₁/P₀ - R_F) – (R_F) for the US, Spain and the UK. Monthly data.}

\textbf{United States 1980-2006}

\[ (\text{IEP}-g) = -0.5523 R_F - 0.5289\%; \quad R^2 = 0.906 \]

\textbf{Spain 1990-2006}

\[ (\text{IEP}-g) = -0.6705 R_F + 0.6596\% \quad R^2 = 0.9473 \]

\(^{30}\) (d₁/P₀ - R_F) is equal to (IEP – g)
Table 15. Regressions with monthly data of Y (IEP – g) on RF (10 year Gov. Bond Yield)

<table>
<thead>
<tr>
<th>Country</th>
<th>Year</th>
<th>Regression Equation</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>1980-2006</td>
<td>Y = -0.5523 R_f - 0.5289%</td>
<td>0.9060</td>
</tr>
<tr>
<td>Germany</td>
<td>1980-2006</td>
<td>Y = -0.7192 R_f + 0.5907%</td>
<td>0.8205</td>
</tr>
<tr>
<td>UK</td>
<td>1980-2006</td>
<td>Y = -0.6833 R_f + 1.2913%</td>
<td>0.9469</td>
</tr>
<tr>
<td>France</td>
<td>1988-2006</td>
<td>Y = -0.9587 R_f + 2.5862%</td>
<td>0.9245</td>
</tr>
<tr>
<td>Italy</td>
<td>1991-2006</td>
<td>Y = -1.0693 R_f + 3.0398%</td>
<td>0.9563</td>
</tr>
<tr>
<td>Spain</td>
<td>1991-2006</td>
<td>Y = -0.6705 R_f + 0.6596%</td>
<td>0.9473</td>
</tr>
</tbody>
</table>

8. How do I calculate the REP?

For calculating the cost of equity (required return to equity cash flows) of a company, a valuator has to answer the following question: which differential rate over current T-Bond yields do I think compensates the risk of holding the shares? If there is only an owner of the shares, we can directly ask him the question. But if it is a traded company, the valuator has to make a prudential judgment. As Grabowski (2006), points out, “the entire appraisal process is based on applying reasoned judgment to the evidence derived from economic, financial and other information and arriving at a well reasoned opinion of value”.

We need the cost of equity to discount the expected equity cash flows of the company. Note that there is a kind of schizophrenic approach to valuation: while all authors admit that different valuers and investors may have different expectations of equity cash flows, most authors look for a unique discount rate. It seems as if the expectations of equity cash flows are formed in a democratic regime, while the discount rate is determined in a dictatorship. In any market, different investors may have different expectations of equity cash flows and different evaluations of its risk (that translate into different discount rates). Then, in the case of a traded company, there are investors that think that the company is undervalued (and buy or hold shares), investors that think that the company is overvalued (and sell or not buy shares), and investors that think that the company is fairly valued (and sell or hold shares). The investors that did the last trade, or the rest of the investors that held or did not have shares do not have a common REP (nor common expectations of equity cash flows).
For calculating the REP, we must answer the same question, but thinking in a diversified portfolio of shares, instead in just the shares of a company. In the valuations that I have done in the 21st century I have used REPs between 3.8 and 4% for Europe and for the U.S. Given the yields of the T-Bonds, I think that an additional 4% compensates the additional risk of a diversified portfolio.

9. Conclusion

The equity premium (also called market risk premium, equity risk premium, market premium and risk premium), is one of the most important, discussed but elusive parameters in finance. Much of the confusion arises from the fact that the term equity premium is used to designate four different concepts (although many times they are mixed): Historical Equity Premium (HEP), Expected Equity Premium (EEP); Required Equity Premium (REP) and Implied Equity Premium (IEP).

In the finance literature and in valuation textbooks, there are authors that claim different identities among the four equity premiums defined above: some claim that \( \text{HEP} = \text{EEP} = \text{REP} \); others claim that \( \text{EEP is smaller than HEP} \); others claim that there is a unique IEP and that \( \text{REP} = \text{IEP} \); others “have no official position”; others claim that \( \text{EEP is near zero} \); others try to find the EEP doing surveys; others affirm “that no one knows what the REP is”.

The HEP is equal for all investors, but the REP, the EEP and the IEP are different for different investors. There is no an IEP for the market as a whole: different investors have different IEPs and use different REPs. A unique IEP requires assuming homogeneous expectations for the expected growth \( g \), but there several pairs \( \text{(IEP, g)} \) that satisfy current prices.

We claim that different investors have different REPs and that it is impossible to determine the REP for the market as a whole, because it does not exist. Heterogeneous investors do not hold the same portfolio of risky assets; in fact, no investor must hold the market portfolio to reach equilibrium.

There is a kind of schizophrenic approach to valuation: while all authors admit that different valuators and investors may have different expectations of equity cash flows, most authors look for a unique discount rate. It seems as if the expectations of equity cash flows are formed in a democratic regime, while the discount rate is determined in a dictatorship. In any market, different investors may have different expectations of equity cash flows and different evaluations of its risk (that translate into different discount rates).

It has been argued that, from an economic standpoint, we need to establish the primacy of the EEP, since it is what guides investors' decisions. However, the REP is more important for many important decisions, among others, valuations of projects and companies, acquisitions, and corporate investment decisions. On the other hand, EEP is important only for the investors that hold the market portfolio.

For calculating the cost of equity (required return to equity cash flows) of a company, a valuator has to answer the following question: which differential rate over current T-Bond yields do I

---

31 And also my clients that are able to answer to that question.
think compensates the risk of holding the shares? If there is only an owner of the shares, we can directly ask him the question. But if it is a traded company, the valuator has to make a prudential judgment. There are investors that think that the company is undervalued (and buy or hold shares), investors that think that the company is overvalued (and sell or not buy shares), and investors that think that the company is fairly valued (and sell or hold shares). For calculating the REP, we must answer the same question, but thinking in a diversified portfolio of shares, instead in just the shares of a company. Recently, I have used REPs between 3.8 and 4% for Europe and for the U.S. Given the yields of the T-Bonds, I think that an additional 4% compensates the additional risk of a diversified portfolio.
References


Smith, E. L. (1926), Common Stocks as Long Term Investments,


