This book is printed on acid-free paper. ♾

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

For general information on our other products and services, or technical support, please contact our Customer Care Department within the United States at 800-762-2974, outside the United States at 317-572-3993 or fax 317-572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

For more information about Wiley products, visit our Web site at *http://www.wiley.com*.

Library of Congress Cataloging-in-Publication Data:

Pratt, Shannon P.
Cost of capital: applications and examples/Shannon P. Pratt, Roger J. Grabowski.—3rd ed.
p. cm.
"Published simultaneously in Canada."
Previous editions had subtitle: Estimation and applications.
Includes bibliographical references and index.
ISBN 978-0-470-17115-8 (cloth: alk. paper)
1. Capital investments. 2. Business enterprises—Valuation. 3. Capital investments—United States.
4. Business enterprises—Valuation—United States. I. Grabowski, Roger J. II. Title.
HG4028.C4P72 2008
658.15'2—dc22
2007020132

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Estimating the ERP is one of the most important decisions you must make in developing a discount rate. For example, the effect of a decision that the appropriate ERP is 4% instead of 8% in the Capital Asset Pricing Model (CAPM) will generally have a greater impact on the concluded discount rate than alternative theories of the proper measure of other components, for example, beta. One academic study looked at sources of error in estimating expected rates of return over time and concluded:

> We find that the great majority of the error in estimating the cost of capital is found in the risk premium estimate, and relatively small errors are due to the risk measure, or beta. This suggests that analysts should improve estimation procedures for market risk premiums, which are commonly based on historical averages.[1]

In ranking what matters and what does not matter in estimating the cost of equity capital, another author categorizes the choice of the ERP as a "high impact decision," likely to make a difference of more than two percentage points and could make a difference of more than four points.[2]

Three driving forces behind the discussions that have evolved on ERP include:

1. What returns can be expected from investments by retirement plans in publicly traded common stocks by retirement plans?
2. What expected returns are being priced in the observed values of publicly traded common stocks?
3. What is the appropriate cost of capital to use in discounting future cash flows of a company or a project to their present value equivalent?

Because of the importance of the ERP estimate and the fact that we find many practitioners confused about estimating ERP, we report on recent studies and report on ERP estimates at the beginning of 2007. We conclude with our recommended ERP.

DEFINING THE EQUITY RISK PREMIUM

The ERP (or notational RP_m) is defined as:

$$RP_m = R_m - R_f$$

where:

RP_m = the equity risk premium
R_m = the expected return on a fully diversified portfolio of equity securities
R_f = the rate of return expected on a risk-free security

What is referred to as the ERP means, in practice, a general equity risk premium using as a proxy for the "market" either the Standard & Poor's (S&P) 500 or the New York Stock Exchange (NYSE) composite stock index. ERP is a forward-looking concept. It is an expectation as of the valuation date for which no market quotes are observable.

In this chapter, we are addressing returns of publicly traded stocks. Those returns establish a beginning benchmark for closely held investments.

[1] Wayne Ferson and Dennis Locke, "Estimating the Cost of Capital through Time: An Analysis of the Sources of Error," *Management Science* (April 1998): 485–500.

[2] Seth Armitage, *The Cost of Capital: Intermediate Theory* (Cambridge: Cambridge University Press, 2005), 319–320.

ESTIMATING THE ERP

While you can observe premiums realized over time by referring to historical data (i.e., realized return approach or *ex post* approach), such realized premiums do not represent the ERP expected in prior periods, nor do they represent the current ERP. Rather, realized premiums may, at best, represent only a sample from prior periods of what may have been the *expected* ERP.

To the extent that realized premiums on the average equate to expected premiums in prior periods, such samples *may be* representative of current expectations. But to the extent that events that are not expected to reoccur caused realized returns to differ from prior expectations, such samples should be adjusted to remove the effects of these nonrecurring events. Such adjustments are needed to improve the predictive power of the sample.

Alternatively, you can directly derive implied forward-looking estimates for the ERP from data on the underlying expectations of growth in corporate earnings and dividends or from projections of specific analysts as to dividends and future stock prices (*ex ante* approach).[3]

The goal of either approach is to estimate the true *expected* ERP as of the valuation date. Even then the expected ERP can be thought of in terms of a *normal* or *unconditional* ERP and a *conditional* ERP based on current prospects.[4] We address issues involving the conditional ERP later.

There is no one universally accepted standard for estimating ERP. A wide variety of premiums are used in practice and recommended by academics and financial advisors.

NOMINAL OR REAL?

Both the expected return on a fully diversified portfolio of equity securities and the rate of return expected on a risk-free security can be stated in nominal (including expected inflation) or real terms (expected inflation removed). ERP should not be affected by inflation. If both returns are expressed in nominal terms, the difference in essence removes the expected inflation; if both returns are expressed in real terms, inflation has been removed, but the difference remains the same. But *ex post* realized returns will be affected by differences between expected inflation and realized inflation.

WHICH RISK-FREE RATE TO USE IN ESTIMATING THE ERP

Any estimate of ERP must be made in relation to a risk-free security. That is, the expected return on a fully diversified portfolio of equity securities must be measured in its relationship to the rate of return expected on a risk-free security. The selection of an appropriate risk-free security with which to base the ERP estimate is a function of the expected holding period for the investment to which the discount rate (rate of return) is to apply. For example, if you were estimating the equity return on a highly liquid investment and the expected holding period were potentially short-term, a U.S. government short-term bond (e.g., Treasury or T-bill) may be an appropriate instrument to use in benchmarking the ERP estimate.

Alternatively, if you were estimating the equity return on a long-term investment, such as the valuation of a business where the value can be equated to the present value of a series of future cash flows over many years, then the yield on a long-term U.S. government bond may be the more appropriate instrument in benchmarking the ERP estimate.

[3] See, for example, Eugene F. Fama and Kenneth R. French, "The Equity Premium." *Journal of Finance* (April 2002): 637–659.

[4] Robert Arnott, "Historical Results," *Equity Risk Premium Forum, AIMR* (November 8, 2001): 27.

Common academic practice in empirical studies of rates of return realized on portfolios of stocks in excess of a risk-free rate is to benchmark stock returns against realized monthly returns of "risk-free" 90-day T-bills or one-year government bonds. A T-bill rate is the purest risk-free base rate because it contains essentially no maturity risk. If inflation is high, it does reflect the inflation component, but it contains little compensation for inflation uncertainty. Problems in using such a risk-free security as a benchmark are that (1) T-bill rates may not reflect market-determined investor return requirements on long-term investments due to central bank actions affecting the short-term interest rates, and (2) rates on short-term securities tend to be more volatile than yields on longer maturities.

Long-term government bonds are free of default risk but are not "risk-free." Long-term government bonds are sensitive to future interest fluctuations. Investors are not sure of the purchasing power of the dollars they will receive upon maturity or the reinvestment rate that will be available to them to reinvest the interest payments received over the life of the bond. As a result, the long-term empirical evidence is that returns on long-term government bonds on the average exceed the returns on T-bills.[5]

The long-term premium of government bond returns in excess of the average expected interest rates on T-bills (average of future forward rates) is commonly referred to as the *horizon premium*. The horizon premium compensates the investor for the maturity risk of the bond. The horizon premium equals the added return expected *on the average* on long-term bonds due to inflation and interest rate risk. As interest rates change unexpectedly in the future, the bond price will vary. That is, bonds are subject to market risk due to unexpected changes in interest rates. The horizon premium compensates investors for that market risk.

MATCHING RISK-FREE RATE WITH ERP

In theory, when determining the risk-free rate and the matching ERP you should be matching the risk-free security and the ERP with the period in which the investment cash flows are expected. For example (where b is a risk measure for the investment):

$$\text{Short-term cash flows: Current T-bill rate} + b \times (RPm \text{ over T-bills})$$

Cash flows expected in:

$$\text{Year 1: 1-year government bond rate} + b \times (RPm \text{ over 1-year bonds})$$

$$\text{Year 2: 2-year forward rate on government bonds} + b \times (RPm \text{ over 2-year bonds})$$

$$\text{Year 3: 3-year forward rate on government bonds} + b \times (RPm) \text{ over 3-year bonds}), \text{ and so on}$$

$$\text{Cash flows expected in the long-term: Current long-term government bond rate} + b \times (RPm \text{ over long-term government bonds})$$

MEASURING THE AVERAGE PERIOD OF THE EXPECTED CASH FLOWS

Can one measure the "average" period of expected cash flows and use an average maturity period for the risk-free security and the ERP? One measure of the length of planning horizon over which cash flows are expected is the *duration* of cash flows. We introduced the concept of duration in Chapter 6 as a measure of the effective time period over which you receive cash flows from bonds.

In a similar manner, you can calculate the expected duration of any stream of expected cash flows for any project. For valuation of a "going-concern" business, for example, assume you expect the cash flow in the first year following the valuation date of $1 million to increase at an average

[5] When short-term interest rates exceed long-term rates, the yield curve is "inverted."

compound rate of 4% per annum. Assume a discount rate of 15%. If you project cash flows each year for 100 years, the calculated duration of the cash flows is approximately 10.5 years.[6]

In practice, few discount each cash flow using a matched maturity risk-free rate and ERP estimate. In valuing going-concern businesses and long-term investments made by businesses, practitioners generally use long-term government bonds as the risk-free security and estimate the ERP in relation to long-term government bonds. This convention both represents a realistic, simplifying assumption and is consistent with the CAPM.[7] If the expected cash flows are risky and follow a random walk, but the risk-free rate and the ERP are expected to be constant over time, then the risk-adjusted discount rate for discounting the risky cash flows is constant as well. Most business investments have long durations and suffer from a reinvestment risk comparable to that of long-term government bonds. As such, the use of long-term government bonds and an ERP estimated relative to long-term bonds more closely matches the investment horizon and risks confronting business managers in capital budgeting decisions and valuators in valuation problems than reference to T-bills.

Therefore, in the remainder of this chapter we have translated all estimates of ERP to estimates relative to long-term government bonds.

REALIZED RISK PREMIUM (*ex POST*) APPROACH

While academics and practitioners agree that ERP is a forward-looking concept, many practitioners use historical data only to estimate the ERP under the assumption that historical data are a valid proxy for current investor expectations. In the realized risk premium approach, the estimate of the ERP is the risk premium (realized return on stocks in excess of the risk-less rate) that investors have, on the average, realized over some historical holding period (realized risk premium).

The underlying theory is that the past provides a reasonable indicator of how the market will behave in the future and investors' expectations are influenced by the historical performance of the market. If period returns on stocks (e.g., monthly stock returns) are not correlated (e.g., this month's stock returns are not predictable based on last month's returns) and if expected stock returns are stable through time, then the arithmetic average of historical stock returns provides an unbiased estimate of expected future stock returns. Similarly, the arithmetic average of realized risk premiums provides an unbiased estimate of expected future risk premiums (the ERP).

A more indirect justification for use of the realized risk premium approach is the contention that, for whatever reason, securities in the past have been priced in such a way as to earn the returns observed. By using an estimated cost of equity capital incorporating the average of realized risk premiums in applying the income approach to valuation, you may to some extent replicate this level of pricing.

MEASURING REALIZED RISK PREMIUMS

The measure of the risk-free rate is not controversial once the proper duration (long term versus short term) of the investment has been estimated since the expected yield to maturity on appropriate

[6] $$\frac{[(1{,}000{,}000 \times 1)/(1.15) + (1{,}000{,}000 \times 1.04 \times 2)/(1.15)^2 + (1{,}000{,}000 \times 1.04^2 \times 3)/(1.15)^3 \ldots]}{[(1{,}000{,}000 \times 1)/(1.15) + (1{,}000{,}000 \times 1.04)/(1.15)^2 + (1{,}000{,}000 \times 1.04^2)/(1.15)^3 \ldots]} = 10.5(\textit{rounded})$$

[7] Carmelo Giaccotto, "Discounting Mean Reverting Cash Flows with the Capital Asset Pricing Model," *The Financial Review* (May 2007): 247–265. This is true for both the textbook CAPM of Sharpe and Linter and the extension of the textbook CAPM, the intertemporal CAPM of Merton.

government securities is directly observable in the marketplace. Differences in approach to estimating the ERP then hinge on the measure of expected return on equity securities.

In applying the realized risk premium approach, the analyst selects the number of years of historical return data to include in the average. One school of thought holds that the future is best estimated using a very long horizon of past returns. Another school of thought holds that the future is best measured by the (relatively) recent past. These differences in opinion result in disagreement as to the number of years to include in the average.

HISTORICAL STOCK AND BOND RETURNS

The highest-quality data are available for periods beginning in 1926 (the year that the forerunner of the current S&P 500 was first published) from the Center of Research in Security Prices (CRSP) at the University of Chicago. The *SBBI Yearbook* contains summaries of returns on United States stocks and bonds derived from that data.[8] The reported returns include the effects from the reinvestment of dividends.

Returns on common stocks have been assembled by various sources and with various qualities for earlier periods. Good stock market data are available back to 1872, and less reliable data are available from various sources back to the end of the eighteenth century. (In the earliest period, the market consisted almost entirely of bank stocks, and by the mid-nineteenth century, the market was dominated by railroad stocks.[9]) Data for government bond yield data have also been assembled for these periods. Exhibit 9.1 presents the realized average annual risk premium for stocks assembled from various sources for alternative periods through 2006.

We measure the realized risk premium by comparing the stock market returns realized during the period to the income return on long-term government bonds (or yield to maturity for the years before 1926).

While some may question looking at averages including early periods for estimating today's ERP, what is striking is that the largest arithmetic average of one-year returns is the 81 years from 1926 to 2006.

Why use the income return on long term government bonds? The income return in each period represented the expected yield on the bonds at the time of the investment. An investor makes a decision to invest in the stock market today by comparing the expected return from that investment to the rate of return today on a benchmark security (in this case the long-term government bond). While the investor did not know the stock market return when one invested at the beginning of each year, he or she did know the rate of interest promised on long-term government bonds. To try to match the expectations at the beginning of each year, we measure historical stock market returns on an expectation that history will repeat itself over the expected return on bonds in each year.

8 *Stocks, Bonds, Bills and Inflation (SBBI) Valuation Edition 2007 Yearbook* (Chicago: Morningstar, 2007).

9 See Lawrence Fisher and James Lorie, "Rates of Return on Investments in Common Stocks," *Journal of Business* 37, no. 1 (1964); J. W. Wilson and C. P. Jones, "A Comparison of Annual Stock Market Returns: 1871–1925 with 1926–1985," *Journal of Business* 60, no. 2 (1987): 239–258; G. W. Schwert, "Indexes of Common Stock Returns from 1802 to 1987," *Journal of Business* 63, no. 3 (1990): 399–425; Roger G. Ibbotson and Gary P. Brinson, *Global Investing: The Professional's Guide to the World Capital Markets* (New York: McGraw-Hill, 1993); J. W. Wilson and C. P. Jones, "An Analysis of the S&P 500 Index and Cowles's Extensions: Price Indexes and Stock Returns, 1870–1999," *Journal of Business* 75, no. 3 (2002): 505–533; S. H. Wright, "Measures of Stock Market Value and Returns for the US Nonfinancial Corporate Sector, 1900–2000," Working paper, February 1, 2002; W. Goetzmann, R. Ibbotson, and L. Peng, "A New Historical Database for NYSE 1915 to 1925: Performance and Predictability," *Journal of Financial Markets* 4 (2001): 1–32; E. Dimson, P. Marsh, and M. Staunton, *Triumph of the Optimists: 101 Years of Global Investment Returns* (Princeton, NJ: Princeton University Press, 2002) with annual updates of their Global Returns database available at http://corporate.morningstar.com/ib; W. Goetzmann and R. Ibbotson, "History and the Equity Risk Premium," *Yale ICF Working Paper No. 05-04*, April 6, 2005.

Exhibit 9.1 Historical Realized Premiums: Stock Market Returns – Treasury Bonds

Period	Arithmetic Average	Standard Error*	Geometric Average
20 years (1987–2006)	6.4%	3.7%	5.2%
30 years (1977–2006)	5.8%	2.8%	4.7%
40 years (1967–2006)	4.8%	2.6%	3.6%
50 years (1957–2006)	5.2%	2.3%	3.9%
81 years (1926–2006)**	7.1%	2.2%	5.2%
107 years (1900–2006)	6.8%	1.9%	4.9%
135 years (1872–2006)	5.9%	1.6%	4.3%
209 years (1798–2006)	5.1%	1.2%	3.6%

*Calculated as standard deviation of realized excess returns divided by square root N, number of years in sample.

**SBBI Valuation Edition 2007 Yearbook.

Source: Data compiled from R. Ibbotson and G. Brinson, *Global Investing* (New York: McGraw-Hill, 1993); W. Schwert, "Indexes of U.S. Stock Prices from 1802 to 1987," *Journal of Business*, 1990; S. Homer and R. Sylla, *A History of Interest Rates*, 3rd ed. (Piscataway, NJ: Rutgers University Press, 1991); and *SBBI, 2007 Yearbook* (Chicago: Morningstar, 2007).

The realized risk premiums vary year to year, and the estimate of the *true* ERP resulting from this sampling is subject to a degree of error. We display the standard errors of estimate for each period in Exhibit 9.1. The standard error of estimate allows you to measure the likely accuracy of using the realized risk premium as the estimate of ERP. That statistic indicates the estimated range within which the *true* ERP falls (i.e., assuming normality, the true ERP can be expected to fall within two standard errors with a 95% level of confidence).

SUMMARIZING REALIZED RISK PREMIUM DATA

The summarized data in Exhibit 9.1 represent the arithmetic and geometric averages of realized risk premiums for one-year returns. That is, the dollars invested including reinvested dividends are reallocated to available investments annually and the return is calculated for each year. The arithmetic average is the mean of the annual returns. The geometric average is the single compound return that equates the initial investment with the ending investment assuming annual reallocation of investment dollars and reinvestment of dividends.

For example, assume this series of stock prices (assuming no dividends):

Period	Stock price	Period Return
1	$10	
2	$20	100%
3	$10	−50%

The arithmetic average of period returns equals (100% + −50%)/2 = 25% while the geometric average equals $(1 + r_1)(1 + r_2)^{1/2} - 1 = (1 = 1.00 \times 1 - .5)^{1/2} - 1 = 0$.

Realized return premiums measured using the geometric (compound) averages are always less than those using the arithmetic average. The geometric mean is the lower boundary of the arithmetic mean, and the two are equal in the unique situation that every observation is identical to every other observation. Further, the more variable the period returns, the greater the difference between the arithmetic and geometric averages of those returns. This is simply the result of the mathematics of a series that has experienced deviations.

The choice between which average to use is a matter of disagreement among practitioners. The arithmetic average receives the most support in the literature,[10] though other authors recommend a geometric average.[11] The use of the arithmetic average relies on the assumption that (1) market returns are serially independent (not correlated) and (2) the distribution of market returns is stable (not time-varying). Under these assumptions, an arithmetic average gives an unbiased estimate of expected future returns assuming expected conditions in the future are similar to conditions during the observation period. Moreover, the more observations available, the more accurate will be the estimate.

> . . .the arithmetic mean equates the expected future value of investment with its present value. This property makes the arithmetic mean the correct return to use as the discount rate or cost of capital.[12]

> . . .the geometric mean measures changes in wealth over more than one period on a buy and hold (with dividends reinvested) strategy. . . . The arithmetic mean would provide a better measure of typical performance over a single historical period.[13]

WHAT PERIODICITY OF PAST MEASUREMENT?

But even if we agree that stock returns are serially independent, the arithmetic average of realized risk premiums based on *one-year* returns may not be the best estimate of future returns. Textbook models of stock returns (e.g., CAPM) are generally single-period models that estimate returns over unspecified investment horizons. For example, assume that the investment horizon equals two years. Then in using realized returns to estimate expected returns, you need to calculate realized returns over two-year periods (i.e., the geometric average over consecutive two-year periods) and then calculate the arithmetic average of the two-year geometric averages to arrive at the unbiased estimate of future returns. For example, assume that the realized one-year returns are:

Year 1 = 10%
Year 2 = 25%
Year 3 = −15%

The geometric averages of the two-year holding periods are:

$$(1.10 \times 1.25)^{1/2} - 1 = 17.3\%$$

$$(1.25 \times 0.85)^{1/2} - 1 = 3.1\%$$

The arithmetic average of *typical* two-year periods is therefore:

$$\frac{(17.3 + 3.1)}{2} = 10.2\%$$

10 See, e.g., Paul Kaplan, "Why the Expected Rate of Return Is an Arithmetic Mean," *Business Valuation Review* (September 1995); *SBBI Valuation Edition 2002 Yearbook,* 71–73; Mark Kritzman, "What Practitioners Need to Know about Future Value," *Financial Analysts Journal* (May/June 1994): 12–15; Zvi Bodie, Alex Kane, and Alan J. Marcus, *Investments* (1989): 720–723.

11 See, for example, Aswath Damodaran, *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset*, 2nd ed. (Hoboken, N.J.: John Wiley & Sons, 2002), 161–162.

12 Roger Ibbotson and Rex Sinquefeld, *Stocks, Bonds, Bills and Inflation: Historical Returns (1926–1987)* (1989), 127.

13 Willard T. Carleton and Josef Lakonishok "Risk and Returns on Equity: the Use and Misuse of Historical Estimates," *Financial Analysts Journal* 41, no. 1 (1985): 39.

The issue then becomes what is the appropriate interval over which average realized returns should be measured (1-year periods as in the case of the returns reported in the *SBBI Yearbook*; 2-year periods; 20-year periods)? When you are valuing businesses, should you compare returns over periods greater than one year? The most likely answer is yes. Practitioners have adopted the use of interest rates on long-term government bonds, typically 20-year bonds, as the appropriate long-term benchmark risk-free rate when valuing businesses. It follows then that a longer investment horizon of, say, 20 years is the appropriate period over which you should calculate realized returns. As the investment horizon increases, the arithmetic average of realized investment returns decreases asymptotically to the geometric average of the entire series.

While Morningstar only reports on the arithmetic average of one-year returns, we calculated the realized risk premiums for various investment horizons using the data from 1926 to 2006 as shown in the next table.[14]

Arithmetic Average of	Realized Risk Premium
1-year returns[1]	7.1%
2-year returns[2]	6.1%
3-year returns[3]	5.8%
4-year returns[4]	5.5%
5-year returns[5]	5.3%
81-year returns (geometric average)[1]	5.2%

[1]*SBBI Valuation Edition 2007 Yearbook.*
[2]Excluding investment period beginning 2006.
[3]Excluding investment periods beginning 2005 and 2006.
[4]Excluding investment periods beginning 2004, 2005, and 2006.
[5]Excluding investment periods beginning 2003, 2004, 2005, and 2006.
Source: Compiled from data in *Stocks, Bonds, Bills, and Inflation 2007 Yearbook.* Copyright

Assuming that you have an investment horizon longer than one year, you can conclude that the realized risk premium that provides the "best estimate" of the ERP is likely between the arithmetic average of one-year returns and the geometric average of the entire series.

In one recent study, the authors show that compounding the arithmetic average of historical one-year returns as a forecaster of cumulative future returns results in estimates of cumulative returns that overstate the future cumulative returns that investors are likely to realize. This is due to the fact that distributions of stock market returns are skewed. The authors show that use of the geometric mean of historical one-year returns result in estimates of cumulative returns that more approximate the median of cumulative returns (50% if investors will realize more than the median cumulative return and 50% will realize less than the median return). They demonstrate that the difference between the median of forecasted cumulative returns obtained from compounding the arithmetic average versus the geometric average of one-year historical returns increases as the expected investment horizon increases.[15]

14 The equity risk premium of each investment horizon was calculated by taking equity returns (S&P 500) less the bond returns (U.S. Long-term Government Bond Income Return) for the respective periods. We calculated a series of rolling returns, one for stocks and another for bonds, for each investment horizon. We then took the arithmetic average of each series of rolling returns for the respective investment horizon. For example, the 2-year return, for equities and bonds, is the arithmetic average of a series of 2-year rolling returns from 1926 to 2006. We performed the same calculation for each investment horizon. We then subtract the bond return from the equity return to estimate the equity risk premium for each investment horizon.

15 Eric Hughson, Michael Stutzer, and Chris Yung, "The Misuse of Expected Returns," *Financial Analysts Journal* (November/December 2006): 88–96.

SELECTING A SAMPLE PERIOD

The average realized risk premium is sensitive to the period chosen for the average. While the selection of 1926 as a starting point corresponds to the initial publishing of the forerunner to the current S&P 500, that date is arbitrary. Regarding the historical time period over which equity risk should be calculated, Morningstar offers two observations[16]:

1. Reasons to focus on recent history:
   - The recent past may be most relevant to an investor.
   - Return patterns may change over time.
   - The longer period includes "major events" (e.g., World War II, the Depression) that have not repeated for over 50 years.
2. Reasons to focus on long-term history:
   - Long-term historical returns have shown surprising stability.
   - Short-term observations may lead to illogical forecasts.
   - Focusing on the recent past ignores dramatic historical events and their impact on market returns. We do not know what major events lie ahead.
   - Law of large numbers: More observations lead to a more accurate estimate.

But the average calculated using 1926 return data as a beginning point may be too heavily influenced by the unusually low interest rates during the 1930s to mid-1950s. For example, the average yield on long-term government bonds was only 2.3% during the 1940s (the lowest decade on record) and under 3% in each year from 1934 through 1955. Yields on government bonds exceeded 4% for most of the nineteenth century and have been consistently higher since the 1960s.

The years 1942 through 1951 were a period of artificial stability in U.S. government bond interest rates. In April 1942, the Federal Reserve publicly committed itself to maintaining an interest rate ceiling on government debt, both long term and short term, to support the financing of World War II. After World War II, the Fed continued maintaining an interest rate ceiling fearing return to the high unemployment of the Great Depression. But postwar inflationary pressures caused the Treasury and the Fed to reach an accord announced March 4, 1951, freeing the Fed of its obligation of pegging interest rates. Including this period in calculating realized returns is analogous to valuing airline stocks today by looking at prices of airline stocks before deregulation.

Some observers have suggested that the period, which includes the 1930s, 1940s, and the immediate post–World War II boom years, may have exhibited an unusually high average realized return premium. The 1930s exhibited extreme volatility while the 1940s and early 1950s saw a combination of record low interest rates and rapid economic growth that led the stock market to outperform Treasury bonds by a wide margin.

> The low real rates on bonds may have contributed to higher equity returns in the immediate postwar period. Since firms finance a large part of their capital investment with bonds, the real cost of obtaining such funds increased returns to shareholders. It may not be a coincidence that the highest 30-year average equity return occurred in a period marked by very low real returns on bonds. As real returns on fixed-income assets have risen in the last decade, the equity premium appears to be returning to the 2% to 3% norm that existed before the postwar surge.[17]

[16] *SBBI Valuation Edition 2007 Yearbook* (Chicago: Morningstar, 2007), 129, 134.
[17] Jeremy Siegel, *Stocks for the Long Run* (New York: McGraw-Hill, 1994), 20.

Exhibit 9.2 Realized Equity Risk Premiums over Treasury Bond Income Returns

Nominal (i.e., without inflation removed)	1926–1955	1956–2006
Arithmetic Average	10.5%	5.1%
Geometric Average	7.5%	3.9%
Standard Deviations		
Stock Market Annual Returns	25.3%	16.5%
Long-term Treasury Income Returns	0.5%	2.4%
Long-term Treasury Total Returns	4.7%	10.9%
Ratio of Equity to Bond Total Return Volatility	5.4	1.5

Source: Compiled from data in *Stocks, Bonds, Bills and Inflation 2007 Yearbook*. For more information on other Morningstar publications, please visit global.morningstar.com/DataPublications. Calculated (or Derived) based on CRSP® data, ©2006 Center for Research in Security Prices (CRSP®), Graduate School of Business, The University of Chicago.

If we disaggregate the 81 years reported in the *SBBI Yearbook* into two subperiods, the first covering the periods before and after the mid-1950s, we get the comparative figures for stock and bond returns shown in Exhibit 9.2.

The period since the mid-1950s has been characterized by a more stable stock market and a more volatile bond market compared to the earlier period. Interest rates, as reflected in Long-term Government Bond Income Return statistics as summarized in the *SBBI Yearbook*, have become more volatile in the later period. The effect is amplified in the volatility of Long-term Government Bond Total Returns as summarized in the *SBBI Yearbook,* which include the capital gains and losses associated with interest rate fluctuations. From these data, we can conclude that the relative risk of stocks versus bonds has narrowed; based on this reduced relative risk, we would conclude that the ERP is likely lower today. As a result, we question the validity of using the arithmetic average of one-year returns since 1926 as the basis for estimating today's ERP.

Evidence since 1871 clearly supports the premise that the difference between stock yields and bond yields is a function of the long-run difference in volatility between these two securities.[18] And if you examine the volatility in stock returns (as measured by rolling 10-year average standard deviation of real stock returns), you find that the volatility beginning in 1929 dramatically increased and that the volatility since the mid-1950s has returned to prior levels.[19] This also suggests that the arithmetic average realized risk premiums reported for the entire data series since 1926 as reported in the *SBBI Yearbook* likely overstate expected returns.

Using historical data may also tend to overstate expected returns given the increasing opportunities for international diversification. International diversification lowers the volatility of investors' portfolios, which in theory should lower the required return on the average asset in the portfolio. This would lower the expected return on U.S. government securities generally and hence would suggest a lower ERP on a forward-looking basis than indicated by historical data. Several authors have studied the influence of increased globalization, and their results suggest that costs of capital for companies operating in the international markets have decreased.[20]

[18] Clifford S. Asness, "Stocks versus Bonds: Explaining the Equity Risk Premium," *Financial Analysts Journal* (March/April 2000): 96–113.
[19] Laurence Booth, "Estimating the Equity Risk Premium and Equity Costs: New Ways of Looking at Old Data," *Journal of Applied Corporate Finance* (Spring 1999):100–112 and "The Capital Asset Pricing Model + Equity Risk Premiums and the Privately-Held Business," 1998 CICBV/ASA Joint Business Valuation Conference (September 1998): 23.
[20] See, e.g., Kate Phylaktis and Lichuan Xia, "Sources of Firm's Industry and Country Effects in Emerging Markets," *Journal of International Money and Finance* (2005): 459–475; and Gikas Hardouvelis, Dimitrious Malliartopulos, and Richard Priestly, "The Impact of Globalization on the Equity Cost of Capital," Working paper, May 9, 2004.

If the average expected risk premium has changed through time, then averages of realized risk premiums using the longest available data become questionable. A shorter-run horizon may give a better estimate if changes in economic conditions have created a different expected return environment than that of more remote past periods. Why not use the average realized return over the past 20-year period? A drawback of using averages over shorter periods is that they are susceptible to large errors in estimating the *true* ERP due to high volatility of annual stock returns. Also, the average of the realized premiums over the past 20 years may be biased high due to the general downward movement of interest rates since 1981.

While we can only observe historical realized returns in the stock market, we can observe both expected returns (yield to maturity) and realized returns in the bond market. Prior to the mid-1950s, the difference between the yield at issue and the realized returns was small since bond yields and therefore bond prices did not fluctuate very much.

Beginning in the mid-1950s until 1981, bond yields trended upward, causing bond prices to generally decrease. Realized bond returns were generally lower than returns expected when the bonds were issued (as the holder experienced a capital loss if sold before maturity). Beginning in 1981, bond yields trended downward, causing bond prices to generally increase. Realized bond returns were generally higher than returns expected when the bonds were issued (as the holder experienced a capital gain if sold before maturity). If we choose the period during which to measure realized premiums beginning from the late 1950s/early 1960s to today, we will be including a complete interest rate cycle.[21]

Even if we use long-term observations, the volatility of annual stock returns will be high. Assuming that the 81-year average gives an unbiased estimate, still a 95% confidence interval for the unobserved *true* ERP spans a range of approximately 3.0% to 11.5%.[22]

IS BIAS INTRODUCED BY USING THE ARITHMETIC AVERAGE IN ESTIMATING ERP?

The issue of bias is important from two different vantage points when using an ERP estimate derived from the arithmetic average of realized risk premium data:

1. In predicting the compound return you might expect for an investment in stocks, will you get an answer that is biased? (i.e., will measurement error be introduced simply due to the mathematics?)
2. In discounting expected cash flows where you develop a cost of equity capital estimate using that ERP estimate, will you get an answer that is biased?

Even if you accept the arithmetic average of annual realized risk premiums as an unbiased estimate of expected *annual* risk premium (i.e., investment horizon equals one year), it is a somewhat stronger assumption to compound this annual average over multiple periods (i.e., investment horizon equals *n* years); you are assuming that the estimate of the expected single-period return is accurate (in other words, that the estimate has no allowance for error). If you introduce measurement error and compound the estimated annual return over multiple periods, you will get a biased estimate of the true expected future value. This upward bias occurs even if the single-period arithmetic average itself is an unbiased estimate. The fact that you get an expected upward bias in future investment results if you project future returns using an arithmetic average is important if you are estimating the returns

[21] Booth, "Estimating the Equity Risk Premium and Equity Costs."
[22] Calculated as two standard errors around the average; 7.1% A+/− (2 × 2.2%).