What Risk Premium Is "Normal"?

Robert D. Arnott and Peter L. Bernstein

The goal of this article is an estimate of the objective forward-looking U.S. equity risk premium relative to bonds through history—specifically, since 1802. For correct evaluation, such a complex topic requires several careful steps: To gauge the risk premium for stocks relative to bonds, we need an expected real stock return and an expected real bond return. To gauge the expected real bond return, we need both bond yields and an estimate of expected inflation through history. To gauge the expected real stock return, we need both stock dividend yields and an estimate of expected real dividend growth. Accordingly, we go through each of these steps. We demonstrate that the long-term forward-looking risk premium is nowhere near the level of the past; today, it may well be near zero, perhaps even negative.

The investment management industry thrives on the expedient of forecasting the future by extrapolating the past. As a consequence, U.S. investors have grown accustomed to the idea that stocks “normally” produce an 8 percent real return and a 5 percent (that is, 500 basis point) risk premium over bonds, compounded annually over many decades. Why? Because long-term historical returns have been in this range with impressive consistency. And because investors see these same long-term historical numbers year after year, these expectations are now embedded in the collective psyche of the investment community.

Both the return and the risk premium assumptions are unrealistic when viewed from current market levels. Few have acknowledged that an important part of the lofty real returns of the past stemmed from rising valuation levels and from high dividend yields, which have since diminished. As we will demonstrate, the long-term forward-looking risk premium is nowhere near the 5 percent level of the past; indeed, today, it may well be near zero, perhaps even negative. Credible studies in and outside the United States are challenging the flawed conventional view. Well-researched studies by Claus and Thomas (2001) and Fama and French (2000) are just two (see also Arnott and Ryan 2001). Similarly, the long-term forward-looking real return from stocks is nowhere near history’s 8 percent. We argue that, barring unprecedented economic growth or unprecedented growth in earnings as a percentage of the economy, real stock returns will probably be roughly 2–4 percent, similar to bond returns. In fact, even this low real return figure assumes that current near-record valuation levels are “fair” and likely to remain this high in the years ahead. “Reversion to the mean” would push future real returns lower still.

Furthermore, if we examine the historical record, neither the 8 percent real return nor the 5 percent risk premium for stocks relative to government bonds has ever been a realistic expectation, except from major market bottoms or at times of crisis, such as wartime. But this topic merits careful exploration. After all, according to the Ibbotson Associates data, equity investors earned 8 percent real returns and stocks have outpaced bonds by more than 5 percent over the past 75 years. Intuition suggests that investors should not require such outsized returns in order to bear equity market risk. Should investors have expected these returns in the past, and why shouldn’t they continue to do so? We examine these questions expressed in a slightly different way. First, can we derive an objective estimate of what investors had good reasons to expect in the past? Second, why should we expect less in the future than we have earned in the past?

The answers to both questions lie in the difference between the observed excess return and the prospective risk premium, two fundamentally different concepts that, unfortunately, carry the same label—risk premium. If we distinguish between past excess returns and future expected risk premiums, the idea that future risk premiums should be different from past excess returns is not at all unreasonable.

Robert D. Arnott is managing partner at First Quadrant, L.P., Pasadena, California. Peter L. Bernstein is president of Peter L. Bernstein, Inc., New York.
This complex topic requires several careful steps if it is to be evaluated correctly. To gauge the risk premium for stocks relative to bonds, we need an expected real bond return and an expected real stock return. To gauge the expected real bond return, we need both bond yields and an estimate of expected inflation through history. To gauge the expected real stock return, we need both stock dividend yields and an estimate of expected real dividend growth. Accordingly, we go through each of these steps, in reverse order, to form the building blocks for the final goal—an estimate of the objective forward-looking equity risk premium relative to bonds through history.

Has the Risk Premium Natural Limits?

For equities to have a zero or negative risk premium relative to bonds would be unnatural because stocks are, on average over time, more volatile than bonds. Even if volatility were not an issue, stocks are a secondary call on the resources of a company; bondholders have the first call. Because the risk premium is usually measured for corporate stocks as compared with government debt obligations (U.S. T-bonds or T-bills), the comparison is even more stark. Stocks should be priced to offer a superior return relative to corporate bonds, which should offer a premium yield (because of default risk and tax differences) relative to T-bonds, which should typically offer a premium yield (because of yield-curve risk) relative to T-bills. After all, long bonds have greater duration—hence, greater volatility of price in response to yield changes—so a capital loss is easier on a T-bond than on a T-bill.

In other words, the current circumstance, in which stocks appear to have a near-zero (or negative) risk premium relative to government bonds, is abnormal in the extreme. Even if we add 100 bps to the risk premium to allow for the impact of stock buybacks, today’s risk premium relative to the more relevant corporate bond alternatives is still negligible or negative. This facet was demonstrated in Arnott and Ryan and is explored further in this article.

If zero is the natural minimum risk premium, is there a natural maximum? Not really. In times of financial distress, in which the collapse of a nation’s economy, hyperinflation, war, or revolution threatens the capital base, expecting a large reward for exposing capital to risk is not unreasonable. Our analysis suggests that the U.S. equity risk premium approached or exceeded 10 percent during the Civil War, during the Great Depression, and in the wake of World Wars I and II. That said, however, it is difficult to see how one might objectively measure the forward-looking risk premium in such conditions.

A 5 percent excess return on stocks over bonds compounds so mightily over long spans that most serious fiduciaries, if they believed stocks were going to earn a 5 percent risk premium, would not even consider including bonds in a portfolio with a horizon of more than a few years: The probabilities of stocks outperforming bonds would be too high to resist. Hence, under so-called normal conditions—encompassing booms and recessions, bull and bear markets, and “ordinary” economic stresses—a good explanation is hard to find for why expected long-term real returns should ever reach double digits or why the expected long-term risk premium of stocks over bonds should ever exceed about 5 percent. These upper bounds for expected real returns or for the risk premium, unlike the lower bound of zero, are “soft” limits; in times of real crisis or distress, the sky’s the limit.

Expected versus “Hoped-For” Returns

Throughout this article, we deal with expected returns and expected risk premiums. This concept is rooted in objective data and defensible expectations for portfolio returns, rather than in the returns that an investor might hope to earn. The distinction is subtle; both represent expectations, but one is objective and the other subjective. Even at times in the past when valuation levels were high and when stockholders would have had no objective reason to expect any growth in real dividends over the long run, hopes of better-than-market short-term profits have always been the primary lure into the game. When we refer to expected returns or expected risk premiums, we are referring to the estimated future returns and risk premiums that an objective evaluation—based on past rates of growth of the economy, past and prospective rates of inflation, current stock and bond yields, and so forth—might have supported at the time. We explicitly do not include any extrapolation of past returns per se, because past returns are driven largely by changes in valuation levels (e.g., changes in yields), which in an efficient market, investors should not expect to continue into the indefinite future. By the same token, we explicitly do not presume any reversion to the mean, in which high yields or low yields are presumed to revert toward historical norms. We presume that the current yield is “fair” and is an unbiased estimator of future yields, both for stocks and bonds.
Few investors subjectively expect returns as low as the objective returns produced by this sort of analysis. In a recent study by Welch (2000), 236 financial economists projected, on average, a 7.2 percent risk premium for stocks relative to T-bills over the next 30 years. If we assume that T-bills offer the same 0.7 percent real return in the future that they have offered over the past 75 years, then stocks must be expected to offer a compounded geometric average real return of about 6.6 percent. Given a dividend yield of roughly 1.5 percent in 1998–1999, when the survey was being carried out, the 236 economists in the survey were clearly presuming that dividend and earnings growth will be at least 5 percent a year above inflation, a rate of real growth three to five times the long-term historical norm and substantially faster than plausible long-term economic growth.

Indeed, even if investors take seriously the real return estimates and risk premiums produced by the sort of objective analysis we propose, many of them will continue to believe that their own investments cannot fail to do better. Suppose they agree with us that stocks and bonds are priced to deliver 2–4 percent real returns before taxes. Do they believe that their investments will produce such uninspired pretax real returns? Doubtful. If these kinds of projections were taken seriously, markets would be at far different levels from where they are. Consequently, if these objective expectations are correct, most investors will be wrong in their (our?) subjective expectations.

What Were Investors Expecting in 1926

Are we being reasonable to suggest that, after a 75-year span with 8 percent real stock returns and a 5 percent excess return over bonds (the Ibbotson findings), an 8 percent real return or a 5 percent risk premium is abnormal? Absolutely. The relevant question is whether the investors of 1926 would have had reason to expect these extraordinary returns. In fact, they would not. What they got was different from what they should have expected, which is a normal result in a world of uncertainty.

At the start of 1926, the beginning of the returns covered in the Ibbotson data, investors had no reason to expect the 8 percent real returns that have been earned over the past 75 years nor that these returns would provide a 5 percent excess return over bonds. As we will describe, these outcomes were the consequence of a series of historical accidents that uniformly helped stocks and/or helped the risk premium.

Consider what investors might objectively have expected at the start of 1926 from their long-term investments in stocks and bonds. In January that year, government bonds were yielding 3.7 percent. The United States was on a gold standard, government was small relative to the economy as a whole, and the price level of consumer goods, although volatile, had been trendless throughout most of U.S. history up to that moment; thus, inflation expectations were nil. It was a time of relative stability and prosperity, so investors would have had no reason to expect to receive less than this 3.7 percent government bond yield. Accordingly, the real return that investors would have expected on their government bonds was 3.7 percent, plain and simple.

Meanwhile, the dividend yield on stocks was 5.1 percent. We can take that number as the starting point to apply the sound theoretical notion that the real return on stocks is equal to

- the dividend yield
- plus or minus any change in the real return (now viewed as participation in economic growth)
- plus or minus any change in valuation levels, as measured by P/E multiples or dividend yields.

What did the investors expect of stocks in early 1926? The time was the tail end of the era of “robber baron” capitalism. As Chancellor (1999) observed, investors were accustomed to the fact that company managers would often dilute shareholders’ returns if an enterprise was successful but that the shareholder was a full partner in any business decline. More important was the fact that the long-run history of the market was trendless. Thoughts of long-term economic growth, or long-run capital appreciation in equity holdings, were simply not part of the tool kit for return calculations in those days.

Investors generally did not yet consider stocks to be “growth” investments, although a few people were beginning to acknowledge the full import of Smith’s extraordinary study Common Stocks as Long-Term Investments, which had appeared in 1924. Smith demonstrated how stocks had outperformed bonds over the 1901–22 period. His work became the bible of the bulls as the bubble of the late 1920s progressed. Prior to 1926, however, investors continued to follow J.P. Morgan’s dictum that the market would fluctuate, a traditional view hallowed by more than 100 years of stock market history. In other words, investors had no trend in mind. The effort was to buy low and to sell high, period.

Assuming that markets were fairly priced in early 1926, investors should have expected little or no benefit from rising valuation levels. Accordingly, the real long-term return that stock investors could reasonably have expected on average, or from
the market as a whole, was the 5.1 percent dividend yield, give or take a little. Thus, stock investors would have expected roughly a 1.4 percent risk premium over bonds, not the 5 percent they actually earned in the next 75 years. The market exceeded objective expectations as a consequence of a series of historical accidents:

- **Historical accident #1: Decoupling yields from real yields.** The Great Depression (roughly 1929–1939) introduced a revolutionary increase in the role of government in peacetime economic policy and, simultaneously, drove the United States (and just about the rest of the world) off the gold standard. As prosperity came back in a big way after World War II, expected inflation became a normal part of bond valuation. This change created a one-time shock to bonds that decoupled nominal yields from real yields and drove nominal yields higher even as real yields fell. Real yields at year-end 2001 were 3.4 percent (the Treasury Inflation-Indexed Securities, commonly called TIPS, yield\(^9\)), but nominal yields were 5.8 percent. This rise in nominal yields (with real yields holding steady) has cost bondholders 0.4 percent a year over 75 years. That accident alone accounts for nearly one-tenth of the 75-year excess return for stocks relative to bonds.

- **Historical accident #2: Rising valuation multiples.** Between 1926 and 2001, stocks rose from a valuation level of 18 times dividends to nearly 70 times dividends. This fourfold increase in the value assigned to each dollar of dividends contributed 180 bps to annual stock returns over the past 75 years, even though the entire increase occurred in the last 17 years of the period (we last saw 5.1 percent yields in 1984). This accident explains fully one-third of the 75-year excess return.

- **Historical accident #3: Survivor bias.** Since 1926, the United States has fought no wars on its own soil, nor has it experienced revolution. Four of the fifteen largest stock markets in the world in 1900 suffered a total loss of capital, a –100 percent return, at some point in the past century. The markets are China, Russia, Argentina, and Egypt. Two others came close—Germany (twice) and Japan. Note that war or revolution can wipe out bonds as easily as stocks (which makes the concept of “risk premium” less than relevant). U.S. investors in early 1926 would not have considered this likelihood to be zero, nor should today’s true long-term investor.

- **Historical accident #4: Regulatory reform.** Stocks have gone from passing relatively little economic growth through to shareholders to passing much of the economic growth through to shareholders. This shift has led to 1.4 percent a year growth in real dividend payments and in real earnings since 1926. This accelerated growth in real dividends and earnings, which no one in 1926 could have anticipated, explains roughly one-fourth of the 75-year excess return.\(^10\)

In short, the equity investors of 1926 probably expected to earn a real return little different from their 5.1 percent yield and expected to earn little more than the 140 bp yield differential over bonds. Indeed, an objective investor might have expected a notch less because of the greater frequency with which investors encountered dividend cuts in those days.

**What Expectations Were Realistic in the Past?**

To gauge what risk premium an investor might have objectively expected in the longer run past, we need to (1) estimate the real return that investors might reasonably have expected from stocks, (2) estimate the real return that investors might reasonably have expected from bonds, and (3) take the difference. From this exercise, we can gauge what risk premium an investor might reasonably have expected at any point in history, not simply an isolated snapshot of early 1926. A brief review of the sources of stock returns over the past two centuries should help lay a foundation for our work on return expectations and shatter a few widespread misconceptions in the process. The sources of the data are given in Appendix A.\(^11\)

**Step I: How Well Does Economic Growth Flow into Dividend Growth?** Over the past 131 years, since reliable earnings data became available in 1870, the average earnings yield has been 7.6 percent and the average real return for stocks has been 7.2 percent; this close match has persuaded many observers to the view (which is wholly consistent with finance theory) that the best estimate for real returns is, quite simply, the earnings yield. On careful examination, this hypothesis turns out to be wrong. In the absence of changing valuation levels, real returns are systematically lower than earnings yields.

Figure 1 shows stock market returns since 1802 in a fashion somewhat different from that shown in most of the literature. The solid line in Figure 1 shows the familiar cumulative total return for U.S. equities since 1802, in which each $100 invested grows, with reinvestment of dividends, to almost $700 million in 200 years. To be sure, some of this growth came from inflation; as the line “Real Stock Return” shows, $700 million will not buy what it...
would have in 1802, when one could have purchased the entire U.S. GNP for less than that sum.\textsuperscript{12} By removing inflation, we show in the “Real Stock Return” line that the $100 investment grew to “only” $37 million. Thus, adjusted for inflation, our fortune is much diminished but still impressive. Few portfolios are constructed without some plans for future spending, and the dividends that stocks pay are often spent. So, the “Real Stock Price Index” line shows the wealth accumulation from price appreciation alone, net of inflation and dividends. This bottom line (literally and figuratively) reveals that stocks have risen just 20-fold from 1802 levels. Put another way, if an investor had placed $100 in stocks in 1802 and received and spent the average dividend yield of 4.9 percent for the next 200 years, his or her descendants would today have a portfolio worth $2,099, net of inflation. So much for our $700 million portfolio!

Worse, the lion’s share of the growth from $100 to $2,099 occurred in the massive bull market from 1982 to date. In the 180 years from 1802 to the start of 1982, the real value of the $100 portfolio had grown to a mere $400. If stocks were priced today at the same dividend yields as they were in 1802 and 1982, a yield of 5.4 percent, the $100 portfolio would be worth today, net of inflation and dividends, just $550. These data put the lie to the conventional view that equities derive most of their returns from capital appreciation, that income is far less important, if not irrelevant.

Figure 2 allows a closer look at the link between equity price appreciation and economic growth. It shows that the growth in share prices is much more closely tied to the growth in real per capita GDP (or GNP) than to growth in real GDP per se. The solid line shows that, compounding at about 4 percent in the 1800s and 3 percent in the 1900s, the economy itself delivered an impressive 1,000-fold growth.
But net of inflation and dividend distributions, stock prices (the same “Real Stock Price Index” line in Figure 1) fell far behind, with cumulative real price appreciation barely 1/50 as large as the real growth in the economy itself.

How can this be? Can’t shareholders expect to participate in the growth of the economy? No. Shareholders can expect to participate only in the growth of the enterprises they are investing in. An important engine for economic growth is the creation of new enterprises. The investor in today’s enterprises does not own tomorrow’s new enterprises—not without making a separate investment in those new enterprises with new investment capital.

Finally, the “Real Per Capita GDP Growth” line in Figure 2 shows the growth of the economy measured net of inflation and population growth. This growth in real per capita GDP tracks much more closely with the real price appreciation of stocks (the bottom line) than does real GDP itself.

Going one step further, Figure 3 shows the internal growth of real dividends—that is, the growth that an index fund would expect to see in its own real dividends in the absence of additional investments, such as reinvestment of dividends.\(^\text{13}\)

Real dividends exhibit internal growth that is similar to the growth in real per capita GDP. Because growth in per capita GDP is a measure of productivity growth, the internal growth that can be sustained in a diversified market portfolio should closely match the growth of productivity in the economy, not the growth in the economy per se. Therefore, the dotted line traces per capita real GDP growth, the “Real Stock Price Index” line shows real stock prices, and the bottom line shows real dividends \((\times 10)\).\(^\text{14}\) Figure 3 reveals the remarkable resemblance between real dividend growth and growth in real per capita GDP.

When we measure the internal growth of real dividends as in Figure 3, we see that real dividends have risen a modest fivefold from 1802 levels. In other words, the real dividends for a $100 portfolio invested in 1802 have grown merely 0.9 percent a year net of inflation. To be sure, the price assigned to each dollar of dividends has quadrupled, which leads to the 20-fold real price gain in the 200 years.

Although real dividends have tracked remarkably well with real per capita GDP, they have consistently fallen short of GDP gains. Not only have real dividends failed to match real GDP growth (as many equity investors seem to think is a minimal future growth rate for earnings and dividends), they have even had a modest shortfall, at an average of about 70 bps a year, relative to per capita economic growth.

In short, more than 85 percent of the return on stocks over the past 200 years has come from (1) inflation, (2) the dividends that stocks have paid, and (3) the rising valuation levels (rising P/Es and falling dividend yields) since 1982, not from growth in the underlying fundamentals of real dividends or earnings.\(^\text{15}\) Furthermore, real dividends and real per capita GDP both grew faster in the 20th century than in the 19th century. Conversely, GDP grew faster in the 19th century than in the 20th century, unless we convert to per capita GDP.

Many observers think that earnings growth is far more important than dividend growth. We respectfully disagree. As noted by Hicks (1946), “... any increase in the present value of prospective net receipts must raise profits.” In other words, properly stated, earnings should represent a proportional share of the net present value of all future earnings.
profits. The problem is that reported earnings often do not follow this theoretical definition. For example, negative earnings should almost never be reported, yet reported operating losses are not uncommon. Furthermore, the quality of earnings reports prior to the advent of the U.S. SEC is doubtful at best; worse, we were unable to find any good source for earnings information prior to 1870. Accordingly, the dividend is the one reliable aspect of stock ownership over the past two centuries. It is the cash income returned to the shareholders; it is the means by which the long-term investor earns most of his or her internal rate of return. Finally, with earnings growth barely 0.3 percent faster than dividend growth over the past 131 years, an analysis based on earnings would reach conclusions nearly identical to our conclusions based on dividends.

Finance theory tells us that capital is fungible; that is, equity and debt, retained earnings and dividends—all should flow to the best use of capital and should (in the absence of tax-related arbitrages and other nonsystematic disruptions) produce a similar risk-adjusted return on capital. Thus, the retained earnings should deliver a return similar to the return an investor could have earned on that capital had it been paid out as dividends. Consider an example: If a company has an earnings yield of 5 percent (corresponding to a P/E of 20), it can pay out all of the earnings and thereby deliver a 5 percent yield to the shareholder. The real value of the company should not be affected by this full earnings distribution (unless the earnings are themselves being misstated), so the 5 percent earnings yield should also be the expected real return. Now, if the company, instead, pays a 2 percent yield and retains earnings worth 3 percent of the stock price, the company ought to achieve 3 percent real growth in earnings; otherwise, it should have distributed the cash to the shareholders. How does this theory stand up to reality?

Over the past 200 years, dividend yields have averaged 4.9 percent, yet real returns have been far higher, 6.6 percent. Since 1870, earnings yields have averaged 7.6 percent, close to the real returns of 7.2 percent over that span. This outcome is consistent with the notion of fungible capital, that the return on capital reinvested in an enterprise ought to match the return an investor might otherwise have earned on that same capital if it had been distributed as a dividend. However, if we take out the changes in valuation levels since 1982 (regardless of whether dividend yields or P/Es are used for those levels), the close match between earnings yield and real stock returns evaporates.

Moreover, with an average earnings yield of 7.6 percent and an average dividend yield of 4.7 percent since 1871, the average “retained earnings yield” has been nearly 3 percent. This retained earnings yield should have led to real earnings and dividend growth of 3 percent; otherwise, management ought to have paid this money out to the shareholders. Instead, real dividends and earnings grew at annual rates of, respectively, 1.2 percent and 1.5 percent. Where did the money go? The answer is that during the era of “pirate capitalism,” success often led to dilution: Company managers issued themselves more stock

Furthermore, retained earnings often chase poor internal reinvestment opportunities. If existing enterprises experienced only 1.2–1.5 percent internal growth of real dividends and earnings in the past two centuries, most of the 3.6 percent economic growth the United States has enjoyed has clearly not come from reinvestment in existing enterprises. In fact, it has stemmed from entrepreneurial capitalism, from the creation of new enterprises. Indeed, dividends on existing enterprises have fallen relative to GDP growth by approximately 100-fold in the past 200 years.

The derring-do of the pirate capitalists of the 19th and early 20th centuries is not the only or even the most compelling explanation for this phenomenon. All the data we used are from indexes, which are a particular kind of sampling of the market. Old companies fading from view lose their market weight as the newer and faster growing companies gain a meaningful share in the economy. The older enterprises often have the highest earnings yield and the worst internal reinvestment opportunities, but the new companies do not materialize in the indexes the minute they start doing business or even the minute they go public. When they do enter the index, their starting weight is often small.

Furthermore, an index need only change the divisor whenever a new enterprise is added, whereas we cannot add a new enterprise to our portfolio without cost. The index changing the divisor is mathematically the same as selling a little bit of all other holdings to fund the purchase of a new holding, but when we add a new enterprise to our portfolios, we must commit some capital to effect the purchase. Whether through reinvestment of dividends or infusion of new capital, this new enterprise cannot enter our portfolio through the internal growth of an existing portfolio of assets. In effect, we must rebalance out of existing stocks to make room for the new stock—which produces the natural dilution that takes place as a consequence of the creation of new enterprises in a world of entrepreneurial capitalism: The same dollar cannot own an existing enterprise and simultaneously fund a new enterprise.
The dynamics of the capitalist system inevitably lead to these kinds of results. Good business leads to expansion; in a competitive environment, expansion takes place on a wide scale; expansion on a wide scale intensifies the competitive environment; margins begin to decline; earnings growth slows; in time, earnings begin to decline; then, expansion slows, profit margins improve, and the whole thing repeats itself. We can see this drama playing out in the relationship between payout ratios in any given year and earnings growth: Since 1984, the payout ratio has explained more than half of the variation in five-year earnings growth rates with a $t$-statistic of 9.51.19

Few observers have noticed that much of the difference between stock dividend yields and the real returns on stocks can be traced directly to the upward revaluation of stocks since 1982. The historical data are muddied by this change in valuation levels—which is why we find the current fashion of forecasting the future by extrapolating the past to be so alarming. The earnings yield is a better estimate of future real stock returns than any extrapolation of the past. And the dividend yield plus a small premium for real dividend growth is even better, because in the absence of changes in valuation levels, the earnings yield systematically overstates future real stock returns.

If long-term real growth in dividends had been 0.9 percent, real stock returns would have been only 90 bps higher than the dividend yield if it were not for the enormous jump in the price-to-dividend ratio since 1982. Even if we adjust today’s 1.4 percent dividend yield sharply upward to include “dividends by another name” (e.g., stock buybacks), making a case for real returns higher than the 3.4 percent currently available in the TIPS market would be a stretch.20

**Step II: Estimating Real Stock Returns.**
To estimate the historical equity risk premium, we must compare (1) a realistic estimate of the expected real stock return that objective analysis might have supported in past years with (2) the expected real bond return available at the time. Future long-term real stock return is defined as21

$$ RSR(t) = DY(t) + RDG(t) + ΔPD(t) + ε, $$

where

- $DY(t)$ = percentage dividend yield for stocks at time $t$
- $RDG(t)$ = percentage real dividend growth rate over the applicable span starting at time $t$
- $ΔPD(t)$ = percentage change in the price assigned to each dollar of dividends starting at time $t$
- $ε$ = error term for sources of return not captured by the three key constituents (this term will be small because it will reflect only compounding effects)

Viewed from the perspective of forecasting future real returns, the $ΔPD(t)$ term is a valuation term, which we deliberately exclude from our analysis. If markets exhibit reversion to the mean, valuation change should be positive when the market is inexpensive and negative when the market is richly priced. If markets are efficient, this term should be random. We choose not to go down the slippery slope of arguing valuation, even though we believe that valuation matters. Rather, we prefer to make the simplifying assumption that market valuations at any stage are “fair” and, therefore, that the real return stems solely from the dividend yield and real growth of dividends.

That said, the estimation process becomes more complex when we consider a sensible estimate for real dividend growth. For example, what real dividend growth rate might an investor in 1814 have expected on the heels of the terrible 1802–14 bear market and depression, during which real per capita GDP, real dividends, and real stock prices all contracted 40–50 percent? How can we objectively put ourselves in the position of an investor almost 200 years ago? For this purpose, we partition the real growth in dividends into two constituent parts, real economic growth and the growth of dividends relative to the economy.

Why not simply forecast dividend growth directly? Because countless studies have shown that analysts’ forecasts are too optimistic, especially at market turning points. In fact, dividends (and earnings) in aggregate cannot grow as fast as the economy on a sustainable long-term basis, in large part because of the secular increase in shares outstanding and introduction of new enterprises. So, long-term dividend growth should be equal to long-term economic growth minus a haircut for dilution or entrepreneurial capitalism (the share of economic growth that is tied to new enterprises not yet available in the stock market) or plus a premium for hidden dividends, such as stock buybacks. So, real dividend growth is given by

$$ RDG(t) = RGDP(t) + DGR(t) + ε, $$

where

- $RGDP(t)$ = percentage real per capita GDP growth over the applicable span starting at time $t$
- $DGR(t)$ = annual percentage dilution of real GDP growth as it flows through to real dividends starting at time $t$
- $ε$ = error term for compounding effects (it will be small)
Basicly, in Equation 2, we are substituting $RGDP(t) + DGR(t)$ for $RDG(t)$ and rolling the $\Delta PD(t)$ term into the error term (to avoid getting into the debates about valuation and regression to the mean). With these two changes, and converting to an expectations model, our model for expected real stock market returns, $ERSR$, becomes

$$ERSR(t) = EDY(t) + ERGDP(t) + EDGR(t),$$

where

- $EDY(t) = \text{expected percentage dividend yield for stocks at time } t$
- $ERGDP(t) = \text{expected percentage real per capita GDP growth over the applicable span starting at time } t$
- $EDGR(t) = \text{expected annual percentage dilution of real per capita GDP growth as it flows through to real dividends starting at time } t$

A complication in this structure is the impact of recessions. In serious recessions, dividends are cut and GDP growth stops or reverses, possibly leading to a decline in even the long-term GDP growth. The result is a dividend yield that is artificially depressed, real per capita GDP growth that is artificially depressed, and long-term dividend growth relative to GDP growth that is artificially depressed, all three of which lead, in recessionary downturns in computing each of the three constituents of expected real stock returns.\textsuperscript{22}

We illustrate how we constructed an objective real stock return forecast for the past 192 years in Figure 4; Panel A spans 1810 to 2001, and Panel B shows the same data after 1945. To explain these graphs, we will go through them line by line.

The easiest part of forecasting real stock returns, the “Estimated Real Stock Return” line in Figure 4, is the dividend yield: It is a known fact. We have adjusted dividends to correct for the artificially depressed dividends during recessions to get the $EDY(t)$ term shown as the “Dividend Yield” line in Figure 4. This step allows us to avoid understating the equity risk premium in recessions when dividends are artificially depressed. This adjustment boosts the expected dividend yield slightly relative to the raw dividend yield because the deepest recessions are often deeper than the average recessions of the prior 40 years. Against an average dividend yield of 4.9 percent, we found an average expected dividend yield of 5.0 percent.

Most long-run forecasts of earnings or dividend growth ignore the simple fact that aggregate earnings and dividends in the economy cannot sustainably grow faster than the economy itself. If new enterprise creation and secondary equity offerings dilute the share of the economy held by the shareholders in existing enterprises, then one sensible way to forecast dividend growth is to forecast economic growth and then forecast how rapidly this dilution will take place.\textsuperscript{23} Stated another way, we want to know how much less rapidly dividends (and earnings) on existing enterprises can grow than the economy at large. The sum of real economic growth less this shortfall is the real growth in dividends.

The resulting line, “Dilution of GDP Growth in Dividends,” in the two graphs of Figure 4 represents the $EDGR(t)$ term in our model (Equation 3). Note the persistent tendency for dividend growth to lag GDP growth: Real dividends have grown at 1 percent a year over the past 192 years, whereas the real economy has grown at 3.8 percent a year, and even real per capita GDP has grown at 1.8 percent a year. Why should real dividends have grown so much more slowly than the economy?

First, much of the growth in the economy has come from innovation and entrepreneurial capitalism. More than half of the capitalization of the Russell 3000 today consists of enterprises that did not exist 30 years ago. The 1971 buy-and-hold investor could not participate in this aspect of GDP growth or market growth because the companies did not exist. So, today’s dividends and earnings on the existing companies from 1971 are only part of the dividends and earnings on today’s total market.

Second, as was demonstrated in Bernstein (2001b), retained earnings are often not reinvested at a return that rivals externally available investments; earnings and dividend growth are faster when payout ratios are high than when they are low, perhaps because corporate managers are then forced to be more selective about reinvestment alternatives.\textsuperscript{24}

Finally, as we have emphasized, corporate growth typically leads to more shares outstanding, which automatically imposes a drag on the growth in dividends per share.

As a sensible estimate of the future dividend/GDP shortfall, the rational investor of any day might forecast dividend growth by using the prior 40-year shortfall in dividend growth relative to per capita GDP or might choose to use the cumulative (by now, 200-year) history. We chose the simple expedient of averaging the two.

The dilution effect we found from the 40-year and cumulative data for real dividends and real per capita GDP averages ~60 bps. So, in the past 40 years, the dilution of dividend growth is almost
exactly the same as the long-term average, –80 bps. With a standard deviation of just 0.5 percent, this shortfall of dividend growth relative to economic growth is the steadiest of any of the components of real stock returns or real bond returns. It has never been materially positive on a long-term sustained basis; it has never risen above +10 bps for any 40-year span in the entire history since 1810.

The history of dividend growth shows no evidence that dividends can ever grow materially faster than per capita GDP. Indeed, they almost always grow more slowly. Suppose real GDP growth in the next 40 years is 3 percent a year and population growth is 1 percent a year. These assumptions would appear to put an upper limit on real dividend growth at a modest 2 percent a year, far below consensus expectations. If the historical average dilution of dividend growth relative to real per capita GDP growth prevails, then the future real growth in dividends should be only about 1 percent, even with relatively robust, 2.5–3.0 percent, real GDP growth.

Now consider the third part of forecasting real stock returns in this fashion—the forecast of long-term real per capita GDP growth, \( ERGDP(t) \) in our model. How much real per capita GDP growth would an investor have expected at any time in the past 200 years? Again, a simple answer might come from the most recent 40 years’ growth rate; another might come from the cumulative record going back as far as we have dividend and GDP data, to 1802. These historical data are shown in the “Real per Capita GDP Growth” line in Figure 4. And again, we chose the simple expedient of averaging the average of the two. Real per capita GDP growth has been remarkably stable over the past 200 years, particularly if we adjust it to correct for temporary dips during recessions. If we examine truly long-term
Note from Figure 4 that the total economy grew faster during the 19th century than the 20th century whereas stock returns (and the underlying earnings and dividends) grew faster in the 20th century than the 19th. Why would the rapid growth of the 19th century flow through to the shareholder less than the slower growth of the 20th century? We see two possible answers. First, the base from which industrial growth started in the 19th century was so much smaller that much faster new enterprise creation occurred then than in the 20th century. Second, with nearly 3 percent growth in the population from 1800 to 1850, the growing talent and labor pool fueled a faster rate of growth than the 1.25 percent annual population growth rate of the most recent 50 years. It is not surprising that the pace of dilution, both from the creation of new enterprises and from secondary equity offerings, is faster when the population is growing faster. Population growth fuels growth in human capital, in available labor, and in both demand and supply of goods and services. As a result, when population growth is rapid, the pace of dilution of growth in the economy (as it flows through to a shareholder’s earnings and dividends) is far more stable relative to real per capita GDP than relative to real GDP itself.

The simple framework we have presented for estimating real stock returns reveals few surprises. As Panels A and B of Figure 4 show, the expected stock return is the sum of the three constituent parts graphed in the other lines. We estimate that expected real stock returns for the past 192 years averaged about 6.1 percent with the following constituent parts: an expected yield averaging 5.0 percent plus real per capita GDP growth of 1.7 percent a year minus an expected shrinkage in dividends relative to real per capita GDP averaging –0.6 percent. Meanwhile, investors actually earned real returns of 6.8 percent. Most of this 70 bp difference from the 6.1 percent rational expectation over the past 192 years can be traced to the rise in valuation levels since 1982; the rest consists of the other happy accidents detailed previously.

Expectations for real stock returns have soared above 6 percent often enough that many actuaries even today consider 8 percent a “normal” real return for equities. Our estimate for real stock returns, however, exceeds 8 percent only during the depths of the Great Depression, in the rebuilding following the War of 1812, the Civil War, World War I, and World War II, and in the Crash of 1877. In the past 50 years, expected real stock returns above 7 percent have been seen only in the aftermath of World War II, when many investors still feared a return to Depression conditions, and in the depths of the 1982 bear market.

When viewed from the vantage point of this formulation for expected real stock returns, the full 192-year record shows that expected real stock returns fell below 3.5 percent only once before the late 1990s, at the end of 1961 just ahead of the difficult 1962–82 span, real stock prices fell by more than 50 percent. Since 1997, expected real stock returns have fallen well below the 1961 levels, where they remain at this writing.

This formulation for expected real stock returns reveals the stark paradigm shift that took place in the 1950s. Until then, the best estimate for real dividend growth was rarely more than 1 percent, so the best estimate for real stock returns was approximately the dividend yield plus 100 bps—considerably less than the earnings yield! From the 1950s to date, as Panel B of Figure 4 shows, the shortfall of dividends relative to GDP growth improved (perhaps because the presence of the SEC discourages company managers from ignoring shareholder interests) and the real return that one could objectively expect from stocks finally and persuasively rose above the dividend yield. Today, it stands at almost twice the dividend yield, but it is still a modest 2.4 percent.

Figure 5 shows the strong correlation between our formulation for expected real stock returns and the actual real returns that stocks have delivered over the subsequent 10-year span. The correlation is good—at 0.62 during the modern market era after World War II and 0.46 for the full 182 years.26 If we test the correlation between this simple metric of expected real stock returns and the actual subsequent 20-year real stock returns (not shown), the correlations grow to 0.95 and 0.60 for the post-1945 period and the full 182 years, respectively.

---

**Figure 5. Estimated and Subsequent Actual Real Stock Returns, 1802–2001**

![Graph showing the relationship between estimated and subsequent actual real stock returns between 1802 and 2001.](image-url)
The regression results given in Panel A Table 1 show that the coefficient in the regression is larger than 1.00. So, that 100 bp increase in the expected real stock return, ERSR, is worth more than 100 bps in the subsequent 10-year actual real stock return, RSR. The implication is that some tendency for reversion to the mean does exist and that it will magnify the effect of unusually high or low expected real stock returns. This suggestion has worrisome implications for the recent record low levels for expected real stock returns.

Because rolling 10-year returns (and expected returns in our model) are highly serially correlated, the t-statistics given in Panel A of Table 1 are not particularly meaningful. One way to deal with overlapping data is to eliminate the overlap by using nonoverlapping samples—in this case, examining only our 19 nonoverlapping samples beginning December 1810. The Panel B results, with a coefficient larger than 1.00, confirm the previous results (and approach statistical significance, even with only 17 degrees of freedom). One worrisome fact, in light of the recent large real stock returns, is that the nonoverlapping real stock returns by decades have a –31 percent serial correlation. Although it is not a statistically significant correlation, it is large enough to be interesting: It suggests that spectacular decades or wretched decades may be considerably more likely to reverse than to repeat.

Evaluating the real returns on stocks is clearly a useful exercise if the metric of success for a model is subsequent actual real returns, but we live in a relative world. The future real returns on all assets will rise and fall; so, real returns are an insufficient metric of success. What is of greater import is whether this metric of prospective real stock returns helps us identify the attractiveness of stocks relative to other assets.

Table 1. Regression Results: Estimated Real Stock Return versus Actual 10-Year Real Stock Return (t-statistics in parentheses)

<table>
<thead>
<tr>
<th>Period</th>
<th>Period</th>
<th>1810–2001</th>
<th>1945–2001</th>
<th>(a)</th>
<th>(b)</th>
<th>(R^2)</th>
<th>Correlation</th>
<th>Serial Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Raw data: (RSR(t) = a + b[ERSR(t - 120)])</td>
<td></td>
<td>(-1.51%)</td>
<td>(-7.80)</td>
<td>1.38%</td>
<td>3.15%</td>
<td>0.214</td>
<td>0.46</td>
<td>0.992</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-4.2)</td>
<td>(-8.8)</td>
<td></td>
<td></td>
<td>(24.4)</td>
<td></td>
<td>0.991</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.35%</td>
<td>0.182</td>
<td></td>
<td></td>
<td>0.182</td>
<td>0.430</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.1)</td>
<td>(1.9)</td>
<td></td>
<td></td>
<td>(19.0)</td>
<td></td>
<td>0.021</td>
</tr>
<tr>
<td>B. Using 19 nonoverlapping samples, beginning December 1810</td>
<td></td>
<td>(-0.35%)</td>
<td>1.22%</td>
<td>0.182</td>
<td>0.430</td>
<td>0.996</td>
<td>0.315</td>
<td>0.996</td>
</tr>
</tbody>
</table>

Step III: Estimating Future Real Bond Returns. On the bond side, real realized returns are equal to the nominal yield minus inflation (or plus deflation) and plus or minus yield change times duration:

\[
RBR(t) = BY(t) - INFL(t) + ΔBY(t)DUR(t) + ε, \quad (4)
\]

where

\[
BY(t) = \text{percentage bond yield at time } t
\]

\[
INFL(t) = \text{percentage inflation over the applicable span starting at time } t
\]

\[
ΔBY(t)DUR(t) = \text{annual change in yield over the applicable span times duration at time } t \text{ (under the assumption that rolling reinvestment is in bonds of similar duration)}
\]

\[
ε = \text{error term (compounding effects lead to a small error term in this simple formulation)}
\]

As with stocks, we prefer to take current yields as a fair estimate of future bond yields. So, we eliminate the variable that focuses on changes in yields, \(ΔBY(t)DUR(t)\). We also need to shift our focus from measuring past real bond returns to forecasting future real bond returns. Therefore, our model is

\[
ERBR(t) = BY(t) - EINFL(t), \quad (5)
\]

where \(BY(t)\) is the percentage bond yield at time \(t\) and \(EINFL(t)\) is the expected percentage inflation over the applicable span starting at time \(t\).

Equation 5 is difficult only in the sense that expectations for inflation in past economic environs are difficult to estimate objectively. How, for example, are we to gauge how much inflation an investor in February 1864 would have expected at a time when inflation had averaged 20 percent over the prior three years because of wartime shortages?
Expectations would depend strongly on the outcome of the war: A victory by the North would have been expected to result in a restoration of the purchasing power of the dollar as wartime shortages disappeared; a victory by the South could have had severe consequences on the ultimate purchasing power of the North’s dollar as a consequence of debt that could no longer be serviced. A rational expectation might have been for inflation greater than 0 (reflecting the possibility of victory by the South) but less than the 20 percent three-year inflation rate (reflecting the probability of victory by the North).

We based the estimate for expected future inflation on an ex ante regression forecast of 10-year future inflation based, in turn, on recent three-year inflation. Figure 6 shows how the expected rate of inflation has steadily become more closely tied to recent actual inflation in recent decades. Bond yields responded weakly to bursts of inflation up until the time of the Great Depression; they responded more strongly as inflation became a structural component of the economy in the past four decades.

Until the last 40 years, inflation was generally associated with wars and was virtually nonexistent—even negative—in peacetime. Figure 6 shows a burst of double-digit inflation on the heels of the War of 1812, in the late stages of the Civil War, during World War I, and in the rebuilding following World War II. And more recently, double-digit inflation characterized the “stagflation” of 1978–1981 that followed the Vietnam War and the oil shocks of the 1970s. The most notable changes since the Great Depression, especially since World War II, involve the magnitude and perceived role of government and loss of the automatic brakes once applied by the gold standard. From the end of World War II to the great inflationary crisis at the end of the 1970s, the dread of unemployment that was inherited from the Great Depression was the driving factor in both fiscal and monetary policy.

With the introduction of TIPS in January 1997, we finally have a U.S. government bond that pays a real return, which allows us to simplify the expected real bond returns to be the TIPS yield itself from that date forward; that is,

\[ ERBR(t) = YTIPS(t), \]

where \( YTIPS(t) \) is the percentage TIPS yield at time \( t \).

Figure 7 shows how the current government bond yield (the “Bond Yield” line) minus expected inflation (“Estimated Inflation”) leads to an estimate of the real bond return and hence the long-term expected real bond return (“Estimated Real Bond Yield”), which is the estimate through March of 1998 and the TIPS yield thereafter. From the Equation 5 (or, more recently, Equation 6) formulation, expected real bond returns averaged 3.7 percent over the full period, a very respectable real yield, given the limited risk of government bonds, and good recompense for an investor’s willingness to bear some bond-price volatility. Investors may not always have viewed government debt as the rock-solid investment, however, that it is generally considered today.

The 3.7 percent real bond return consists of an average nominal bond yield of 4.9 percent minus an expected inflation rate of 1.2 percent. For comparison, the average actual inflation rate has been 1.4 percent. In the years after World War II, the rate of peacetime inflation embedded in investors’ memory banks was essentially zero, perhaps even slightly negative. Consequently, bond investors kept expecting inflation to go away, despite its persistence at a modest rate in the 1950s and early 1960s and an accelerating rate thereafter. As a result, bonds were badly priced for reality during most of

---

**Figure 6. Estimating Future Inflation, 1810–2001**

![Chart showing estimated inflation over time](chart.png)
these two decades; they turned out to be certificates of confiscation for their holders until people finally woke up in the 1970s and 1980s. Actual inflation exceeded expected inflation with few exceptions from the start of World War II until roughly 1982; as can be seen in Figure 7, our model captures this phenomenon. Expectations are lower than actual outcomes during this span.

Figure 7 also shows several regimes of real yield with distinct structural change from one regime to the next. From the time the United States was in its infancy until the end of Reconstruction in the late 1870s, investors would not have viewed U.S. government bonds as a secure investment. They would have priced these bonds to deliver a 5–7 percent real yield, except during times of war. The overall stability of the yields is impressive: Unlike the history of stock prices, the surprise elements have been small.

Once the United States had survived the Civil War and the security of U.S. government debt had been demonstrated repeatedly, investors began to price government debt at a 3–5 percent real yield. As Figure 7 shows, this level held, with a brief interruption in World War I, until the country went off the gold standard in 1933. This record is remarkable in view of the high rate of economic growth, but revolutionary technological change in those days, especially in transportation and agriculture, led to such stunning reductions in product costs that inflation was kept at bay except for very brief intervals.

For the next 20–25 years, the nation struggled with the Great Depression, World War II, and the war’s aftermath. Investors slowly began to realize that deflationary price drops did not rebound fully after the trough of the Depression and that inflationary price increases did not retreat after the end of the war. The changed role of government plus the end of the gold standard had altered the picture, perhaps irrevocably. During this span, investors priced bonds to offer a 2–4 percent notional yield but a rocky −3 percent to +3 percent real yield. As Figure 7 shows, bond investors woke up late to the fact that inflation was now a normal part of life.

From the mid-1950s to date, investors have struggled with more structural inflation and more inflation uncertainty than ever before. Although investors sought to price bonds to deliver a real yield, inflation consistently exceeded their expectations. Only during the down cycle of the inflation roller coaster of 1980–1985 did bonds finally provide real yields to their owners. After this experience, bond investors developed an anxiety about inflation far greater than objective evidence would support. The result was a brief spike in real bond returns in 1984, as Figure 7 shows, with bond yields still hovering at 13.8 percent, even though three-year inflation had fallen to 4.7 percent (and our regression model for future inflation would have suggested expected inflation of 4.6 percent). The “expected” real yield was a most unusual 9.2 percent because investors were not yet prepared to believe that double-digit inflation was a thing of the past.

Another interesting fact is evident in Figure 8: The expected real bond returns produced by our formulation are highly correlated with the actual real returns earned over the subsequent decade. For 1810 to 1991, the expected real bond return has a 0.52 correlation with the actual real bond return earned over the next 10 years; from 1945 to date, the correlation rises to an impressive 0.63. Panel A of Table 2 shows that the coefficient is reliably positive but not reliably more than 1.00, which suggests that, unlike expected real stock returns, no powerful tendency for reversion to the mean is at work in real bond yields. When we used the 19 available nonoverlapping samples (Panel B), we found the resulting correlation to be 0.64, which is a statistically significant relationship.
Why is the bond model a better predictor, when raw data are used, than the stock model for the two-century history? Two reasons seem evident. First, stocks have been more volatile than bonds for almost all 200 years of U.S. data. Therefore, any model for expected real stock returns should have a larger error term. Second, stocks are by their very nature longer term than bonds: A 10-year bond expires in 10 years; stocks have no maturity date. The bond market correlations would be even better were it not for the negative real yields during times of war, when people tend to consider the inflation a temporary phenomenon. These episodes show up as the “loops” to the left of the body of the scatterplot in Figure 8. At these times, many U.S. investors apparently subordinated their own interests in a strong real yield to the needs of the nation: Long Treasury rates were essentially pegged during World War II and up to 1951, but that did not stop investors from buying them.

Step IV: Estimating the Equity Risk Premium. If we now take the difference between the expected real stock return and the expected real bond return, we are left with the expected equity risk premium:

$$ERP(t) = ERSR(t) - ERBR(t),$$

where $ERSR(t)$ is the expected real stock return starting at time $t$ and $ERBR(t)$ is the expected real bond return starting at time $t$.

Figure 9 shows the results of this simple framework for estimating the risk premium over the past 192 years. Many observers may be startled to see that this estimate of the forward-looking risk premium for stocks has rarely been above 5 percent in the past 200 years; the exceptions are war, its aftermath, and the Great Depression. The historical average risk premium is a modest 2.4 percent, albeit with a rather wide range. The wide range is more a result of the volatility of expected real bond returns than the volatility of expected real stock returns, which are surprisingly steady except in times of crisis.

Over the past 192 years, our model (Equation 3) suggests that an objective evaluation would have pegged expected real stock returns at about 6.1 percent on average, only 120 bps higher than the average dividend yield. Investors have earned fully 70 bps more than this objective expectation, but they did not have objective reasons to expect to earn as much as they did. Our model suggests that an objective evaluation would have pegged expected real bond returns at about 3.7 percent. Investors have earned 20 bps less because of the inflationary shocks of the 1960s to 1980s; they expected more than they got.

The difference between the expected real returns for stocks and bonds reveals a stark reality. An objective estimate of the expected risk premium would have averaged 2.4 percent (240 bps) during this history (6.1 percent expected real stock returns minus 3.7 percent expected real bond returns), not the oft-cited 5 percent realized excess return that

---

Table 2. Regression Results: Estimated Real Bond Return versus Actual 10-Year Real Bond Return (t-statistics in parentheses)

<table>
<thead>
<tr>
<th>Period</th>
<th>$a$</th>
<th>$b$</th>
<th>$R^2$</th>
<th>Correlation</th>
<th>Serial Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Raw data: $RBR(t) = a + b(ERBR(t - 120))$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1810–2001</td>
<td>0.45%</td>
<td>0.81%</td>
<td>0.266</td>
<td>0.52</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>(3.5)</td>
<td>(28.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1945–2001</td>
<td>−0.74</td>
<td>1.05</td>
<td>0.399</td>
<td>0.63</td>
<td>0.997</td>
</tr>
<tr>
<td></td>
<td>(−4.0)</td>
<td>(19.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Using 19 nonoverlapping samples, beginning December 1810</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1810–2001</td>
<td>−1.81%</td>
<td>1.31%</td>
<td>0.4120</td>
<td>0.64</td>
<td>0.182</td>
</tr>
<tr>
<td></td>
<td>(−1.1)</td>
<td>(3.5)</td>
<td></td>
<td></td>
<td>0.677</td>
</tr>
</tbody>
</table>
much of the investment world now depends on. Investors have earned a higher 3.3 percent (330 bps) excess return for stocks (6.8 percent actual real stock returns minus 3.5 percent for bonds), but the reason is the array of happy accidents for stocks and one extended unhappy accident for bonds.

All of this analysis is of mere academic interest, however, unless we can establish a link between our estimated risk premium and actual subsequent relative returns. Indeed, such a link does exist. The result of our formulation for the equity risk premium has a 0.79 correlation with the actual 10-year excess return for stocks over bonds since 1945 and a 0.66 correlation for the full span. This strong link is clear in Figure 10, for 1810–2001, and Table 3 (where, for convenience, we have defined the 10-year excess return of stocks relative to bonds as $ERSB$); each 100 bp change in the equity risk premium is worth modestly more than 100 bps in subsequent annual excess returns for stocks relative to bonds over the next 10 years. As with the expected stock return model (Equation 3), the link for 20-year results is stronger, with correlations over the full span and since 1945 of, respectively, 0.64 and 0.95.

This strong link between objective measures of the risk premium and subsequent stock–bond excess returns is also clear for the 1945–2001 period shown in Figure 11, in which every wiggle of our estimate for the risk premium is matched by a similar wiggle in the subsequent 10-year excess return that stockholders earned relative to bondholders. Figure 11 shows that the excess returns on stocks relative to bonds became negative in the late 1960s on a 10-year basis, following low points in the risk premium, and again touched zero 10 years after the 1981 peak in bond yields.

We can also see in Figure 11 how the gap in 10-year results opened up sharply for the 10 years of the 1990s; it opened to unprecedented levels, even wider than in the early 1960s. Prior to this gap opening, the fit between the risk premium and subsequent excess returns is remarkably tight. The question is whether this anomaly is sustainable or is destined to be “corrected.” History suggests that such anomalies are typically corrected, especially when the theoretical case to support them is so weak. This reminder should be sobering to investors who are depending on a large equity risk premium.
As with the models for real stock returns and for real bond returns, we also used nonoverlapping spans to take out the effect of the strong serial correlation in the estimated risk premium. For the 19 nonoverlapping spans (Panel B of Table 3), the correlation for the full period jumps to 0.70, with a highly significant $t$-statistic of 4.0.32

## Conclusions

We have advanced several provocative assertions.

- The observed real stock returns and the excess return for stocks relative to bonds in the past 75 years have been extraordinary, largely as a result of important nonrecurring developments.
- It is dangerous to shape future expectations based on extrapolating these lofty historical returns. In so doing, an investor is tacitly assuming that valuation levels that have doubled, tripled, and quadrupled relative to underlying earnings and dividends can be expected to do so again.
- The investors of 75 years ago would not have had an objective basis for expecting the 8 percent real returns or 5 percent risk premium that stocks subsequently delivered. The estimated equity risk premium at the time was above average, however, which makes 1926 a better-than-average starting point for the historical risk premium.
- The real internal growth that companies generated in their dividends averaged 0.9 percent a year over the past 200 years, whereas earnings growth averaged 1.4 percent a year over the past 131 years.
- Dividends and earnings growth was slower than the increase in real per capita GDP, which averaged 1.6 percent over the past 200 years and 2.0 percent over the past 131 years. This internal growth is far less than the consensus expectations for future earnings and dividend growth.
• The historical average equity risk premium, measured relative to 10-year government bonds as the risk premium investors might objectively have expected on their equity investments, is about 2.4 percent, half what most investors believe.

• The “normal” risk premium might well be a notch lower than 2.4 percent because the 2.4 percent objective expectation preceded actual excess returns for stocks relative to bonds that were nearly 100 bps higher, at 3.3 percent a year.

• The current risk premium is approximately zero, and a sensible expectation for the future real return for both stocks and bonds is 2–4 percent, far lower than the actuarial assumptions on which most investors are basing their planning and spending.

• On the hopeful side, because the “normal” level of the risk premium is modest (2.4 percent or quite possibly less), current market valuations need not return to levels that can deliver the 5 percent risk premium (excess return) that the Ibbotson data would suggest. If reversion to the mean occurs, then to restore a 2 percent risk premium, the difference between 2 percent and zero still requires a near halving of stock valuations or a 2 percent drop in real bond yields (or some combination of the two). Either scenario is a less daunting picture than would be required to facilitate a reversion to a 5 percent risk premium.

• Another possibility is that the modest difference between a 2.4 percent normal risk premium and the negative risk premiums that have prevailed in recent quarters permitted the recent bubble. Reversion to the mean might not ever happen, in which case, we should see stocks sputter along delivering bondlike returns, but at a higher risk than bonds, for a long time to come.

The consensus that a normal risk premium is about 5 percent was shaped by deeply rooted naiveté in the investment community, where most participants have a career span reaching no farther back than the monumental 25-year bull market of 1975–1999. This kind of mind-set is a mirror image of the attitudes of the chronically bearish veterans of the 1930s. Today, investors are loathe to recall that the real total returns on stocks were negative for most 10-year spans during the two decades from 1963 to 1983 or that the excess return of stocks relative to long bonds was negative as recently as the 10 years ended August 1993.

When reminded of such experiences, today’s investors tend to retreat behind the mantra “things will be different this time.” No one can kneel before the notion of the long run and at the same time deny that such circumstances will occur in the decades ahead. Indeed, such crises are more likely than most of us would like to believe. Investors greedy enough or naive enough to expect a 5 percent risk premium and to substantially overweight equities accordingly may well be doomed to deep disappointment in the future as the realized risk premium falls far below this inflated expectation.

What if we are wrong about today’s low equity risk premium? Maybe real yields on bonds are lower than they seem. This chance is a frail reed to rely on for support. At this writing, at the end of 2001, an investor can buy TIPS, which provide government-guaranteed yields of about 3.4 percent, but inflation-indexed bond yields are a relatively recent phenomenon in the United States. So, we could not estimate historical real yields for prior years directly, only through a model such as the one described here. If we compare our model for real stock returns, at 2.4 percent in mid-2001, with a TIPS yield of 3.4 percent, we get an estimate for the equity risk premium of −100 bps.

Perhaps real earnings and dividend growth will exceed economic growth in the years ahead, or perhaps economic growth will sharply exceed the historical 1.6 percent real per capita GDP growth rate. These scenarios are certainly possible, but they represent the dreams of the “new paradigm” advocates. The scenarios are unlikely. Even if they prove correct, it will likely be in the context of unprecedented entrepreneurial capitalism, unprecedented new enterprise creation, and hence, unprecedented dilution of shareholders in existing enterprises.

The recurring pattern of history is that exceptionally poor or exceptionally rapid economic growth is never sustained for long. The best performance that dividend growth has ever managed, relative to real per capita GDP, is a scant 10 bp outperformance. This rate, the best 40-year real dividend growth ever seen, fell far short of real GDP growth: Real dividend growth was some 2 percent a year below real GDP growth during those same 40-year spans. So, history does not support those who hope that dividend growth will exceed GDP growth. This evidence is not encouraging for those who wish to see a 1.4 percent dividend yield somehow transformed into a 5 percent (or higher) real stock return.

The negative risk premium that precipitated the writing of “The Death of the Risk Premium” (Arnott and Ryan) in early 2000 was not without precedent, although most of the precedents, until recently, are found in the 19th century. In 1984 and again just before the 1987 market crash, real bond yields rose materially above the estimated real return on stocks. How well did this development
predict subsequent relative returns? Stated more provocatively, why didn’t our model work? Why didn’t bonds beat stocks in the past decade? After all, with the 1984 peak in real bond returns and again shortly before the 1987 crash, the risk premium dipped even lower than the levels seen at the market peak in early 2000. Yet, stocks subsequently outpaced bonds. For an answer, recall that the context was a more than doubling of stock valuations, whether measured in price-to-book ratios, price-to-dividend ratios, or P/E multiples. If valuation multiples had held constant, the bonds would have prevailed.35

Appendix A. Estimating the Constituents of Return

An analysis of historical data is only as good as the data themselves. Accordingly, we availed ourselves of multiple data sources whenever possible. We were encouraged by the fact that the discrepancies between the various sources led to compounded rates of return that were no more than 0.2 percent different from one another.

Long Government Bond Yields, BY(t). Our data sources are as follows: for January 1800 to May 2001, 10-year government bond yields from Global Financial Data of the National Bureau of Economic Research (NBER) (data were annual until 1843 and were interpolated for monthly estimates); for June 2001 to December 2001, Bloomberg; and for January 1926 to December 2000, Ibbotson Associates, long-term government bond yields and returns. In cases of differences, we averaged the available data. Ibbotson data were given primary (two-thirds) weighting for 1926–1950 because the NBER data are annual through 1950.


Dividend Yield in Month t, DY(t), and Return on Stocks in Month t, RS(t). For January 1802 to December 1925, G. William Schwert (1990); for February 1871 to March 2001, Robert Shiller (2000); for January 1926 to December 2000, Ibbotson Associates (2001); and for April 2001 to December 2001, Bloomberg. In cases of differences, we averaged the available data. In Shiller’s data, monthly dividend and earnings data are computed from the S&P four-quarter data for the quarter since 1926, with linear interpolation to monthly figures. Dividend and earnings data before 1926 are from Cowles (1939), interpolated from annual data.

Notes

1. The “bible” for the return assumptions that drive our industry is the work of Ibbotson Associates, building on the pioneering work of Ibbotson and Sinquefield (1976a, 1976b). The most recent update of the annual Ibbotson Associates data (2001) shows returns for U.S. stocks, bonds, bills, and inflation of, respectively, 11.0 percent, 5.3 percent, 3.8 percent, and 3.1 percent. These figures imply a real return for stocks of 7.9 percent and a risk premium over bonds of 5.7 percent (570 bps), both measured over a 75-year span. These data shape the expectations of the actuarial community, much of the consulting community, and many fund sponsors.

2. Fischer Black was fond of pointing out that examining the same history again and again with one new year added each passing year is an insidious form of data mining (see, for example, Black 1976). The past looks best when nonrecurring developments and valuation-level changes have distorted the results; extrapolating the past tacitly implies a belief that these nonrecurring developments can recur and that the changes in valuation levels will continue.

3. We strongly suggest that the investment community draw a distinction between past excess returns (observed returns from the past) and expected risk premiums (expected return differences in the future) to avoid continued confusion and to reduce the dangerous temptation to merely extrapolate past excess returns in shaping expectations for the risk premium. This habit is an important source of confusion that, quite literally, (mis)shapes decisions about the management of trillions in assets worldwide. We propose that the investment community begin applying the label “risk premium” only to expected future return differences and apply the label “excess returns” to observed historical return differences.

4. To see the effect of compounding at this rate, consider that if our ancestors could have earned a mere 1.6 percent real return on a $1 investment from the birth of Christ in roughly 4 B.C. to today, we would today have enough to buy more than the entire world economy. Similarly, the island of Manhattan was ostensibly purchased for $24 of goods, approximately the same as an ounce of gold when the dollar was first issued. This modest sum invested to earn a mere 5 percent real return, as stocks earned from 1926 to 2001 in the Ibbotson data, this $24 investment would now suffice to buy more than the entire world economy.
5. No rational investor buys if he or she expects less than 1 percent real growth a year in capital, but objective analysis will demonstrate that this return is what stocks have actually delivered, plus their dividend yield, plus or minus any profits or losses from changes in yields. As Asness pointed out in “Double Logic” (2000), few buyers of Cisco would have expected a 1 percent internal rate of return at the peak, although the stock was priced to deliver just that, even if the overly optimistic consensus earnings and growth forecasts at the time were used. These buyers were focused on the view that the stock would produce handsome gains, as it had in the past, rather than on pursuing an objective evaluation, by using IRR or similar objective valuation tools, of expected returns. Such a focus plants the seeds of major disappointment.

6. The Welch study investigated an expected arithmetic risk premium for stocks relative to cash, not bonds. The difference between arithmetic and geometric returns is often illustrated by someone earning 50 percent in one year and –50 percent in the next. The arithmetic average is zero, but the person is down 25 percent (or 13.4 percent a year). Most practitioners think in terms of compounded geometric returns; in this example, practitioners would focus on the 13 percent a year loss, not on the zero arithmetic mean. If stocks have 16 percent average annual volatility (the average since World War II), the result is that the arithmetic mean is 130 bps higher than the geometric mean return (the difference is approximately half the variance, or 16 percent × 16 percent/2). Such a difference might be considered a “penalty for risk.” If we add a 70 bp real cash yield (the historical average) plus a 720 bp risk premium minus a 130 bp penalty for risk, we find 6.6 percent to be the implied consensus of the economists for the geometric real stock return.

7. Such a return could easily fall to 0–2 percent net of taxes, especially in light of government’s taxes on the inflation component of returns.

8. Smith’s work even won a favorable review from John Maynard Keynes (for Keynes’ approach, see his 1936 classic).

9. TIPS is the acronym for Treasury Inflation-Protected Securities, which have been replaced by Treasury Inflation-Indexed Securities.

10. In fairness, growth is now an explicit part of the picture. Dividend payout ratios are substantially lower than in the early 1920s and the 19th century as a result, at least in part, of corporate desires to finance growth. That said, our own evidence would suggest that internal reinvestment is not necessarily successful: High payout ratios precede higher growth than do low payout ratios.

11. We are indebted to G. William Schwert and Jeremy Siegel for some of the raw data for this analysis (see also Schwert 1990 and Siegel 1998). Although multiple sources exist for data after 1926 and a handful of sources provide data beginning in 1855 or 1870, Professor Schwert was very helpful in assembling these difficult early data. Professor Siegel provided earnings data back to 1870. We have not found a source for earnings data before 1870.

12. The U.S. Bureau of Labor Statistics maintains GDP data from 1921 to date; the earlier data are for GNP (gross national product). Because the two were essentially the same thing until international commerce became the substantial share of the economy that it is today, we used the GNP data from the Bureau of Labor Statistics for the 19th century and the first 20 years of the 20th century.

13. We stripped out reinvestment in the measure of real dividend growth shown in Figure 3 because investors are already receiving the dividend. To include dividends in the real dividend growth would double-count these dividends. What should be of interest to us is the internal growth in dividends stemming from reinvestment of the retained earnings.

14. We multiplied the real dividends by 10 to bring the line visually closer to the others; the result is that on those few occasions when the price line and dividend line touch, the dividend yield is 10 percent.

15. The fact that growth in real dividends and earnings is closer to per capita GDP growth than it is to overall GDP growth is intuitively appealing on one fundamental basis: Real per capita GDP growth measures the growth in productivity. It is sensible to expect real income, real per share earnings, and real per share dividends to grow with productivity rather than to mirror overall GDP growth.

16. This history holds a cautionary tale with regard to today’s stock option practices.

17. This fall in dividends of existing enterprises is not surprising when one considers that the companies that existed in 1802 probably encompass, at most, 1 percent of the economy of 2001. The world has so changed that, at least from the perspective of the dominant stocks, today’s economy would be unrecognizable in 1802.

18. Another way to think about this idea is to recognize the distinction between a market portfolio and a market index. The market portfolio shows earnings and dividend growth that are wholly consistent with growth in the overall economy (Bernstein 2001a). But if one were to unitize that market portfolio, the unit values would not grow as fast as the total capitalization and the earnings and dividends per unit (per “share” of the index) would not keep pace with the growth in the aggregate dollar earnings and dividends of the companies that compose the market portfolio. (When one stock is dropped and another added to a market index, typically the added stock is larger in capitalization than the deletion, which increases the divisor for constructing the index.) Precisely the same thing would happen in the management of an actual index fund. When a stock was replaced, the proceeds from the deleted stock would rarely suffice to fund the purchase of the added stock. So, all stocks would be trimmed slightly to fund that purchase, so consequence is implied by the change in the divisor for an index. It is this mechanism that drives the difference between the growth of the aggregate dollar earnings and dividends for the market portfolio, which will keep pace with GDP growth over time, and the growth of the “per share” earnings and dividends for the market index that creates the dilution we attribute to entrepreneurial capitalism. After all, entrepreneurial capitalism creates the companies that we must add to the market portfolio, thus changing our divisor and driving a wedge between the growth in market earnings or dividends and the growth in earnings and dividends per share in a market index.

19. See Bernstein (2001b). Over the past 131 years, the correlation between payout ratios and subsequent 10-year growth in real earnings has been 0.39; over the past 50 years, this correlation has soared to 0.66. Apparently, the larger the fraction of earnings paid out as dividends, the faster earnings subsequently grow, which is directly contrary to the Miller–Modigliani maxim (see Miller and Modigliani 1961 and Modigliani and Miller 1958).

20. To produce a 3.4 percent real return from stocks, matching the yield on TIPS, real growth in dividends needs to be 1.9 percent (twice the long-term historical real growth rate) while valuation levels remain where they are. Less than twice the historical growth in real dividends, or a return to the 3–6 percent yields of the past, will not get us there.

21. We have made the simplifying assumption that “long term” is a 10-year horizon. Redefining the long-term returns over a 5-year or 20-year horizon produces similar results.

22. Because this adjusted dividend is always at or above the true dividend, we have introduced a positive error into the average dividend yield. We offset this error by subtracting the 40-year average difference between the adjusted dividend and the true dividend. In this way, EDY(t) is not overstated, on average, over time.
Of course, stock buybacks increase the share of the economy held by existing shareholders.

Arnott and Asness (2002) have shown that since 1945, the payout ratio has had a 77 percent correlation with subsequent real earnings growth. That is, higher retained earnings have historically led to slower, not faster, earnings growth.

Throughout this article, when we refer to a 10-year average or a 40-year average, we have used the available data if fewer years of data were available. For instance, for 1820, we used the 20-year GDP growth rate because 40 years of data were unavailable. We followed a convention of requiring at least 25 percent of the intended data; so, if the analysis was based on a 40-year average, we tolerated a 10-year average if necessary. To do otherwise would have forced us to begin our analysis in about 1840 and lose decades of interesting results. Because data before 1800 are very shaky and we required at least 10 years of data, our analysis begins, for the most part, in 1810.

We cannot know the 10-year returns from starting dates after 1991, so 192 years of expected return data lead to 182 years of correlation with subsequent 10-year actual returns.

Another way to deal with serially correlated data is to test correlations of differenced data. When we carried out such tests, we found that over the full span, the $R^2$ actually rose to 0.446 from the 0.214 shown in Panel A of Table 1; moreover, since 1945, the differenced results showed a still impressive 46 percent correlation. These results are available from the authors on request.

In an ex ante regression, the model is respecified for each monthly forecast with the use of all previously available data only.

We made the simplifying assumption that “long term” is a 10-year horizon. Redefining the long-term returns over a 5-year or 20-year horizon produced similar results.

Even when we considered successive differences to eliminate the huge serial correlation of real bond yields and 10-year real bond returns, the result from 1945 to date (available from the authors) was identical to the result for the raw data—a correlation of 0.63.

For investors accustomed to the notion that stock returns are uncertain and bond returns are assured over the life of the bond, this result will come as a surprise. But conventional bonds do not assure real returns; their expected real returns, therefore, should be highly uncertain. Stocks do, in a fashion, pass inflation through to the shareholder. So, nominal returns for stocks may be volatile and uncertain, but expected real stock returns are much more tightly defined than expected real bond returns.

Differencing caused the correlation for the full 182-year span to fall from 0.66 to 0.61 and, for the span following World War II, caused it to fall from 0.79 to 0.48.

For the taxable investor, the picture is worse, of course. In the United States, investors are even taxed on the inflation component of returns. From valuation levels that are well above historical norms, a negative real after-tax return is not at all improbable.

The excess return of stocks over bonds was negative also in the decades ended September 1991, November 1990, most 10-year spans ending August 1977 to June 1979, and the spans ending September 1974 to January 1975.

Consider the 10 years starting just before the stock market crash in September 1987. This span began with double-digit bond yields. The bond yield of 9.8 percent minus a regression-based inflation expectation of 3.6 percent led to an expected real bond return of 6.2 percent. The stock yield of 2.9 percent plus expected real per capita GDP growth of 1.6 percent minus an expected dividend shortfall relative to per capita GDP of 0.4 percent led to an expected real stock return of 4.0 percent. The risk premium was –2.0 percent. But stocks beat bonds by 4.9 percent a year over the next 10 years ending September 1997. What happened? The dividend yield plunged to 1.7 percent. This plunge in yields contributed 5.8 percent a year to stock returns; in the absence of this revaluation, stocks would have underperformed bonds by –0.9 percent. So, the –2.0 percent forecast was not bad; dividends rose a notch faster than normal, and more importantly, the price that the market was willing to pay for each dollar of dividends nearly doubled.

References


