Stock Market Charts You Never Saw

Edward F. McQuarrie*

Professor Emeritus

Leavey School of Business

Santa Clara University

September 2017

Revised March 2021

Contact: emcquarrie@scu.edu

Abstract

Investors have seen countless charts of US stock market performance which start in 1926 and end near the present. But US trading long predates 1926, and the foreshortened perspective that results from a focus on post-1926 data can be misleading. To compound the problem, visual and arithmetic frailties, as catalogued in behavioral finance, make it difficult for investors to draw appropriate inferences from long-term records of performance. As a partial corrective, this paper displays a novel set of charts, with some rooted in the 19th rather than the 20th century, and others ending well before the present. The goal is to challenge shibboleths about the expected outcomes of buy-and-hold stock market investing, and to raise questions about the expected performance of stocks versus bonds over long periods. [*This paper has been partially replaced by later work. See revision notes that follow this abstract.*]

Revision notes

In the years since this paper was posted I went on to collect a large amount of new data on 19th century stock and bond returns. Accordingly, better data for the stock returns before 1897 charted in this paper can be found in SSRN #3805927; likewise, for bond returns before 1974. A better take on the international data can also be found there, based on the annual updates published by Credit Suisse, kindly provided to me by Mike Staunton. Points made in this paper using more recent US data remain largely undisturbed.

I leave this paper posted because the general points made are ratified in the newer work, and because it contains some thought experiments not published elsewhere. It also develops the idea of visual heuristics more fully than was possible in later work.

I caution the reader not to rely on this paper as a full representation of my views on the early history of the stock and bond markets, especially stocks before 1871 and bonds before 1897, periods where new data collection has changed the picture in important ways. The new SSRN #3805927 has a complete record of US stock and bond returns from 1793 to 2019. It also develops the idea of regime change, which was only nascent in this paper.

Investors have seen hundreds of charts showing how the stock market has fared over time. The more historically-minded investor expects to see the date of 1926 somewhere on the chart, because, based on repeated exposures, that is when stock market history truly begins (Ibbotson SBBI 2016). A few may opine that the true point of beginning was 1896, when the DJIA became available (Pierce 1991). And a smaller number still may recall seeing charts with an earlier start date of 1871 (e.g., Shiller 2015). There will be a vague sense that stocks probably traded in the United States even before 1871, but the average investor has to assume that early data must be lost, scattered, or otherwise insufficient for charting.

This foreshortened view of US stock market history misleads investors as to how stocks might perform over lengthy intervals. Today, investor expectations are anchored by charts like Figure 1, which shows the total return offered by the S&P 500 Index from year-end 1925 through year-end 2006. The overall import of Figure 1 is unmistakable. Stocks go up. And up. And up, returning over $3000 for each $1 initially invested. Yes there were hiccups along the way; stocks didn't go up every year. But during this lengthy period, declines always proved temporary. Lows were never re-touched; highs were followed by higher highs. Panic attacks, such as October 1987, shrink with distance. Visually, these don't amount to much after some decades—the Crash of 1987 is hard to spot in Figure 1.

A plausible interpretation of Figure 1 is that when it comes to stocks, investors should buy and hold. Attempts at market timing will fail. Investors who sold one share of the S&P 500 at the outset of the crash of October 19th, 1987 protected their $280 from over $50 of losses that day; but if they never got back in, then by the present day, those investors had foregone an eight-fold increase in wealth. Stalwarts who bought the day before the crash, set their jaw, and held and held and held, ended up with eight times as much wealth. And that's not counting dividends!

The preceding paragraph tries to capture the conventional wisdom about stock investing, as imbibed by ordinary investors and their financial planners in recent decades, when mass participation in investing first developed (Nocera 1994). That wisdom is supported by virtually any chart of the post-1926 US stock market—especially one that stops short of 2008, as Figure 1 deliberately does. Figure 1 demonstrates, at a glance, that severe stock market declines are few and far between. Plus, lows reached during those rare declines are never re-touched. New highs are significantly higher than prior highs. Time heals all stock market wounds.

Figure 2 looks very different. It too covers an eighty-one year span of the US stock market, but it is not the same time span, overlapping for only a few years with Figure 1. This chart looks more discouraging. During this eighty-one year period stocks went nowhere, ending slightly below where they began. Or more exactly, stocks did go up, again and again, but always came back down, again and again. Lows were revisited, sometimes decades later, as in 1921. Sideways movements also occurred, several lasting for over a decade. True, successive highs were higher, but in part, that's because the fluctuations became more extreme with time: higher highs were also followed by steeper plunges. Two declines of over 70% occurred toward the end of the period, in contrast to the smaller declines on the order of 50% seen earlier. After holding on for eighty-one years, a stock investor had gained less than nothing.

Figure 2 does not argue for buy and hold; time does not heal any wound in this period. There is fluctuation but no appreciation. Figure 2 argues for market timing, if possible; else, avoidance of stocks in favor of some less volatile means of wealth preservation. No one looking at Figure 2 would counsel retirees that it was important to continue always to hold a substantial portion of their wealth in stocks, even at age 65, 75, or 85.

Figure 2 is a chart you've never seen.

Lessons of History?

Which of these lengthy periods from the past provides the best guide to what will happen over the next few decades of stock market investing: the period charted in Figure 1 or that in Figure 2? Trick question. Technically, the future course of the stock market can't be predicted from any time sample of how stocks moved in the past (Malkiel 2007). The next eighty-one years might produce a chart that looks different from either Figure 1 or Figure 2. That's the meaning of random walk.

Unfortunately, that lesson is not the inference investors typically draw from historical stock market charts. As the work of Dimson, Marsh and Staunton (2002), Ibbotson (2016), and Siegel (2014) began to coalesce, Peter Bernstein (1997) was moved to wonder whether stock market returns, examined over a long enough period, might follow a natural law, converging on a positive return generally estimated, following Siegel, at a compounded 6-7% in real terms. Might what appeared to be a random walk in the short term, when viewed over longer periods, turn out to be an inexorable rise?

When truly long term charts are examined—Figure 5-4 in Siegel (2014, p. 82) covers 210 years, and Figure 1 in Goetzmann (1993) covers 289 years—an inexorable rise, or at least, a sustained and relatively unvarying upward movement, does seem to emerge. The longer the time period covered, the more the chart comes to look like my Figure 1, marching from the lower left corner to the upper right, and the less it resembles Figure 2, oscillating aimlessly around its beginning point.

But there are problems with estimating stock market return over centuries. The obvious problem is that no individual investor has a century-long planning horizon. The typical retiree has to be coaxed to push their time horizon out even to thirty years.

A less obvious problem is that the longer the period of estimation, the less prescriptive its mean return can be about stock market returns over much shorter periods of, say, thirty years. In other words, over the time span that concerns a retiree. A mean return of 6-7%, estimated over 300 years, is consistent with there being at least one 30 year period whose return deviates substantially up or down from that estimated mean. There are, after all, 270 other years to average out the difference.

Figure 3 is another chart you never saw, because it starts in 1919, outside the Ibbotson SBBI horizon, and ends well before the present day. Nonetheless, it qualifies as a long-term chart, covering thirty years. It shows the full dimensions of the great boom of the 1920s, followed by the familiar sharp plunge, after 1929. Subsequent rallies during this period are vigorous, but conform to a pattern of lower highs. A small tidbit: the 1932 low is here revealed not to be a one-off, as it always appears in post-1926 charts; rather, it simply retouched the 1921 low, which itself revisited still earlier lows (see Figure 2). US stocks did not plumb all-time lows in 1932; they merely cancelled out the great boom of the 1920s, returning stocks to an old low seen a decade earlier.

These thirty years are part of the 210 year record in Siegel (2014), and are thus consistent with a long term return of 6-7% real, or 9-10% nominal, per Siegel. But few investors would expect to see this sort of stock market performance over the next thirty years. The possibility that such a period might recur is not what most investors infer from reading Siegel (2014).

Nominal, Real, and Total Returns

To get the most dramatic contrast between the conventional wisdom backed up by Figure 1, and the more daunting outcomes charted in Figures 2 & 3, I charted different sorts of return. Figure 1 used nominal total returns; the next two figures used real (deflated) price returns.

Neither of these is the same as the nominal price returns that appear in daily newspapers, as in the statement, "Yesterday the S&P 500 rose 22.43 points." Total returns add dividends and assume their re-investment. That re-investment powerfully boosts returns when stocks plunge and then stay down for years, as in the 1930s, but ultimately rise to new highs, as in the 1950s. Likewise, during periods of inflation—i.e., most of the 20th century—nominal returns, especially compounded over long periods, will visually appear to be much, much greater than real returns, especially real price-only returns.

As a rule, to get the sunniest and most optimistic take on potential stock market returns, the analyst should chart: (1) total (2) nominal returns (3) during a period of sustained and substantial inflation, (4) ending at a point where stocks were near all-time highs. Figure 1 was constructed according to this rule. While it appears to be entirely objective, if the goal had been to construct a selective and visually biased portrayal intended to promote the inference that stocks promise high returns, while posing little risk of loss over the long run, it would be difficult to improve on Figure 1.

Conversely, to get the most pessimistic portrayal of what might be gained from buying and holding a broad index of the stock market, one simply reverses each of these desiderata, to obtain a chart like Figure 2. During the 1852 to 1932 period, inflation was intermittent, with long spells of deflation; dividends accounted for almost all of total return; and of course, 1932 was no kind of top for stocks.

But perhaps the 19th century is too far in the past to provide a good guide to potential stock market outcomes here in the 21st century? In rejoinder, Figure 4 reproduces the more recent data from Figure 1 but adds, for contrast, the deflated price return. For additional contrast, the charting period is shifted slightly so that it begins at the end of 1928 and ends at the close of

2008—almost, but not quite the same period as in Figure 1. This small shift in start and end points—this slight change in selection bias—has marked effects on the estimated stock market return, as will emerge.

The solid top line, corresponding to the performance charted in Figure 1, does not much change its aspect, continuing to show an ever-rising escalator, albeit now with some backing and filling toward the end of the period. An investor who bought near the top of the 1920s bull market, but re-invested all dividends, grew an initial $10,000 investment into about $9.3 million dollars of (nominal) final wealth—even when evaluated near the bottom of the 2007-08 bust. Sounds great—until one realizes that the small shift in time period has cut the estimate of final wealth accumulated by two-thirds, relative to the period charted in Figure 1.[1]

Conversely, when valued in constant dollars, and with dividends spent along the way—as a retiree must perforce do—that same $10,000 investment produced a bit more than $33,000 in real final wealth. The three hundred-fold difference captures the effect of omitting dividends and adjusting for inflation. Put another way, since 1928 dividends plus inflation accounted for 99.7% of the nominal wealth produced, as of 2008, by investing in stocks.

Figure 4 illuminates how much of the long-term return on stocks since 1926 has been due to sustained high inflation on the one hand, and to the favorable enhancement from re-investing dividends on the other. Under the one depiction, the portfolio returned about 9% compounded, from near the high in the Twenties to near the low in the Oughts; under the other, only about 1.5%.

Few contemporary investors expect a multi-decade return on their stock portfolios of 1 – 2% per year. They have no reason to expect such poor results, because most investors have never

[1] From over 3000 to 1, when starting at the conventional point of beginning, at the end 1925, to just over 900 to 1, when beginning three years later and extending two years further.

seen a post-1926 chart of inflation-adjusted, price-only returns, and have rarely seen any charts extending back past 1896.

More chilling yet: investors have regularly been told that even if they bought at the end of August 1929, but re-invested all dividends, they would have been made whole after not too long. And indeed, the top line in Figure 4 shows that as early as February 1937, stock investors were down only 15%, and that by the end of the war, they had been made whole—on a nominal, total return basis. But the bottom line, showing the deflated price-only return, tells a grimmer tale. It was mid-1955 before the portfolio had regained its real value as of 1929, spending almost 28 years in the dumps. Even worse, because the post-war boom petered out in the 1960s, and due to the galloping inflation of that era, by 1975 all the gains of the 1950s and 1960s had been given up, and the real value of the portfolio dropped *below* what it had been at the end of August 1929. Not until 1985, some fifty-six years after the crash on 1929, did the real value of the portfolio again recover to match that peak; and shortly thereafter, the crash of 1987 drove the real value of the portfolio most of the way back to the 1929 level. In short: almost all the real price appreciation of the S&P 500 since the late 1920s occurred in the past 30 years, after the boom of the 1980s was well under way. Investors who bought in summer of 1929, and held only until the summer of 1982, spending dividends along the way, would have seen the real value of their portfolios *drop* by 40%.

Finally, even Figure 3 can too readily be dismissed as "long ago;" no adult in 1919 is still investing today. Figure 5 shows two more recent thirty-year periods where US stocks also proved not very rewarding. These have been extracted from the red performance line charted in Figure 4, and re-scaled to highlight the movements up and down. From mid-1952 to mid-1982, the real value of a stock portfolio went up, up, up, and then down, down, down, almost to the

point of beginning. Likewise, beginning at the end of 1964, the real price of stocks went down and down, then irregularly up and up, finally recovering the point of beginning in 1994. The culprit here was not the sort of severe decline seen in 1929, although the 1972-74 bear market was bad enough; rather, it was the sustained and high level of inflation that held back real price appreciation in stocks.

It might be objected in response to Figures 3 and 5 that their endpoints, 1949, 1982, and 1994, were terrific times to put money into stocks. And that is precisely the point: for an investor who knows how to time the market, stocks can be an extremely lucrative investment. But if investors eschew market timing, and spend their dividends, then they must be prepared to endure long decades of poor real returns, as in 1919-1949, 1952-1982, and 1964-1994.

The Frailty of Investor Visual Perception

Before proceeding further with the argument, I need again to produce a familiar chart, one that investors have seen before, in this case, a version of the graph featured in Siegel (2014, p. 82, his Figure 5-4). Figure 6 charts inflation-adjusted total returns for the US stock market over the 185 years from end 1830 through 2015.[2] The dotted line represents the estimated real geometric mean return of 6.6% per Siegel (2014).

Now consider an ordinary investor's first visual impressions of Figure 6. Looking at this chart, most investors will see a nearly straight ascending line, an unvarying progression sustained for almost two centuries. Of course there are small squiggles around the line, temporary

[2] The Siegel data from before 1825 consists of Schwert's (1990) compilation of stock price data from Smith and Cole (1935), with dividends interpolated by Siegel. There is a weighting error in Schwert's compilation, see McQuarrie (2017). The Goetzmann, Ibbotson and Peng (2001) data, which mostly replaces the Schwert compilation in the recent edition of Siegel (2014), does not have dividend data before 1825, indicating that interpolated dividend data have been supplied prior to that date. In addition, the Goetzmann et al. data are based on as few as five stocks in some months in the 1840s, are missing months in the 1850s and 1860s, appear to omit numerous dividends, and are upwardly biased by the Harlem corners in the 1860s (details in McQuarrie 2017). Hence, for the period before 1871, Figure 6 uses the return series of McQuarrie (2017), which corrects for these deficiencies. It can be compared to Figure 9, which reproduces Siegel (2014, Figure 5-4), which is based on Goetzmann et al. from 1825 through 1870.

deviations above and below it. But the operative terms are small and temporary; overall, the advance is remarkably uniform. Visually speaking, an investment in stocks, held for any multi-decade period, *with all dividends re-invested*, and even adjusted for inflation, goes up and up and up. No wonder Peter Bernstein (1997) could speak of long-term stock market returns in terms of mean reversion to an underlying constant rate.

Figure 6 tends to cast doubt on Figures 2, 3, and 5. The argument would be that total return, charted over the longest period for which good data exist, provides the best possible metric for expected stock market performance.

Unfortunately, however optimal its metric, Figure 6 has been so constructed as to take advantage of the frailty of investor visual perception.[3] Upon closer inspection, the impression of a uniform advance will prove to be a visual illusion, fostered by the way the chart has been prepared.

The discipline of behavioral finance uncovers and analyzes investors' tendency to process textual information using faulty and misleading heuristics, often to their detriment (Kahneman 2011; Shefrin 1999; Statman 2017). A newer strand shows that investors also apply misleading heuristics when confronted with *visual* representations of stock performance (e.g., Raghubir and Das 2009). Essentially, investors misread charts in predictable ways, as they misapply heuristics which might be serviceable in other contexts.

Looking again at Figure 6, the first visual trap stems from the very long period of time covered, and the vast appreciation that had occurred by the end (stocks returned almost 130,000-to-1 in real terms over these 185 years). Especially with the mean return also placed on the chart

[3] Lest there be any misunderstanding, although my Figure 6 is modelled on Siegel's (2014) Figure 5-4 (see Figure 9 below), this sentence does not cast any aspersion on Professor Siegel, his research, his conclusions, or his charting technique. It describes the rhetorical potential of a chart that *I* constructed with a view to maximizing its potential to mislead.

as a straight line, it is easy to miss the sagging of the performance line in the 1840s. Those years exhibit one of the largest deviations below the mean return line for the entire period. But the eye easily passes over it, because it is located on the lower left side of the chart, and because it appears small.[4] In any chart where there is so obviously a substantial rise to much higher levels by the end of the period, early developments will shrink to visual insignificance. Because every value on the left side is low, the eye passes over the quite low values of the early 1840s. The trend's the thing.

But the most misleading visual element in Figure 6 is its horizontal compression. Printed on a book page, the entire investing horizon of a retiree, some 30-40 years, would span less than one inch. The eye soon forgets that a "small" squiggle on the 185-year line of performance may represent a substantial period of little or no return.

To overcome the frailty of investor visual perception, it helps to disaggregate the 185 years charted in Figure 6 into smaller chunks, and to size these so that the time horizon of an individual investor occupies a large rather than small portion of the horizontal axis. We can then better assess just how small—or large—those squiggles in Figure 6 really are, when viewed from the perspective of a human being (Statman 2017), rather than from the god-like view of history.

To that end, Figure 7 has four panels. These perform both a vertical and a horizontal dis-aggregation of the stock market performance charted in Figure 6. In each panel the logarithmic vertical axis covers potential appreciation of 100:1. The horizontal axis in each panel covers relatively short periods of time, 31 to 50 years, and has been scaled on the page so that the panel covering fifty years is about 1.6 times wider than the panel covering thirty-one years. The

[4] The dip is even easier to miss in Siegel's (2014) Figure 5-4, for three reasons. First, his chart starts in 1802, and there was good appreciation in the period from 1802 through the beginning of the Figure 6 period, so that the dip below the geometric mean line in the 1840s does not appear so large. Second, in the Siegel graph the dip is partly obscured by the bond line (see the reproduction in Figure 9 below). Third, the vertical scale is greater in the Siegel graph, which tends to shrink the apparent size of any deviation above or below the trend line.

logarithmic scaling of the vertical axis, and the constant 100:1 range across panels, entails that a vertical movement of, say, one inch up or down on the page always indicates the same percentage change across the panels. The consistent horizontal scaling similarly insures that slope and run have the same interpretation across panels. Finally, placing the panels on *different* 100:1 portions of the vertical axis defeats the visual temptation to dismiss early deviations from trend as squiggles (again, compare Figure 5-4 from Siegel 2014, on which my Figure 6 is modelled; it uses a single vertical scaling to cover the entire two century period with a ratio of 100 *million* to 1).

As seen in Figure 7, Panel A, the years 1831 through 1861 do not argue "stocks for the long run." The compounded real total return for this 31 year period was less than 4%. But that average is misleading. There was one good bull run in the middle, from 1842 to 1852, in which stocks tripled; but this was preceded by a decade in which stocks went nowhere, and followed by another in which stocks barely advanced.

Panel B, which charts the 1861 to 1906 period, looks very different. During this 45 year period, stocks soared with only brief interruptions, providing a *real* annual total return of over 9.5%. The investor never had to wait more than five years to see the old high point in wealth surpassed by a new high; there were no decade-long periods of stagnation, as seen in Panel A. The post-Civil War era was an extraordinarily lucrative period to be a buy-and-hold investor in stocks.

Panel C, charting the 1906 to 1949 period, looks different again. As in Panel A, overall returns are weak, about 4%. There was one good bull run in the middle, where wealth almost sextupled: the great boom of the 1920s. But in Panel C, the periods of stagnation are longer, lasting about fifteen years, and the overall course of stocks is more volatile, with larger

fluctuations sustained for longer periods of time. Panel C is important because it supplies the 20 years that precede 1926, when Ibbotson SBBI data begin. Those years, seldom seen on any chart today, highlight the fact that the crash of 1929 and its aftermath represented not some epic, one-time destruction of value, but merely returned investor wealth to levels prevailing two decades earlier, prior to WWI. In fact, with the geometric mean return line re-based to 1906, it took the entire boom of the 1920s to bring cumulative real wealth back to its expected value (note how the stock line just regains the expected mean return line in summer 1929). The vicissitudes surrounding WWI were that bad. Hence, datasets constrained to begin in 1926 provide a misleading context for understanding the crash of 1929 within a long-term perspective.

Finally, panel D in Figure 7 re-charts much of the Figure 1 period, our familiar modern times, now in terms of real total return and using the same metrics as Panels A through C. Panel D looks somewhat like Panel B, again showing real total returns north of 9%, albeit not quite so high as in the post-Civil War period charted in Panel B. Part of the shortfall is due to the marked belly in the curve beginning in the middle 1960s, a fifteen to twenty-year period of stagnation (see Figure 5). Panel D, charting real (deflated) total returns, when compared to Figure 1 with its nominal price return, highlights how bad the decline that climaxed in 1972-74 really was. The key difference relative to Figure 1 is the use of real returns in Panel D. Adjusting for inflation produces a more negative portrayal of the period. Then again, post-war returns had been so high up until the 1960s that all the vicissitudes of the 1970s, and the great surge in inflation, had only taken real wealth back to trend line by summer of 1982 (the mirror image of what had happened in the 1920s, see preceding discussion of Figure 7, Panel C).

The Uses of History

Next, Figure 8 contains a single panel scaled to the same metrics as those in Figure 7, covering the period since 1999. In charts that start in 1926, this recent period can appear anomalous, witnessing two severe declines of 50% or more in close succession. Two successive declines of that magnitude had never previously occurred in the post-1926 period. However, with Panels A and C of Figure 7 in hand, investors can appreciate that in a longer view of stock market history, there's nothing unusual about the 1999 to 2016 period. Revisiting a prior low, and going fourteen years before notching a new high, are events that have occurred before—a point easily grasped once the historical record is pushed back past 1926.

Next, all too often investors want something more than history can give, and are willing to torture charts until they yield it. With the four panels of Figure 7 in view, and examining the short time span of Figure 8, the temptation to extrapolate will be almost irresistible. Visually speaking, it appears that the period beginning in 1999, once concluded in 2035 or 2045 or thereabouts, will likely resemble the period charted in Panels A and C more than those charted in Panels B and D; will, in other words, end up being a multi-decade period where stocks provided comparatively weak total returns.

But that extrapolation is illegitimate, a classic instance of the misleading heuristics catalogued in behavioral finance. Either stock prices follow a random walk, or they don't. If stocks do follow a random walk, then Figures 7 and 8 cannot reflect a cyclic process. By definition, cyclic means not random. Given a random walk, stocks cannot oscillate *predictably* through periods of strong and weak returns. Neither kind of period has any set duration, nor is the periodicity fixed. In any case, Panel C does not recapitulate Panel A; they are similar only in their weak performance. Similarly, Panel D does not recapitulate Panel B; there was no belly similar to that from 1966 to 1981 following the Civil War.

History documents but does not constrain. The stock market history charted in this paper is useful in documenting the number and variety of occasions where stock returns would have disappointed an investor who expected to obtain a 6-7% real return over *every* multi-decade interval. The theoretical possibility of such deviations follows directly from the fact that that 6-7% return represents a *mean* of dispersed values rather than a constant. But the foreshortened view of history promoted by data that stop at 1926, combined with the arithmetic frailty of investors, make this point difficult to grasp, absent a vivid visual demonstration.

History, especially on the century scale, is also useful for documenting which stock market events have been one-offs and which have already occurred multiple times. Technically, under a random walk hypothesis, an event which had occurred only once in centuries past could nonetheless recur at some much later date; and an event that had recurred multiple times, might cease to occur, for centuries to come. But this mathematical truth can be devilishly difficult for investors to keep in mind. The difficulty is greatest when the event the investor wants to dismiss as a one-off also presents a deeply unpleasant prospect which he or she would very much not like to see repeated.

The period after 1929 provides a case in point. In my experience investors not alive then are too quick to dismiss this decline, as something that could never happen to their portfolio within their lifetime. Foreshortened history from 1926 supports this conclusion; there have been no parallels to 1929-1932 in the time since. One context where this dismissal becomes tempting is the process for determining the "safe" rate of withdrawal from a retirement portfolio. The 4% rate on which Bengen (2004) settled was designed to withstand the worst case of which he knew—i.e., the decline following 1929. Contemporary financial planners have questioned whether that 4% rate might be too conservative (Guyton 2004). Does the retiree really need to

plan to withstand an event that occurred only once, and so long ago to boot? If the 4% rate is too low, because forged against a one-off that will never be repeated, then retired investors will sacrifice a portion of the well-being that could have been derived from their accumulation.

Although investors committed to a random walk theory of the market may know, intellectually, that a decline on the order of 1929-32 could happen again, it's hard to stay true to this knowledge. Charts like Figure 2, and Panels A & C of Figure 7, can help. They show, visually and vividly, that the decline of 1929-32 had precursors; that stock prices merely re-touched the lows seen in 1921; and that wealth was only dialed back by two decades to levels seen just before WWI. No one-off there.

Arithmetic Frailty of Investors

The other use of history, especially graphic history in charts, is to help investors overcome what might be called their arithmetic frailty. In this regard, I question how many investors really understand the behavior of the large-sample geometric mean, calculated using the product function (Π). Too often, investors apply an inappropriate frame based on small-sample arithmetic means and the behavior of the more familiar summation function ($\sum$).

To illustrate, consider how an ordinary investor might respond to this statement in Siegel (2014, p. 81): "The compound annual real return on stocks is approximately 6.6% … Despite the addition of 20 years of stock market data since the first edition … this return is just one-tenth of a percentage point lower than the 6.7% return that I reported in 1994". On the next page Siegel re-iterates the point, noting "the extraordinary stability of the real return on stocks" as against changes in time frame or expansion of data.

On their face, these remarks declare that when measured over any decent interval of time, the stock market will yield a total real return somewhere in the vicinity of 6.7%. But Panels A and C of Figure 7 show this not to be the case. How to reconcile?

I would counter that the geometric mean, once it has been estimated over many years, is itself extra-ordinarily stable against new observations, in a way that a small-sample arithmetic mean is not. Thus, Siegel appears to have taken an intrinsic property of any large-sample geometric mean, and projected it onto the stock market in particular. In turn, investors reading Siegel may misattribute a mathematical property of the product function (Π) to an underlying dynamism of stocks. After all, the arithmetic mean is not so stable. When new, much lower observations are added to an existing sample, the arithmetic mean will drop, and the greater the discrepancy of these new observations, the more marked the drop. But the large-sample geometric mean does not behave the same.

A thought experiment will dramatize the stability of estimates of the geometric mean (Table 1). Assume a 200-year period has elapsed in which the geometric mean return on stocks was calculated as 6.7% (Siegel's initial 1994 estimate). That means that the 200th root of the ending portfolio value of $429,422, assuming the investor began with $1, comes out to 1.067. For the thought experiment, assume that in the 201st year, the market suffers a significant drop, ranging from a punishing 50% loss to a wealth-destroying 99% decline. Next, take the 201st root of the shrunken portfolio values so obtained. Now, before looking at the top panel of Table 1, check your intuition: by about how much will the 201st root fall, relative to the 200th root of 1.067?

An investor who uses a mental frame extrapolated from calculations of the arithmetic mean might assume that the loss of half the portfolio in the 201st year might cut the geometric

mean approximately in half, from 6.7% to something less than 4%; and that a ninety-nine percent wipeout might reduce the geometric mean to a negligible 0.07% or thereabouts. But the top panel of Table 1 shows how far off the mark such intuitions would be. A fifty percent drop in the 201st year barely budges the re-calculated geometric mean, slightly lowering it to 6.5%; a ninety percent drop leaves the 201-year compounded return at 5.45%; and even the ninety-nine percent calamity produces a 201-year return that is still more than 4%--greater than the long term return on government bonds as estimated in Siegel (2014).

To continue the thought experiment, next imagine not one bad year following the first 200, but a bad thirty year stretch. Consider alternatives in which market returns during the ensuing thirty years range from a mediocre +3% compounded, down to a demoralizing minus 3%. As seen in the rightmost columns of Table 1, top panel, a disappointing 30 year stretch of 3% returns still leaves the geometric mean, estimated now over 230 years, at 6.21%. A calamitous 30 year stretch of -3% returns, in which 60% of the portfolio value erodes away, would still leave the 230 year geometric mean at 5.38%.

Many investors, again due to arithmetic frailty, would probably say that a return of 5.38% is "not too far" below 6.7%;[5] and many might consider a return of 6.5% "within rounding error" of 6.7%. But these intuitions go far astray when dealing with geometric mean returns estimated over centuries. To bring about even those small decreases in the estimated geometric mean requires a major shortfall in subsequent stock market performance.

To continue the thought experiment, Amsterdam had a stock market from the early 1600s, and it is not difficult to imagine a future scholar building on Goetzmann (1993) to construct a *400-year history* of stock market returns. The middle panel of Table 1 again assumes

[5] Or, to use Siegel's (2014, p. 90) words: "[a World portfolio launched in 1900] would have produced a compound real return of 5.4 percent, *very close to* the 6.2 percent that is found in the United States" (emphasis added).

that this extended history estimates the geometric mean return to be 6.7%, and considers by how much a shortfall in the 401st year, or a bad thirty-year stretch through the 430th year, would change this initial estimate of the (very) long term return on stocks. Answer: by very little. Most of the scenarios previously examined in the top panel of Table 1 now leave re-estimated geometric means above 6%.

Given that inertia, the middle panel of Table 1 adds some truly nightmarish scenarios: drops in portfolio value of 99.9% and 99.99% in the 401st year, or negative returns of -10% and -20% compounded over the ensuing thirty years. All these calamities still leave the re-estimated geometric mean well above the long term return on bonds of 3.6%, as estimated in Siegel (2014).

From a behavioral finance perspective, the longer the period over which stock market returns are estimated, the more misleading that estimate may be, given investors' visual and arithmetic frailty. Rather, the best stock market history might focus on 30-40 year stretches, corresponding to the length of investor time horizons; should not confine itself to periods ending exactly at the present day, or falling within the lived experience of today's investors; and should probably search out—cherry pick—time periods with outcomes maximally discrepant from those which investors have recently experienced. History of that sort may be required if investors are to absorb the full implications of the thesis that stock prices follow a random walk, and if investors are to accept and truly heed the bromide that "past performance provides no guarantee of future results."

Another benefit of the 400-year thought experiment is to question the fitness of total return as a metric for long-term investor outcomes. Had an investor in Amsterdam in 1617 put $150 into the stock market, and left it to compound at 6.7%, he or she would now have a

portfolio equal in value to the entire US stock market.[6] As no one investor now owns all the US market, we can say with certainty that no real investor, individual or institutional, received that return over that period on such a sum.[7]

Last, the bottom panel of Table 1 runs the thought experiment once more, this time with the geometric mean estimated over a short interval of but 40 years, a duration more congruent with investor time horizons,. Estimated on this much smaller sample, the geometric mean behaves more as expected. When the 41st year sees a severe decline, or when years 41 to 70 bring a long period of weakness, the re-estimated geometric mean drops considerably, more consistent, here in a small sample, with expectations based on how small-sample arithmetic means behave.

More Visual Frailty: Mis-Anchoring

Siegel is also noted for this claim, made in earlier editions but dropped from the most recent 2014 edition: "You have to go back more than 1 ½ centuries to the period from 1831 through 1861 to find any 30-year period where the return on either long- or short-term bonds exceeded that on equities."[8] In this vein, most investors have seen charts documenting the long-term out-performance of stocks over bonds. An example is Siegel (2014, Figure 5-4, p. 82), reproduced here as Figure 9 (compare Exhibit 2.1 in the Ibbotson 2016 SBBI). Charts like Figure 9 show the final wealth gained from a stock portfolio soaring far above that of a portfolio invested in bonds. This is visible on the right side of the chart as a huge gap. The gap on the right side, between stock and bond outcomes, is so large that a logarithmic scaling on the vertical axis must be used, lest the bond performance line be reduced to a few smudges on the bottom axis (as

[6] That is, $150 multiplied by the portfolio value in the middle panel yields about $27 trillion, corresponding to recent estimates of total US market capitalization.

[7] Aficionados of popular fiction may also take from Table 1 an appreciation for why vampires are often depicted as very rich: only the undead can leave their wealth to compound for centuries. The prospect of an investor receiving such outcomes is about as real as they.

[8] This quote is from p. 15 of the 3rd edition in 2002. After 2008, Siegel caught grief for this claim, from Rob Arnott among others, who showed that it no longer applied to multi-decade returns ending in 2008 or shortly thereafter. The controversy is well-summarized by Tony Keller, in *The Financial Post* of 9/13/2011.

in Ibbotson 2016 SBBI, Exhibit 2.2b, which compares logarithmic and regular scaling). Figure 9, like Figure 1, is a chart that investors have seen before.

Charts such as Figure 9, with their accompanying commentary, and combined with the distinctive behavior of the product function, may lead investors to mis-anchor their expectations about the future performance of bond and stock investments. Faced with a yawning visual gap, and apprised of the numerical dominance of stock returns (in Siegel 2014, estimated at 6.6% real versus 3.6% for bonds), an investor readily infers that bonds are *never* going to out-perform stocks over any lengthy period.[9] The large gap on the right side of the chart provides compelling visual proof.

Unfortunately, that gap on the right side of the chart does not mean what investors all too often assume it does. Consider: if two exponential lines have been intertwined for a period, and then the performance charted by one hiccups and briefly falls behind the other, the vertical separation between them will continue thereafter, unless the temporary setback is followed by a sustained period of outperformance. Case in point: the stock and bond lines in Figure 9, which intertwine from 1802 to about 1861, until the great inflation associated with the introduction of greenbacks in 1862 sharply sets the bond line back, through the end of the war.[10]

Once two exponential lines have separated, if the performance setback shortly ceases, with rough equality of performance resuming, but with no matching period of outperformance,

[9] The technical term for an enduring or intrinsic advantage of stocks over bonds is the equity premium. Conventional post-1926 market history assumes that premium must be constrained to a positive value.

[10] There are important differences between the Siegel chart reproduced in Figure 9, which uses the Goetzmann et al. (2001) data to represent stock performance from 1825 through 1870, and the corresponding figures from earlier editions (e.g., Figure 1-4 on p. 11 of the 2002 edition), where Siegel used data compiled by Schwert (1990), with dividends interpolated rather than observed. Most notably, figures in older editions showed the stock performance line as always lying *above* the bond performance line, receding only so far as to touch it from time to time. The newer Siegel data charted in Figure 9, which are superior in being more complete, show the stock performance line falling *below* the bond performance line in some earlier periods, most notably around 1842, indicating a negative equity premium for the forty-year period from 1802 to 1842. Examined closely, the stock shortfall in 1842 relative to the long term trend line is among the deepest of any in the chart. See also note 3 above.

then what the eye sees is a continuing shortfall. Thus, in Figure 9 reproduced from Siegel (2014), the stock and bond lines are almost parallel for the thirty years from 1860s to the 1890s, same as had occurred for the sixty years from 1802 and 1862. In other words, following the Civil War the performance of (government) bonds resumed their previous behavior of roughly matching the performance of stocks, perhaps because after 1873 the next decades saw a sustained *deflation*, cumulatively almost 50% (per Schiller 2015). But this resumption is not easily grasped visually, and will be missed in a casual perusal of Figure 9. To the eye, the bond line continues to fall measurably short of the stock line, and that's that.

Next, if ever there should occur a period of sustained and substantial under-performance, as happened for bonds after about 1940, a wide gap will quickly open between the stock and bond performance lines, never again to narrow, even should equivalence someday resume. In sum, in very long-term charts, one bad period at any point, if not balanced by an equally superlative period before or after, suffices to open a large gap between the performance lines by the time the right side of the chart is reached. The visual impression is compelling: on a long term basis, bonds stink relative to stocks; and the longer the investment horizon, the greater the shortfall.

Returning to Figure 9, after the mid-1890s, deflation ceased and the great bond rally seen in the latter half of the 19th century tailed off. Stocks, by contrast, began one of their periods of above-trend performance, lasting until WWI. Accordingly, in the later 1890s the bond line begins again to fall further behind. The result: visually, heuristically, it appears that bonds have been trailing stocks by greater and greater margins, decade by decade after the Civil War; when in fact, numerically, there was only a sharp setback during the Civil War, and a trailing off after

1897, with rough equality over the thirty years in between. But it is not easy to grasp that fact visually in a chart like Figure 9.

Continuing with Figure 9, the bond and stock outcomes are again roughly equivalent between WWI and 1940, but this is almost impossible to see, due to the volatility of stocks during this era and the leftover vertical separation from before. Then, with WWII and the years following, the great period of bond under-performance takes hold. No doubt about it: relative to stocks, bonds were a terrible choice during those decades. The performance shortfall during this period accounts for much of the long-term under-performance of bonds seen on the right side of Siegel's 210 year chart. Or, to assert the same fact from a different point of view, this prolonged stretch of deep bond under-performance was *unprecedented*: nothing like it had ever occurred in the preceding 140 years. And yet, it was so severe and so prolonged that this unique period dominates long-term estimates of bond performance—especially visually. You can see it on the chart: as a long term investment, bonds fall woefully short.

More difficult to see on Siegel's chart: over the past few decades, bonds have resumed their older historical pattern of roughly matching the returns on stocks. That parallelism, crammed as it is into the upper right corner of the chart, and given the great vertical separation obtaining after several horrible decades for bonds, is almost impossible to grasp visually. As before, the longer the period charted, the less illuminating that chart may be.

Which Bonds?

Next, the demonstration thus far can be improved, and made to favor bonds even more. The key is to substitute less familiar 19th century data. Siegel (2014) used *government* bonds, and the data for these bonds before 1926 require a great deal of reconstruction based on scanty or absent information. For instance, the federal debt was paid off for a period in the 1830s, so that

there were no (Federal) government bonds. At other points in the 19th century, Federal bonds afforded unique privileges in banking, requiring an adjustment to the observed coupon.[11] Hence, although not a traditional procedure within the equity premium literature, a practicing investor might well ask: what if one compared stock returns to *observed* corporate bond returns, rather than *theoretically re-constructed* government bond returns?[12]

Fortunately, Ibbotson SBBI data for the post-1926 period include corporate bond returns. Less well known is that Macaulay (1938) provides data on corporate bonds from 1926 back through 1857 (Snowden 1990).[13] Combining these two sources provides a 158-year period over which stock and (corporate) bond returns can be compared (see Figures 10 & 11); and we know from Figure 9 and Siegel (2014) that the preceding 52 years saw nearly equivalent performance for (theoretically reconstructed, government) bonds and stocks. Figures 9, 10 & 11 taken together allow assertions to be made about the relationship between stock and bond performance over the entire period from 1802.

Figure 10 is a chart no investor has seen before. It dis-aggregates the 1857 to 1926 period from the total period charted by Siegel and reproduced here as Figure 9, and also substitutes corporate bonds for government bonds. The results are strikingly at variance with

[11] These considerations are laid out in Siegel (1992); see also Officer (2017). Siegel had to use government bonds because the scholarly question of interest was the equity premium, which is properly defined against a debt instrument free of default risk. Investors deciding between bonds and stocks are free to consider corporate bonds instead, in a way that scholars investigating the equity premium puzzle are not.

[12] To reiterate: in an academic context, this substitution would not be legitimate. The equity premium is defined in terms of a comparison between a risk asset and a risk-free asset. But my concern here is with the practical choice facing real investors, who are free to combine two risky assets if they wish. Put another way: investors can choose to own shares in a private corporation, or to be owed fixed payments by that private corporation; and that comparison is appropriate for elucidating whether the performance of "stocks" must be superior to that of "bonds" over any lengthy interval.

[13] Although the US corporate bond market, defined as a dozen or more issues trading regularly, appears to predate 1857 by a decade or more, to my knowledge no scholar has attempted to go past the Macaulay horizon. Thus Homer and Sylla (2005) do not estimate interest rates for corporate bonds past this point. This neglect may change as the bond data newly compiled by Sylla, Wilson and Wright (2006) become more familiar.

conventional wisdom concerning the long-term performance of stocks versus bonds.[14] After the dramatic but brief out-performance of stocks during the Civil War period of greenback inflation, these corporate bonds proceeded to outperform stocks by a significant margin for the thirty-one years from 1865 to 1896, especially after 1873. A bond investment turned $1.00 into about $16.28, while a dollar put into stocks only produced $8.67. This was not a bad period for stocks by any means (compare Panel B of Figure 7); but corporate bonds did better, with a compounded real return of 9.4%.

One reason these results may surprise is that few stock market charts extend back past the advent of the Dow Jones Industrial Average in 1896,[15] while Macaulay's (1938) work on corporate bond returns has been largely forgotten.[16] Returning to Figure 10, after 1896 stocks again enjoyed a period of out-performance relative to bonds. To make this superiority more visible, the dashed red line re-presents corporate bond performance, now re-based to 1896 to match the value of the stock performance line. From 1896 that dashed bond line behaves in accordance with conventional wisdom, falling far behind the stock line over the course of eighteen years, producing the familiar widening fan, until the vicinity of WWI. After the war, stocks again begin to soar far above bond returns, with the gap widening until August 1929 (the years after 1925 are not shown on Figure 10; see Figure 9).

---

[14] Note that Figure 10 uses the Goetzmann et al. (2001) data up through 1870, like Figure 9, but unlike Figures 6 & 7, which use the McQuarrie (2017) data. Since Figure 10 is intended as a re-do of Figure 9 and Siegel (2014), it seemed best to revert to the Goetzmann et al. data used by Siegel. McQuarrie has stocks performing more strongly than Goetzmann in the 1866-1870 period, because McQuarrie incorporates stock dividends, rights issues and merger premia, which were absent from the sources used by Goetzmann et al. Hence, had the McQuarrie data been used, the bond line would fall behind that of stocks from 1866 to 1870, rather than tracking it. However, all the charts in this paper and in Siegel use Cowles' data from 1871, so the pattern seen in Figure 10 would not change after 1870, regardless of whether McQuarrie or Goetzmann et al. were used prior to that point.

[15] And it is likewise unfortunate that Dimson et al. (2002) chose to terminate their historical data at 1900—just after this significant stretch of bond out-performance in the US came to an end. The next section will consider the world-wide data on stocks and bonds.

[16] More exactly, while Macaulay is not unfamiliar to scholars, most citations are to his yield data, as in Homer and Sylla (2005). To my knowledge, Snowden (1990) stands alone in his effort to convert the Macaulay yield data into holding period returns.

The story continues in Figure 11, using now the more familiar Ibbotson SBBI dataset which compares large company stocks to long term corporate bonds. Again in the spirit of visual disaggregation, in Figure 11 the bond and stock lines have been set equal at the outset of 1926, in contrast to the large separation already seen at that point in Siegel's aggregated chart that covers the entire 210 year period (Figure 9). Hence, the left side of Figure 11 should appear familiar to any investor who has perused the Ibbotson SBBI, which benchmarks all its series to that start point. The 1940s were well along before stocks recovered sufficiently from the Crash of 1929 to surpass bond performance; in fact, had the bond and stock lines been set equal as of August 1929, rather than the SBBI convention of December 1925, the figure would show 22 years of bond out-performance, out through August 1951.[17]

After WW II, stock performance surged while bonds stagnated, producing again the widening fan shape so prominent in Figure 9, evident in the middle of Figure 11. This period provides the bedrock supporting conventional wisdom that bonds *must* under-perform stocks over the long term.

Interestingly, although bond performance in absolute terms is generally held to have been awful from 1940 all the way through the early 1980s, Figure 11 suggests that relative to stocks, bond under-performance tailed off much earlier, toward the end of 1968. The fan stops widening at that point, because stock performance also began to lag by about the middle of the 1960s, as seen in Figure 5B and Figure 7D.

To help the frail investor eye see that the period of bond under-performance does in fact end in the later 1960s, Figure 11 resets the bond line to equal the stock line as of February 1980,

[17] The values charted come from Appendix A of the 2016 SBBI. Appendix C was also helpful in locating stretches where corporate bonds had, in fact, matched or exceeded the returns on large company stocks; but Appendix C values only cover yearly intervals beginning exactly at the beginning of each year. The claim in the text and in Figure 11, concerning August 1951, rests on examination of the monthly data in Appendix A.

and then extends this reset bond line back to late 1968 and forward to February 2009. With that reset, the stock and bond lines can be seen to intertwine rather than diverge, until the great boom of the 1990s takes hold. For the ensuing decade, stocks again soar above bonds, as seen earlier in the 1920s, only to come crashing back down, crossing below the bond line in February 2009. Over the 40-year period from February 1969, stocks and bonds returned about the same. The widening fan shape disappears. As everyone knows, stocks performed terrifically in the 1980s and 1990s; but as many fewer realize, by the end of the Oughts, bonds had caught up (and not just for endpoints in 2008). In short, bonds can match stocks when both perform terrifically, as in recent decades, and bonds can match stocks when both perform terribly, as in the late 1960s through the late 1970s.

The post-1926 record for corporate bonds, with their higher yield relative to government bonds,[18] confirms the picture gleaned from careful scrutiny of Figure 9 reproduced from Siegel (2014): there have been several multi-decade stretches, in the 20th century no less than the 19th century, where an investment in bonds produced roughly the same final wealth as an investment in stocks. These become visible once the terrible period for bonds beginning after 1940 is set aside. Long term charts make it difficult to perform this mental set aside. Disaggregated charts are needed if a compelling visual case is to be made. Re-based charts help to combat mis-anchoring, by pulling out and making visible periods of renewed equivalence between stocks and bonds.

To summarize: a disaggregated view of the historical performance of stocks relative to bonds calls into question the thesis that stocks must always out-perform bonds when the investment horizon is long. The 210-year record of Figure 9, partly broken out in Figures 10 & 11, instead reveals a range of possible outcomes.

[18] In the 2016 SBBI, the advantage is estimated at +40 basis points relative to long term government bonds.

1) There was one lengthy period, beginning after 1940, where bonds dramatically under-performed stocks.
2) There was another lengthy period, beginning after 1873, where bonds significantly out-performed stocks.
3) There were several lengthy periods, ranging in length from 30 to 50 years, where the performance lead swung back and forth, but in the end, bonds produced about the same final wealth as stocks; examples include periods beginning in 1802, 1865, 1906, and 1968. Collectively, these periods of rough equivalence cover about two-thirds of the 210 years.
4) There were other periods, some shorter and others longer, where stocks out-performed bonds, as during the Civil War, or at the time of the great industrial consolidations in the late 1890s, or during the 1920s or 1990s.
5) There were occasional short periods, mostly of twenty years or less, and beginning precisely at a stock market peak, where bonds out-performed stocks; or more precisely, where stocks dramatically under-performed bonds, as after 1929—and 1999.

The best one sentence summary of the 210 year record would be that sometimes, stocks out-performed bonds, but at other times, bonds out-performed stocks; while much of the time, stocks and bonds performed about the same. Needless to say, this flies in the face of conventional wisdom. The illegitimacy of the assertion that stocks *must* dominate bonds cannot readily be seen in long-term charts dominated by the period after 1940, which is to say, almost any SBBI chart of the past thirty years. The truth about the equity premium, or lack thereof, can't be visually grasped in any very long-term chart such as that reproduced in Figure 9, because the eye

can't visually disaggregate the multiple regime changes that comprise it. That frail investor eye needs the help which Figures 10 & 11 provide.

A simple explanation for the multiple regime changes captured in the summary stands on three legs: 1) bonds provide a relatively low but relatively stable return; 2) stocks provide an even lower but almost as stable return from dividends, while stock prices follow a random walk which can be very volatile; 3) consumer prices move through long waves of inflation and deflation, punctuated by sharp surges and retreats. Putting the three together, a substantial rise in inflation, coinciding with an up phase in the random walk of stock prices, produces investment results like those seen after 1940: very unfavorable for bonds, and very favorable to stocks. Conversely, a sustained fall in prices, such as that seen after the Civil War through the end of the 19th century, combined with the random walk of stock prices, produces investment results like those seen after 1873: very favorable to bonds, but also favorable to stocks. Inflationary shocks, like that seen in the Civil War, may give stocks a temporary advantage, while deflationary shocks, like that seen after 1929, may give bonds a temporary advantage. Finally, the intersection of these three factors can also produce roughly equivalent performance over multi-decade intervals of no set length.

There remains an important asymmetry: the much greater volatility of stocks relative to bonds. Stocks can go up much faster and reach much higher peaks than bonds. But this fact only helps investors who are capable of market timing: investors who know how to get out at the top, before the flip side of volatility takes hold, and stocks plunge further and faster than investment grade bonds ever could.

The International Evidence

As suggested by its title (i.e., *Triumph of the Optimists*), the work of Dimson et al. (2002) is generally interpreted as consistent with and supportive of the thesis in Siegel (2014): that stocks are an excellent investment if held for the long run, certainly superior to bonds over any lengthy interval. And most of the individual country charts in Dimson et al. show the same widening fan shape as Figure 9, with returns on stocks climbing farther and farther above returns on bonds as the 20th century proceeds. Alas, the same visual frailty that undercuts interpretation of Figure 9 prevents investors from seeing the many exceptions to "stocks for the long run" that a more patient examination of the Dimson et al. data might reveal.

Table 2 catalogues international instances of: 1) low or negative real returns on equities that persisted for twenty years or more; 2) negative equity premia that persisted for twenty years or more; and 3) instances where the equity premium was negligible for twenty years or more. The table begins by presenting Dimson et al. results for the United States, for use as a benchmark (top row).[19] Consistent with conventional wisdom, in the Dimson et al. accounting, as in the Ibbotson SBBI, there was never a twenty year or longer period where US stocks showed a negative real return, nor was there ever such a period that produced a negative equity premium.

Continuing down Table 2, the next panel shows results for other markets in the Anglo-Saxon cultural tradition (Australia, Canada, South Africa, and the United Kingdom). These results are generally as positive as those in the US, with no lengthy period where equities show a negative real return. It may be this ensemble of results across the US and kindred nations, and the restriction to 20th century data, which has so boosted confidence in equities among US investors.

[19] Dimson et al. use a slightly different index of US stocks for the 1900 to 1957 period than either Siegel or Ibbotson; see Wilson and Jones (2002).

A very different picture emerges in the third section of Table 2, where the experiences of nations on the European continent are grouped. Here there are multiple instances where equities provided a negative real return over a prolonged period, in some cases quite severely negative, and in others, prolonged for fifty years or more. There are also almost a dozen cases of negative equity premia, lasting for as long as forty years. In most cases, the explanation for these disappointing returns is straightforward: the problematic stretches either include WWI, or the Great Depression, or WW II. More surprising is the recurrence of periods that begin about 1960. The ensuing two decades were more disappointing for investors in Europe than the conventional wisdom based on US returns might suggest (but see Figure 5B and Figure 7, Panel D).

The final section shows two countries that don't fit either of the preceding categories, Japan and Switzerland,[20] showing once again either very negative outcomes, or very prolonged periods of weak or negative results. Table 2 indicates that the well-known slump in the Japanese stock market beginning after 1989 was not without precedent. The fate of the Japanese stock investor has often not been a happy one.

Again, it is virtually impossible for the eye to extract these periods of negative performance by glancing at a Dimson et al. chart covering 101 years. The opening fan shape, in which stock and bond returns sharply diverge at some point after 1940, is visually compelling, and repeated country after country. By contrast, the Dimson et al. tables, from which these negative real returns and negative equity premia have been extracted, are visually pallid, and require careful scrutiny to pick out the occasional low or negative number. After all, in the 20th century, especially for periods ending 2000 or before, stocks often were a terrific investment, and far superior to government bonds: most of the stock returns in the Dimson et al. tables for individual countries are positive in absolute terms, and relative to bond returns. Alas, visually

[20] On the Continent but not of it, given its neutrality during both great wars of the 20th century.

frail investors, subject to premature closure and other misleading heuristics, too easily elide "often" into "always."

Conclusion

When investment advisors counsel that stocks are the best bet for a long investment horizon, they should append the acknowledgement: "if my market timing is good." When advisors argue for stocks over bonds, they should append the caveat "as long as you are not French, or Italian, or Japanese, or Swiss, and provided that the 20th century is a better guide to the future than the 19th century." For real investors with their limited time horizons, who may reside anywhere in the world, there have been times when both stock recommendations were bad.

This paper wouldn't be the first to declare that means and averages can be misleading; but it may be unprecedented in its declaration that when it comes to investments, the longer the time period covered, the more misleading the stated mean return. As always, the solution to mean blindness is attention to variance. To this end, disaggregated charts can help investors to overcome misleading heuristics and to adjust for arithmetic frailty.

Finally, this paper is also unusual in calling into question the fitness of the performance metric favored by Ibbotson and Siegel, and many others: total return, especially when computed over periods of a century or more. Total return measured on the century scale presumes an investor who never needs to spend the dividends or interest received. No real investor, individual or institution, has that luxury. And there is one class of individual investor, now of growing importance within the financial planning literature as the Baby Boom generation ages, for whom the total return metric is particularly malaprop: retirees. Once portfolio accumulation ceases with retirement, portfolio income must be spent to live. Under those circumstances real

price return, over short periods lasting two or three decades, becomes an important metric. By that measure, an investment in stocks has been dicey indeed.

References

Bengen, W.P., 2004. Determining withdrawal rates using historical data. *Journal of Financial Planning*, 17(3), pp. 64-73.

Bernstein P. L. 1997. What Rate of Return Can You Reasonably Expect … or What Can the Long Run Tell Us about the Short Run? *Financial Analysts Journal*, 52 (2), 20-28.

Cowles, A., 1939. *Common-Stock Indexes*. Principia Press.

Dimson, E., Marsh, P. and Staunton, M., 2002. *Triumph of the Optimists: 101 Years of Global Investment Returns*. Princeton University Press.

Fisher, L. and Lorie, J. H., 1964. Rates of Return on Investments in Common Stocks. *The Journal of Business*, *37*(1), pp. 1-21.

Giesecke, K., Longstaff, F. A., Schaefer, S. and Strebulaev, I., 2011. Corporate Bond Default Risk: A 150-year Perspective. *Journal of Financial Economics*, *102*(2), pp. 233-250.

Goetzmann, W.N., 1993. Patterns in three centuries of stock market prices. *Journal of Business*, 66(2), pp. 249-270.

Goetzmann, W. N., Ibbotson, R.G. and Peng, L., 2001. A New Historical Database for the NYSE 1815 to 1925: Performance and Predictability. *Journal of Financial Markets*, 4(1), pp.1-32.

Guyton, J.T., 2004. Decision Rules and Portfolio Management for Retirees: Is the 'Safe'Initial Withdrawal Rate Too Safe? *Journal of Financial Planning*, 17(10), pp. 54-62.

Homer, S. and Sylla, R., 2005. *A History of Interest Rates*, 4th edition. John Wiley & Sons.

Ibbotson, R., 2016. *Stocks, Bonds, Bills and Inflation Yearbook (SBBI).* Wiley.

Kahneman, D. (2011), *Thinking, Fast and Slow*. Macmillan.

Macaulay, F. R., 1938. *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields, and Stock Prices in the United since 1856*. National Bureau of Economic Research.

Malkiel, B. G., 2007. *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing (Revised and Updated),* New York: WW Norton & Company.

McQuarrie, E. F., 2017. Introducing a New Database of 19th Century US Stocks: Returns from 1831 to 1870, Capitalization-Weighted, with Dividends. *SSRN Working Paper* #3011486.

Nocera, J. (1994), *A Piece of the Action: How the Middle Class Joined the Money Class*. New York: Simon and Schuster.

Officer, L. H., 2017. "What Was the Interest Rate Then?" *MeasuringWorth*, URL: http://www.measuringworth.com/interestrates/

Officer, L. H. and Williamson, S. H., 2017. "The Annual Consumer Price Index for the United States, 1774-2015," *MeasuringWorth*, URL: http://www.measuringworth.com/uscpi/.

Pierce, P. S., 1991. *The Dow Jones Averages 1885-1990*. Business One Irwin.

Raghubir, P. and Das, S.R., 2009. The Long and Short of it: Why are stocks with Shorter Runs Preferred? *Journal of Consumer Research*, 36(6), pp.964-982.

Schwert, G.W., 1990. Indexes of US Stock Prices from 1802 to 1987. *Journal of Business*, pp. 399-426.

Shefrin, H.,1999. *Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing*. Harvard University Press.

Shiller, R. J., 2015. *Irrational Exuberance,* revised and expanded third edition. Princeton University Press.

Siegel, J. J., 1992. The Equity Premium: Stock and Bond Returns Since 1802. *Financial Analysts Journal*, pp. 28-46.

Siegel, J. J., 2014. *Stocks for the Long Run*, fifth edition. McGraw-Hill.

Smith , W. B. and Cole, A. H., 1935. *Fluctuations in American Business 1790-1860*. Russell & Russell.

Snowden, K.A., 1990. Historical Returns and Security Market Development, 1872–1925. *Explorations in Economic History*, *27*(4), pp.381-420.

Statman, M., 2017. *Finance for Normal People: How Investors and Markets Behave*. Oxford University Press.

Sylla, R. E., Wilson, J. W. and Wright, R. E., 2006. "Price Quotations in Early United States Securities Markets, 1790–1860." *Inter-University Consortium for Political and Social Research*.

Wilson, J. W. and Jones, C. P., 2002. An Analysis of the S&P 500 Index and Cowles's Extensions: Price Indexes and Stock Returns, 1870–1999. *The Journal of Business*, *75*(3), pp.505-533.

Table 1
The Unintuitive Behavior of Exponentiation over Long Periods

Value of $1.00 invested at 6.7% for the stated number of years:	After a decline in 201st year of:	… the re-estimated geometric mean return would then be:	Or, if the next 30 years showed a compounded return of:	…the portfolio would then be worth:	… and the re-estimated geometric mean return over the 230 years would then be:
200 years = $429,422	50%	*6.30%*	3%	$1,042,319	*6.21%*
	60%	*6.18%*	2%	$777,838	*6.07%*
	70%	*6.03%*	1%	$578,795	*5.94%*
	80%	*5.81%*	0%	$429,422	*5.80%*
	90%	*5.45%*	-1%	$317,643	*5.66%*
	95%	*5.09%*	-2%	$234,243	*5.52%*
	99%	*4.25%*	-3%	$172,201	*5.38%*

	After a decline in 401st year of:	… the re-estimated geometric mean return would then be:	Or, if the next 30 years showed a compounded return of:	…the portfolio would then be worth:	… and the re-estimated geometric mean return over the 430 years would then be:
400 years = $184,402,909,814	50%	*6.50%*	3%	$447,594,262,570	*6.44%*
	60%	*6.44%*	2%	$334,020,346,834	*6.37%*
	70%	*6.36%*	1%	$248,547,261,977	*6.29%*
	80%	*6.26%*	0%	$184,402,909,814	*6.22%*
	90%	*6.07%*	-1%	$136,402,901,243	*6.14%*
	95%	*5.89%*	-2%	$100,588,895,752	*6.07%*
	99%	*5.46%*	-3%	$73,946,870,295	*5.99%*
	99.90%	*4.86%*	-10%	$7,817,052,936	5.44%
	99.99%	*4.26%*	-20%	$228,279,745	4.58%

	After a decline in 41st year of:	… the re-estimated geometric mean return would then be:	Or, if the next 30 years showed a compounded return of:	…the portfolio would then be worth:	… and the re-estimated geometric mean return over the 70 years would then be:
40 years = $13.38	50%	*4.75%*	3%	$32.49	*5.10%*
	60%	*4.18%*	2%	$24.24	*4.66%*
	70%	*3.45%*	1%	$18.04	*4.22%*
	80%	*2.43%*	0%	$13.38	*3.78%*
	90%	*0.71%*	-1%	$9.90	*3.33%*
	95%	*-0.97%*	-2%	$7.30	*2.88%*
	99%	*-4.79%*	-3%	$5.37	*2.43%*

Note. The table was generated in Microsoft Excel using the power function. Taking the top row as an example, and reading across, a 50% decline in the 201st year would reduce the portfolio value to $214,711. The Excel function [=power(214711, 1/201)] generates the 201st root of the new portfolio value, which is approximately 1.06298; rounding and then subtracting 1.0 yields the cell entry of 6.30%, which is the new geometric mean return estimated now over 201 years. Continuing across that row, the results of compounding $429,422 for 30 years at 3% is generated from [=power(1.03, 30]. Taking the 230th root of the resulting portfolio value of $1,042,319 yields approximately 1.06209; rounding and then subtracting 1.0 yields the next cell entry of 6.21%.

All returns in the table assume a 6.70% rate of return for the initial period, and are to be read in terms of the magnitude of the drop, after new observations, relative to that initial rate.

Table 2

Periods of Negative or Low Real Returns on Equities and of Negative Equity Premia: The International Evidence 1900-2000

	Lowest real equity returns for periods of 20+ years (longest & deepest)	**Negative equity premia lasting 20+ years**	**Equity premia between 0 – 1% for periods of 20+ years**
I. Benchmark:			
United States	2.4%, 1960-1979	--	0.9%, 1920-1939
III. Anglo-Saxon diaspora			
Australia	1.8%, 1970-1989	--	--
Canada	3.3%, 1900-1919	-1.1%, 1980-2000	0.2%, 1930-1949
South Africa	3.6%, 1910-1929	--	--
United Kingdom	0.2%, 1900-1919	-1.1%, 1920-1939	0.0%, 1920-1949
II. Continental Europe:			
Belgium	-1.7%, 1900-1949 -3.5%, 1900-1919	-0.5%, 1920-1939 -0.5%, 1930-1949	0.4%, 1960-1979
Denmark	0.7%, 1960-1979	-1.9%, 1920-1949	0.5%, 1920-1979
France	-0.5%, 1900-1949 -5.8%, 1930-1949	-1.8%, 1960-1979	0.8%, 1960-89
Germany (pre-1920 & post-1930 only)	-2.2%, 1930-1949 -4.9%, 1900-1919	-1.8%, 1960-1979	
Ireland	-2.6%, 1900-1919	-0.6%, *1900-1939*	0.6%, 1900-1949
Italy	0.8%, *1900-1979* -6.0%, 1960-1979	**-4.0%**, 1960-1979	0.2%, 1960-2000
Netherlands	0.4%, 1960-1979	-0.9%, 1910-1939	0.2%, 1900-1939
Spain	0.4%, 1910-1979 -4.3%, 1930-1949	-0.6%, 1930-1959	0.1%, 1900-1949
Sweden	1.0%, 1960-1979	-0.3%, 1910-1949	--

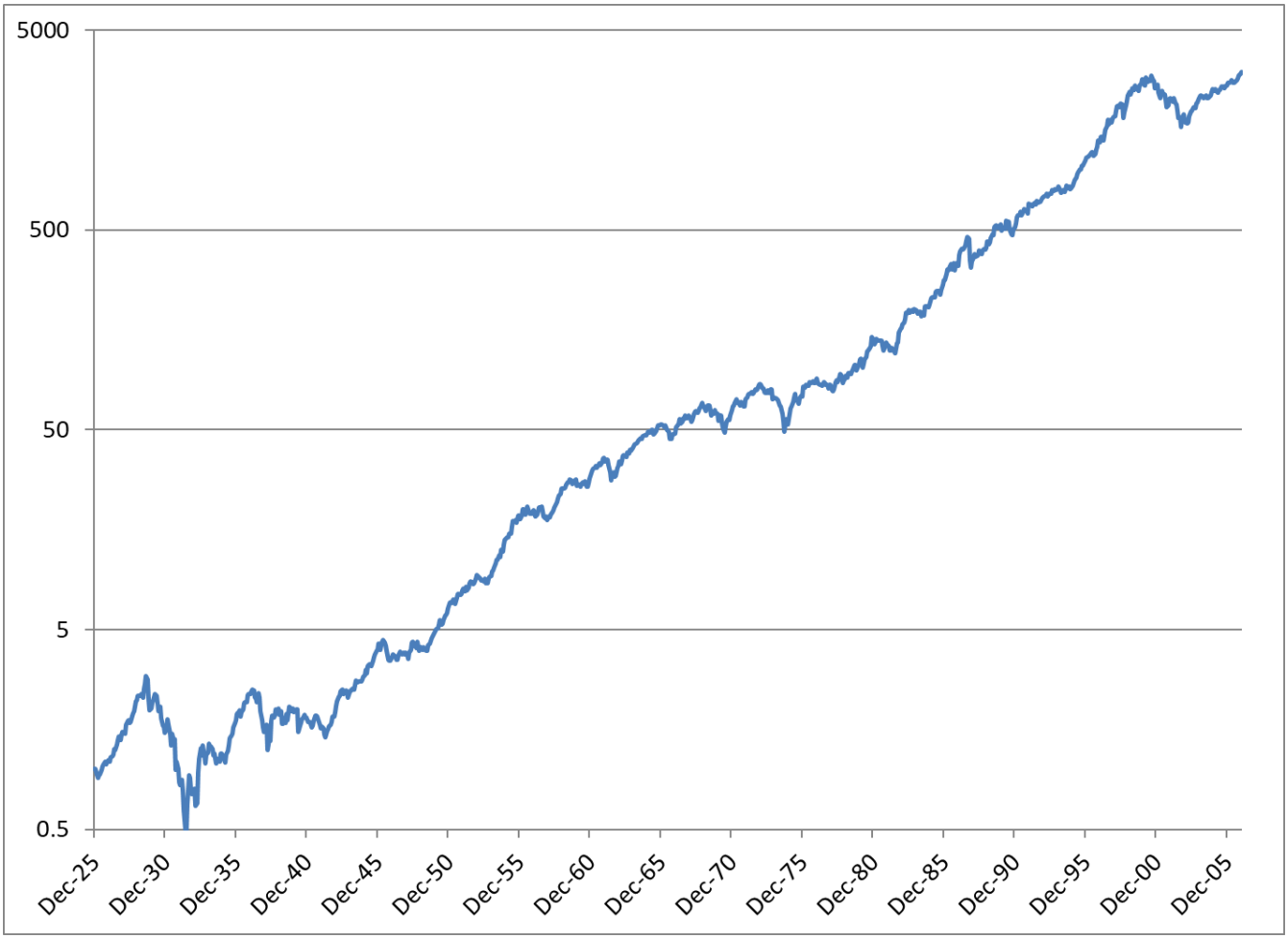
IV. Unique cases			
Japan	-2.6%, 1910-1949 **-9.5%**, 1930-1949	-1.8%, 1980-2000	0.6%, 1960-2000
Switzerland	1.6%, 1910-1939	-0.5%, *1910-1949*	

Note. Values selected and extracted by the author from the tables published for individual nations in Section II of Dimson et al. (2002). Equity premia are relative to government bonds (not bills). In the first and second columns, the greatest negative value is bolded and the longest period is italicized.

The rationale for including the third column is that if one were not constrained to start the period at the exact end of each decade, as the Dimson tables do, and if one were able to substitute corporate bond data, then a period similar in length to that tabled, and overlapping that period, could be found that would show a *negative* rather than a small positive equity premium. (Dimson et al. estimate the incremental return for corporate bonds over government bonds as 0.48%.)

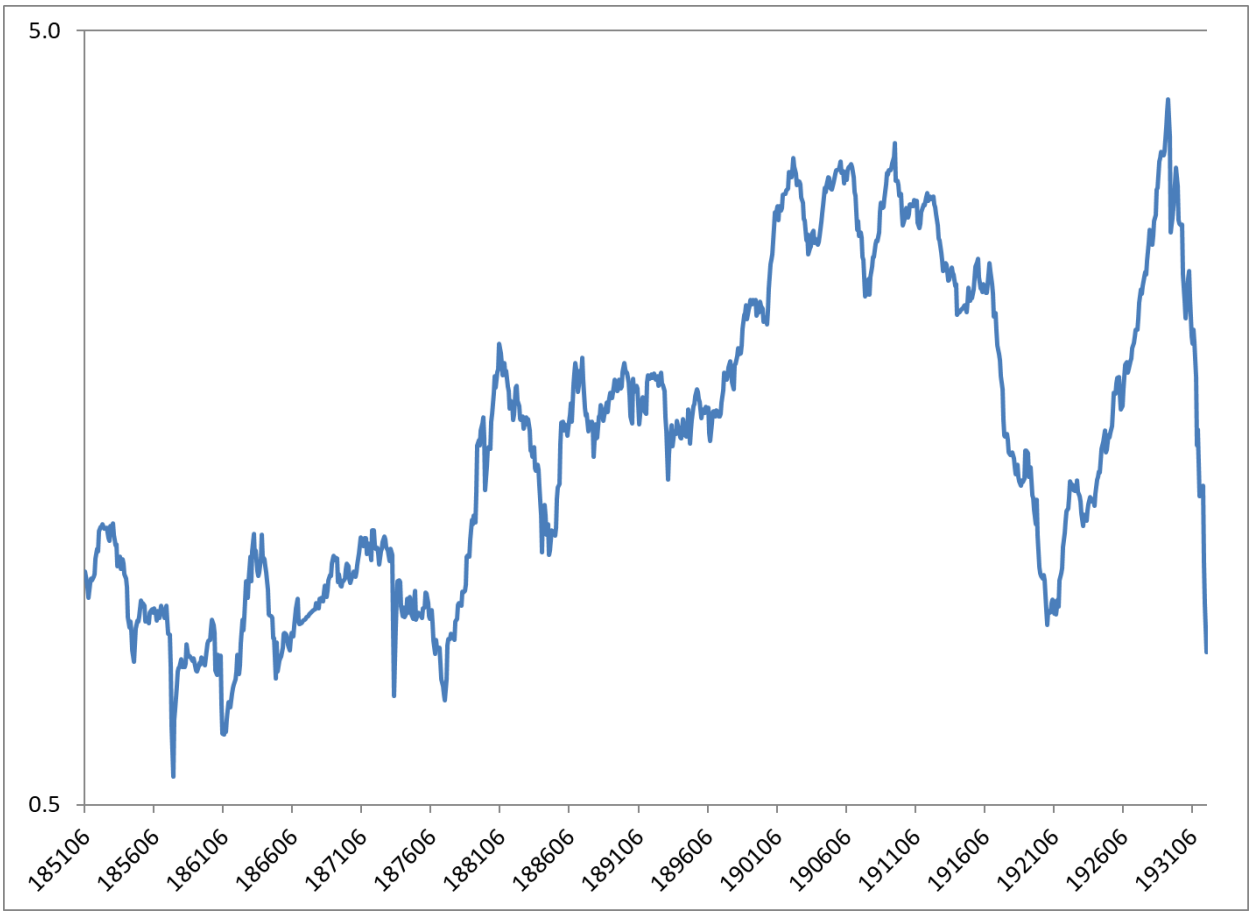
For example, consulting the 2016 Ibbotson SBBI Appendix C, one finds a period from the beginning of 1929 to the end of 1948 where US stocks did show a negative equity premium, relative to corporate bonds, in contrast to the 0.9% equity premium over government bonds for 1920-1939 shown here in Table 2, which is subject to the constraints of the published Dimson et al. record.

Figure 1
A Familiar Chart: The Long-Term Total Return on Large Company Stocks



Note: data from Ibbotson SBBI, chart construction by the author. Month-end close with all dividends reinvested. Logarithmic scale with start value of 1.0 at the end of 1925.

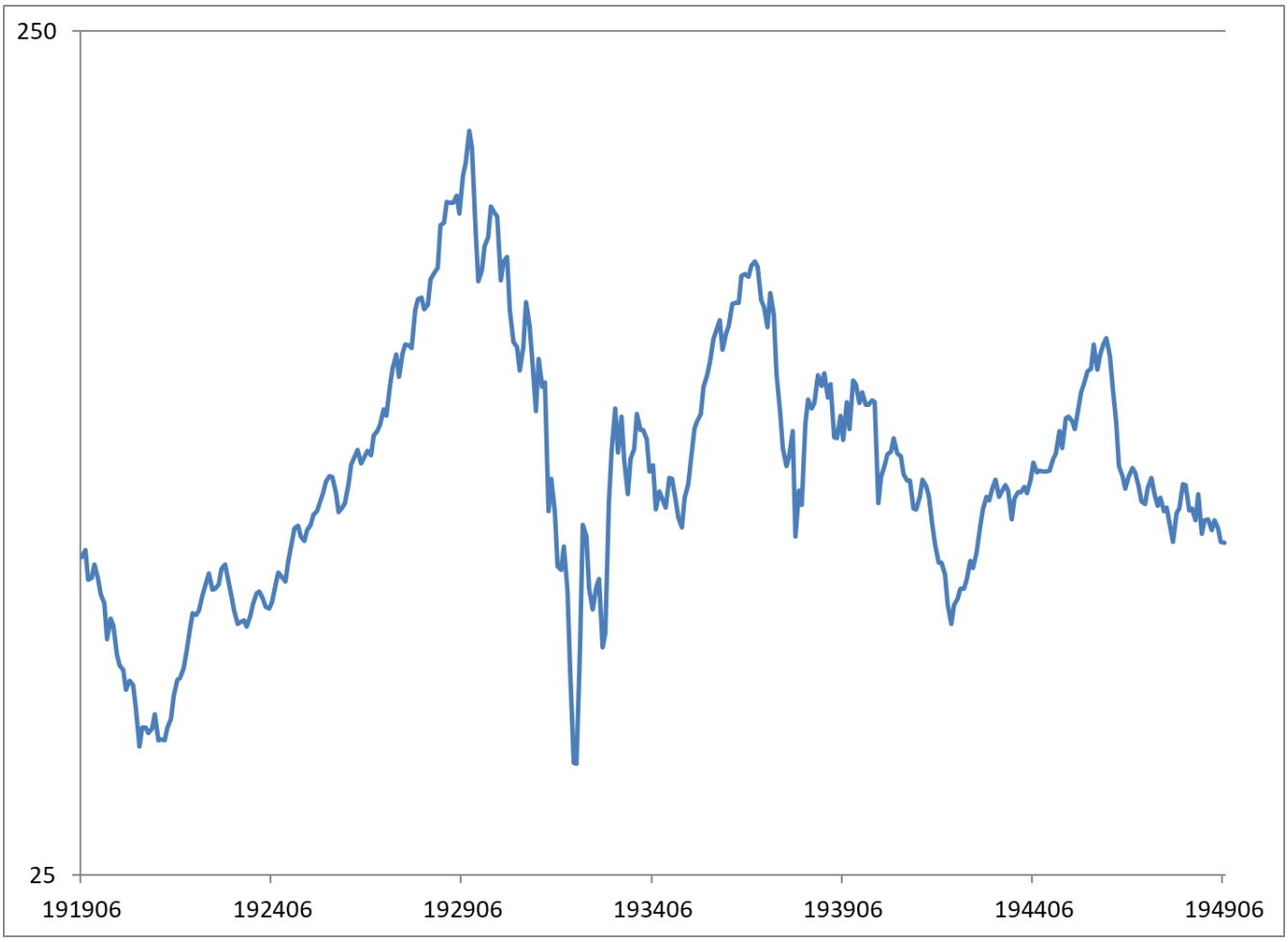
Figure 2
An Unfamiliar Chart: Price Return on Stocks from an Earlier Period



Note: Monthly data chaining the price index of Goetzmann, Ibbotson and Peng (2001) to Cowles' (1938) All Stocks Price index, deflated first by the CPI data of Officer and Williamson (2017) and then by the Consumer Price Index after 1926. Data from the named sources, chart construction by the author. Start value set to 1.0 in June of 1851; deflated price return was just under zero percent per annum for this eighty-one year period (total drop of about 21% as of June 1932).

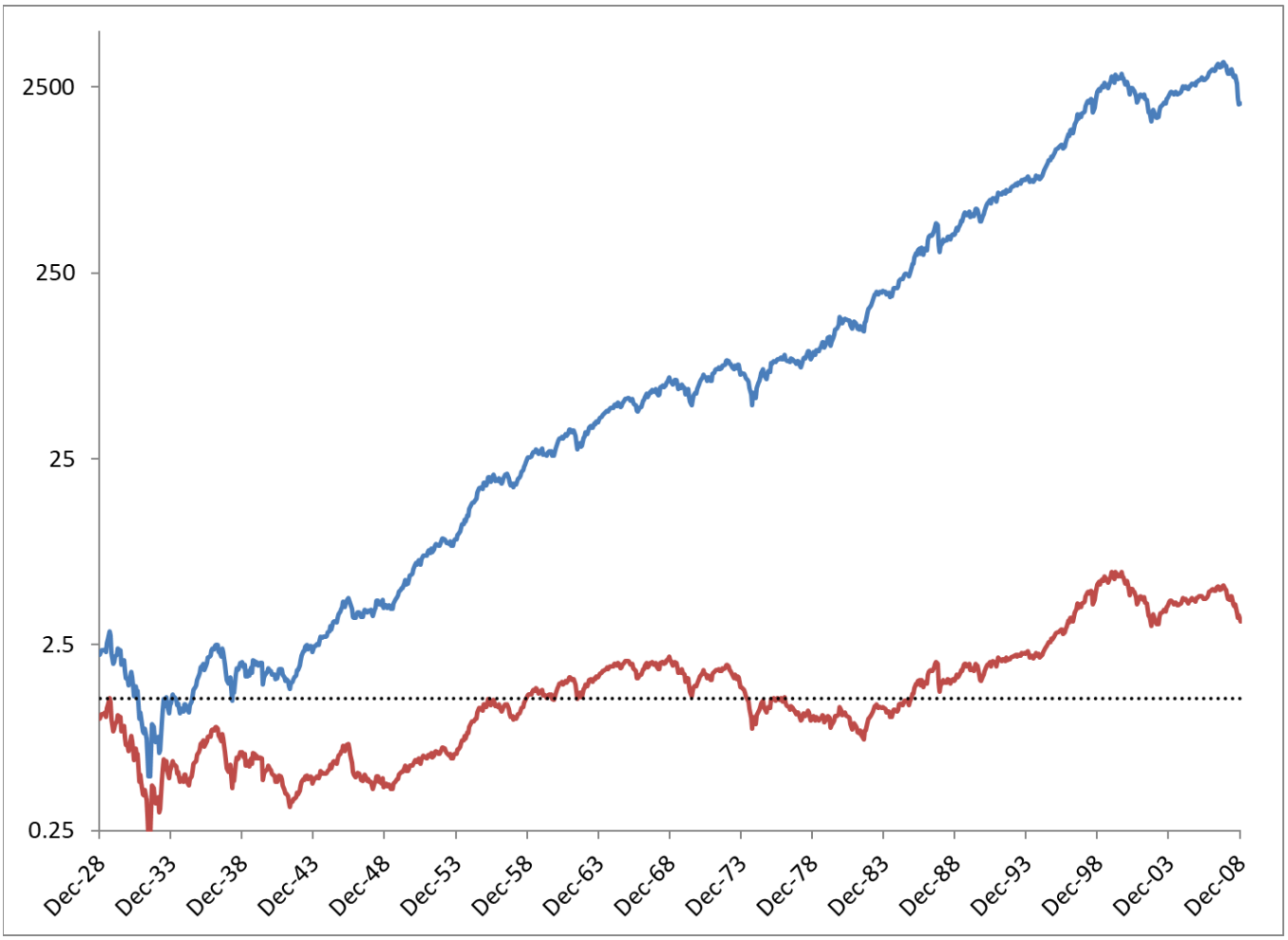
The vertical axis on this chart is 1/1000 the scale on Figure 1. Had the same scale been used, the entire movement charted here would fit into the bottom fourth of that chart.

Figure 3
Thirty Years of Disappointment in US Stocks: 1919 to 1949



Note. Cowles' (1938) All Stocks price index chained to Ibbotson SBBI large company stock price index, deflated by Officer and Williamson (2017) before 1926 and the Consumer Price Index after. The Cowles data for this chart from 1919 to 1926 are based on the Standard Statistics Services index (a corporate predecessor of the Standard & Poor's company); the SBBI data from 1926 are based on a predecessor to the S&P 500 (see Wilson and Jones 2002). Data from the named sources, chart construction by the author.

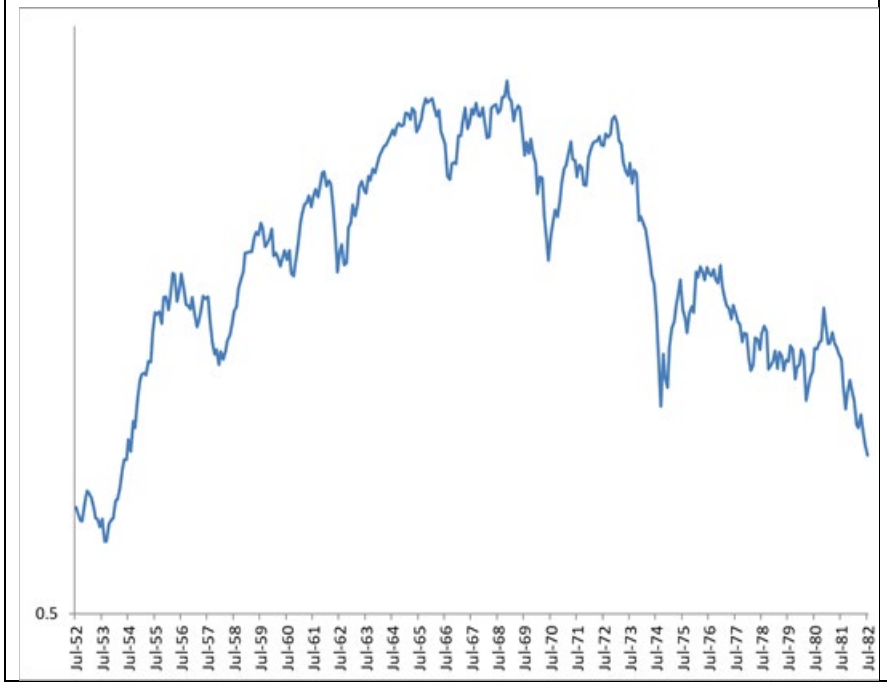
Figure 4
Nominal Total Return and Deflated Price Return on US Stocks

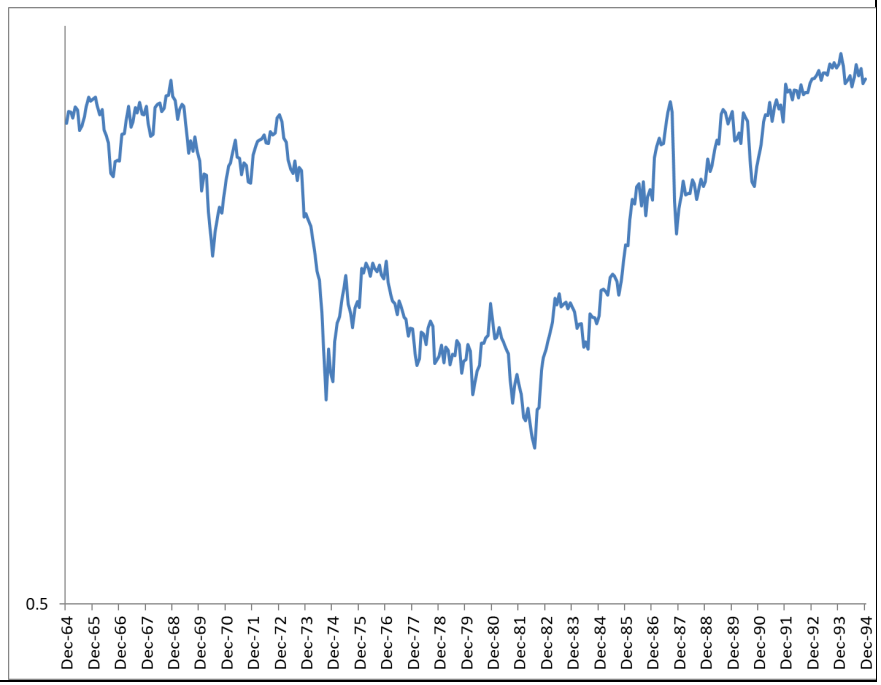


Note. Monthly total return and price return are from the Ibbotson SBBI, as is the Consumer Price Index data used to deflate price return (total return line shows nominal total return). Dotted line shows the (deflated) closing price for August 1929. Start value set at 1.0 in December 1925 for the total return line (off-chart, see Figure 1), and in December 1928 for the deflated price line, to obtain some visual separation in the early years. Calculations and chart construction by the author.

Figure 5
Two More Recent Examples Where Stock Performance Was Disappointing over a Thirty-Year Interval

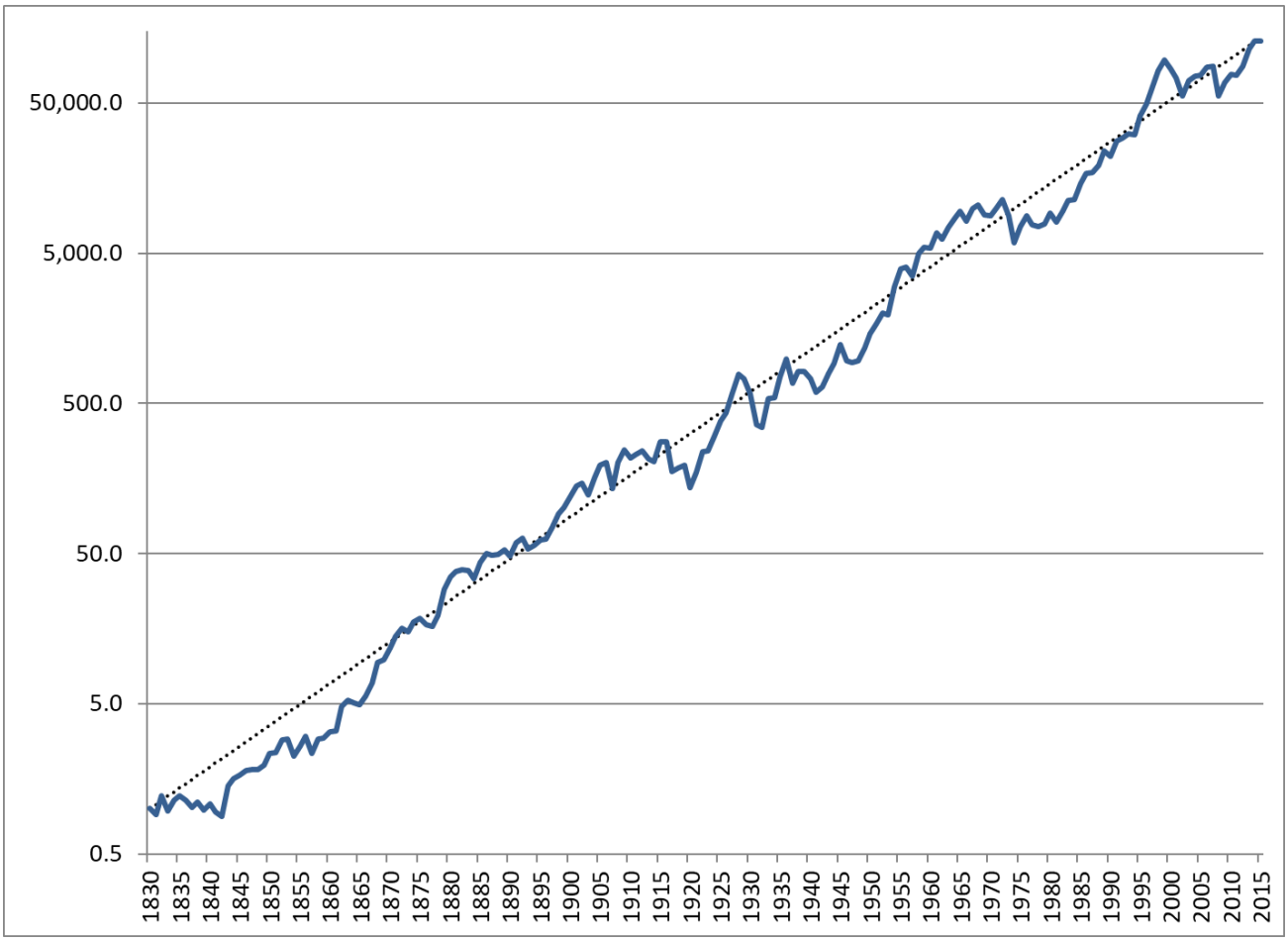
A: July 1952 to July 1982, up then back down	*B: December 1964 to December 1994, down then back up*





Note. Monthly data from Ibbotson SBBI for large company stock price index, deflated by the CPI, also taken from the SBBI. Calculations and chart construction by the author. The two charts have an overlap of 17 years and six months. Each is a subset of the data charted by the red line in Figure 4. Scaling of the vertical axis covers a much smaller range than in Figure 4, to magnify the apparent movement of prices here.

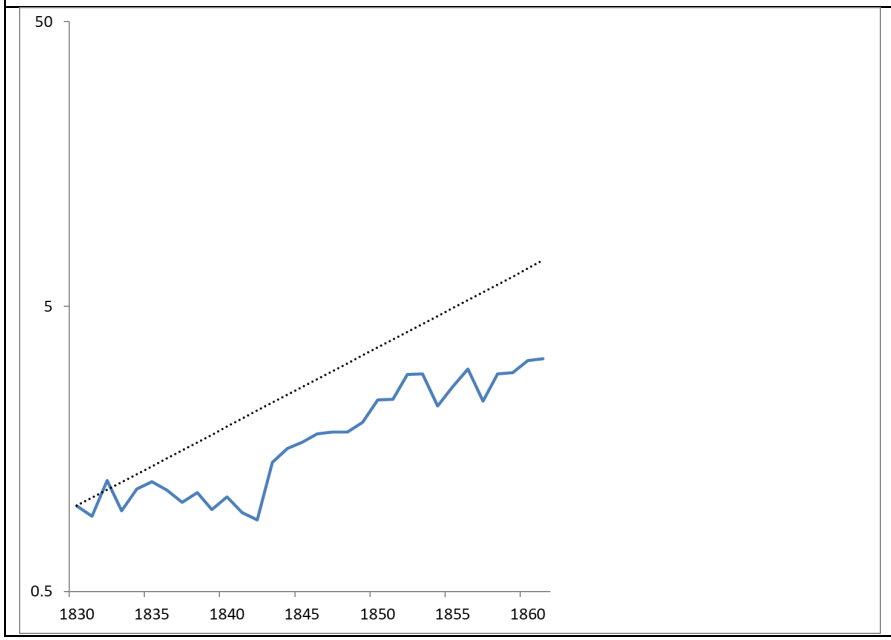
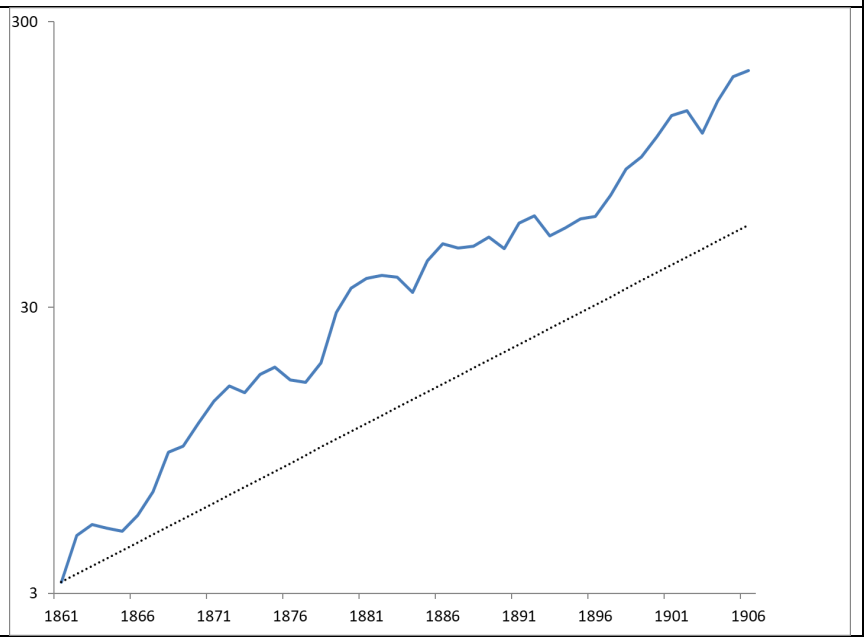
Figure 6
Total Real Return on US Stocks Over the Very Long Term

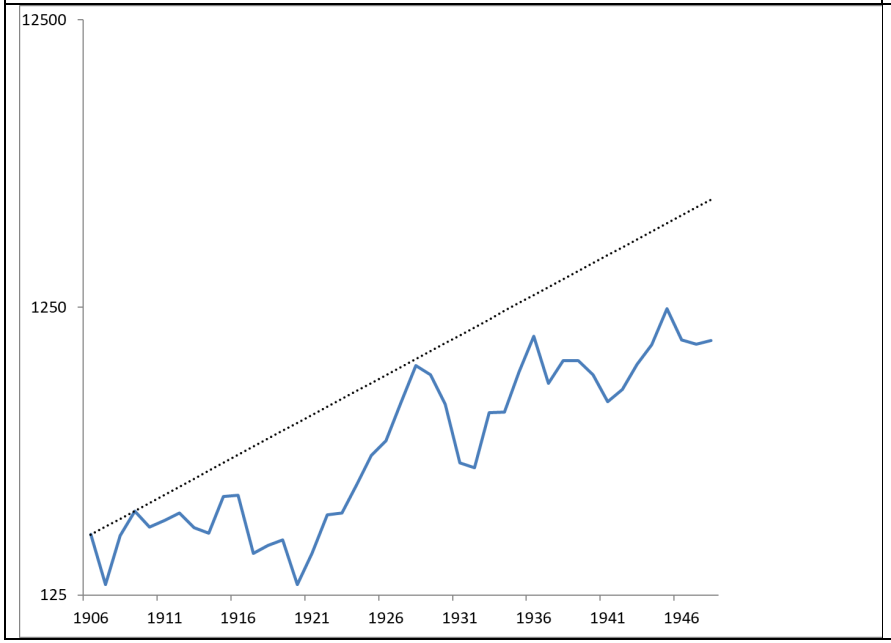
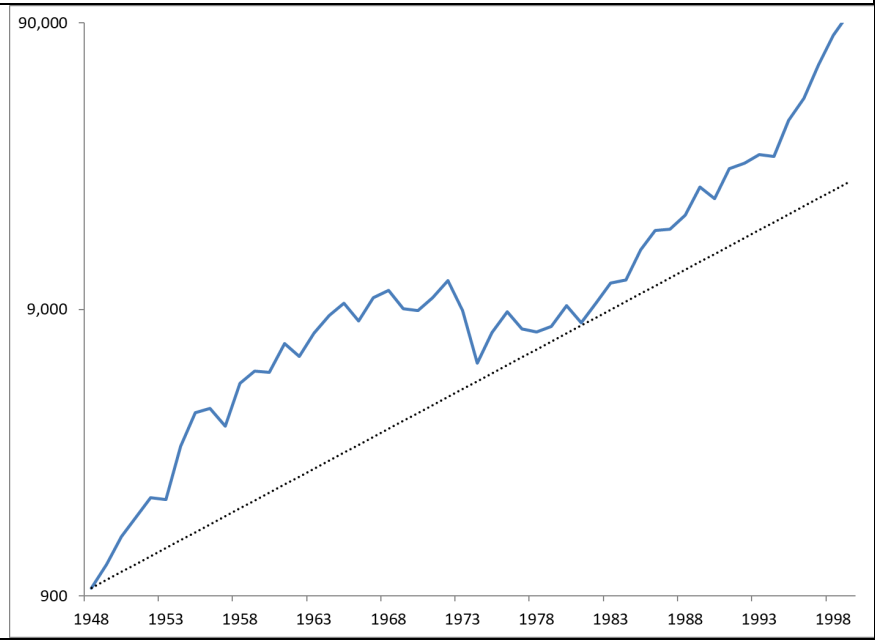


Note. Annual data from the following sources: before 1871 from McQuarrie (2017); data from 1871 through 1925 from Cowles' (1938) All Stock index with cash dividends re-invested; and data from 1926 show the total return on large company stocks per the Ibbotson SBBI. Chart shows deflated values using Officer and Williamson (2017) before 1926 and the CPI afterwards. Dotted line shows a constant 6.6% real return compounded, per the latest estimate of the long-term return on stocks from Siegel (2014). Calculations and chart construction by the author.

See also Figure 9, which reproduces the chart from Siegel (2014) on which this chart is modeled.

Figure 7
How Stock Returns Have Varied by Period, Sometimes Lagging, Sometimes Surging

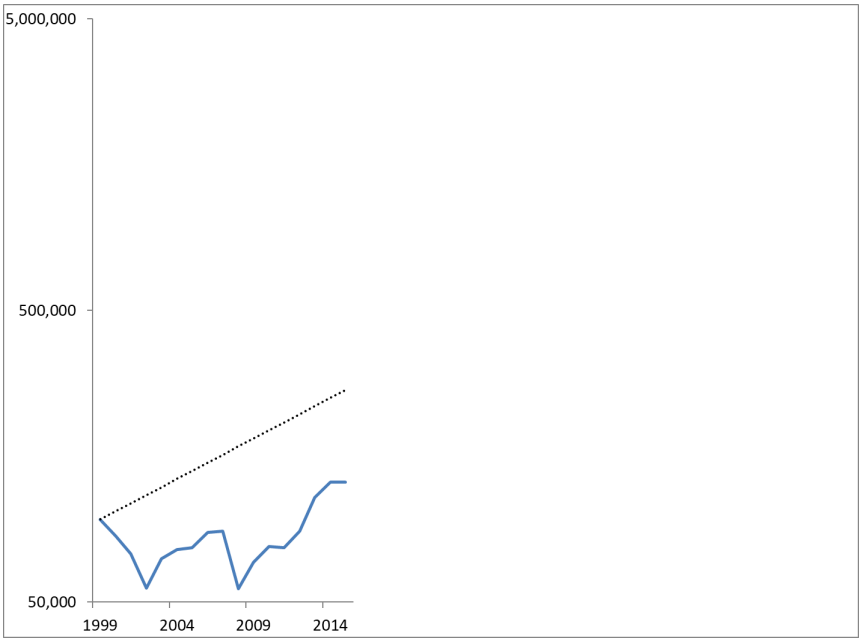
Panel A: Lagging from 1830 to 1861	*Panel B: Surging from 1861 to 1906*
	

Panel C: Lagging from 1906 to 1949	*Panel D: Surging from 1949 to 1999*
	

Note. All panels show monthly real total return, with deflation by Officer and Williamson (2017) before 1926 and then by the CPI. Data sources are as follows: 1) Panel A based on McQuarrie (2017); 2) Panel B based on McQuarrie (2017) chained at the start of 1871 to Cowles' (1938) All Stock index with cash dividends re-invested; 3) Panel C based on that Cowles index chained at the end of 1925 to the total return on large company stocks per the SBBI; 4) Panel D based on that SBBI data. Calculations and chart construction by the author. See text for description of how the panels were scaled.

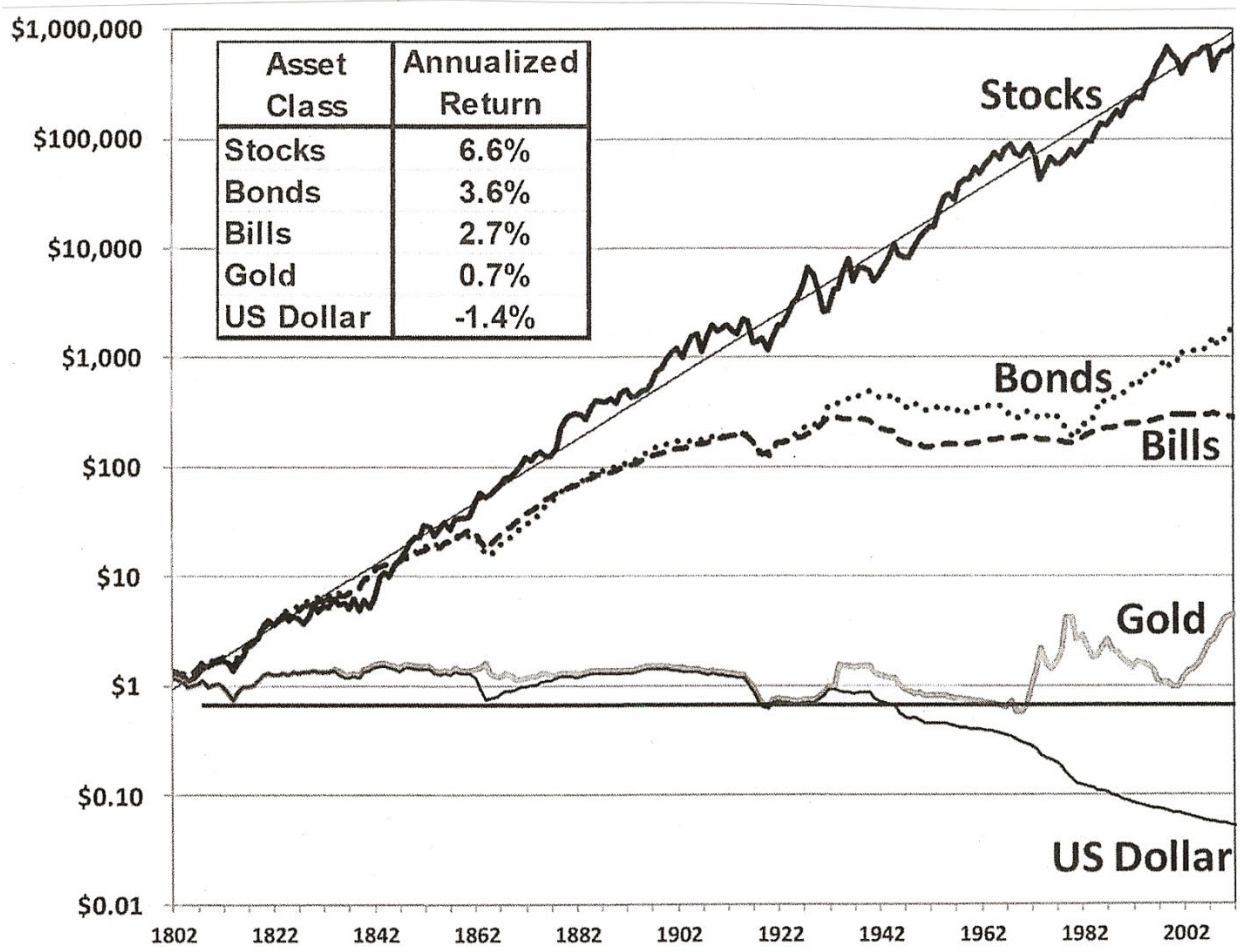
Each panel in Figure 7 is a snippet of the line charted in Figure 6. Dotted line is again the 6.6% geometrical mean return per Siegel (2014), now re-based in each panel to the stock value at the outset of the period charted.

Figure 8
Very Recent Stock Performance in Historical Perspective



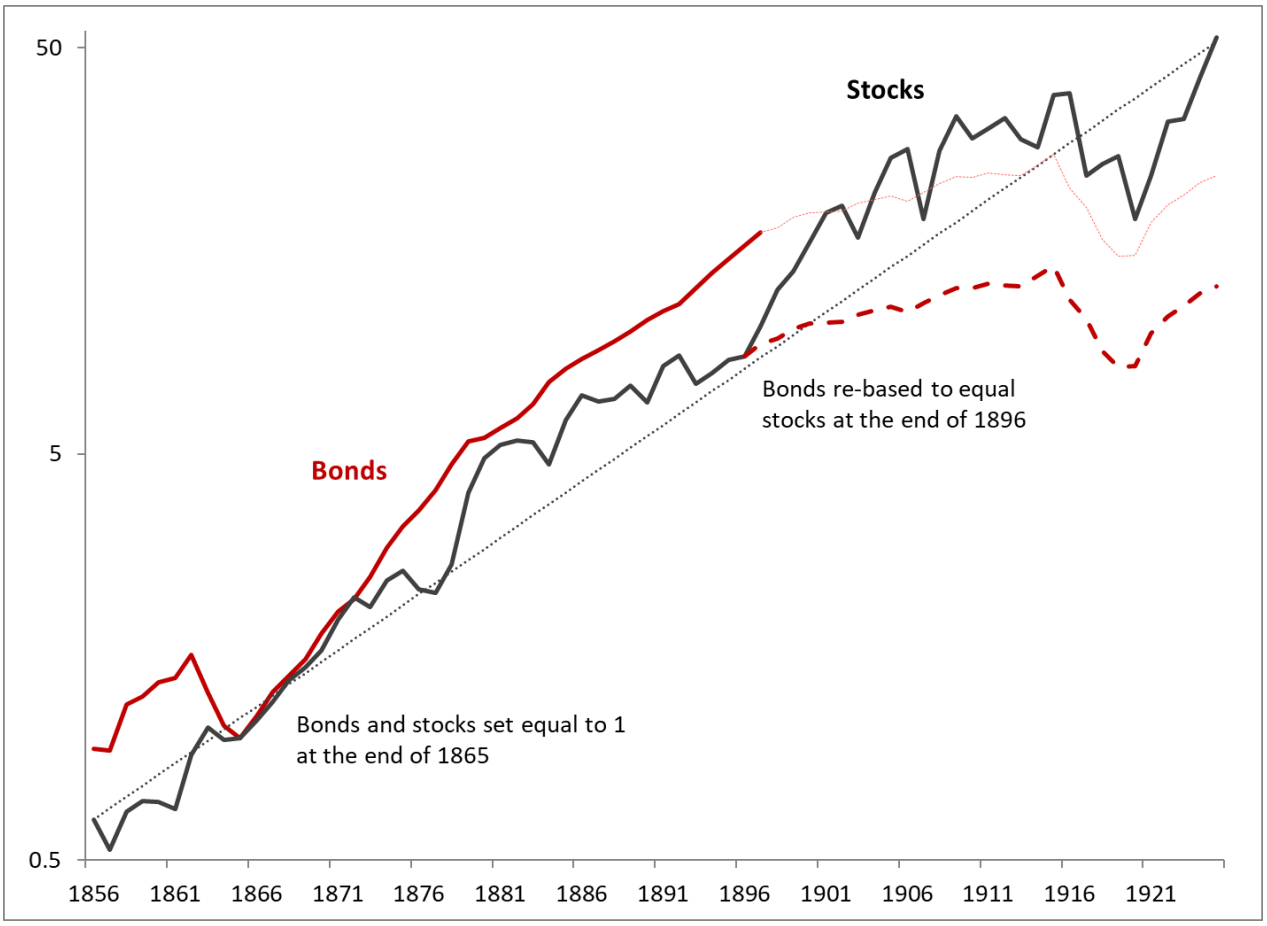
Note. Data from Ibbotson SBBI for large company stocks, deflated by the CPI. Calculations and chart construction by the author. Figure 8 is the final snippet from Figure 6, drawn to the same scale as the panels in Figure 7, with the dotted line representing 6.6% compounded from the end of 1999. See text for an explanation of that scaling.

Figure 9
Long Term Return on Stocks and Other Assets Reproduced from Siegel (2014)



Note. Reproduction of Figure 5-4 on page 82 of Siegel (2014), scanned, cropped and laid out by the author. After 1871, the stock line in this Figure uses the same data as appears in Figure 6, but the inflation adjustment is different up to 1926, after which the stock performance lines in Figure 6 and Figure 9 derive from the same sources and should be identical. Before 1871, Siegel (2014) uses a mix of Goetzmann et al. (2001) and Schwert (1990) to estimate stock performance, with interpolated dividends in the period before 1825, while Figure 6 is based on McQuarrie (2017). The bond performance line here in Figure 9 is not based on public data, but on proprietary data and calculations by Siegel. Compare Figures 10 & 11 below.

Figure 10
Stocks Versus Corporate Bonds Before 1926

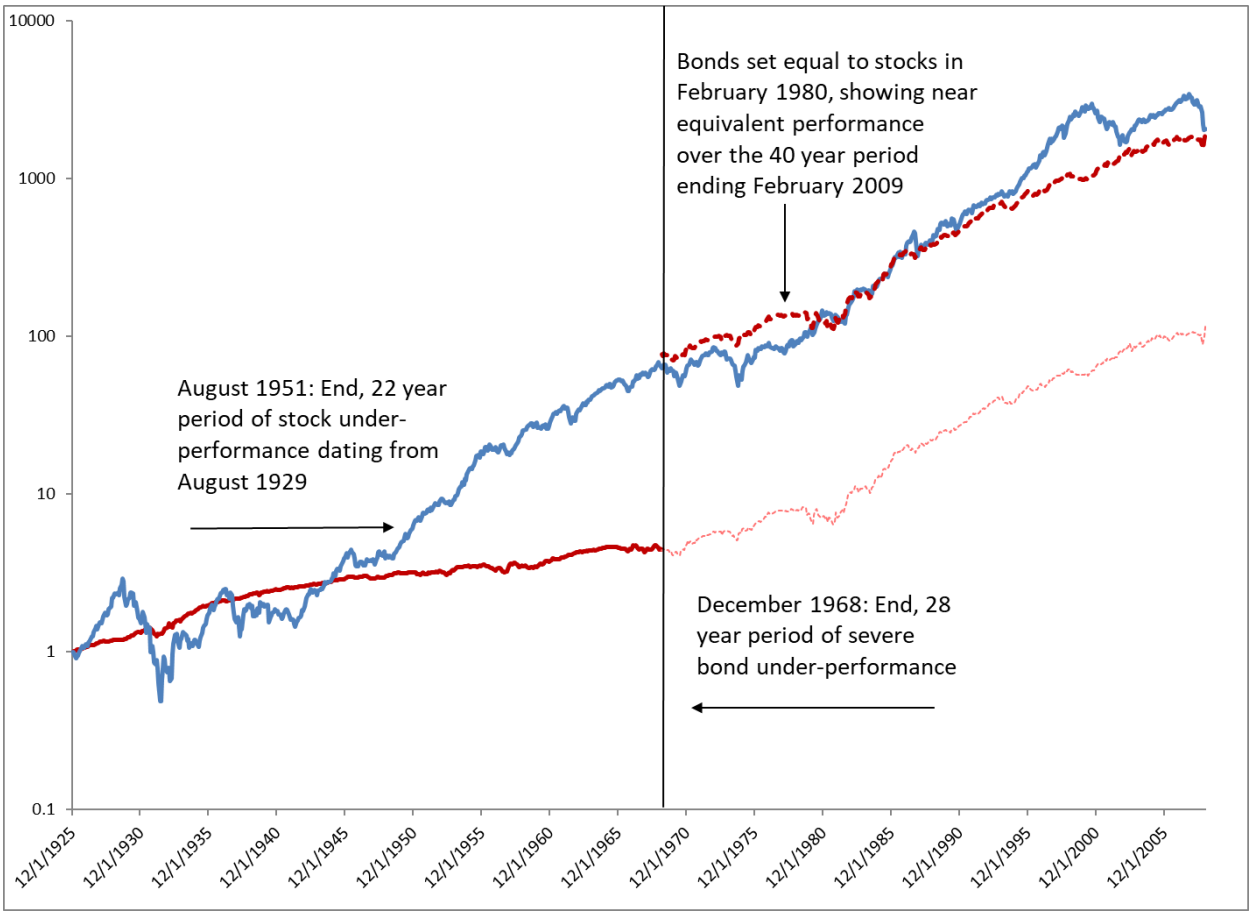


Note. Stock return is per Goetzmann et al. (2001) and Cowles (1938), for comparability with Figure 9. Corporate bond returns from 1871 are from Snowden (1990), who derived holding period returns from yields tabled in Macaulay (1938). From the end of 1856 through 1871, the bond returns represent the author's derivation of holding period returns during that interval, calculated directly from the bond price data given in Macaulay (1938, Table 2, pp. A21ff.).

Macaulay used high grade railroad bonds, none of which defaulted during the period (cf. Gieseke et al. 2011). Up until the early 1900s, railroad stocks also accounted for most of the market capitalization in the stock indexes of Goetzmann et al. and Cowles. Therefore, for that period the stocks and bonds charted were issued largely by the same corporations.

Both series have been deflated using Officer and Williamson (2017). The black dotted line shows a constant real return of 6.6%, per Siegel's (2014) estimate of the long-term return on stocks. Calculations and chart construction by the author.

Figure 11
Stocks Versus Corporate Bonds from 1926



Note. These are nominal total returns for large company stocks and long term corporate bonds from the Ibbotson SBBI. Start values set to 1.0 in December 1925 and continued through February 2009. Calculations and chart construction by the author. Dashed red line is the corporate bond return rebased to equal that of stocks in February 1980. If the bond line had been set equal to stocks in August 1929, the stock line would not break back above the bond line until August 1951, as labelled on the chart.