Revision notes November 2023

A downstream version of this paper has been published in the *Financial Analysts Journal*. The link is: https://www.tandfonline.com/doi/full/10.1080/0015198X.2023.2268556.

The published version is both a subset and a superset of this working paper. It deletes some material on inflation and dividends and shortens the international section.

Conversely, the published version has a Conclusions section with new material not appearing here. The opening section also goes into more detail about data collection. There is a new Appendix on the evolution of international data, and the current Appendix is expanded. Consistent with academic journal norms, the published version is more terse and uses less vivid language. The paper is also positioned somewhat differently, focused more on lack of stationarity and less on where Siegel went awry.

I've changed the title of this working paper to help keep the two versions separate. Otherwise, what follows is the text posted here on SSRN in July of 2021.

Where Siegel Went Awry: Outdated Sources & Incomplete Data

Edward F. McQuarrie*

Abstract

When Jeremy Siegel published his Stocks for the Long Run thesis, little information was available on stocks before 1871 or bonds before 1926. But today, digital archives have made it possible to compute real total return on stock and bond indexes back to 1793. This paper presents that new market history and compares it to Siegel's narrative. The new historical record shows that over multi-decade periods, sometimes stocks outperformed bonds, sometimes bonds outperformed stocks, and sometimes they performed about the same. More generally, the pattern of asset returns in the modern era, as seen in the Ibbotson SBBI and other datasets that begin in 1926, emerges as distinctly different from what came before. Contrary to Siegel, the pattern of asset returns seen in the 20th century does not generalize to the 19th century. A regime perspective is introduced to make sense of the augmented historical record. It argues that both common stocks and long bonds are risk assets, capable of outperforming or underperforming over any human time horizon.

Must stocks beat bonds over the long run? Can buy-and-hold investors expect a stock portfolio to compound in real terms at 6% to 7% per year?

Over the last three decades Jeremy Siegel of Wharton, in a series of papers and across five book editions,[1] has defended both claims under the rubric of Stocks for the Long Run. In financial analysis both tenets have become articles of faith. Planners believe that bonds will necessarily lag stocks over long horizons. And even adjusted for inflation, planners expect stocks to double about every decade.

In support, Siegel marshalled over two centuries of historical data. In this paper I revisit the historical record and use newly available data to challenge both claims. Prior attempts to critique the historical record assembled by Siegel have been few.[2]

I begin with an account of Siegel's sources for stock and bond performance and explain how these have been superseded by new data. I go on to present a revised and more comprehensive historical record of US stock and bond performance from 1793. I then show how recent international data also challenge Stocks for the Long Run. The paper proceeds to argue for the prevalence of regime change in the pattern of asset returns. I add examples showing how inflation regimes have come and gone, how the role of dividends has changed, and how stock-bond correlations and volatility have varied over a wide range. Net of all the examples, the post-1926 period, what I will call the modern era, emerges as highly distinct within a broader historical view. Finally, I discuss why Siegel's assertions may have escaped criticism for so long, and the implications of the new historical record for 21st century investors.

Parsing the historical record

The Appendix tables the sources of historical data used by Siegel, calls out flaws and limitations, and details enhancements introduced by the new data.** The major revisions involve stocks before 1871 and bonds before 1897. Hence, it is mostly the 19th century where Siegel went astray—not through any error of his own, but because the limited sources available to him three decades back proved to be faulty. Siegel's key sources were published over eighty years ago. Those compilations were the best available when Siegel first advanced his thesis. But that is no longer the case.

* Professor Emeritus, Santa Clara University; contact: emcquarrie@scu.edu. This revision dated July 2021. I would like to acknowledge the pathbreaking work of Richard Sylla, head of the research team whose data compilation made this new market history possible. I thank him, Jason Zweig, James Grant, Evan Lorenz, Jeremy Siegel, Mark Hulbert, Mike Staunton and Bryan Taylor for helpful comments on earlier iterations of this work. Special thanks to Bryan for providing access to GFD data. I would also like to acknowledge hathitrust.org, books.google.com, Readex, Mergent, Fraser and all the other digitized repositories that made it possible to go back centuries to extract dividends and other data not previously accessible to financial historians.

** The Appendix begins on page 30; endnotes can be found on the last three pages. Tables and figures are in line.

Toward the end of the 1990s a team led by Richard Sylla of New York University examined hundreds of antebellum newspapers looking for stock and bond quotes.[3] The Sylla team discovered that widespread stock and bond trading dated back to the 1790s, not only in New York but in Boston and Philadelphia.[4] By the 1820s there were half a dozen active exchanges trading dozens of stocks. By the Panic of 1837, quotations on 200 stocks could be compiled.

Shortly after the Sylla research concluded, the digital revolution revamped historical newspaper archives, even as Google began to digitize older books that had gone out of copyright. These searchable digital archives allowed me to compile a dividend record going back to 1793. Siegel had to estimate dividends before 1871; he did not observe them.[5] I was also able to compile share counts from these archives to compute the capitalization-weighted total return on stocks. Siegel's sources provided only equal-weighted or price-weighted indexes.

For the new bond record I again used the Sylla compilation to obtain price quotes and coupons. Few investors realize that Siegel did not observe bond prices before 1926. Nor do investors recognize that Siegel did not sample from the universe of available bonds, as would be done today in constructing a bond index. Instead, he applied a formula to transform successive annual yields, often from a single bond, into imputed capital appreciation for the bond market as a whole.[6] Likewise, no part of Siegel's bond record includes corporate bonds, even during the decades when these dominated the fixed income markets. His theoretical commitments required him to use the best available proxy for a risk-free instrument. In the absence of a deep and liquid Treasury market, Siegel approximated the risk-free bond by always selecting the lowest yielding Federal or municipal bond available, even when that yield was depressed by tax exemptions and other privileges.[7]

I observe bond prices throughout to discover how an investor in a hypothetical bond index mutual fund would have fared across the centuries. I also assumed a bond investor who wanted to maximize risk-adjusted total return over the long run. Hence, I had no reason to exclude corporate bonds once these became available in the middle of the 19th century.[8]

The new history of US stock and bond performance

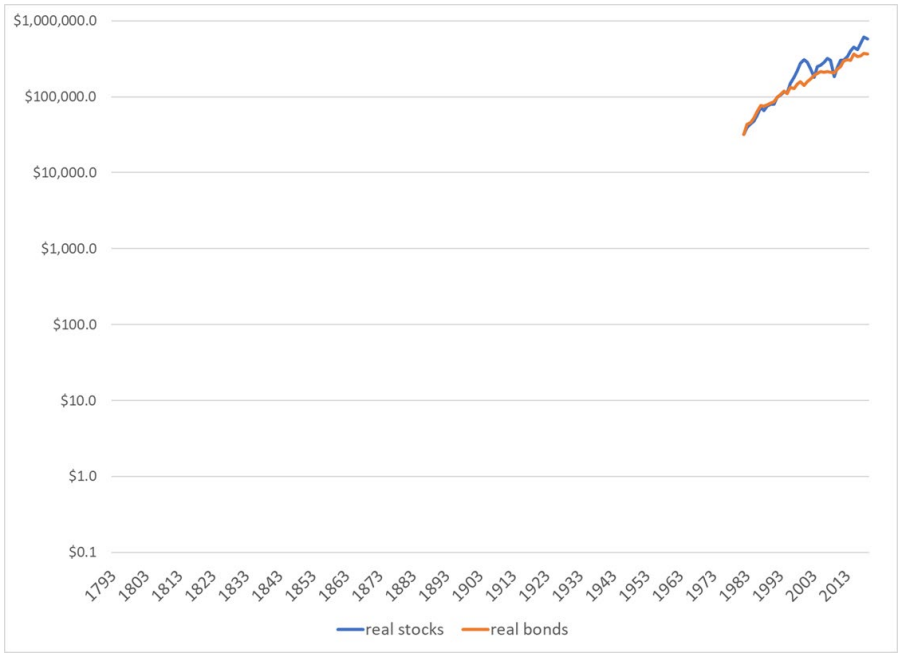
Figure 1 shows recent stock and bond performance, from January 1982 to January 2019.[9] This period will represent lived history for many readers.

But Figure 1 will strike many as unfamiliar. It shows that stocks and bonds performed about the same over the most recent four decades—not what most investors expect to hear. Occasionally, stocks soared above bonds, as during the dotcom boom; but as late as January 2009, stocks had fallen behind bonds in what by then had been a twenty-seven-year horse race.

Figure 1 will also look unfamiliar because the 1982 starting value for both stocks and bonds is set equal to the wealth achieved by a stock investor over the first one hundred and eighty-nine years of US financial market history, as will emerge. And Figure 1 is also scaled to accommodate the addition of earlier data, leaving recent data crowded into the upper right corner, making it appear a minor coda to an extended history. That is a deliberate reversal: most published multi-century historical charts have presented the earliest returns as a minor preface to the main act.

Figure 1 offers little to support the Stocks for the Long Run thesis. Stocks did very well in the boom following 1982—but so did bonds. True, there were short periods where stocks soared ahead of bonds; but these periods were offset by plunges that took stocks back down to the bond line. Most of Figure 1 postdates Siegel's earliest efforts and can be approached as a failure of replication: the pattern Siegel thought he saw in

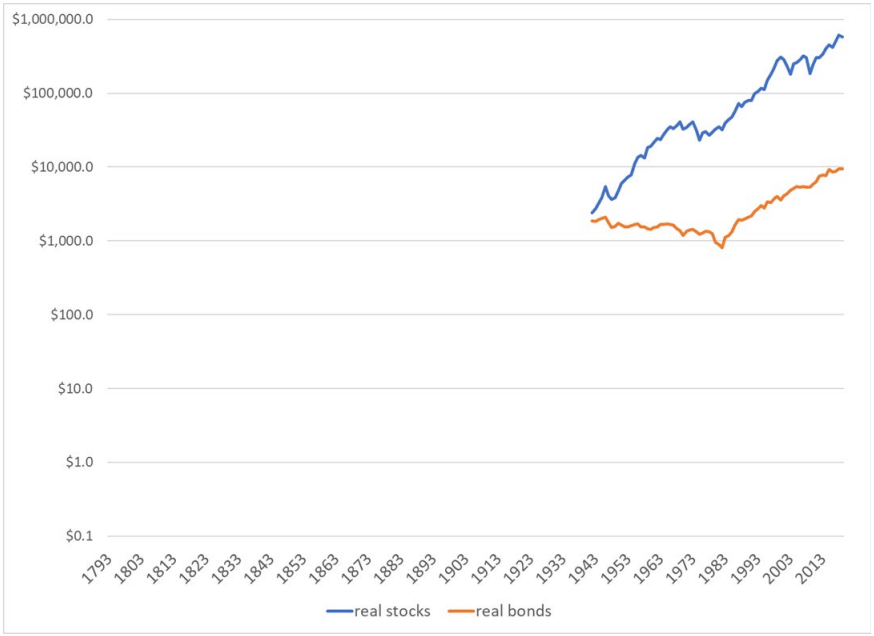
Figure 1
Stock and Bond Performance Reset in 1982
and Scaled to Accommodate Earlier Data



the historical record was not sustained after he published.

The next figure looks quite different. It adds another four decades, extending the record back to January 1942, just after the United States entered World War II. Here the starting values are the actual wealth achieved in stocks and in bonds to that point, over the first one hundred and forty-nine years of US investment history. Figure 2 shows terrible long-term performance by bonds. Stocks rose far ahead of bonds during the decades immediately following World War II and never fell back by much. Conversely, real bond returns were negative across the newly added period. Figure 2 would appear to provide strong support for Stocks for the Long Run.

Figure 2
Stock and Bond Performance since 1942 Set in Historical Context



But Figure 2 is highly misleading. Not one investor in a thousand can look at this chart and detect that following 1982, the stock and bond performance lines are essentially parallel, as demonstrated in Figure 1. What investors see in Figure 2 is a steadily widening gap between stock and bond performance: an opening fan. The fact that the fan stops widening after 1982 is nearly impossible to discern.

Investors reading charts are prone to misleading visual heuristics, parallel to the many cognitive errors identified in the behavioral finance literature. If chart lines begin at the same point, then spread apart, then stay spread, investors readily infer that the one asset has outperformed the other, *and* that the outperformance has been consistent and ongoing. The eye observes no qualification: the gap on the right side of Figure 2 becomes proof positive of the Stocks for the Long Run thesis. The untutored eye cannot see that almost all the out-performance occurred prior to 1982; and the greater part of it, before 1969.

To destroy the visual illusion of consistent outperformance, the analyst has to break out sub-periods. Hence the need to begin with something like Figure 1. Conversely, charts similar to Figure 2 have been appearing in the Ibbotson SBBI yearbooks from the 1980s, before Siegel first published his thesis. Correctly described, Figure 2 shows a performance gap that widens dramatically in the decades just after the war, and then stops widening. But how many investors have encountered such commentary: "the period of sustained stock outperformance was a postwar phenomenon, which had mostly run out of steam by the end of the 1960s"?

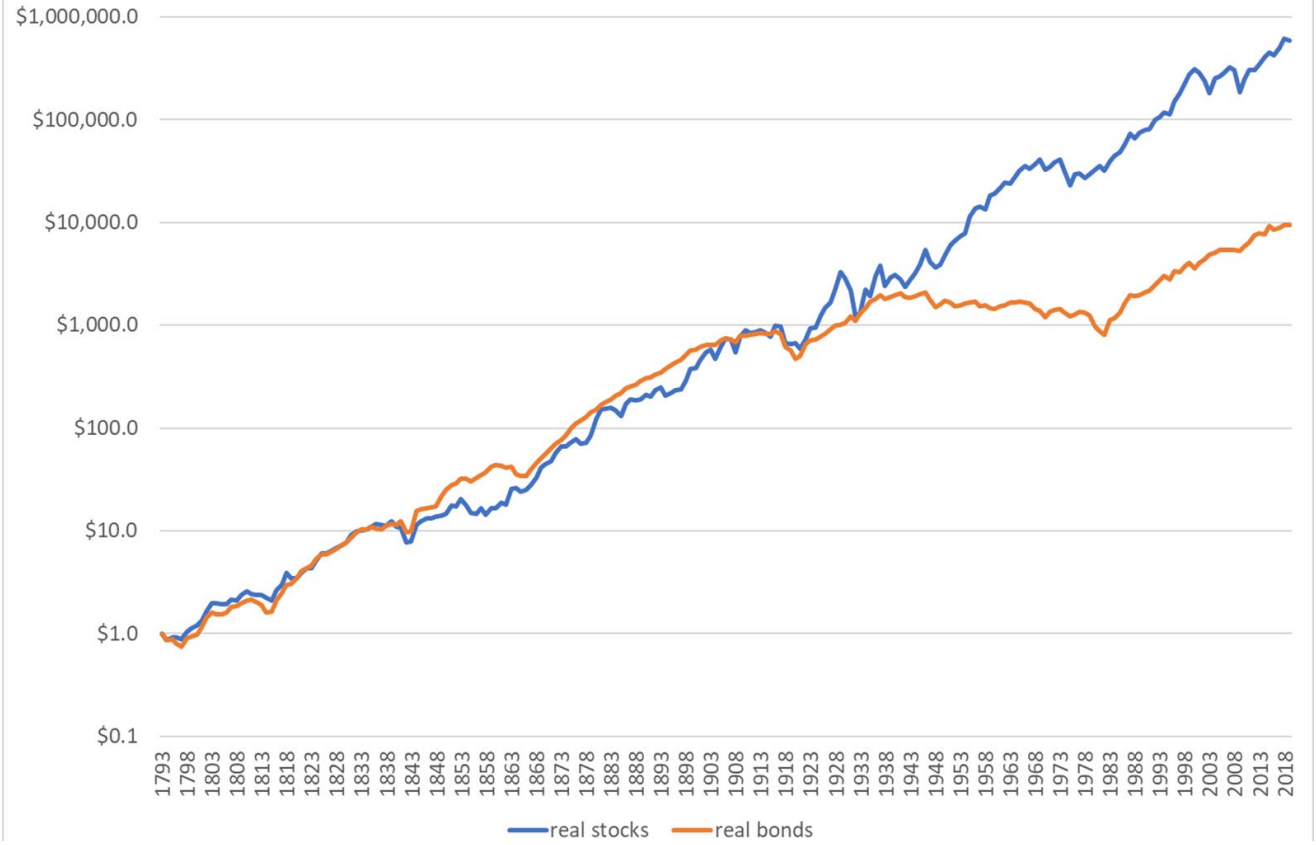
Figures 1 and 2 capture two different regimes in the relative performance of stocks and bonds: a recent era of parity performance, and an earlier era in which stocks soared far above bonds. These preliminary figures set the stage for the much longer span of market history that stretches back from World War II to the beginnings of the Republic. That period falls outside the lived history of today's investors.

Figure 3 fills out the complete record by adding stock and bond performance from January 1793 to January 1942: the first century and a half of US financial markets. Turns out, performance before 1942 looks more like Figure 1 and not so much like Figure 2: it shows mostly parity performance, with stocks and bonds producing about the same wealth accumulation by 1942. As seen in Figure 1, and parallel to the dot-com boom, the extended historical record also shows short spans of extreme stock outperformance, as in the 1920s. But there is no period, going back to 1793, that looks anything like the sustained postwar surge of stocks over bonds seen in Figure 2. The postwar period emerges as a one-off.

However, neither does the inverse of Figure 2 appear anywhere in the first century and a half. True, bonds did move ahead of stocks during the Panic of 1837. And afterwards, bonds stay ahead in the horse race for over seventy years. But the magnitude of the bond advantage, at its greatest point just before the Civil War, is only about 2.5X in terms of the cumulative wealth index. In stark contrast, from 1942 to 1982 the wealth index for stocks grew to be over 30X larger than that of bonds.

Nonetheless, had Siegel written in 1942 rather than 1992, he would have had no occasion to advance the Stocks

Figure 3
New Historical Record of US Stock and Bond Performance since 1793

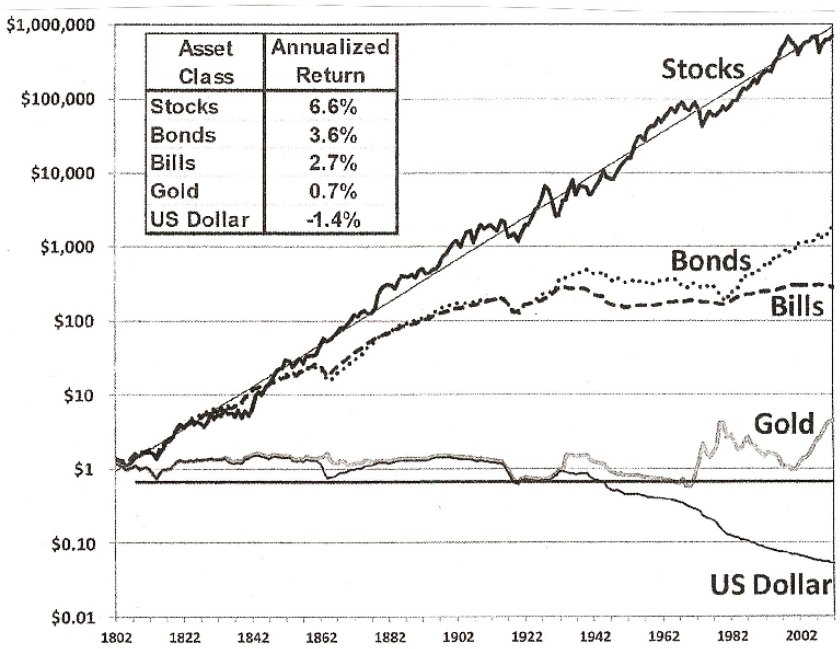


for the Long Run thesis. Market history to that point would have indicated that over long intervals stocks and bonds typically ran neck and neck, with now the one in the lead when measured from 1793, now the other. Stocks and long bonds were brother and sister, two different ways to invest for long term returns.[10]

Now it is time to compare the new historical record to the older rendition per Siegel (Figure 4). Examining the Siegel chart, another misleading visual heuristic becomes evident: that in a semi-log chart extending across centuries, the viewer tends to ignore everything in the lower left corner. Visually, that portion of Figure 4 almost disappears. But if this heuristic can be overcome, what becomes apparent there? Answer: Siegel found that stocks and bonds performed about the same over the first six decades of his data series.

More commentary investors have never seen: "according to Siegel, the

Figure 4
The Old Market History According to Siegel



Note. This is a scan made from my copy of Siegel (2014), his Figure 5-4, p. 82. In this chart the series end at the close of 2012.

advantage of stocks over bonds made a late appearance, only commencing after six decades of market history had passed." Siegel does not give a blow-by-blow account of what I have called the horse race between stocks and bonds, and generally does not look at sub-periods except at important transitions in his data sources (1871 and 1926, see Appendix). And Siegel's readers have typically not been able to overcome the "ignore the lower left corner" heuristic, especially in face of the ever-widening fan between stocks and bonds that dominates the right two-thirds of his chart. A glance at Siegel's chart compels the familiar response: stocks have *always* substantially out-performed bonds. But that is not correct, even using Siegel's flawed and outdated data sources.

Turning back to the new historical record in Figure 3, the contrast with the older history can be summed up simply: where Siegel showed parity performance between stocks and bonds extending over the first six decades, the new record more than doubles the span of parity performance, to about 150 years. Sustained stock outperformance emerges as a postwar phenomenon, not seen before (or since).

The next section drills down to show where the new historical record departs from the old, to identify what caused these departures, and to establish the superiority of the new record, as a better rendition of market history than what Siegel had available thirty years ago.

Changes in the new history

Figure A.1 in the Appendix compares the new bond record directly to a reconstruction of the Siegel bond record. The two show nearly identical returns up to the beginning of the Civil War, with most of the variation relative to Siegel coming in the decades following. Explanation: I observe municipal bond prices from 1857 to 1897 as compiled by the Boston stockbroker, Joseph G. Martin. Siegel relied on a tertiary source, the summary table in Sidney Homer's *History of Interest Rates*. Homer does not state there that the municipal bond index he had tabled represents a theoretical construction erected by Macaulay upon Martin's compilation rather than an observed portfolio. I also adjust for the greenback price of interest paid in gold coin—a distinguishing feature of Federal and some municipal bonds between 1862 and 1879. Finally, I use an aggregate bond portfolio up to 1897, and corporate bonds after that point. Corporate bonds returned more than the government bonds used by Siegel, especially when government bond yields were depressed by tax and other privileges.

Figure A.2 compares the new stock record to Siegel. Here most of the variation comes before the Civil War, and this deviation is readily explained: Siegel's sources omitted the largest single stock that traded before the Panic of 1837, the 2nd Bank of the United States. At the peak before the Panic hit, the 2nd BUS accounted for almost 30% of total market capitalization. It failed spectacularly as the Panic proceeded, with shares dropping in price from $120 to $1.50, and never recovered. To duplicate this omission in the contemporary stock market, it would be necessary to drop Microsoft, Apple, Amazon, Alphabet/Google, and Facebook from the S&P 500; and even these five would not account for as high a percentage of S&P 500 capitalization as did the 2nd BUS at its peak. Omission of the BUS is the single most glaring error I found in Siegel's stock market sources.

More generally, I found and corrected survivorship bias. Banks failed during panics, turnpikes and canals succumbed to railroads, and struggling railroads went bust in the 1840s and 1850s to an extent not previously understood. In short, Siegel's sources had left out the bad parts, producing an overly rosy picture of antebellum stock returns.

To sum up, the argument for the new historical record stands on three legs. First, that it is better to observe bond prices across a portfolio than to use mathematical constructions erected upon black box yield series. Second, that value-weighting gives a more accurate view of true wealth accumulation. In part it was the sheer size of the BUS that drove down total market returns after the Panic of 1837. Third, that accounts of the distant past must stay vigilant against survivorship bias. Good market history needs to include the losers: the banks that failed, the railroads that never paid a dividend, the bonds that were downgraded or defaulted.

Rolling returns

Figures 1 and 2 manipulated the starting date to make a point about regime change. When return series are found to be sensitive to the point of beginning, a favored analytic solution is to compute rolling returns over some fixed interval. Counts of how many rolls each asset "won" can then be tallied and treated as the odds that stocks will beat bonds over such a holding period. The roll procedure can be further enhanced by systematically varying the interval, for instance, by examining in turn five-, ten-, twenty-, thirty-, and fifty-year rolls. A pattern across roll length can be interpreted as a change in odds as the holding period grows longer.

Siegel includes such a rolls analysis,[11] and his data indicated that the longer the holding period, the greater the odds that stocks will outperform bonds. Specifically, he found that the odds in favor of stocks rose from 3-in-5 over any one year to 9-in-10 over thirty-year rolls, offering strong support for Stocks for the Long Run.

The new historical record tells a different story. Table 1 provides overall results with a breakout by subperiod, using 1862 and 1942 as the break points. Recall that the new stock record shows the greatest deviation from the Siegel record before the Civil

Table 1
Odds that Stocks Beat Bonds Over Rolls of Increasing Length

	Holding Period in Years					
	One	*Five*	*Ten*	*Twenty*	*Thirty*	*Fifty*
All data:	56.2%	61.7%	65.0%	65.7%	68.0%	68.4%
Counts:	226	222	217	207	197	177
Sub-periods:						
Rolls ending 1794 - 1862	47.8%	41.5%	36.7%	12.0%	0.0%	0.0%
Rolls ending 1863 - 1942	53.8%	63.8%	67.5%	71.3%	71.3%	55.0%
Rolls ending 1943 - 2019	66.2%	76.6%	84.4%	94.8%	100.0%	100.0%

Note. Annualized real returns were calculated for stocks and bonds and then a count was taken of the number of rolls where the stock return exceeded the bond return. The percentage stated is that count divided by the total number of rolls. That total declines with the length of holding period as shown in the second row. In the sub-periods the one-year counts are 69, 80, and 77 respectively; if desired, use the pattern for the total sample to derive the counts for longer rolls within sub-periods.

War, and that World War II marked the beginning of the worst bond bear market in US history.

I think it fair to say that the results for the entire 226 years in Table 1 are weakly supportive of Siegel's thesis: the odds that stocks will outperform bonds *do* increase as the holding period lengthens from one to fifty years. But using the new historical data, these odds never get much higher than 2-in-3 and increase only slowly as the holding period stretches out.

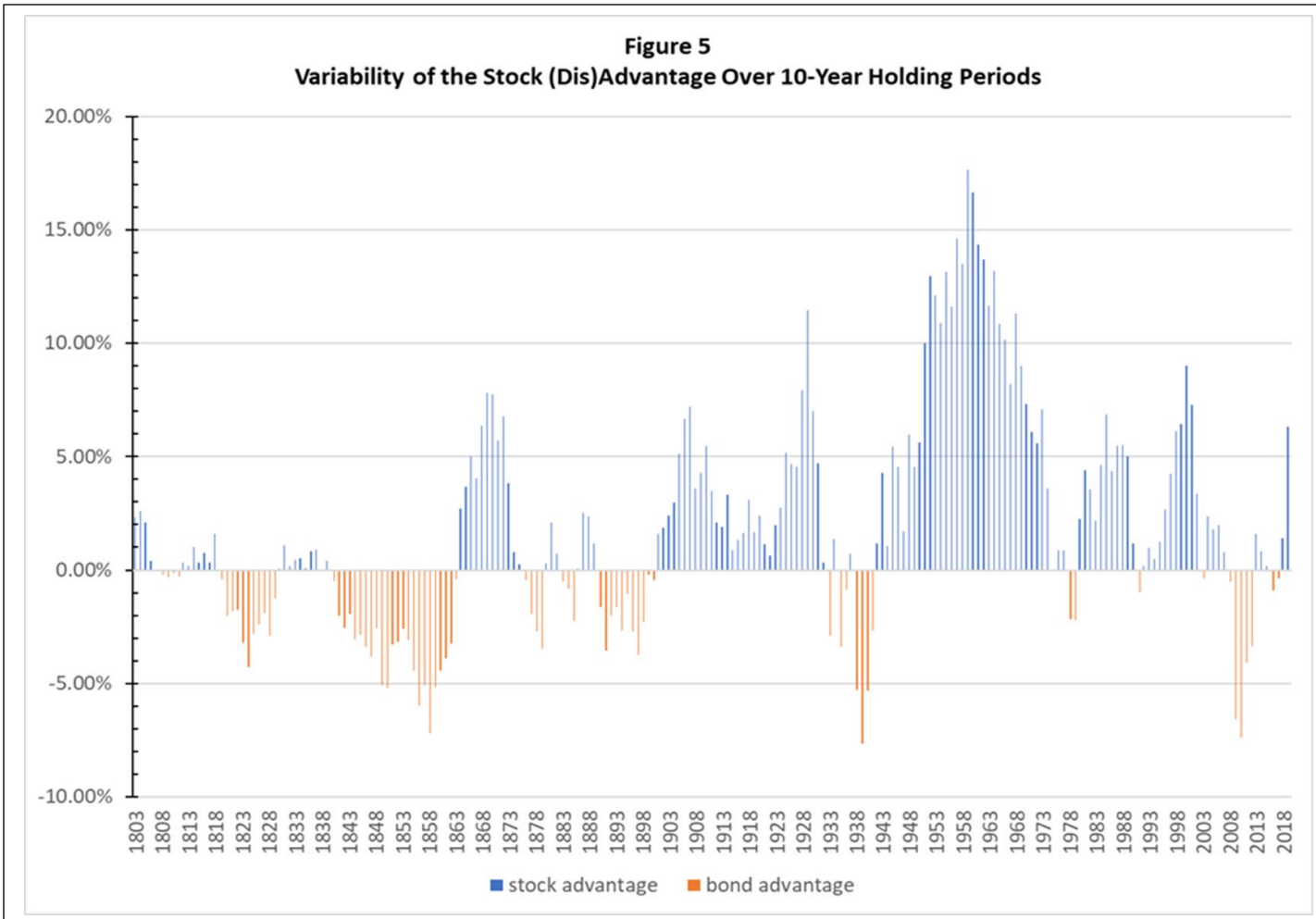
However, the subperiod results in Table 1 do not support Stocks for the Long Run. Prior to the Civil War the pattern was the reverse: the longer the holding period, the greater the odds that stocks would *underperform* bonds. In fact, for holding periods of thirty and fifty years, in that era stocks *always* lost to bonds. Conversely, the most recent period shows complete vindication of Siegel's thesis: after 1942 the odds that stocks outperform start high and increase with holding period, until by thirty years the odds reach 100%. Finally, results for the middle period are confounding. These fit an inverted U-shape, with stock odds increasing up to twenty years and then declining back toward the one-year level for fifty-year holding periods.

No enduring, lawful relationship between asset outperformance and length of holding period can be extracted from Table 1. The results are better interpreted as showing changes in regime. Prior to 1942 a regime of parity performance held sway: sometimes stocks out-performed, sometimes bonds; hence, no matter the holding period length, the odds that the one or the other will outperform hover either side of 50%.[12] Then for a few decades after World War II, a new regime of extreme stock outperformance held sway. For post-1942 holding periods the longer the holding period the more likely stocks are to outperform—not because there is a lawful relationship, but because longer holding periods are more likely to encompass more of the scattered huge annual wins that occurred for stocks in the postwar epoch.

The rolls analysis broken out by sub-period is peculiarly destructive of Siegel's thesis. It supports instead the idea of time-varying regimes in asset returns. Sometimes stocks outperform bonds, sometimes bonds outperform stocks. What matters is not the length of the holding period, but which regime is in operation: the antebellum regime when stocks languished, or the regime following World War II when bonds fell into an abyss.

Next, Figure 5 converts the tabulation of odds into a chart of the stock – bond advantage over ten-year rolls specifically.[13] This figure goes beyond the count of rolls where "stocks won" to lay out the magnitude and durability of the stock advantage at different points in market history. In addition to long-term shifts, the regime perspective can accommodate a more rapid alternation of regimes, as developed next.[14]

Figure 5 shows that the stock advantage has regularly reversed, even in the postwar period. In fact, across the two centuries there have been eight episodes where stocks underperformed bonds over multiple consecutive ten-year rolls. Unsurprisingly, in view of Table 1, two of the worst runs occurred before the Civil War. The stock disadvantage has several times been worse than five percentage points annualized, most recently in 2008 and 2009. Figure 5 also shows how unusual the immediate postwar decades were when viewed through a longer lens. These decades marked the first time that the ten-year annualized stock advantage exceeded twelve percentage points and the first occasion where runs showing a substantial stock advantage extended beyond a decade. Neither of these patterns had occurred before, nor has either recurred since the late 1960s. Likewise, the sustained stock disadvantage visible after the Panic of 1837, which



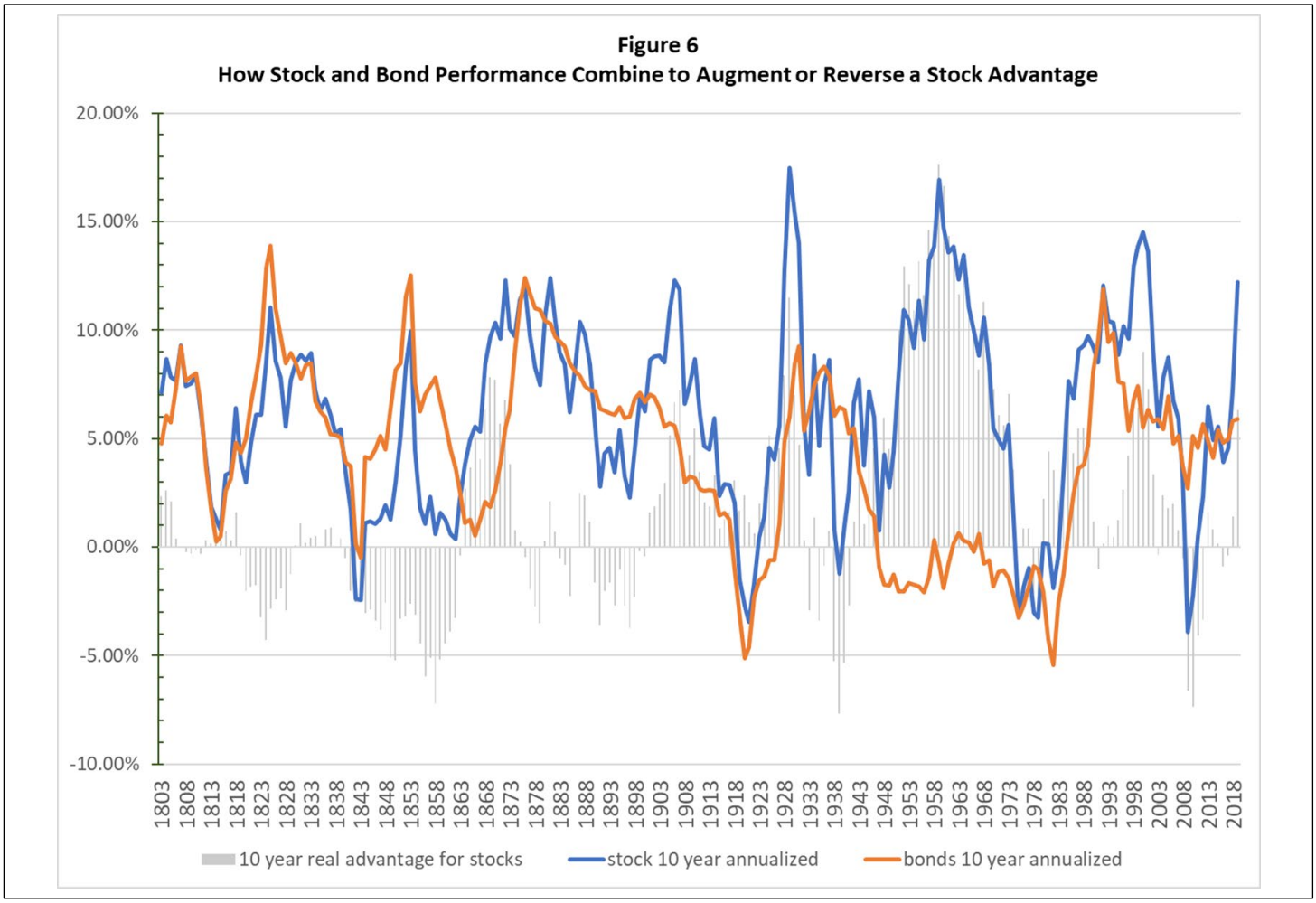
lasted for about two decades, has not recurred.

The pattern visible in Figure 5 is best described as alternation without periodicity and with few constraints on magnitude. Sometimes stocks win, sometimes they lose. Sometimes stocks prevail only briefly, sometimes for runs of a decade or more. Sometimes stocks fall behind bonds only briefly, at the depths of a bear market, while at other times the disadvantage is sustained well past that bottom. Most of the time, the stock advantage or disadvantage has been within a range of +/- 5%, but larger values have also irregularly occurred. Most important: Figure 5 does not show a stock advantage that waxes and wanes, but a stock advantage that regularly reverses into a deficit. Across the centuries there does not seem to be a positive mean value to which the stock advantage consistently reverts.

Next, Figure 6 superimposes the stock and bond return series themselves, also as ten-year rolls. Figure 6 makes it clear that a stock (dis)advantage can be created through a variety of possible combinations and permutations. For instance, the decades following WW II are the only example where extreme positive returns on stocks coincided with sustained negative returns on bonds to produce a large and enduring stock advantage. It is this unique and relatively recent event, fatefully combined with his faulty 19th century sources, that led Siegel to extrapolate that in the long run, stocks would always beat bonds by a substantial amount.

For a rather different permutation, consider the decades following the Panic of 1837. After plunging together at the outset, both stocks and bonds recovered but the bond recovery accelerated faster and was sustained for longer. The combination produced a run of stock disadvantage that lasted for more than a decade. As one final example among many possible permutations, the inflationary surge during World War I and the subsequent mini-depression of 1921 - 1922 caused the ten-year performance of both stocks and bonds to plunge toward the lowest levels for each ever seen in the record. The net effect was a tiny stock advantage of <1% for the roll ending in January 1922.

It seems unlikely that two centuries suffice to field all the possible permutations of stock and bond performance that may occur, and two centuries is certainly far too short an interval to see whether any one permutation more frequently recurs. But two



centuries are sufficient to fatally undermine the idea that an advantage of stocks over bonds, of +2 to +4% annualized, represents any kind of lawful regularity, or any mean to which values must revert. Rather, whether there is a stock advantage at all depends on the regime in place.

The rolls data in Table 1 provide one more instance of how the new historical record explodes comfortable certainties encouraged by Siegel and ratified in orthodox financial thinking. Next I want to probe the most comforting belief of all: that over longer periods, stock investors have always done well.

How bad can it get—and for how long?

In one sense the international data to be discussed in a later section has already answered this question: it is possible for stock investors to lose *everything*, as did Russian investors in 1917 and Chinese investors in 1949. But most of the bottom dwellers internationally are found in nations that either lost a major war or suffered overthrow by revolution. The market nadirs in thriving countries that suffered neither fate, such as the United States since 1793, have seldom been explored. It is not a subject to which Siegel gives much space. Rather, he chooses to focus on how quickly the buy-and-hold investor who reinvested dividends recovered from even the Crash of 1929.[15]

A second purpose of this section is to reexamine the Crash of 1929 through the lens of the new historical record. In the modern era it stands alone as by far the worst decline for the US stock market, with a peak to trough plunge, measured monthly, of almost 80% in terms of real total return from the end of August 1929 to June of 1932.[16] But where does it stand in a two-century perspective? In a 20th century *world* perspective, however devastating it might have been initially, recovery from the Crash of 1929 was astonishingly swift: measured again in terms of real total return, US stock investors had got back to about even by the end of the war, a peak-to-peak span of not much more than fifteen years.[17] By contrast, Japanese investors who bought at the top in December 1989 are nowhere close to recovering their starting wealth, even now, after more than three decades have passed.[18]

An additional metric will be introduced here: doldrums and their duration. It is not necessary to book a negative return over a lengthy interval for stock investors to be sorely disappointed—a prolonged period of returns close to zero, or so minimal as to be on a par with the return on short-term bills, would be hard to

endure for an investor educated by Siegel to expect far more from stocks.

Worst case scenarios for fixed rolls

Figure 7 charts the real total return on stocks for twenty-, thirty- and fifty-year rolls; Table 2 gives a rank order of the worst cases for each. On first glance, Figure 7 is reassuring: even with the new history in hand, I could not find a negative twenty-year return,[19] much less a negative thirty- or fifty-year return. And Table 2 confirms that the worst thirty-year return was stronger than the worst twenty-year return, while the worst fifty-year return was better still and not that bad in absolute terms, at just under 4%. It appears that longer holding periods do tend to raise the floor for stock returns, at least in the US thus far; but the international returns described subsequently will indicate that these US results may be adventitious.

On second glance, Figure 7 tells a darker story. Fifty years, after all, is longer than the accumulation period for most retirement savers—cold comfort, if fifty years is the minimum for always earning a decent return in stocks. Over the more commonly achieved holding period of twenty years, there have been repeated instances where the real total return has been barely one percent. The span from 1929 to 1949 was the worst, but only by a hair; other twenty-year intervals with similarly low returns have recurred throughout the two centuries examined. Likewise, the thirty- and fifty-year returns booked at the depths of the Great Depression also had analogues scattered through the centuries.

Consider next the 1929 peak and 1933 trough. Neither emerges as a best or a worst case on the time scales used in Figure 7. Across all three holding periods, 1929 emerges as only one peak among others of similar magnitude, even as 1933 shapes up as but one notable trough among several. The implication: when viewed over longer time scales, investor experience for multi-decade holding periods terminating in the 1930s was not unique. Accordingly, there is no reason to see the Crash of 1929 as a one-off; a decline of similar magnitude, perhaps less rapid and / or configured differently, could recur. Stocks have disappointed investors again and again. Going forward, there is no reason to expect otherwise.

Worst most prolonged doldrums

I define a doldrum as an interval longer than ten years where the real total return on stocks comes in below +1.0% annualized. A perusal of

Figure 7
Best- and Worst-Case Scenarios for Stocks over 20-, 30-, and 50-year Holding Periods

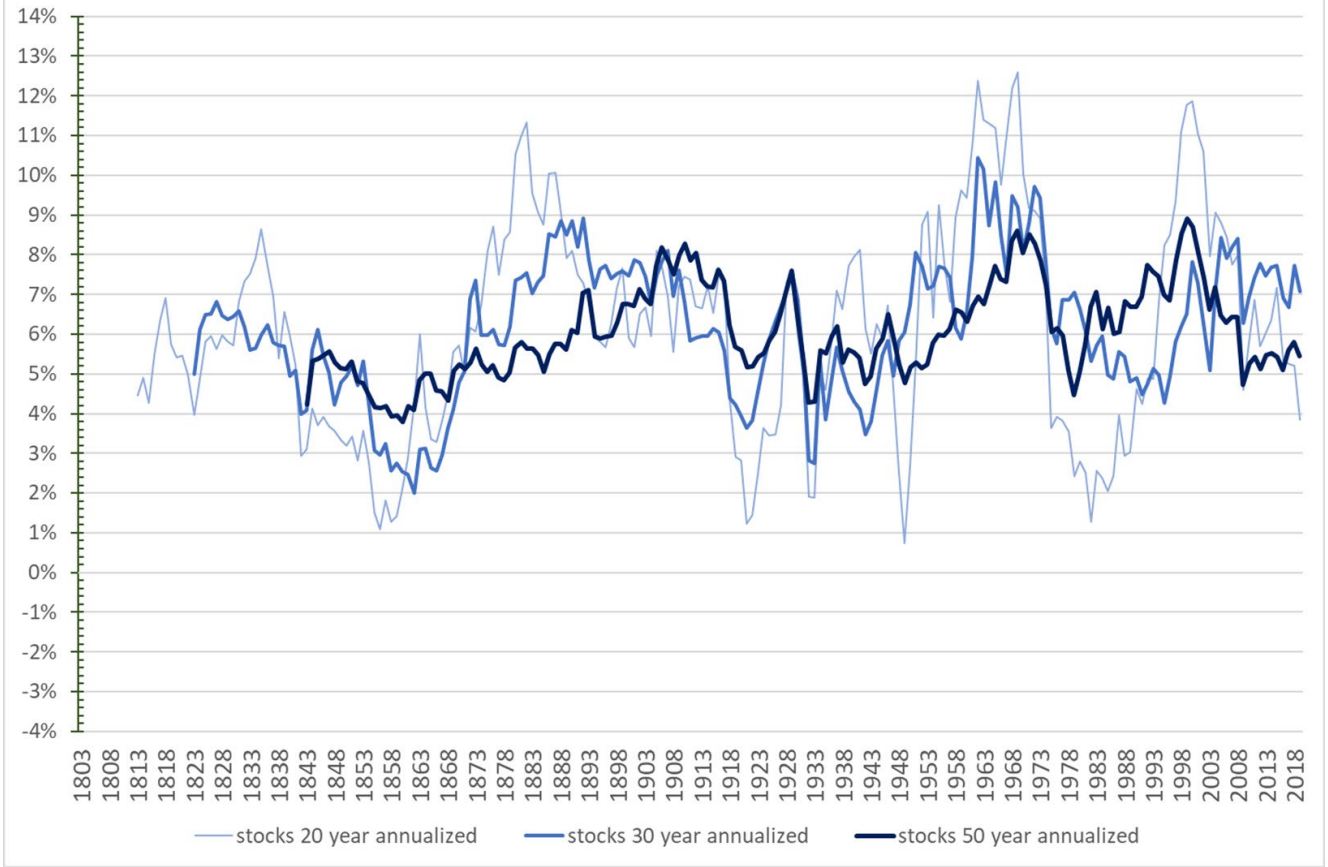


Table 2
How Bad Can It Get for Stocks?
Worst Rates of Return Over Lengthy Holding Periods

20 years		30 years		50 years	
Endpoint	Annualized return	Endpoint	Annualized return	Endpoint	Annualized return
1949	+0.75%	1862	+2.00%	1860	+3.80%
1856	+1.09%[a]	1933	+2.75%	1932	+4.29%
1921	+1.22%	1921	+3.64%	1979	+4.48%
1982	+1.29%	1995	+4.27%[b]	2009	+4.72%

Note. These are real total returns measured January to January and ending in the stated year. These data are for the US; instances of much lower returns for each of these holding periods, for markets outside the US, can be found in Table 5.

[a] There were other similarly low returns in the later 1850s, but I consider these to be part of a single complex, so only the worst one is given here.

[b] Rolls ending in 1935 and 1943 have a lower return, but I consider these to be part of the 1933 complex and have omitted them here in favor of the next lowest roll ending in 1995.

Figure 3 will help the reader pick out candidate stretches; specific dates and numbers are in Table 3. In terms of length the twenty years following 1929 do represent one of the worst doldrums, but the very worst case by that metric was the twenty-two-year interval following 1836. Other notable doldrums include the nineteen years preceding 1921 and the sixteen years preceding 1982. The most recent doldrum began in January 2000 while the first one to show up in the record began in 1803. In short: US stocks have recurrently gone nowhere over periods averaging somewhat less than twenty years.

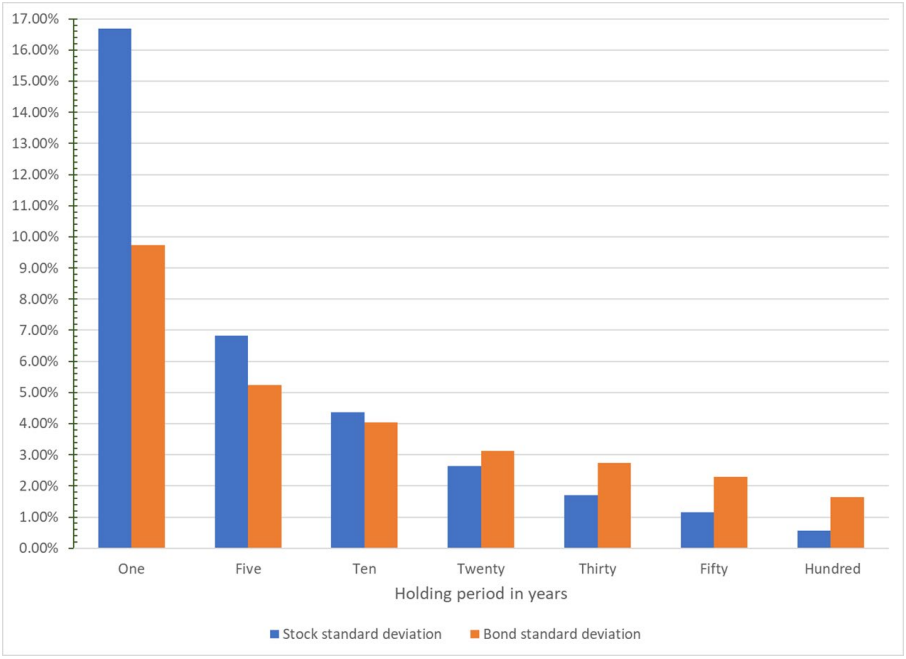
Table 3
Long Doldrums for Stocks

Period (in chronological order)	Length in years	Annualized real return
1803 - 1815	12	0.51%
1836 - 1858	22	0.95%
1903 - 1921	18	0.09%
1929 - 1949	20	0.75%
1966 - 1982	16	-0.63%
2000 - 2013	13	0.77%

Which one is the risk asset?

Next, Figure 8 harvests one more insight from the rolls data that will later help to underwrite the challenge posed by the regime perspective. The chart shows the standard deviation of the annualized returns making up the stock and bond rolls for each holding period.[20] It is not surprising to find that returns over longer holding periods exhibit lower standard deviations. However, there is an artifactual component to that decline in volatility, best illustrated by the hundred-year rolls.[21] Each of these overlaps by 99% with the prior and succeeding rolls. Only a truly discrepant annual return could budge the returns on adjacent hundred-year rolls by much. Accordingly, the decline in the absolute level of the standard deviations in Figure 8 is not meaningful. What matters is how the relative size of the stock and bond standard deviations changes as roll length increases. Specifically, stock volatility falls more rapidly. Consistent with conventional expectation, at the annual level the volatility of stocks has been much higher than that of bonds. But for holding periods

Figure 8
How the Volatility of Stock and Bond Performance Changes as Holding Period Lengthens



twenty years or longer, *bonds have been more volatile than stocks*.[22] That risk reversal is sustained and accentuated as the holding period lengthens further.

Figure 8 provides a useful preface to the examination of worst-case rolls data for bonds. These are given in Figure 9, which parallels Figure 7, now showing twenty-, thirty-, and fifty-year rolls for bonds. The investor unfamiliar with statistical formulae may find a side-by-side comparison puzzling: the stock rolls in Figure 7 appear more bumpy and jagged than the bond rolls in Figure 9, with more zigs and zags; how can bonds have the higher standard deviation, per Figure 8? Answer: the calculation of standard deviation involves squaring differences.[23] Extreme departures from the mean count exponentially more in that calculation; zigs and zags in the vicinity of the mean do not count for as much. To focus on the jaggedness of the stock returns in Figure 7 is to fall prey to yet another misleading visual heuristic—for it is not the jaggedness, but the magnitude of the peaks and troughs which distinguish the bond rolls as the more volatile.[24] When compared to stocks, over longer holding periods bond investors are more likely to experience outcomes that deviate farther from the mean. And as Table 4 reveals, it is specifically downside risk that bond investors incur. As holding period lengthens, negative extremes on bonds come to exceed negative extremes on stocks to a greater and greater degree, even as positive departures converge to the same magnitude as for stocks. Bonds are risky, and as holding period lengthens bonds get riskier as compared to stocks.

In other respects, Figure 9 holds few surprises. The worst twenty-, thirty-, and fifty-year bond returns all come after World War II and prior to the bull market in bonds that began in 1982. Unfortunately, these are the bond returns that dominated SBBI yearbooks in the 1980s, which made investors that much more receptive to Siegel's thesis when he introduced it in the early 1990s. Before the war, US bond investors could have reassured themselves the way US stock investors can still do: that there had never been a negative return on bonds over a thirty- or fifty-year holding period. The worst fifty-year return before the war, of 3.78% real annualized, was not bad in absolute terms, and not much different from the worst fifty-year return for stocks (Table 2). Then things changed: what had never happened before began to occur repeatedly.

Rolls summary

Received wisdom among financial analysts can be laid out as follows: 1) investors demand to be compensated for bearing extra risk; 2) stocks are riskier than bonds; 3) therefore, when computed over long intervals, stock returns are expected to be higher than

Figure 9
Best- and Worst-Case Scenarios for Bonds over 20-, 30-, and 50-year Holding Periods

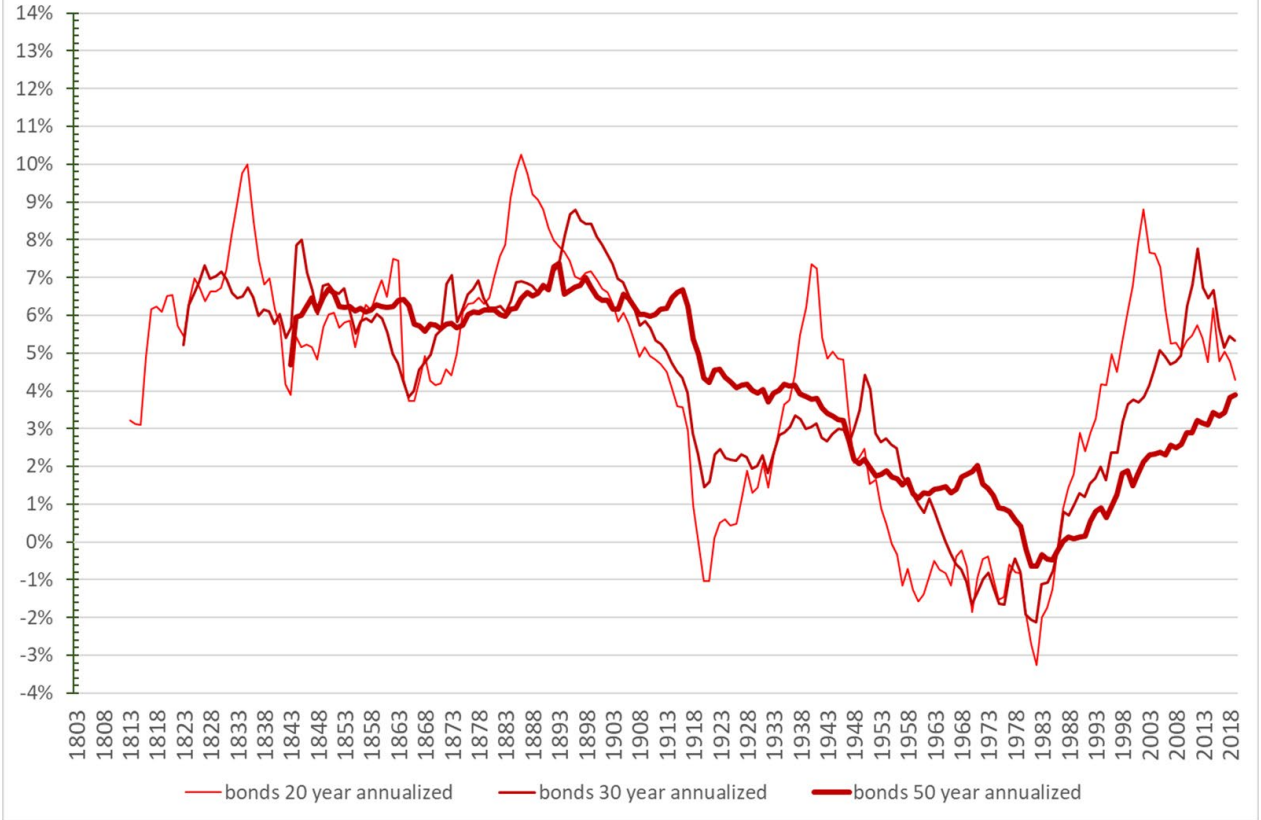


Table 4
How Bonds Become Riskier as Holding Period Lengthens

	Holding period in years							
	Ten		*Twenty*		*Thirty*		*Fifty*	
	Stocks	*Bonds*	*Stocks*	*Bonds*	*Stocks*	*Bonds*	*Stocks*	*Bonds*
Mean Return	6.18%	4.33%	6.13%	4.22%	6.15%	4.17%	6.06%	3.94%
Positive extreme deviation	11.31%	9.55%	6.45%	6.03%	4.29%	4.62%	2.85%	3.42%
Negative extreme deviation	-10.08%	-9.74%	-5.39%	-7.48%	-4.14%	-6.30%	-2.26%	-4.58%

Note. Each number is a percent. Mean is the arithmetic mean of the annualized returns over all rolls of the stated length. Deviation is with respect to the mean; thus, the ten-year positive stock deviation of 11.31%, the maximum found in the record, indicates that the best ten-year roll for stocks had an annualized return of 17.49%. Note the pattern: as holding period lengthens, the magnitude of the extremes first converges for stocks and bonds, then reverses, with the magnitude of the negative bond extremes diverging to a greater and greater degree when expressed as a percent of the stock extremes.

bond returns. But if stocks are consistently *less* risky than bonds over longer holding periods, then why should stocks consistently return *more* than bonds over these long intervals? I will return to this conundrum after bolstering my critique of Siegel's thesis by introducing recent data on international returns.

The evolving international record

In book editions appearing after 2002, Siegel claimed to find out-of-sample support for his thesis in the international data record, drawing on the work of Elroy Dimson, Paul Marsh and Mike Staunton in their book *Triumph of the Optimists*. That volume gave returns for multiple markets outside the US from 1900 through 2000, and indeed, reported overall conclusions consonant with Stocks for the Long Run.

Unfortunately, data collection for *Triumph* had concluded just at the peak of one of the great bull markets, a period when stock markets worldwide hit new records. Fortunately, the authors have begun to produce annual updates in the form of a yearbook published by Credit Suisse.[25] Beyond updates, the team improved the database in two ways. First, they now calculate a return for the World *ex-USA*; the "World" return in *Triumph* had encompassed the US return with about a 50% weighting. Second, they have added more countries, including international markets analogous to the 2nd Bank of the United States—the losers who had been left out of the earlier compilation.

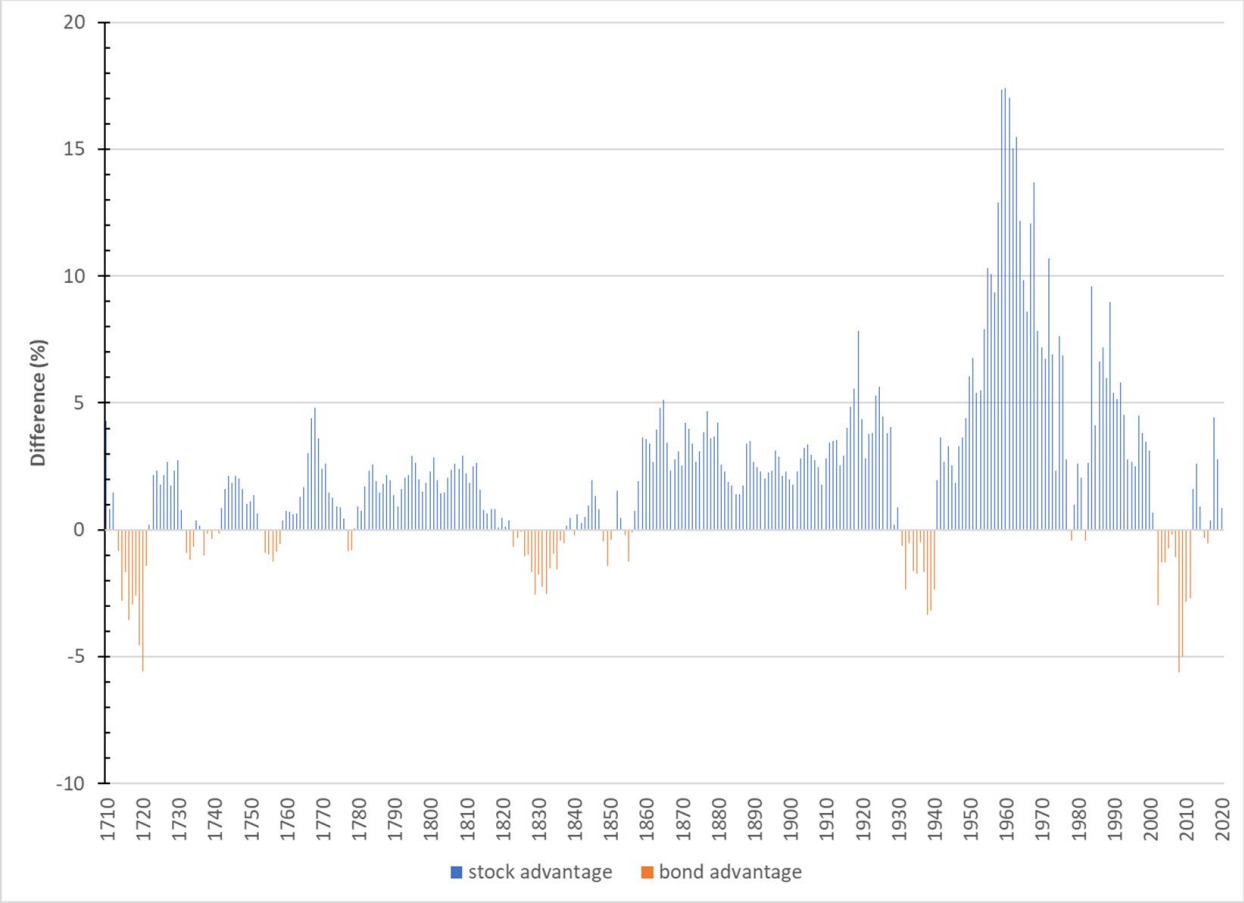
With the bear markets following 2000 and 2007 in hand and with survivorship bias reduced, the 120-year annualized real return, on stock markets outside the US, is now estimated by the Dimson team as 4.4%—not the 6.6% estimated by Siegel as the true long-term expected value.[26] That difference doesn't sound like much but as the 2020 Credit Suisse yearbook notes: "A dollar invested in US equities in 1900 resulted in a terminal value of USD 1937 … An equivalent investment in stocks from the rest of the world gave a terminal value of USD 179 … less than a tenth of the US value."[27]

The vanishing equity premium

Next, the 2020 Credit Suisse yearbook calls out returns over the trailing fifty years. These more recent data address a concern sometimes expressed about long run historical data, for example, that returns from the antebellum era are just too far past to be relevant. For the World ex-USA, for 1970 through 2019, the authors now report *parity performance* for equities versus government bonds, at 5.1% and 5.0% annualized. Outside the US, the equity premium has been just that small in recent decades.

At the country level the yearbook format makes it easy to find instances where the equity premium was a *deficit* across a multi-decade period. Japan since December 1989 is a well-known counterexample to Siegel's thesis.[28] And indeed, over the past three decades the equity deficit in Japan came to a stunning -5.4% annualized. But there is no need to cherry-pick the exact ending of the great 1980s boom in Japan to make the equity premium there disappear. From 1960 to the present the equity premium in Japan has been precisely … 0.0%. Investors

Figure 10
The Equity Premium in Great Britain from 1700



Note. Compare to Figure 5. Ten-year rolls are used; the value charted is the annualized real return on stocks minus the annualized real return on bonds. Data from GFD, charting by the author.

fared even worse in Austria, France and Italy over the past six decades, with the equity deficits in those markets equal to -0.6%, -0.6%, and -1.9% annualized, respectively.

Next, Global Financial Data has assembled stock and bond returns for Great Britain extending back to 1700. This dataset allows an examination of the equity premium across three centuries, the longest span I could locate. Figure 10 for Great Britain is structured the same as Figure 5 for the US, showing the equity premium over ten-year rolls. Figure 10 confirms multiple elements seen in the US chart: 1) the recurrence of regimes where bonds outperform stocks, in the 18th century as well as the 19th and 20th centuries; 2) the unusual character of 20th century returns following WW II, where stock returns in GB, as in the US, briefly soared far above bond returns in a manner not seen before (or since); and 3) the magnitude of the equity deficit, when present.

But figure 10 is also important for what it does *not* show. Prior to WW II, registration is poor relative to the US results in Figure 5—some peaks and troughs are aligned, others not. This suggests that regime change need not be driven by external, global economic factors. The two figures in conjunction also place a check on an inference that might otherwise be pushed based on Schmelzing (2020). He showed that real interest rates have been declining for centuries. Equivalently, a great bull market in bonds has been ongoing for centuries, and has only recently flattened out. Therefore, had negative equity premia been most prominent in the 18th century results for Britain, and minimal thereafter, just as, until recently, US negative premia had appeared most prominent in the 19th century, rather than the 20th century, then one might have reasoned from Schmelzing's data that negative equity premia are a relic of an earlier phase of economic development, concluded by 1800 for Britain and by 1900 for the US; an artifact, that is, of the maturation from an emerging to a developed market. But that is not what Figure 10 and Figure 5 show. Negative equity premia recur, century after century, country after country.

Finally, Figure 11 is constructed parallel to Figure 6, showing ten-year rolls for the underlying stock and

bond returns for Great Britain superimposed over the equity premium. Figure 11 reinforces the point that there are multiple permutations and combinations that can produce an equity deficit: e.g., very strong bond returns, as around 1718, or conversely, relatively weak stock returns, as around 1830. Figure 11 also confirms a key US finding: the combination of very strong stock returns and negative bond returns, seen in the decades after the war, had never previously occurred across three centuries of British financial market history. In the US, those postwar decades have dominated the SBBI, and the truncated history found in the SBBI has misled a generation of US investors about the normal, customary, and usual relationship between stock and bond returns.

In short: contrary to Siegel, once the historical lens is widened to include the 19th century in the US or swung out to include markets outside the US, it is straightforward to find instances of "bonds for the long run."

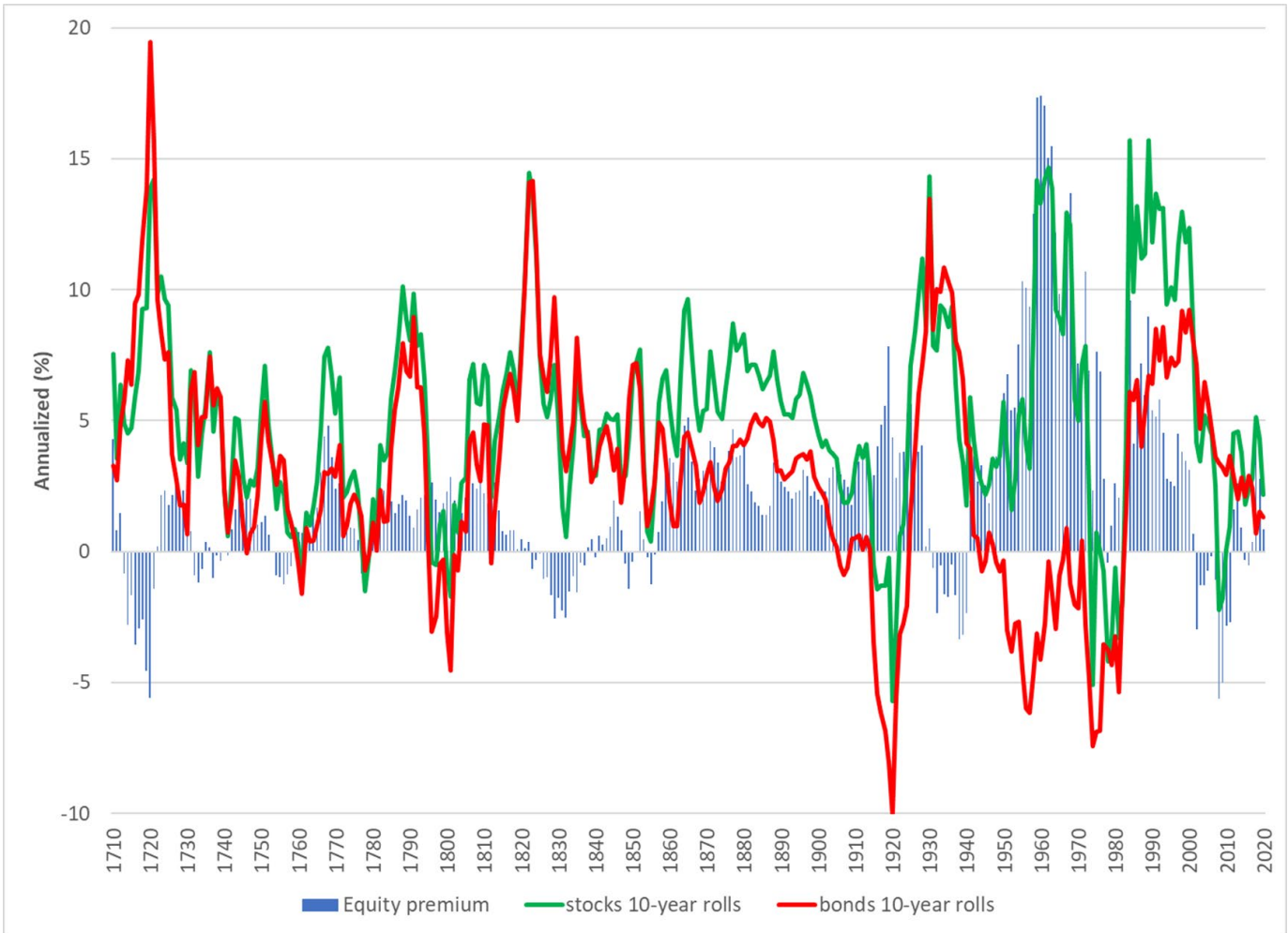
Worst case stock returns ex-USA

Independent of the Dimson team, Bryan Taylor at Global Financial Data has also compiled international returns, which in many cases extend back well beyond the 1900 cut-off used in the Credit Suisse yearbooks. Annual data obtained from GFD will help to put the US results in Tables 2 and 3 in context.[29]

For countries in the GFD database I identified the worst 20-, 30-, and 50-year real stock returns. These are compiled alphabetically in Table 5. My US results from Table 2 are included to provide a reference point.

In contrast to the US, most of the markets in Table 5 *did* experience a negative twenty-year real return; more than half experienced a negative thirty-year return; and one third had experienced a negative fifty-year return. Clearly, holding periods of this length are insufficient to guarantee a

Figure 11
Permutations and Combinations of Stock and Bond Returns Observed in Great Britain



Note. Compare to Figure 6. These are the ten-year rolls for stocks and bonds that underlie the previous chart, with the equity premium (deficit) reproduced in monochrome here. Data from GFD, charting by the author.

Table 5
Worst Case Scenarios for Stocks Outside the United States

	Holding period					
	20 years	**End:**	**30 years**	**End:**	**50 years**	**End:**
A. Reference						
United States	0.75	1949	2.0	1862	3.8	1860
B. Countries						
Australia	2.0	1990	4.0	1980	4.8	2020
Austria	-11.9	1930	-7.94	1953	-4.8	1950
Belgium*	-6.21	1948	-2.5	1940	-1.32	1948
Canada	1.7	1920	3.25	1932	4.30	1921
Denmark*	-4.82	1925	-2.32	1926	-0.14	1948
Finland*	-1.8	1930	-0.69	1946	0.85	1966
France*	-6.58	1949	-1.87	1949	-1.04	1950
Germany	-6.94	1922	-2.9	1950	-3.4	1950
Ireland	-6.60	1826	-2.48	1834	0.18	1854
Italy	-7.34	1979	-2.35	1991	-1.41	1978
Japan	-11.72	1947	-7.72	1947	-1.63	1950
Netherlands*	-3.19	1931	-0.01	1931	1.87	1952
New Zealand	0.43	1990	2.34	1990	4.1	1980
Norway*	0.2	1930	1.5	1970	2.5	1960
Portugal	-3.2	1980	1.0	1930	2.5	2020
South Africa	-1.40	1920	0.51	1920	2.81	1921
Spain*	-4.3	1950	-1.07	1940	-1.6	1980
Sweden	-5.17	1932	-1.10	1932	1.61	1948
Switzerland	-4.39	1981	-0.78	1991	2.2	1950
UK	-1.27	1920	0.61	1920	2.79	1781
C. Summary						
World ex-USA	-3.25	1920	-0.64	1920	1.13	1949
Median	-4.34		-0.92		1.23	

Note. Table uses the lower of the Credit Suisse 2020 yearbook decade returns (selection by the author) or rolls based on the GFD annual series provided to the author by Bryan Taylor. Single decimal place returns for years ending in zero are from the yearbook; all others from GFD. All values are real total returns expressed as an annualized percentage. See also Figure 12 for a visual rank order of these returns. Asterisks mark invaded countries that are colored orange in Figure 12, and italicization indicates countries defeated in war that are colored red in Figure 12. The worst 20 and 30 year returns for Italy did not include the war years, hence there are only three red dots in Figure 12 for those holding periods.

positive real total return on an investment in stocks; the US experience does not generalize. It turns out that US investors have been comparatively lucky—the worst stretches of market history in the US have been rather more favorable than the worst cases in most of the rest of the world.

Next, Figure 12 provides a visual representation of the rank order of the worst-case outcomes in Table 5. Entries are color-coded to assist interpretation. Specifically, red dots show countries defeated in war during the period when the worst-case return was incurred; orange dots show returns for countries invaded during the 20th century or that suffered a civil war (Portugal, Spain); blue dots show former colonies of Great Britain, including the US; all others are in black.

The color coding shows a clear stacking pattern: with rare exceptions, all the worst returns were recorded when a country was defeated in war. To some extent, markets in defeated nations represent cheap shots with respect to assessing Stocks for the Long Run, except insofar as they offer a salutary reminder that when a nation loses a war, that nation's investors lose their shirts. The next lowest returns were recorded for countries whose territory was invaded or occupied, reminding again that stocks have no magical power to overcome national disaster. Finally, all the highest minima have been recorded in former colonies of Great Britain; these were the only markets where the worst case was more favorable than the US worst case. As can be seen from the scatter of black dots, the favorable experience of stock investors in British ex-colonies has not generalized to other countries, even when these were fortunate enough not to be defeated or invaded.

Next, the curved line in Figure 12 shows the return that would produce a 50% drop in real wealth over the period. As can be seen, to suffer this great a loss from a buy-and-hold investment has been a common

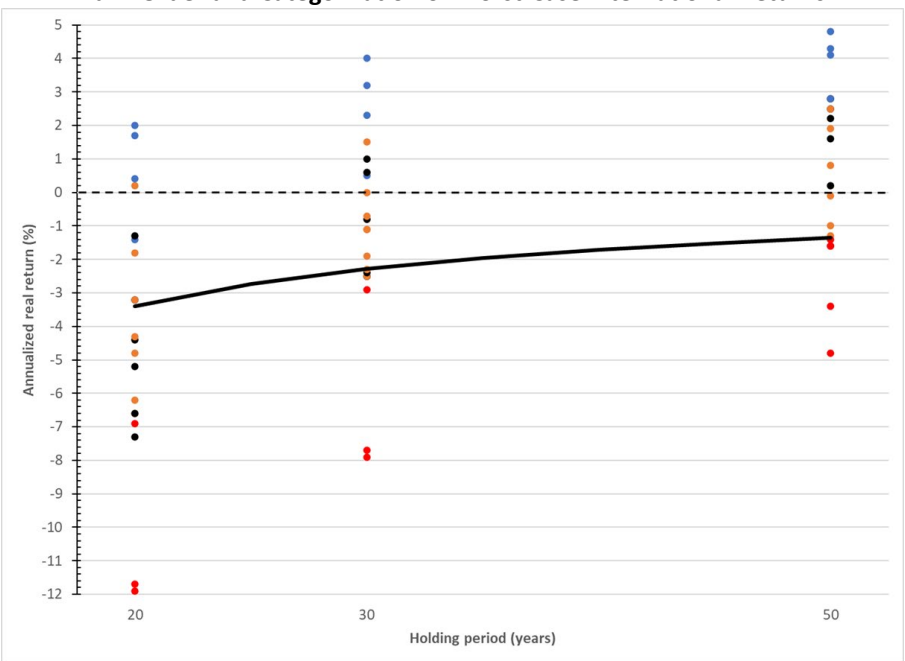
occurrence at the twenty-year mark for international stock investors, has not been unusual over thirty years, but requires defeat or invasion to occur when the holding period stretches out to fifty years. In this regard, Figure 12 tends to confirm the US finding, that stocks pose less risk of minimal or negative returns as the holding period lengthens. This pattern also holds for the GFD World ex-USA index, a capitalization-weighted amalgam of international markets (bottom of Table 5).[30] On the other hand, twenty years emerges as nowhere near long enough to achieve a reliable reduction in the risk of loss from holding stocks.

Reexamining Table 5, it is surprising how recent many of the dates are. Likewise, although GFD has returns extending into the 19th century for most European markets, most of the worst periods in Table 5 do *not* terminate before 1920. Accordingly, the mediocre 19th century US results in Table 2 ought not to be dismissed as too old to be relevant; even worse results have been experienced by international investors across quite recent decades. There was nothing special about the 19th century that drove US results lower; for most international markets, the 20th century was worse.

Returning to the Credit Suisse yearbooks, although the decade tabulations there do not lend themselves to a doldrum analysis parallel to Table 3, the Dimson et al. team has conducted a similar analysis using a stricter standard: they identified the longest span where a nation's stock market return stayed *negative*.[31] After excluding the losers of World War I and II, there were seven countries where returns stayed negative over a span ranging from twenty to forty years, and five more where the doldrums lasted for between forty and sixty years. Thus, twelve out of sixteen qualifying countries experienced stock market doldrums that lasted for longer than anything seen in the US. On a global view, stocks remain risky even when the holding period stretches out over multiple decades.

Summing up, the larger sample of markets outside the US compiled in the Credit Suisse yearbooks and by Global Financial Data shows diverse outcomes for stocks and bonds. In some periods in some markets for some length of time, stocks have enjoyed strong returns equal to those estimated by Siegel. In other periods in other markets over other intervals stocks have performed weakly and/or bonds have outperformed stocks. World outcomes and 19th century US outcomes expose a lack of uniformity that undermines the hypothesis of Stocks for the Long Run.

Figure 12
Rank Order and Categorization of Worst-Case International Returns



Note. Values are from Table 5. See text for an explanation of the color coding. Curved line shows the annualized negative return that would produce a 50% decline in wealth at the end of the period. US and World ex-USA values not charted; see Table 5 for their values and for median returns.

The regime thesis

In a more typical financial history, greatest weight would be placed on values computed over the longest interval available, on the theory that sampling over time follows the same logic as sampling people from groups. The argument: longer time samples provide more precise estimates of the true expected return on an asset, just as a larger sample of people would produce a better estimate of any differences in height or weight across two groups.

The regime thesis denies that longer time samples are like larger samples of people. There is no population of asset returns existing outside of time from which larger samples can be drawn by lengthening the series. There are only the asset returns that have been recorded in history thus far. These do not predict future asset returns because their pattern is specific to the regime that prevailed at the time. No analysis of US bond returns from 1793 to 1942 could have predicted the bond returns seen following the war—such a bond abyss had never occurred before. And no extrapolation of the stock and bond record from 1926 to 1982 could have

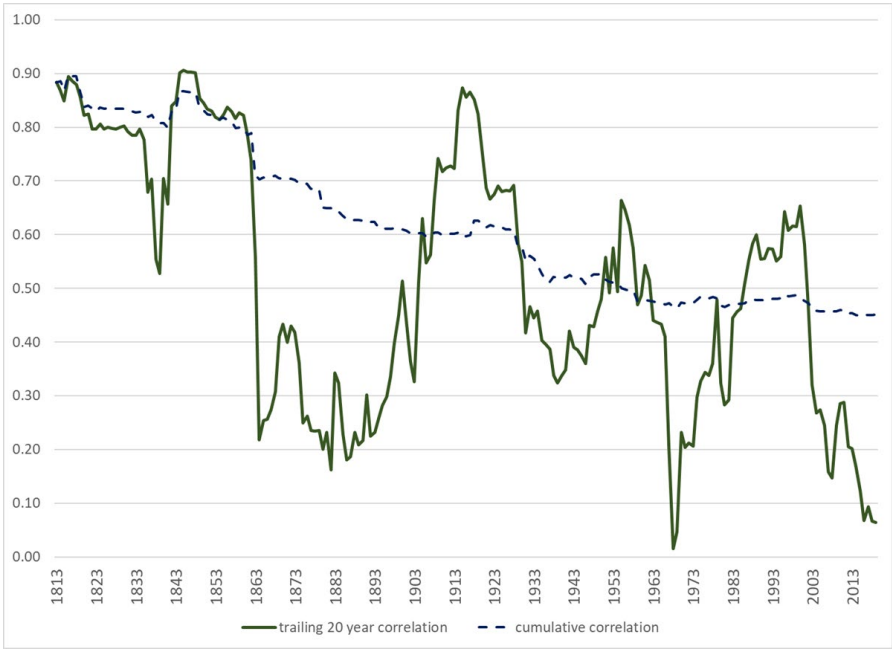
predicted the parity performance seen after 1982.

A regime is a temporary pattern of asset returns that persists for an arbitrary interval that may be decades in length. The pattern may implicate multiple assets, as with stocks and bonds, or a single asset, as with the contribution of dividends to stock returns discussed subsequently. I use the encompassing term 'pattern' to drive home the point that regimes need not be toggles, nor cycles, nor trendlines. The extremity of the out- or under-performance of an asset also varies by regime. The regime thesis entails no assertion of any regular periodicity nor any necessary reversion. Rather, it expects to find ceaseless variation, as the various permutations and combinations of individual asset regimes play out, as seen in Figure 6.

More examples of such variation can be found by examining two quantities that play an important role in asset allocation and portfolio management: the correlation between stocks and bonds, and their respective standard deviations. Correlation, or rather the lack of it, indicates the potential for diversification when combining assets. Because they are supposed to be such different assets, analysts expect the stock-bond correlation to be low; and indeed, over the trailing nine decades the 2020 SBBI has it as 0.23 for corporate bonds and 0.09 for long government bonds. But as with so many other quantities, those values, measured over the arbitrary interval that I call the modern era, prove a poor guide to what a longer historical record might reveal.

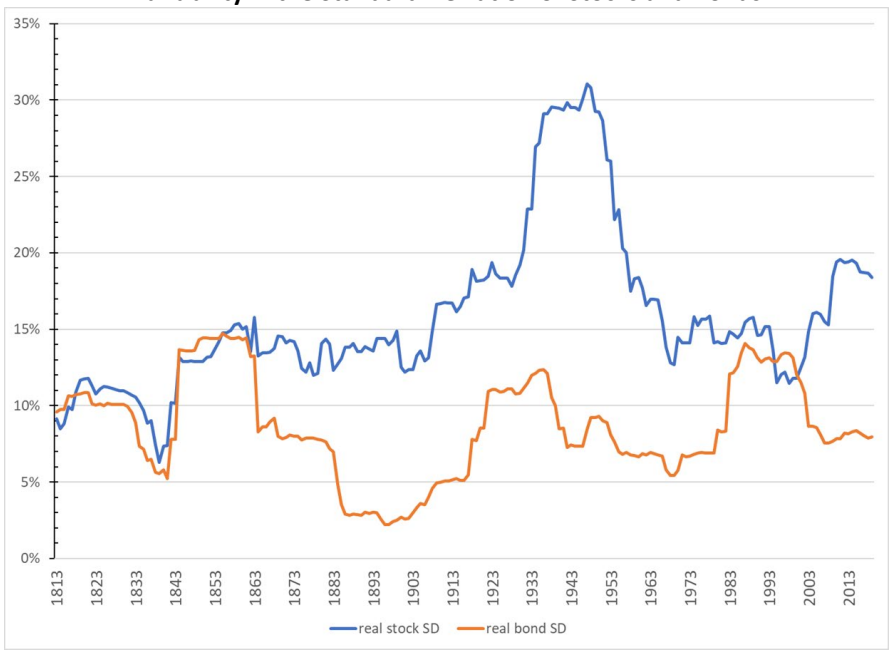
Figure 13 charts the correlation over twenty-year rolls. Figure 13 shows first, that the correlation has been highly variable over twenty-year intervals, ranging from about 0.00 to 0.90. Second, sizable fluctuations have occurred within the modern era as well as beforehand. However, correlations in the modern era did reach low points never seen before 1926. The overall correlation for the modern era is 0.335, significantly lower than the .609 for the preceding one hundred thirty-three years ($p < .01$).[32] In short: correlations in the modern era give a misleading impression when viewed in isolation. The diversification

Figure 13
How the Correlation Between Stocks and Bonds Has Fluctuated



Note. Each value represents the correlation of annual returns over the trailing twenty years. The cumulative correlation uses all the returns from 1794 to the date on the x-axis.

Figure 14
Variability in the Standard Deviation of Stocks and Bonds



Note. Each value is computed on the trailing twenty years of annual returns.

potential from adding bonds to a stock portfolio appears to have been lower in prior decades.

It is also sobering to note the two points where the stock-bond correlation dropped toward the strongly diversifying level of zero. The first instance represents the twenty years to 1968, near the end of the bond abyss, decades where stocks had consistently gone up and bonds had consistently gone nowhere. The second occurred recently, after stocks had soared in the dotcom boom, plunged, recovered, plunged again in the financial crisis, and recovered to soar once more, even as bonds trundled along throughout. The latter case shows the hoped-for diversification benefit from adding bonds to a stock portfolio, i.e., a steadying influence without a substantial reduction in return. The former case assuredly is not what is expected from diversification—decade after decade of depressed returns achieved by adding uncorrelated but ever-tanking bonds to ever-soaring stocks. In short: stock-bond correlations have fluctuated through a wide range, and the implications of a low (high) correlation have also varied. Mean reversion has not occurred.

Variability has also characterized the standard deviation of returns. And that variability too has been hidden by the truncated modern record. Since at least the early 1990s when Siegel wrote, the SBBI has estimated the annualized standard deviation of stocks to be about 20% and of long bonds to be about 10%. These unchanging estimates suggest a lack of variability in this important metric over time—at least, after markets settled down following the Depression and the war. But the hypothesis that stock volatility has not been particularly volatile needs to be tested over a longer time frame.

Figure 14 shows rolling twenty-year standard deviations on stocks and bonds. Taking the stock line first, in this case the augmented history reveals that the period around 1929 was indeed without precedent. Stocks had never had a standard deviation as high as 20% before rolls that included 1929, and have never exceeded that level since, once rolls containing the last convulsion of the 1930s in 1937-38 rolled off. However, that level was approached in rolls that included both the post-2000 and post-2007 declines. The gist of Figure 14 is that outside the Great Crash and its aftermath, 20% has been closer to the *ceiling* on the standard deviation of stocks than to its average. Furthermore, it is difficult to describe the stock line as mean reverting. The levels seen around the Great Crash had no precedent, and the very low levels seen before the Panic of 1837 did not recur. The volatility of stocks has itself been volatile.

The bond line better exhibits mean reversion, repeatedly rising and falling through a tighter range in the vicinity of 10%. However, because the variability of stock market variability has been so great, and because bond variability has also fluctuated through a wide range, from time to time the historical record shows the bond standard deviation to be on a par with, or even larger than, the stock market deviation over the same trailing twenty years. This reinforces a key implication of the new historical record: that both stocks and long bonds have to be approached as risk assets.

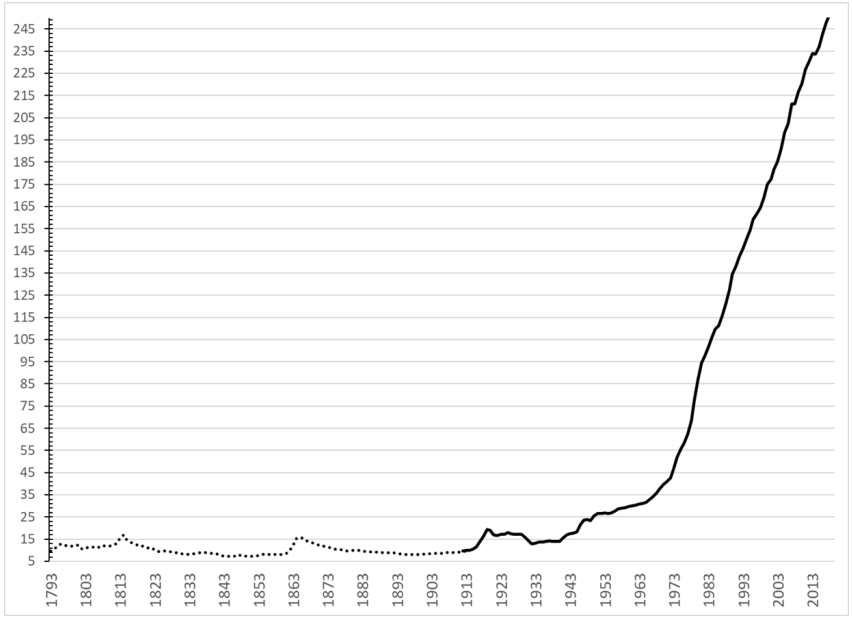
But not a random walk

Although the regime perspective expects many financial quantities to be highly variable over time, the regime thesis should not be mistaken for a random walk. Asset pricing reflects macroeconomic conditions in a way that makes sense—after the fact. But there are so many potentially relevant macroeconomic factors, which can combine in such diverse ways, and change so rapidly, that predictions of which regime will hold over ensuing decades are unlikely to succeed.

Here is an illustration of how the regime thesis differs from the hypothesis that asset returns follow a random walk. One of the most familiar examples of a change of regime is the course of consumer prices in the United States.[33] Siegel expresses the

Figure 15
How to Lie with Figures: A Misleading versus a Rectified Account of the Course of Inflation

A: Designed to mislead the viewer about the scale of 19th century inflation



conventional wisdom when he states: "the price level at the beginning of World War II was essentially the same as it was 150 years earlier." He also gives a common explanation for the one-time switch seen following the war, in which prices began to rise steadily and then at an increasing rate through 1980: Siegel points to the lapse of the gold standard.[34]

Investors often take this conventional wisdom to mean that there was no inflation while the gold standard held.[35] Figure 15a manipulates visual heuristics to compel the false belief that there was negligible price inflation in the US before World War II. It exploits the "ignore the lower left corner" heuristic. Figure 15a also exploits the inability of ordinary humans to adjust for scale effects and to translate linear representations into their logarithmic equivalents unless given a crutch, i.e., plotting on a logarithmic vertical axis. Given only a linear scale, there is no way to visually interpret Figure 15a except as confirmation that inflation had never been a problem in the US before the middle of the 20th century.[36]

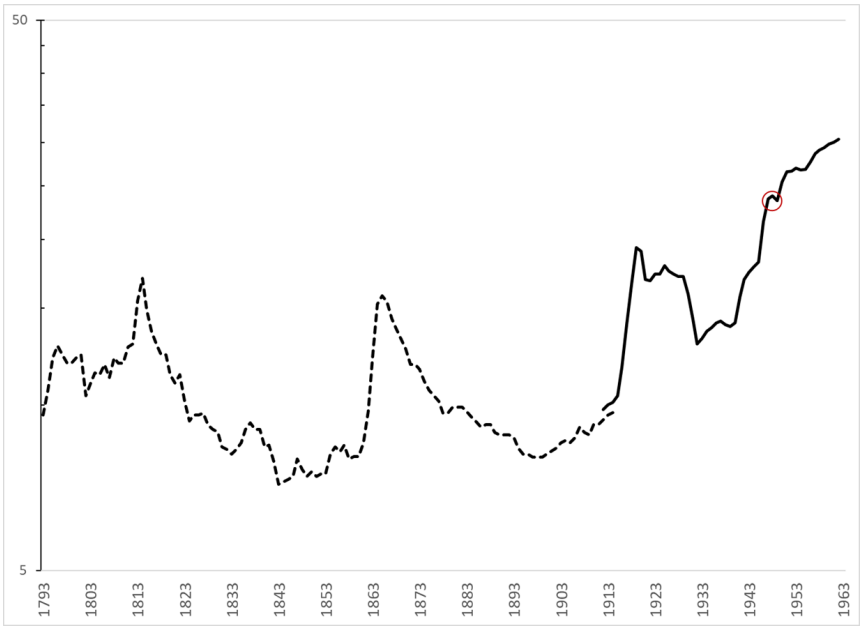
Figure 15b was constructed to overcome these heuristics. It tells a much more interesting story about change across inflationary regimes than just "paper money." Figure 15b shows that there was plenty of inflation before WW II. Prices doubled in the 1810s, again in the 1860s, and again in the late 1910s. In fact, when Milton Friedman wrote in the early 1950s, he characterized the price increase during and after WW II as *on a par with* those seen during the Civil War and WW I.[37] Prices doubled during wartime; it's just what they do. That seemed to be a law of nature.

In fact, what the 19th century had, that the post-war years have lacked, is sustained bouts of *deflation*. The law governing prices had two tenets: that prices will double over a few years during war time, and that prices will fall by half in the decades following the war. The pattern for the War of 1812 and the Civil War looks the same: prices doubled sharply and fell back slowly, doubled again and fell back again. And the first part of the pattern continued to hold for World War I. What changed in the 20th century was the breaking of the second tenet. Prices did *not* fall all the way back after World War I: the fall in prices was halted in 1933 before it had reached even a one-third decline. And the first part of the pattern held once again through World War II: prices had approximately doubled by 1948, the point where Friedman, not then knowing what was to come, marked the culmination of wartime inflation.[38]

The actual change in inflationary regime did not occur during the 1930s when the gold standard lapsed; rather, the change in regime became visible some years after World War II when, for the first time in US market history, deflation did *not* occur following wartime. In the US prices were never again to show a sustained drop of any magnitude. The occasional monthly fall in prices continued to occur; and rarely, an annual decline on the order of a single percentage point might intrude. But the inflationary regime that had held through the 19th century had changed in a specific way: deflation had been banished.

The point of this historical excursion has been to develop the complexity of regimes. Regime change need not take the form of a toggle between two positions, as in gold standard on / off. The regime that had prevailed up to the end of World War II was a complicated affair involving paired episodes of deflation and inflation driven by the exogenous factor of war.

B: Rescaled to make 19th century inflation visible



Note. The rescaled version uses the same data but applies a log scale and truncates the x-axis, which in turn reduces the scale required for the y-axis, both of which accentuate changes in price level. The inflationary surge of the 1960s and 1970s did not gather steam until after 1963. The circled point shows what Friedman (1952) believed to be the peak of World War II inflation. Dashed line is the Officer and Williamson (2020) price series, solid line is the Bureau of Labor Statistics CPI series. Gap represents a slight misregistration of the two circa 1913, which was not allowed to affect the annual price relatives used to deflate stock and bond returns.

After WW II, a more straightforward regime took hold for a time. The new law of fiat money was that prices always go up—the only question was by how many percentage points.

That may have changed with the financial crisis of 2008-09. Today one hears authorities express concern about the ongoing lack of inflation, despite the continued abeyance of the gold standard that supposedly was necessary to restrain prices. Has a new inflationary regime taken hold, or is the recent slowdown in price increases a temporary blip? The answer will only emerge decades hence.

The magnitude of regime variation

Next I report statistical tests designed to expose the futility of expecting longer time samples to improve precision of estimate. Table 6 presents means and standard deviations. I distinguish the first century from the second century, and then break out the twenty-six years thus far available for the third century of US market history. These initial comments focus on the arithmetic means; geometric (annualized or compounded) means are discussed later.

For the entire 226 years annual inflation averaged 1.58%. But this value is not much help in predicting either of the two centuries. In the first century to 1893, inflation averaged almost zero (0.07%), while in the second century to 1993 it averaged a more familiar 2.93%. The two centuries are significantly different from one another at $p < .001$.[39] Taking the 226-year mean as an errorless point, as would be appropriate for a true value, then the inflation rate in the first century falls almost three standard errors below that estimate, while the second century rate falls about two and a half standard errors above. Neither century's inflation rate predicts the other's, and their amalgam does not predict either. There has been no stable long-term rate of inflation whose value at the century level can be grasped by taking a longer sample. Rather, there were different regimes.

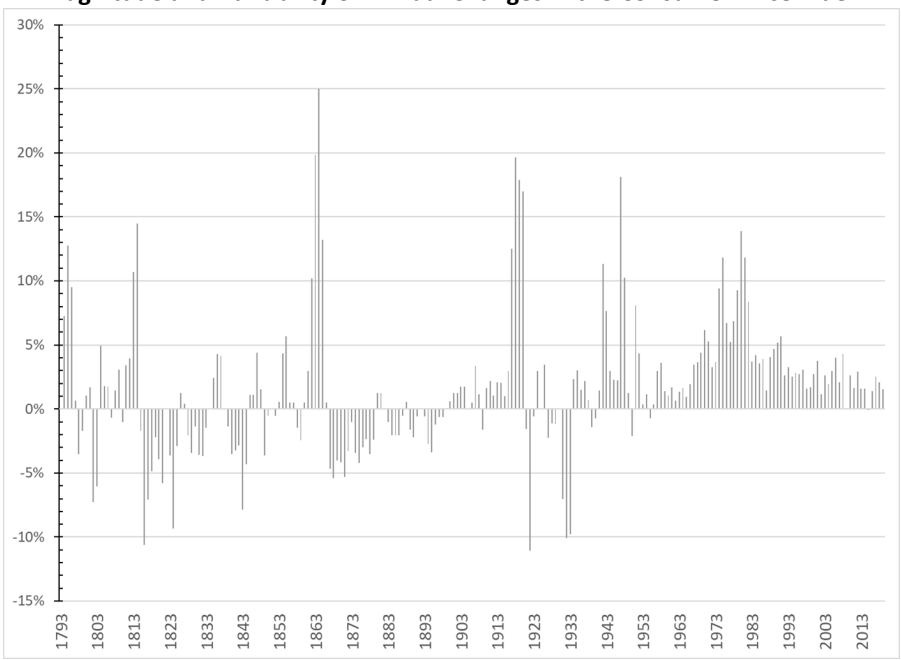
Table 6
Means and Standard Deviations Overall and by Century

	Stocks		Bonds		Inflation	
	A	*G*	*A*	*G*	*A*	*G*
Overall: 1793 – 2019	7.38 *(16.69)*	6.05	4.58 *(9.74)*	4.13	1.58 (*5.21*)	1.42
1st century: 1793 – 1893	6.38 (*12.47*)	5.66	6.45 (*9.72*)	6.02	0.07 (*5.46*)	-0.07
2nd century: 1893 – 1993	8.11 (*19.72*)	6.25	2.55 (*9.67*)	2.09	2.93 (*5.21*)	2.80
3rd century: 1993 – 2019	8.43 (*17.87*)	6.77	5.17 (*8.40*)	4.84	2.21 (*1.02*)	2.21

Note. Real total returns expressed as a percent. *A* stands for arithmetic means, *G* for geometric means (compounded returns). Standard deviations are in parentheses. Standard errors for each of the two centuries can be computed by shifting the decimal one place to the left in the SD; for the partial third century, divide the SD by five to approximate the standard error. The third century inflation means appear identical only because rounded.

It is important to emphasize that the change in inflation regimes was not just a change in level, a toggling from low to high. The bar chart in Figure 16 helps to make this clear. The *average* rate of inflation in the first century was so low because episodes of severe inflation were cancelled out by episodes of severe deflation. In fact, the rate of inflation in the worst years of the War of 1812, of the Civil War, of World War I and of World War II

Figure 16
Magnitude and Variability of Annual Changes in the Consumer Price Index



was higher than in any single year during the great inflation of the 1970s. But the first century saw episodes of *deflation* of a magnitude never again seen in the modern record after the early 1930s. And prior to the second century, there had never been a run of price increases longer than half a dozen without interruption by a deflationary year of -1.0% or worse. It is not the magnitude of price increases but the sustained absence of price *decrease* that distinguishes the postwar regime.

The same pattern holds for real bond returns. The arithmetic mean for the 226 years is 4.58%. But this value does not predict either of the two centuries, with the first century showing strong bond returns of 6.45% and the second century showing weak returns of 2.55%. The two centuries differ at $p < .01$.[40] Of what use is a more precise estimate, based on the longest available time span, that does not predict asset returns over spans of a century? Mashing apples and oranges together cannot give a better grasp of how different fruits taste.

Turning now to the geometric means, the total sample would seem to confirm conventional thinking about the advantage of owning stocks. Over the entire 226 years the new historical record shows that stocks returned a real annualized 6.05% and bonds 4.13%. Taking the ratio,[41] the stock advantage comes out to be 184 bp, a reduction of just over 100 basis points from the Siegel estimate (using his values of 6.6% and 3.6% for stocks and bonds). The reduction comes about because the new record has annualized stock performance down 55 basis points relative to Siegel and bond performance up 53 bp. Again, these tiny-seeming numbers must not deceive. In terms of final wealth, the new record shows a reduction over 226 years of almost 70% in stock wealth relative to the Siegel estimate and a multiplication of over 3X in bond wealth.

But as before, the century breakdowns are more illuminating. The first century shows a stock *disadvantage* of over 30 bp, with US bonds outperforming stocks over those hundred years. The second century shows a strong stock advantage that exceeds the overall estimate in Siegel by over 100 bp. The stock advantage has not been stable across centuries within US market history.

Under finance theory, there *must* be an equity premium; risk must be rewarded. Stocks are risk assets, bond are risk-less assets, therefore stocks must beat bonds over the long term. Scientifically speaking, the relation between risk and return is a lawful regularity. For it to fail to hold over spans as long as a century calls into question whether there is in fact any lawful relationship between the returns on stocks and bonds over arbitrarily selected holding periods of considerable length.

The regime thesis offers an escape from this dilemma. It argues that lawful relationships among asset returns, insofar as these obtain, run between each asset and prevailing macroeconomic factors, not between pairs of assets. There is no law concerning the relationship of stock and bond returns per se. Rather, each of these assets consists of a different basket of heterogeneous risks. Some of these risks gets rewarded under some macroeconomic conditions, other risks get rewarded under other conditions. Depending on the circumstances, there may be no reward for a particular type of risk for a long time. The regime perspective denies that risk is one thing. There are multiple potential risks bundled in any type of investment, and in any one period only some types of risk will get rewarded.

Here is an example. Perhaps there is a law that relates comparatively stronger returns on bonds to macroeconomic conditions where bouts of significant deflation recur. Such macroeconomic conditions might also depress stock returns, insofar as lack of pricing power drives down profitability. Inasmuch as the first century in US market history saw such bouts of deflation, that would explain why bonds outperformed in that century. The lawful character of asset returns gets restored, at the cost of simplistic generalizations about the comparative returns available from assets when risk is measured as one single thing present in one asset and absent in the other.

Finance theory can be rescued from the jaws of the new historical record by dropping the one-dimensional characterization wherein risk means one thing, stocks are fully a risk asset, and bonds are a minimally risky or risk-less asset. No. *Every* investment is a risk asset. Each represents a basket of different risks, and the contents of the basket differ across asset types.

With the new historical record in hand, and in conjunction with the regime perspective, another shibboleth of financial planning can be exposed as a flawed extrapolation from limited data rather than a general law holding universally across time. The next section examines how the contribution of dividends has varied. Its purpose is to reinforce two key points about regimes: that short time frames beginning in 1926 mask the true diversity of market history and give a misleading impression of how assets might perform in a different era; and 2) that assets have no consistent returns profile. There is no one return to risk that comes into more precise focus as longer and longer time intervals are amassed.

Dividend regimes

Most authorities believe that dividends have accounted for a significant fraction of the long-term total return on stocks. However, the proper method for estimating that fraction has not been established. To set up the question, here is a quote from index fund pioneer John Bogle, who

gives two possible metrics, while foregrounding the more dramatic.

> "An investment of $10,000 in the S&P 500 Index at its 1926 inception … with all dividends reinvested would by the end of September 2007 have grown to approximately $33,100,000 (10.4 percent compounded) ... If dividends had not been reinvested, the value of that investment would have been just over $1,200,000 (6.1 percent compounded)—an amazing gap of $32 million. Over the past 81 years, then, reinvested dividend income accounted for approximately 95 percent of the compound long-term return earned by the companies in the S&P 500."[42]

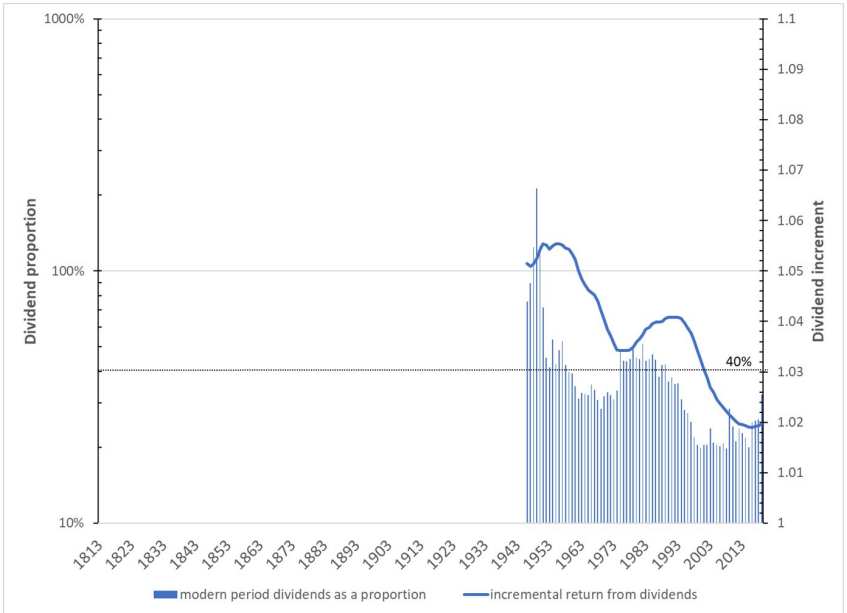
Impressive … but unfortunately, by the nature of exponential math, *any* difference in compounding rate must eventually produce a large arithmetic difference in wealth accumulation. The drama can be drained from Mr. Bogle's demonstration by computing wealth accumulations using the same compounding rates but shortening the holding period. Thus, after forty years at those rates re-invested dividends will only have accounted for 80% of total wealth. And if the holding period shortens further to twenty years, then dividends will only have accounted for 55% of final wealth. Conversely, if the holding period extends further, say to 226 years, at those rates reinvested dividends would by then have accounted for 99.99% of cumulative wealth. In fact, any compounding interval much longer than a century will reduce the contribution of the slower compounding item to negligible levels. That's just math.

A metric based on proportion of cumulative wealth contributed by dividends cannot be the best choice in long-term historical studies. An alternative metric can be derived from the geometric roots underlying the compounding rates. In the Bogle example the total return root is 1.104 and the price appreciation root is 1.064. The total return root is thus about 4.05% greater (1.104 / 1.061). I will refer to values calculated this way as the incremental return from dividends. If the contribution of dividends must be expressed as an arithmetic proportion, then things can be turned around to show that since 1926 dividends have contributed about 40% of the rate of wealth accumulation (4.05% / 10.4%).

With these metrics it becomes possible to probe more deeply for change over time in the contribution of dividends to total return. For reasons that will emerge, it seemed best to focus on twenty-year rolls in this exploration. Figure 17 shows the results for the modern period post-1926. It reveals an unsuspected feature of the chosen metric: when price appreciation over the holding period is negative, as occurred around 1949, twenty years after the Crash, then dividends must account for *more than 100% of total return.* But was this a one-time event, a unique concomitant of the worst crash in US market history? There is no way to tell without extending the historical record past 1926.

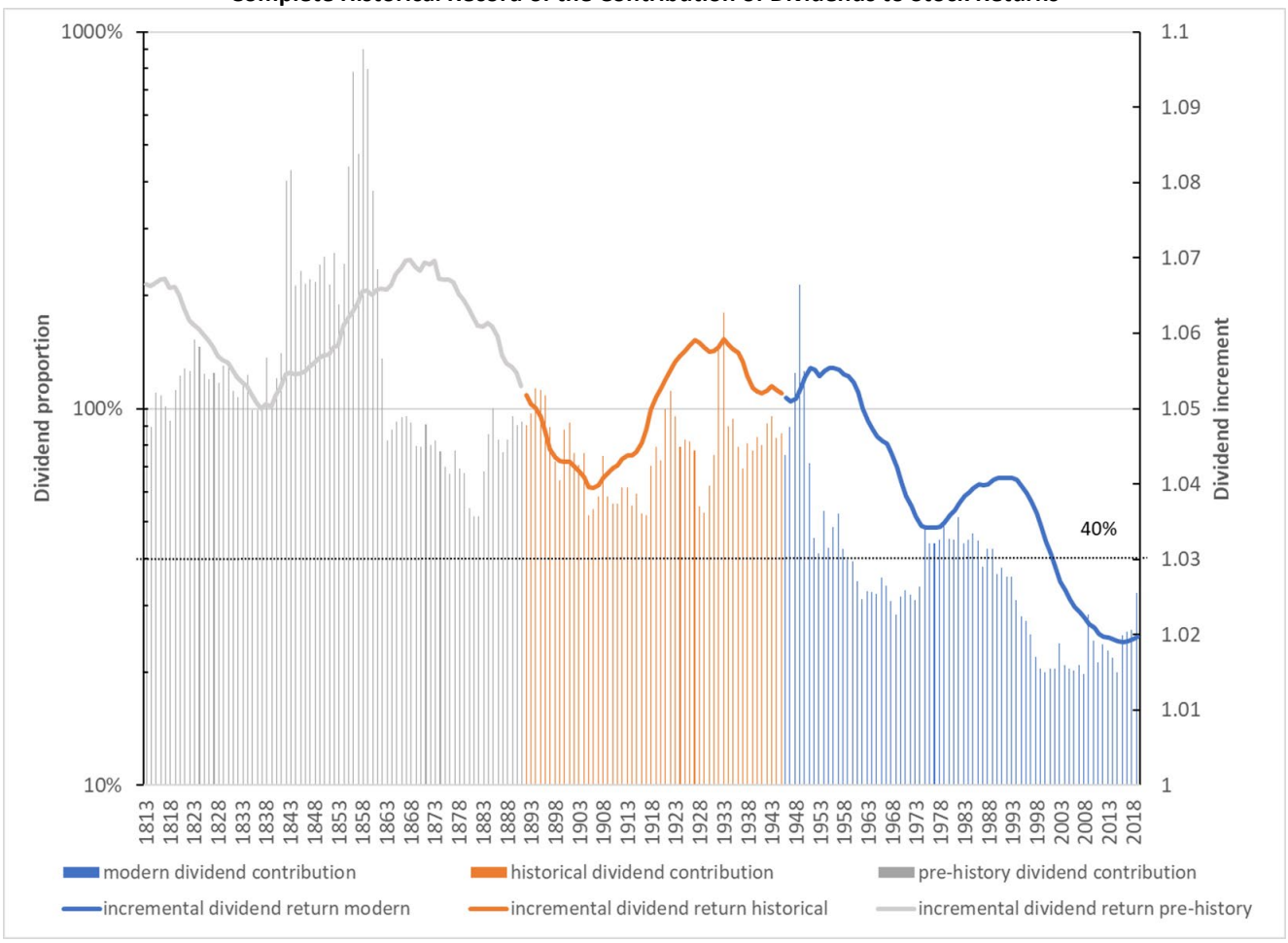
Setting aside these anomalous first few rolls, through the early 1990s one might have imagined that the proportion of total return coming from dividends did indeed fluctuate around the 40% seen in the Bogle example, i.e., was a mean-reverting process. But with the record extended to 2019, and with incremental dividend income charted as a line, a more apt description would be that the contribution of dividends fell across two great waves, first falling during the post-war rally ending in the mid-1960s, then backing up during the turbulent 1970s, and then falling still further in the boom of the 1980s and 1990s. Incremental return from dividends fell from 5.5% just after the war to 2% in recent years. The proportion of total return from dividends fell from just over 50% in the years after 1950 to just under 20% by the 2010s, moving in the same two waves.

Figure 17
The Contribution of Dividends in the Modern Era Post-1926



Note. Each bar and each line value pertain to the twenty years ending on the date shown. This chart has been scaled to accommodate earlier data once added. Bar values charted on the left axis, line values on the right axis. See text for an explanation of the two metrics. The dotted line shows which rolls exceeded or fell short of the 40% contribution estimated from the Bogle example.

Figure 18
Complete Historical Record of the Contribution of Dividends to Stock Returns



Note. When price returns are low, as around the Panic of 1857 (see next chart), the contribution of dividends climbs into the hundreds of percent. Dotted line marks a dividend contribution of 40% as in the Bogle example.

In short: even the abbreviated modern record reveals that the 40% contribution from dividends that falls out of the Bogle example was just a happenstance number obtained across an arbitrary beginning and ending point. It has no special standing as a mean to which fluctuating values will revert.

Figure 18 extends the record back to 1793. I will discuss it in two stages, beginning with the Cowles data featured in Robert Shiller's work and also in Siegel's account of dividends.[43] The orange portion of Figure 18 reveals that there was nothing unique about the negative price appreciation that occurred around 1949. Instances where dividends accounted for more than 100% of total return recur in the post-1871 Cowles period. Specifically, the proportion tends to top 100% in rolls that encompass major bear markets, i.e., rolls that include the early 1930s, or the years around 1921, or the 1890s.

Figure 18 also shows that Cowles' data indicate that the contribution of dividends has been declining almost since the beginning of the 20th century (i.e., rolls ending 1920 and after). Incremental return verged as high as 6% for rolls ending in the 1920s. Even when it was somewhat lower, as in rolls ending around 1900, the proportion of total return contributed by dividends fluctuated between 50% and 100%, much higher than in the modern period. Contribution proportions as low as 40% were never seen prior to rolls ending after 1950.

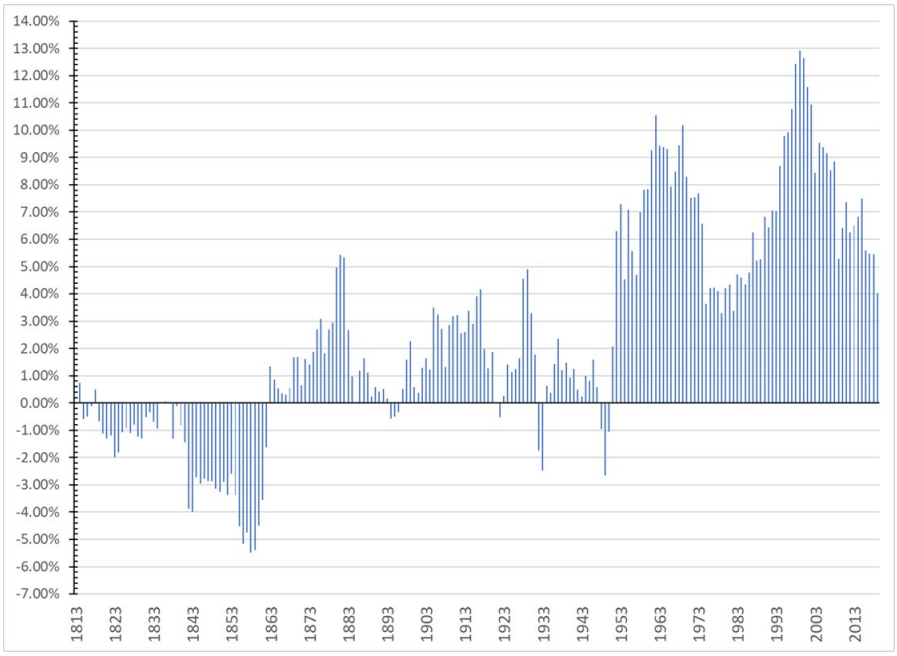
The enhanced historical record in Figure 18 suggests a discontinuity: that something fundamental changed after the Great Depression bottomed out in the 1930s. Before that point, dividends regularly accounted for 50% to 100% of total return, and price appreciation over twenty years was recurrently negative. After 1930 rolled off there was never a twenty-year period with negative price appreciation, and the contribution of dividends shrank both in absolute terms and as a proportion of total return.

Next, the grey portion of Figure 18 reflects new data unavailable to Siegel nor to any commentator who chooses

to stop at 1871. A straightforward interpretation of the complete 226-year record would be that the contribution of dividends has been declining in fits and starts for a long, long time. A more nuanced account would differentiate three broad regimes. Prior to the Civil War the contribution of dividends stayed above 100% for most rolls. After the Civil War an era of fluctuating price appreciation set in, with dividends only sometimes accounting for more than 100% of total return, while at others contributing as little as 50%. Following the Great Depression the modern regime began, with price appreciation growing stronger over time, reducing the contribution of dividends more and more. Finally, the absolute magnitude of the incremental return from dividends has been declining since at least the years following the Civil War, falling from levels once as high as 7% to current levels of 2%.

To this point the separate contribution of price appreciation to stock wealth has been inferred rather than directly portrayed. But the contribution of price appreciation is the flip side of the contribution of dividends and worth examining directly. Figure 19 shows nominal price appreciation rates for twenty-year rolls for the entire period. It roughly corresponds to the preceding figure in marking out three sustained regimes: the period before the Civil War, when twenty-year price appreciation was almost always nearly zero or substantially negative; the period after the Civil War through rolls that include prices from 1930, when price appreciation was mostly moderately positive but episodes of negative price appreciation recurred; and the period up to the present day when nominal price appreciation has *always* been positive over twenty-year intervals. The distinctiveness of recent times emerges clearly: only in the period following World War II have investors been able to count on receiving positive price appreciation from stocks if they bought

Figure 19
Nominal Stock Price Appreciation over 20-Year Holding Periods

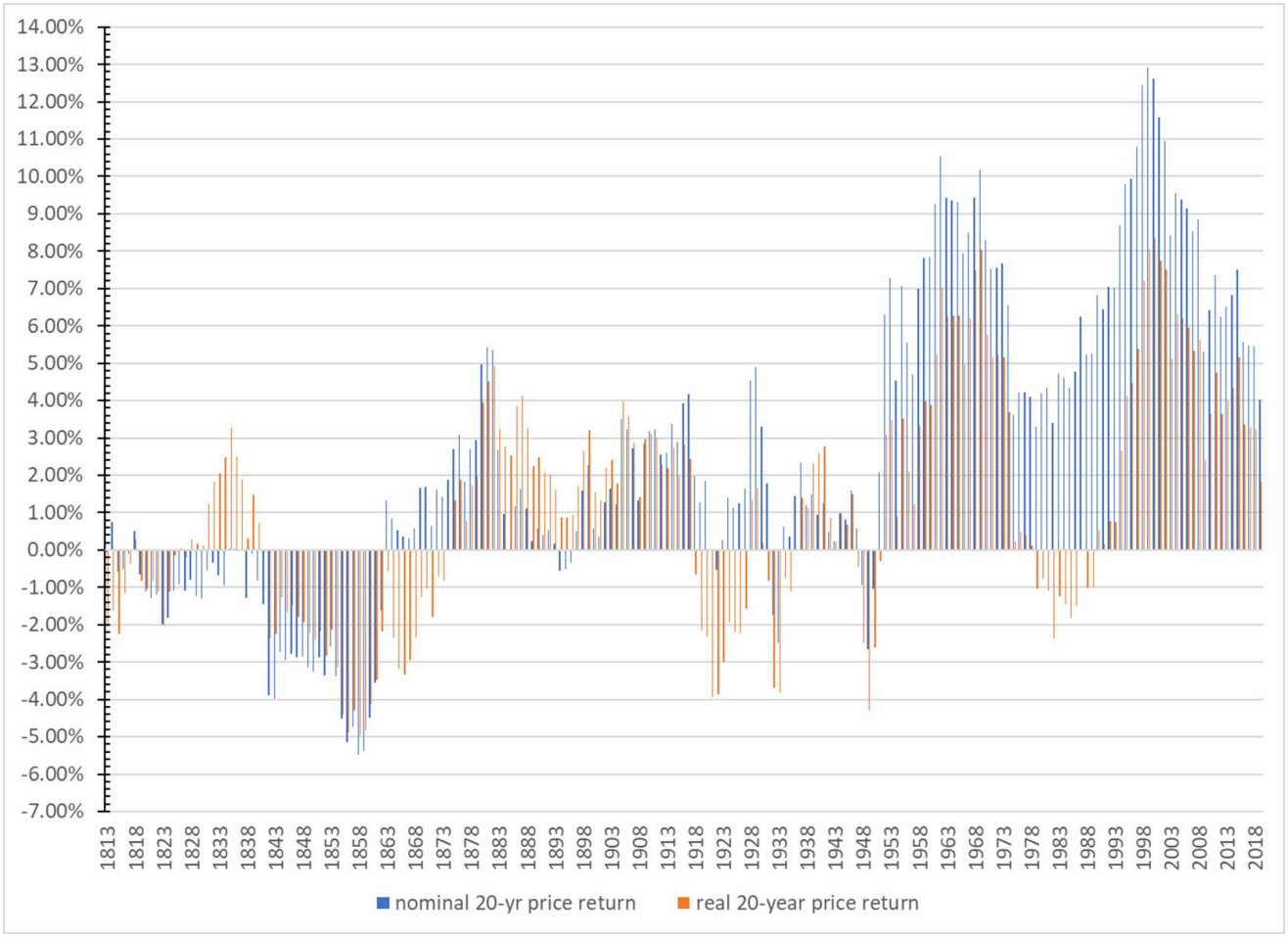


and held for decades. Before 1950, a holding period of twenty years afforded no guarantee that a stock investment would increase in price.

Figure 20 adds nuance by superimposing real price appreciation over the same rolls. This figure provides an important corrective because as noted earlier, the decades following World War II marked the point when inflation became endemic. And indeed, Figure 20 shows that in terms of *real* stock price appreciation, the modern post-war era has not been uniform but can be divided into three sub-periods: there was first the post-war boom that ended in the late 1960s and later the boom of the 1980s and 1990s, when in each case both real and nominal price appreciation hit rates never before seen in US market history. But there was also an interlude following the well-known peak in stock prices that occurred around 1966, when real price appreciation returned to being negative, as so often had been the case in the past. Elsewhere Figure 20 reinforces the importance of the deflation that recurred during the 19th century, showing that there were episodes of *positive* real price appreciation following the war of 1812 and again following the Panic of 1873, appreciation not visible in nominal returns. Finally, Figure 20 shows the boom of the 1920s in a new light: through most of that decade, consequent to high levels of inflation surrounding World War I, twenty-year rolls show negative real price appreciation, year after year, up until the roll ending in January 1928. From the perspective of real stock price appreciation, until the very end the 1920s were a period of recovery, not boom. And the peak around 1929 is reduced to naught but a foothill once dividends are set aside and price returns are rolled over twenty years.

The next figure will reinforce the point about how unusual the sustained price appreciation received in the modern era has been when seen in historical context. Figure 21 shows the cumulative real wealth achieved by

Figure 20
Real versus Nominal Price Appreciation Over 20-Year Holding Periods



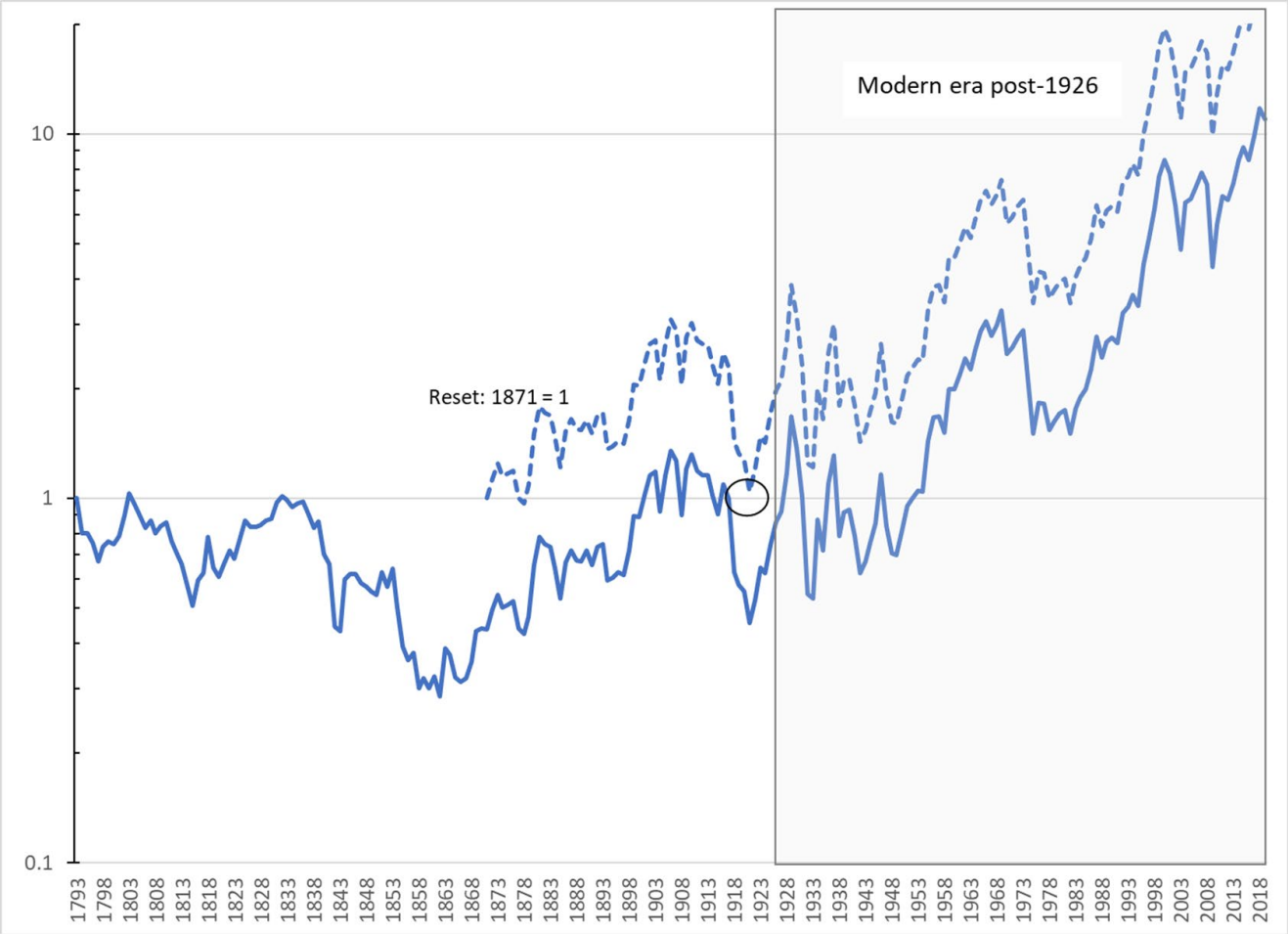
investing one dollar in price appreciation alone (no dividend reinvestment). I have also re-indexed a second version of this wealth index to 1.0 in 1871 to show what could and could not be glimpsed when the detailed historical record stopped at the beginning of the Cowles data extensively discussed by Robert Shiller. Finally, a shaded portion marks out the modern era, showing the limits to what could be learned from an historical record that commences only in 1926.

Taking the latter first, it is easy to see how an investor whose time horizon stops in 1926 could imagine that real price appreciation was the normal or natural condition for stocks. The fluctuations near the outset can be mentally dismissed as a one-off due to the Great Depression. From that point on, stocks appear to be on a never-ending escalator. True, there was quite a sag in rolls ending after 1966, and another after 2000, but when the post-1926 years stretch out to occupy the entire horizontal axis of a chart, as in any SBBI yearbook (e.g., Exhibit 3.1 in the 2020 SBBI), it becomes all too easy for the eye to dismiss these as fluctuations about the trendline, which so clearly goes up, up and up.

Turning now to the 1871 starting point, anyone who cared to look would have found that real price appreciation was essentially zero from 1871 to 1921 (circled); but the major users of the Cowles data, Shiller and Siegel, have tended not to be interested in calling out data-driven sub-periods.[44] And in any case, the fifty years to 1921 would have fallen prey to the ignore-the-lower-left-corner heuristic in charts extending from 1871 to the present.

Only by extending the historical record decades further back to 1793 does it become possible to discern how unusual sustained real price appreciation on stocks has been. Prior to the Civil War there had never been much of any (solid line). The first great wave of sustained price appreciation did not occur until after the Civil War, as the industrial revolution combined with postwar deflation to more than quadruple the real stock wealth obtained from price appreciation alone in the decades to 1907. However, sustained is a bit of a misnomer: much of this gain was given back, first following World War I and then again a decade later after the Crash of 1929. Accordingly, the two-century pattern

Figure 21
Real Wealth Accumulation through Stock Price Appreciation Alone: What the New History Shows



Note. Dashed line is perfectly parallel, and represents a reset to 1 in January 1871, to show what a history that began with Cowles would indicate. Circle shows the point where cumulative price appreciation dropped to near zero after 50 years in January 1921. It was likewise below zero for the fifty-three years from 1929 to 1982, and for the sixteen decades from 1793 to 1951.

of real stock price appreciation provides one of the clearest examples of how misleading a historical record truncated at 1926 can be.

To conclude, here are some specific returns. Through January 1982, the first 189 years of the record, cumulative real wealth from stock appreciation alone was negligible, amounting to a gain of 0.22% annualized—just twenty-two basis points. Real total return to that point had been 5.64% annualized. Hence, dividends had contributed 96% of it. In the thirty-seven years following, real price return jumped to 5.53% annualized; with real total return now at 8.19% for 1982 - 2019, dividends only accounted for 31% of it.

Assets as fuzzy sets

Here is another reason to expect regime change in asset returns: there may not exist any single real entity, "stocks," that has permanent properties such that this asset delivers a consistent returns profile over time. The dividend results come into sharper focus once this point is grasped.

For instance, before the 1840s most market capitalization resided in banks. Bank equity is an unusual creature, as investors learned to their cost in 2008-09. Banks can go along for decades paying generous and steady dividends, and then go bust in a day.[45] Banks before the 1840s also acted like super-REITs, expecting to distribute 100% of their profits as dividends. And prior to today's Federal Reserve, a bust bank's stock never came back: hence, the price depreciation it contributed to market averages was permanent. The conjunction of REIT-style dividends while successful, and permanent price depreciation upon failure, account in part for the dividend and price appreciation

patterns seen before the Civil War. In the next phase of stock market development, up until about 1900, railroads dominated market capitalization, and were also expected to distribute 100% of profits as dividends; but the big railroads that dominated the indexes did not go bust. Nor did they enjoy sustained price appreciation. Cowles found that nominal price appreciation on railroad stocks from 1871 to 1938 had been exactly 0.0%. The modern regime had not yet taken hold.

At some point after 1900 the REIT mindset lapsed, and market capitalization began to shift to industrial firms. After about 1990 capitalization began to shift again, now residing more and more in software and services firms. The distinct business models underlying such firms, as compared to antebellum banks or 19th century railroads, may explain the further change in dividend regimes.[46] In short: in the US it has been possible to invest in a portfolio of dividend-paying common stocks since 1793, but the nature of those "stocks" has changed more than once.

Once it is accepted that "stocks" are not a real entity with one unchanging set of properties, it becomes easier to understand why dividend regimes might have changed so much.

Summary:
The New Market History

The augmented historical record reveals substantial variation in almost any quantity of interest to financial analysts and planners. Stock returns, bond returns, the correlation between them, their standard deviation, and the contribution of dividends have all been more variable than could be glimpsed from a record truncated at 1926. Measurement spanning even one hundred years need not bring convergence. Moreover, the 20th century produced phenomena never seen before, like the great volatility of stocks around 1929, or the bond abyss after the war. There is every reason to suppose that the 21st century will also surprise.

Why Siegel's narrative went unchallenged

Few questioned Siegel's data series because the story these told was so reassuring. Yes stocks can be risky over short horizons, but investors who buy and hold will be handsomely rewarded for enduring that risk. Stocks might be risky in the short run but not in the long. The Siegel version of market history appeared to show this hoped-for outcome to be documented fact.

Furthermore, Siegel's two-century history appeared to provide empirical proof for one of the most fundamental theorems in finance. Investors, it is theorized, demand extra compensation for taking on the risk of stocks. Siegel's data, with its dramatically widening gap between stock and bond performance, appeared to show that this compensation had been received. Few noticed that in his telling, stock market risk had *not* been compensated through the first six decades. And no one could know in the early 1990s how little extra compensation would be received in the decades following. Nor had the international data been developed to any great degree.

In reading Siegel investors decided that they no longer had to bow to that old nostrum so often seen in the mutual fund prospectus: that past performance provides no guarantee of future results. Siegel's market history seemed to offer just such a guarantee. Readers of Siegel neglected to ask how stocks could be a risk asset if there was no longer a risk of underperformance once the holding period was made suitably long. And if stocks are not risky over the long term, why should investors receive extra compensation for holding stocks for long periods?

The regime thesis attempts to resolve this conundrum. The new market history demonstrates that the riskier asset can underperform the safer asset over arbitrarily long intervals.[47] Accordingly, with risk restored, one can reasonably expect the riskier asset to outperform the safer asset over other long intervals.

Glass half-full

I have spent most of this paper using the new historical record to rein in expectations that had been inflated by Siegel's exposition of the faulty data available to him. But it bears emphasis that not once in this paper did I point to any error made by Siegel himself. All my criticisms have been leveled at his antiquated sources and the limits of the data they supplied him.

Nonetheless, it may be appropriate to end on a more upbeat note, one not inconsistent with some elements of Siegel's account. Even in mediocre cases stock investors have made large profits over century-long intervals. A comparatively weak return of 4.25% to 4.50% real annualized, as seen in the US before the Civil War, or the World ex-USA since 1900, would still double wealth about six times over one hundred years, turning an investment of $10,000 into a portfolio worth over $650,000 in real terms. Likewise the bond investor, who gets the 226-year US estimate of only 4%, would still double wealth more than five times over a century, turning that $10,000 into over half a million real dollars. Whether buying stocks or bonds, investors, which is to say risk-takers, have mostly triumphed.

But again, what of the 21st century? How will stock and bond investors fare going forward? No one knows—and that is the real point of the new history. The new data show that US results for the 20th century did not generalize to the 19th century; there is no reason to expect them to generalize to the 21st century either. Sometimes both stock and bond investors do poorly, as in intervals that ended near 1843 or 1921. Sometimes both assets enjoy strong returns, as after the Civil War and again after 1982. But many other permutations are possible.

All that is certain: both remain risk assets over any human investing horizon. Both stocks and bonds can disappoint over an arbitrarily long interval, even as either asset can bestow handsome rewards.

Caveat investor.

References

Arnott, R., 2009. Bonds: Why Bother? *Journal of Indexes*, 12(3), pp. 6-17.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T.J. and Pedersen, L.H., 2018. Size matters, if you control your junk. *Journal of Financial Economics*, 129(3), pp.479-509.

Cowles, A., 1939. *Common-Stock Indexes*. Principia Press. [available at Yale School of Management web site]

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

Dimson, E., Marsh, P.R., Staunton, M., 2020. *Credit Suisse Global Investment Returns Yearbook 2020*. Credit Suisse Research Institute.

Durand, D., 1942. *Basic Yields of Corporate Bonds, 1900-1942*. National Bureau of Economic Research.

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

Friedman, M., 1952. Price, Income and Monetary Changes in Three Wartime Periods. *The American Economic Review*, 42 (2), pp. 612-625.

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. [data can be downloaded from Yale School of Management web site]

Hilt, E., 2009. Wall Street's First Corporate Governance Crisis: The Panic of 1826. *National Bureau of Economic Research*, No. w14892.

Homer, S. 1963. *A History of Interest Rates: 2000 B.C. to the Present*. Rutgers University Press.

Homer, S. and Sylla, R., 2005. *A History of Interest Rates*, fourth edition. NY: Wiley.

Ibbotson, R. G., 2020. *Stocks, Bonds, Bills & Inflation*. Wiley.

Ibbotson, R.G. and Sinquefield, R.A., 1976. Stocks, Bonds, Bills, and Inflation: Year-by-Year Historical Returns (1926-1974). *The Journal of Business*, *49*(1), pp.11-47.

Jordà, Ò., Knoll, K., Kuvshinov, D., Schularick, M. and Taylor, A.M., 2019. The rate of return on everything, 1870–2015. *The Quarterly Journal of Economics*, *134*(3), pp.1225-1298.

Kaplan, P. D., 2012. *Frontiers of Modern Asset Allocation*. Wiley.

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. [download at NBER.com]

Martin, J. G., 1898. *A Century of Finance*. Published by the author.

McQuarrie, E. F., 2009. The Stock Market's Little Shop of Horrors: And You Thought the Aftermath of 1929 Was Grim. *Journal of Investing*, 18 (2), pp. 6-17.

Napier, R., 2005. *Anatomy of the Bear: Lessons from Wall Street's Four Great Bottoms*. CLSA Books.

Officer, L. H. and Williamson, S. H., 2020. "The Annual Consumer Price Index for the United States, 1774-2015," *MeasuringWorth*, URL: http://www.measuringworth.com/uscpi/.

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.

Rousseau, P. L., 1999. The Share Liquidity and Industrial Growth in an Emerging market: The Case of New England, 1854 - 1897. *National Bureau of Economic Research,* Working Paper H0117.

Schmelzing, Paul, 2019. Eight Centuries of Global Real Interest Rates, R-G, and the 'Suprasecular' Decline, 1311–2018). Available at SSRN: 3485734

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

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

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

Siegel, J. J., 1992b. The Real Rate of Interest from 1800–1990: A Study of the US and the UK. *Journal of Monetary Economics*, 29(2), pp. 227-252.

Siegel, J. J., 2005. *The Future for Investors*. Random House.

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.

Sobel, R., 1975. *A History of the New York Stock Exchange*. Weybright and Talley.

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. [data also downloadable at EH.net]

Werner, W. and Smith, S.T., 1991. *Wall Street*. Columbia University Press.

White, R., (2011). *Railroaded: The Transcontinentals and the Making of Modern America*. W. W. Norton.

Zweig, J., 2009. Does Stock Market Data Really Go Back 200 Years? *Wall Street Journal*, July 11th.

Appendix

Old and New Data Sources for Determining the Historical Performance of Stocks, Bonds, and Inflation

This appendix has four parts:

1. A guide to the working papers underlying this summary effort;

2. A tabulation and comparison of the sources available to Siegel versus the newly available sources used to reconstruct the historical record;

3. Two figures, detailing how and where the new histories of stock and bond returns depart from those of Siegel;

4. A table containing the 226 annual stock and bond returns analyzed in the paper, along with the inflation measure used to deflate these returns.

The new historical record was assembled in a series of efforts, each chronicled in a working paper that can be downloaded from SSRN.com.[1] These archived working papers together constitute an extended online appendix to this paper and can be consulted as desired to vet sources and methodology, or for micro-analyses too detailed to be pertinent to this overview (e.g., how banks in different regions fared during the Panics of 1819 and 1837). The papers also contain many appendices which probe key sources in detail, e.g., the problematic construction of Macaulay's (1938) railroad and municipal bond indexes. Appendices are called out by capital letter in the table below; each paper has a Table of Contents to facilitate their perusal.

The identification numbers and periods covered by each paper are as follows. To avoid clutter, for the most part these papers do not otherwise appear in the notes or references sections of this paper.

SSRN ID#	Period covered	Appendices & Sub-Topics of note
Stocks		
https://papers.ssrn.com/abstract=3480838	1793 to 1871	***A***, problems with insurance stocks, ***B***, issues with the Goetzmann et al. dataset; compares bank performance by region, showing the uniqueness of New York; compares turnpike, canal, and early railroad stock performance.
https://papers.ssrn.com/abstract=3564496	1866 to 1897	***A, B,*** Cowles' reliance on Macaulay & potential holes in coverage, ***C***, prevalence of stock dividends and rights issues in this era; quantifies stock market dominance of the railroad sector before 1900; finds scant effect of the Panics of 1873 & 1893 on real total return for stocks.

[1] The earlier paper SSRN #3209440 discusses issues associated with the Sylla et al. data compilation; SSRN #3050736 expands on misleading visual heuristics.

SSRN ID#	Period covered	Appendices & Sub-Topics of note
Bonds		
https://papers.ssrn.com/abstract=3260733	1793 to 1857	***C***, reconstruction of Siegel's bond returns for this period from his published methodology; develops the distinctiveness of early bonds (e.g., no fixed maturity); compares Federal, municipal, and corporate yield levels; compares New England yields to those of other municipalities.
https://papers.ssrn.com/abstract=3269683	1857 to 1897	***D***, explains why Macaulay's bond indexes provide a misleading picture, ***E***, explains need to calculate the greenback price of gold in calculating Federal bond returns, ***L***, probes how Siegel's own stock return estimates changed over time; develops history of bond trading in the US after the Civil War.
https://papers.ssrn.com/abstract=3633277	1897 to 1927	Compares characteristics of railroad, utility, and industrial bonds and how these changed; supplements Hickman (1958) with respect to the near absence of bonds rated below investment grade in these decades.
https://papers.ssrn.com/abstract=3740190	1927 to 1974	***A***, evolution of US bond markets in the 20th century, ***E***, problems with pre-war yield averages published by Moody's and others, ***F***, how distressed bonds performed during the Depression, ***G***, evolution of Moody's rating grades and the advent of the investment grade distinction; develops regime thesis in the context of the corporate bond premium.

All the newly collected individual stock and bond prices used to calculate the index returns analyzed in this paper will be placed on the public record; when posted, a link may be found at edwardfmcquarrie.com.

For a less terse, more relaxed narrative of the data series and periods in the following table, please see my posts here: https://www.bogleheads.org/forum/viewtopic.php?f=10&t=353607

[table footnotes follow section C]

*A. Data sources for historical **stock** performance*

Period	Siegel's sources	Source problems	New historical record
1793 - 1802	Not in Siegel	--[a]	Adds nine years of market history past what could be obtained from Siegel's sources.[b]
1802 - 1825	Smith and Cole (1935) via Schwert (1990)[c]	Index limited to 7 stocks; excluded dozens of other stocks; excluded the largest single stock; equal-weighted; no dividend information.	Includes 3X to 5X the count of stocks; includes the largest single stock; share counts observed and used to weight returns by capitalization; includes a comprehensive record of dividends.
1825 - 1866	Goetzmann, Ibbotson and Peng (2001) [Replaces Schwert series in the most recent Siegel book edition]	Restriction to NYSE; many exclusions, including the largest single stock; very thin coverage in the 1840s with index price changes based on as few as four stocks; some years missing altogether, requiring interpolation; numerous errors in price and dividend data; price-weighted index.	Includes exchanges outside of New York (these accounted for a majority of market cap); includes the largest single stock; no gaps or reductions in coverage for any year or sub-period; more comprehensive coverage of dividends; adjusted for stock rights and stocks splits. NOTE: the greatest shortfall relative to Siegel's estimate is found here.
1866 - 1897	Goetzmann et al. (2001) [for 1866 – 1871] Cowles (1939)[d] and Macaulay (1938)[e] [for 1871 to 1897]	Restriction to NYSE and omission of immediate post-Civil War returns on NYSE industrials and utilities; exclusion of dozens of industrials; survivorship bias evident in railroad index carried over from Macaulay	Adds 1866 – 1871 returns on industrials and utilities on the NYSE. Adds industrials traded in Boston using Rousseau (1999) data (in the early years these had capitalization greater than NYSE industrials). Observes dividends and share counts and adjusts for rights and splits. Corrects survivorship bias by adding failed railroads excluded by Macaulay.
1897 - 1926	Cowles (1939) and Macaulay (1938)	Believed to be few; no new data collected	Not re-examined, given results for 1871 – 1897, which showed only slight variations in performance relative to Cowles. Hence, for this period the new record is the same as in Siegel.
1926 to present	CRSP total market index (NYSE only until 1962)	Omission of lower-quality stocks that traded on the AMEX	Not re-examined. Same record as in Siegel.[f]

*B. Data sources for historical **bond** performance*

Period	Siegel's sources	Source problems	New historical record
1793 - 1802	Not in Siegel	--[a]	Adds nine years of market history past what could be obtained from Siegel's sources[b]
1802 - 1857	Homer (1963) drawing on selected bond prices compiled by Martin (1898)[g]	Bond returns were back-calculated by Siegel from selected yields in Homer rather than observed prices in Martin.[h] Municipal coverage restricted to one or two small bond issues from Massachusetts. Only one of the several Federal bonds available was used. Siegel selected the lowest yielding bond in calculating returns.	Assembles a portfolio of available bonds. When multiple Federal bonds were available, all were included. From 1826 adds municipal bonds from outside of New England, covering New York, Pennsylvania and Maryland. Adds corporate bonds from 1832.[i]
1857 - 1897	Homer (1963) drawing on Macaulay (1938), who drew on Martin (1898) for selected New England municipal bond prices	Returns again back-calculated from yields in Homer rather than the prices compiled by Martin. After 1865 Siegel used exclusively the New England municipal bond yield tabled in Homer. Unbeknownst to Siegel, these yields reflected not an observed average but theoretical values estimated by Macaulay. Interest payments made in gold coin not adjusted for the greenback price of gold.	An aggregate long bond index was constructed from observed prices of dozens of municipal, Federal and corporate bonds through 1897, using multiple sources in addition to Martin. Interest payments corrected for the gold premium during the greenback era. NOTE: Bond performance in these decades found to be much stronger than in Siegel and far above Siegel's long-term estimate.
1897 - 1926	Homer (1963) drawing on Macaulay (1938) to 1900, Bond Buyer municipal index to 1920, and Federal Reserve publications for Federal bond index after 1920	Exclusion of the largest bond market sector (corporate). Still no observed bond prices. Coupons and maturities of the Bond Buyer municipal index unknown, and count and identity of bonds unknown. Siegel imputed coupons and maturities throughout, including the Federal index. Actual effective maturities for the Federal index were lower than assumed by Siegel.	Substitutes an index constructed from investment grade corporate bonds traded on the NYSE, with a minimum maturity of fifteen years, and returns weighted by par value (contains 50 – 100 bonds).

1926-1946	Ibbotson SBBI (single Treasury bond with approximately 20 years to maturity) [not used by Siegel: SBBI LT Corporate bonds] SBBI source: Ibbotson and Sinquefield (1976)	Treasury market not well-developed in these years. Maturity ladder not filled. Effective maturity of the SBBI bond fluctuates. Diverse contract terms caused other long Treasury bonds sometimes to perform differently than the single SBBI bond. SBBI corporate returns not based on observed prices, rather, bond returns were back-calculated from a yield series with imputed coupon and maturity. Downgrades and defaults in the 1930s not reflected in the SBBI corporate series.	Extends the 1897 – 1926 index of investment grade corporate bonds, again using observed prices. Downgrades and defaults affected returns in the 1930s, so that the new index shows lower corporate bond performance than the SBBI corporate index.
1946 - 1974	Ibbotson SBBI (single Treasury bond with approximately 20 years to maturity) [not used by Siegel: SBBI LT Corporate bonds]	Same issues with the Treasury market as in 1926 - 1946. SBBI corporate returns not based on observed bond prices but on yield series compiled by Homer at Salomon Brothers. Returns not value weighted but based on selected coupon ranges. Only utility bonds rated Aa were used through 1969.	Extends the new corporate bond index further to January 1974. Shows lower performance in the 1960s and early 1970s relative to the SBBI corporate bond index.
1974 to present	Ibbotson SBBI (single Treasury bond with approximately 20 years to maturity) [not used by Siegel: SBBI LT Corporate bonds]	Few issues—modern Treasury market structure with well-populated maturity ladder falling into place. No problems found—from 1973 the SBBI corporate index maps well onto Lehman Brothers (now BloombergBarc) index of long Aaa and Aa bonds.	Uses the SBBI long corporate bond index rather than the SBBI long government index used by Siegel. This adds just under 20 basis points to the annualized returns. Rationale: these top-grade bonds (Aaa, Aa) provide a proxy for the entire investment grade space, if their returns are taken as near the midpoint of Treasury, agency, mortgage, and medium grade corporate bond returns (A, Baa).

C. Data sources for historical ***inflation***

Period	Siegel's sources	Source problems	New historical record
1793 - 1913	Snyder – Tucker series to 1850; thereafter, various Bureau of Labor Statistics series[j]	Oliver and Williamson (2020) discuss issues with older measures of consumer prices before 1913.	Uses the Oliver and Williamson consumer price index (annual).
1913 to present	BLS CPI-U index (CPI before 1978)	No problems identified	Uses January values of the CPI-U downloaded from the BLS website.

General notes

In this table "Siegel's estimate" refers to the long run estimates of real return tabled in Siegel (2014): 6.6% annualized for stocks and 3.6% for bonds.

Detailed table notes

[a] Data had not been available for this early period before the work of Sylla et al. (2006), who tapped a much larger archive of early newspapers than any previous compiler (Smith and Cole had stopped in 1802, setting Siegel's starting point, while Martin stopped in 1798, setting Homer's starting point). I had the additional benefit of searchable digital archives, which did not become widespread until after the Sylla et al. project was complete.

[b] One can find in Sylla et al. earlier bond prices (from late 1790) and stock prices (from mid-1791) than my chosen start point of January 1793. However, the earliest quotes of both stock and bond prices are not very stable, have gaps, and are sometimes not readily interpretable (regional currencies and quoting practices had not settled down). January 1793 was selected as the first January where I could assemble trustworthy quotes on three or more bank stocks; by that point, three Federal bonds from the Alexander Hamilton refunding of 1790 were also trading regularly.

[c] In his 1992 papers and early editions of the book Siegel used the Schwert data all the way through 1871. At later points Schwert had supplemented the Smith and Cole bank indexes, first with their railroad index and later with Macaulay's railroad index. Schwert had no dividend data (because Smith and Cole and Macaulay reported none); dividends were estimated by Siegel. The stock performance in Siegel (1992a) is lower than in the book editions because Siegel switched to a higher dividend estimate in later work. The substitution of the Goetzmann et al. series altered the relation of stock and bond performance before the Civil War, but to my knowledge Siegel has not published a commentary on these changes. The Goetzmann et al. spreadsheets can be downloaded from a site maintained by the Yale School of Management. Although I found errors in their dividend compilation, I was also able to complement my own digital searches, and dispense with some, by incorporating dividends they compiled.

[d] Most contemporary investors encounter Cowles by way of Shiller (2015). Cowles' data can be downloaded at the Yale School of Management site.

[e] The text of Macaulay with complete data tables can be downloaded from the NBER site. Selected data series can also be downloaded from the NBER site, not always filed under Macaulay's name; but pretty much any yield series or railroad series that begins in 1857 will have come from Macaulay.

[f] Note, however, the CRSP total market returns used by Siegel and me are lower in the 1930s than the large company returns used in the SBBI.

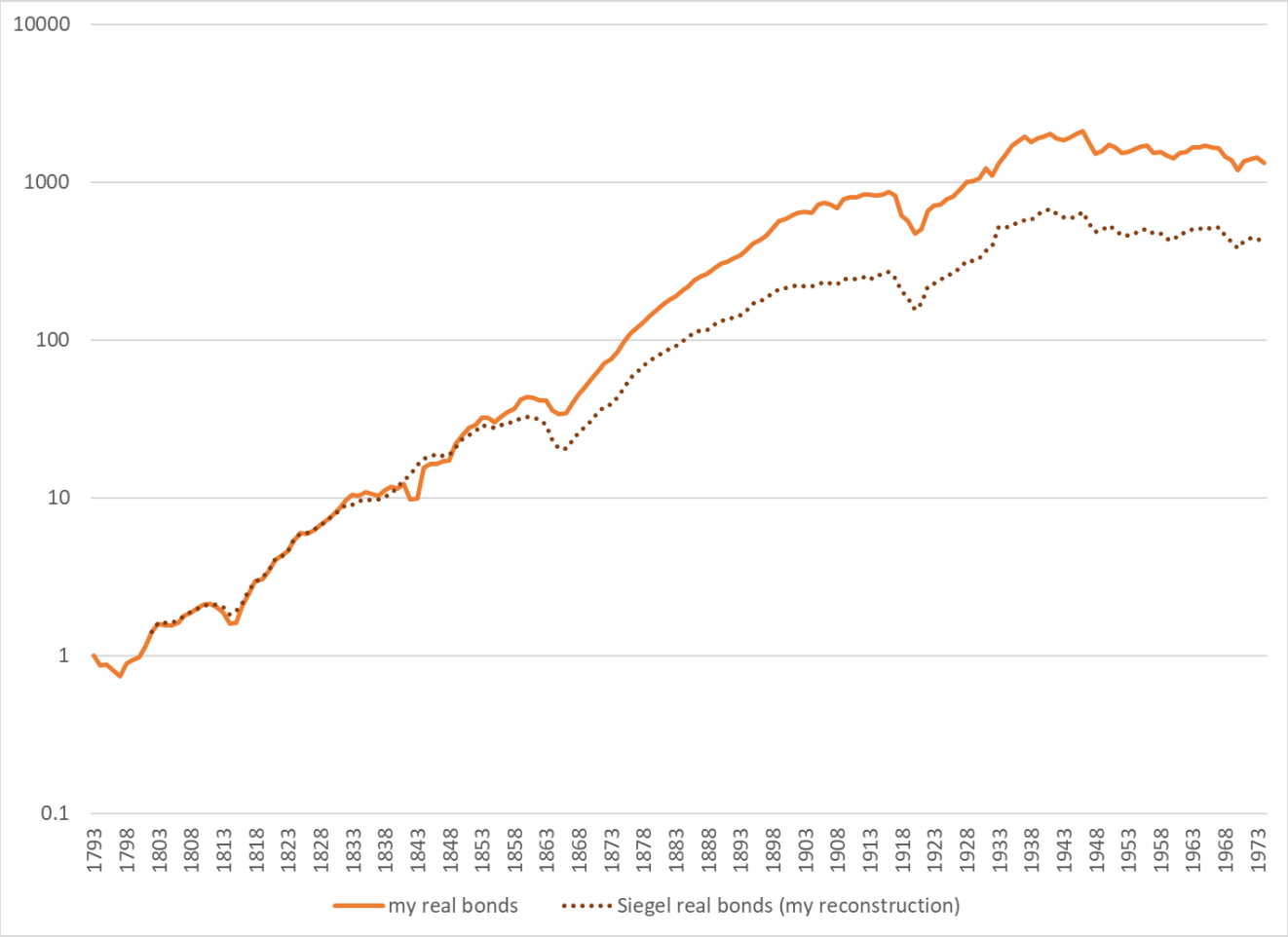
[g] Martin (1898), a stockbroker in Boston, published a large number of tables of stock and bond prices along with dividends and share counts in several editions across the last half of the 19th century. His final compilation, *A Century of Finance*, can be downloaded from books.google.com.

[h] Siegel uses the method of Ibbotson and Sinquefield (1976) to extract capital appreciation from successive yields in a series.

[i] Homer / Martin's results were subject to the Working (1960) effect because Homer used the midpoint of the annual high and low bond prices in Martin to calculate the yields from which Siegel extracted returns. In most cases the new record observes January prices each year rather than averaging annual highs and lows. Spot prices, plus inclusion of bonds from states outside of New England, cause the new historical record to show a much more volatile bond series in the vicinity of the Panic of 1837 relative to Siegel (see Figure A.1).

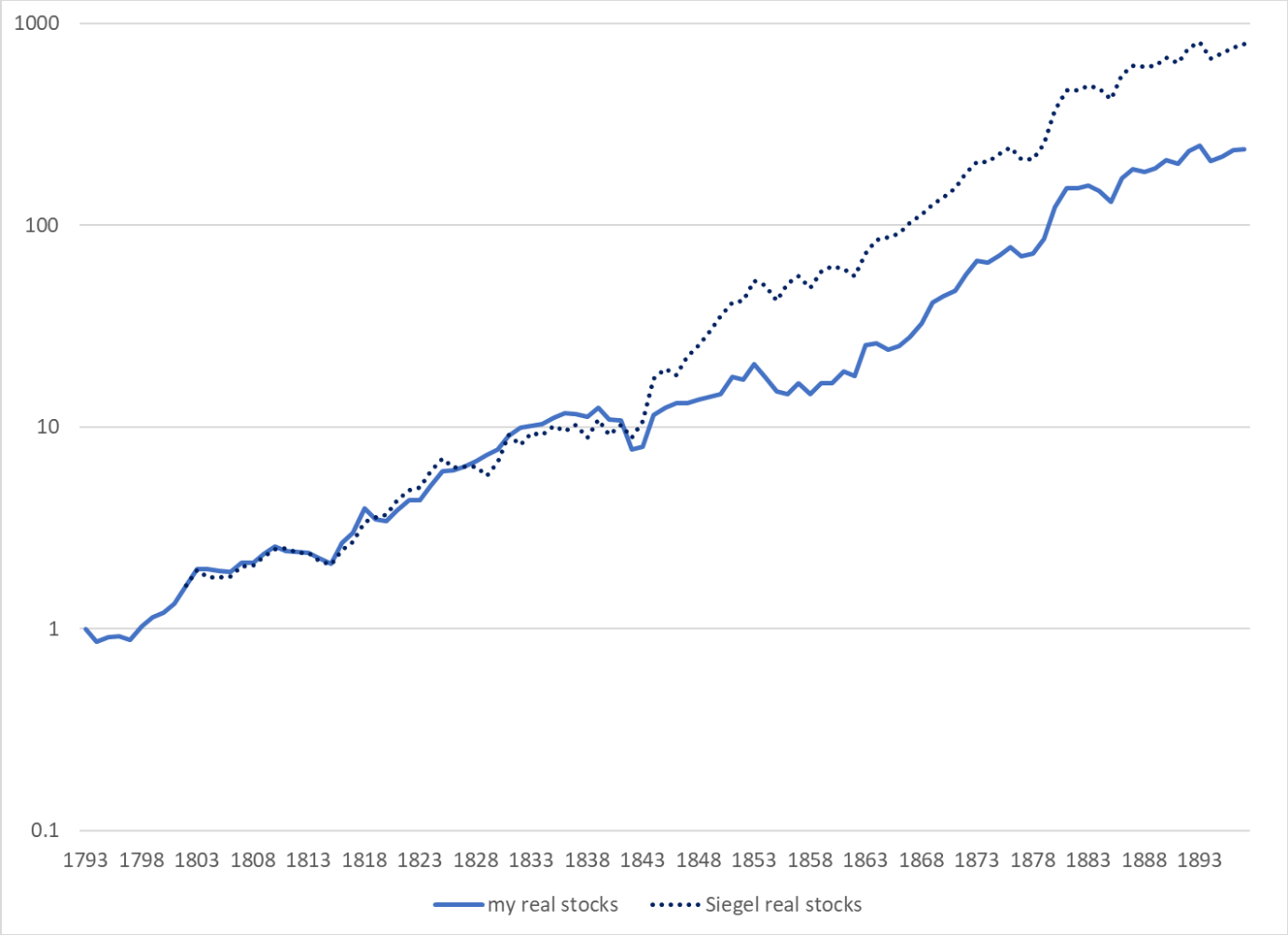
[j] The inflation measures listed in this table are from Siegel (1992b); the inflation measures charted in Siegel (1992a) differ. Book editions do not footnote which price index was used for real return charts in the book.

Figure A.1
Detail of Where the New Historical Record of *Bond* Performance Differs from Siegel's Account



Note. The Siegel bond line before 1926 is my reconstruction based on the methodological details given in the Appendix to Siegel (1992b). After 1926 the comparison is between my corporate index and the Ibbotson SBBI long government bond used by Siegel. Also, I use a different measure of inflation prior to 1926, and applied this to the reconstruction as well as my own series. This chart is truncated at the end of my new data collection, to reduce the scale and let differences emerge more clearly.

Figure A.2
Detail of Where the New Historical Record of *Stock* Performance Differs from Siegel's Account



See the note to Figure A.1, which applies here as well. The Siegel stock line here reflects his substitution of the Goetzmann et al. (2001) index in recent editions of the book. The Schwert-based stock series used initially by Siegel showed stronger returns before the Civil War, see Appendix L of SSRN #3269683. Exception: the Goetzmann series crosses below the new stock line after 1826, due to conditions unique to the New York market, where widespread fraud in bank and insurance stocks was exposed, see Hilt (2009). Differences after 1871 are small, as after reexamination, I found the Cowles series used by Siegel to be not too far off the mark.

Table A.1

The New Historical Record: Stocks, Bonds and Inflation

To January of:	Inflation	Nominal stocks	Nominal bonds
1794	7.3%	-7.64%	-7.24%
1795	12.8%	19.81%	15.32%
1796	9.5%	9.90%	-1.03%
1797	0.6%	-3.50%	-6.10%
1798	-3.5%	13.33%	15.46%
1799	-1.7%	8.50%	3.23%
1800	1.0%	6.10%	5.62%
1801	1.7%	13.64%	18.78%
1802	-7.3%	13.22%	15.00%
1803	-6.0%	14.21%	5.51%
1804	4.9%	4.32%	2.63%
1805	1.8%	0.31%	1.01%
1806	1.8%	0.30%	6.06%
1807	-0.7%	10.63%	11.41%
1808	1.4%	0.50%	4.77%
1809	3.1%	14.74%	10.06%
1810	-1.0%	7.49%	5.20%
1811	3.4%	-2.11%	3.94%
1812	3.9%	2.82%	-1.49%
1813	10.7%	10.30%	3.26%
1814	14.5%	7.38%	-2.76%
1815	-1.7%	-7.44%	-0.17%
1816	-10.6%	12.89%	14.68%
1817	-7.1%	4.73%	9.65%
1818	-4.9%	24.84%	15.66%
1819	-2.2%	-13.71%	-0.12%
1820	-3.9%	-5.16%	8.69%
1821	-5.8%	6.40%	10.41%
1822	0.0%	11.89%	6.38%
1823	-3.6%	-3.96%	2.84%
1824	-9.3%	7.25%	5.85%
1825	-2.9%	14.34%	7.85%
1826	1.3%	2.30%	0.60%
1827	0.4%	5.30%	5.49%
1828	-2.1%	3.94%	6.13%
1829	-3.4%	4.14%	3.02%
1830	-1.4%	5.12%	6.74%
1831	-3.6%	12.40%	6.01%
1832	-3.7%	4.92%	8.61%
1833	-1.4%	1.52%	6.10%
1834	0.0%	1.14%	-0.82%
1835	2.4%	10.20%	8.60%
1836	4.3%	10.86%	1.13%
1837	4.2%	2.11%	1.68%
1838	0.0%	-2.49%	8.03%
1839	-1.4%	8.79%	3.85%
1840	-3.5%	-15.42%	-6.39%
1841	-3.2%	-4.61%	4.47%
1842	-2.9%	-29.84%	-22.88%
1843	-7.9%	-5.47%	-6.96%
1844	-4.3%	38.25%	49.92%
1845	1.1%	9.86%	6.47%
1846	1.1%	6.24%	1.90%
1847	4.4%	4.56%	8.28%
1848	1.6%	5.31%	3.30%
1849	-3.6%	-0.41%	21.26%
1850	-0.5%	2.99%	14.32%
1851	0.0%	21.99%	10.99%
1852	-0.5%	-3.53%	3.93%
1853	0.5%	19.66%	11.40%
1854	4.3%	-9.54%	4.00%
1855	5.7%	-11.27%	-1.57%
1856	0.5%	-1.74%	9.10%
1857	0.5%	13.85%	7.98%
1858	-1.4%	-13.78%	4.04%
1859	-2.4%	10.87%	10.83%
1860	0.5%	0.86%	4.71%
1861	3.0%	17.95%	2.23%
1862	10.2%	4.27%	5.79%
1863	19.8%	70.71%	19.99%
1864	25.0%	26.90%	7.97%
1865	13.2%	5.76%	6.77%
1866	0.5%	4.29%	1.61%
1867	-4.7%	5.56%	10.01%
1868	-5.4%	11.01%	8.52%
1869	-4.0%	21.93%	6.43%
1870	-4.2%	2.37%	7.61%
1871	-5.3%	1.23%	6.32%
1872	-3.3%	16.53%	8.27%
1873	-1.0%	15.77%	6.59%
1874	-3.4%	-5.17%	6.98%
1875	-4.2%	3.99%	10.14%
1876	-3.0%	5.95%	9.89%

1877	-2.3%	-11.55%	5.42%
1878	-3.5%	-0.78%	4.42%
1879	-2.4%	14.76%	7.59%
1880	1.2%	45.70%	8.55%
1881	1.2%	25.87%	12.16%
1882	0.0%	0.54%	5.93%
1883	-1.0%	2.14%	4.88%
1884	-2.0%	-8.54%	6.27%
1885	-2.0%	-13.66%	4.72%
1886	-2.1%	28.45%	7.40%
1887	-0.5%	10.31%	5.06%
1888	0.5%	-2.16%	4.00%
1889	-1.6%	1.88%	6.63%
1890	-2.2%	7.99%	4.43%
1891	-0.6%	-5.25%	2.00%
1892	0.0%	16.26%	5.34%
1893	-0.6%	5.31%	4.29%
1894	-2.7%	-18.47%	4.76%
1895	-3.4%	2.64%	6.47%
1896	-1.2%	5.35%	3.39%
1897	-0.6%	0.53%	5.92%
1898	-0.6%	20.29%	10.97%
1899	0.0%	29.81%	11.28%
1900	0.6%	4.06%	2.91%
1901	1.2%	20.68%	7.54%
1902	1.2%	19.23%	5.29%
1903	1.7%	8.30%	2.00%
1904	1.7%	-17.20%	0.85%
1905	0.0%	31.53%	12.08%
1906	0.5%	21.73%	4.04%
1907	3.4%	0.61%	0.62%
1908	1.1%	-24.42%	-3.92%
1909	-1.6%	39.20%	12.33%
1910	1.6%	17.00%	3.43%
1911	2.2%	-3.81%	3.10%
1912	1.0%	3.52%	4.32%
1913	2.1%	7.08%	2.68%
1914	2.0%	-4.89%	1.00%
1915	1.0%	-5.77%	1.42%
1916	3.0%	31.44%	7.73%
1917	12.5%	9.21%	5.96%
1918	19.7%	-17.38%	-9.49%
1919	17.9%	16.93%	7.92%
1920	17.0%	19.05%	-2.51%
1921	-1.6%	-14.26%	4.66%
1922	-11.1%	9.28%	16.37%
1923	-0.6%	29.29%	8.15%
1924	3.0%	5.46%	3.93%
1925	0.0%	26.87%	8.45%
1926	3.5%	25.79%	8.08%
1927	-2.2%	10.05%	8.78%
1928	-1.1%	32.11%	8.48%
1929	-1.2%	47.14%	0.62%
1930	0.0%	-14.74%	4.35%
1931	-7.0%	-28.54%	6.88%
1932	-10.1%	-48.27%	-18.33%
1933	-9.8%	-6.05%	6.84%
1934	2.3%	75.65%	15.42%
1935	3.0%	-11.49%	18.08%
1936	1.5%	60.94%	8.93%
1937	2.2%	27.82%	9.13%
1938	0.7%	-36.28%	-6.81%
1939	-1.4%	19.87%	3.94%
1940	-0.7%	5.76%	2.48%
1941	1.4%	-9.12%	5.26%
1942	11.3%	-5.33%	3.52%
1943	7.6%	24.17%	5.16%
1944	3.0%	21.37%	7.83%
1945	2.3%	21.62%	6.81%
1946	2.2%	44.95%	6.39%
1947	18.1%	-10.85%	-0.39%
1948	10.2%	-1.91%	-5.73%
1949	1.3%	6.48%	6.04%
1950	-2.1%	21.87%	6.69%
1951	8.1%	35.73%	3.43%
1952	4.3%	16.18%	-2.99%
1953	0.4%	11.25%	1.99%
1954	1.1%	5.85%	5.12%
1955	-0.7%	43.86%	2.66%
1956	0.4%	21.06%	1.91%
1957	3.0%	8.13%	-7.00%
1958	3.6%	-2.78%	5.29%
1959	1.4%	39.19%	-4.59%
1960	1.0%	4.23%	-2.08%
1961	1.7%	15.32%	9.36%
1962	0.7%	15.02%	2.67%
1963	1.3%	-2.12%	8.10%
1964	1.6%	17.89%	1.83%
1965	1.0%	17.76%	3.49%
1966	1.9%	11.46%	-0.78%

1967	3.5%	-2.07%	1.59%
1968	3.6%	14.13%	-8.12%
1969	4.4%	17.68%	-0.35%
1970	6.2%	-16.81%	-9.05%
1971	5.3%	13.60%	21.36%
1972	3.3%	13.59%	6.23%
1973	3.6%	10.97%	6.35%
1974	9.4%	-16.09%	1.09%
1975	11.8%	-18.10%	3.26%
1976	6.7%	35.48%	10.23%
1977	5.2%	8.50%	12.93%
1978	6.8%	-4.87%	3.95%
1979	9.3%	20.58%	2.68%
1980	13.9%	26.32%	-12.00%
1981	11.8%	20.02%	2.60%
1982	8.4%	-2.60%	-1.23%
1983	3.7%	28.84%	43.07%
1984	4.2%	16.18%	10.16%
1985	3.5%	13.49%	17.49%
1986	3.9%	22.23%	26.56%
1987	1.5%	29.14%	21.89%
1988	4.0%	-5.71%	2.67%
1989	4.7%	19.94%	7.39%
1990	5.2%	12.02%	11.75%
1991	5.7%	5.95%	10.49%
1992	2.6%	27.24%	16.07%
1993	3.3%	10.58%	14.10%
1994	2.5%	13.66%	12.65%
1995	2.8%	-1.82%	-5.26%
1996	2.7%	36.71%	24.20%
1997	3.0%	24.16%	0.98%
1998	1.6%	24.29%	14.82%
1999	1.7%	26.39%	10.60%
2000	2.7%	15.85%	-8.77%
2001	3.7%	-3.83%	17.15%
2002	1.1%	-16.01%	8.68%
2003	2.6%	-21.44%	14.57%
2004	1.9%	39.46%	7.02%
2005	3.0%	7.52%	9.68%
2006	4.0%	14.65%	2.06%
2007	2.1%	13.90%	3.67%
2008	4.3%	-1.32%	3.29%
2009	0.0%	-39.22%	-1.70%
2010	2.6%	37.02%	14.90%
2011	1.6%	24.59%	9.17%
2012	2.9%	2.32%	22.67%
2013	1.6%	15.78%	5.17%
2014	1.6%	20.06%	-0.89%
2015	-0.1%	10.83%	20.31%
2016	1.4%	-4.70%	-5.98%
2017	2.5%	22.16%	5.77%
2018	2.1%	24.02%	9.75%
2019	1.6%	-2.92%	0.34%

Note. Stock and bond returns are January to January. Post-1925, these will look different than SBBI or CRSP returns, which are normally calculated and reported December to December.

Nominal returns are given so that if desired, the reader can calculate real return using a different measure of inflation. Stock returns after 1897 come from extant sources, as do bond returns after 1974, see the data sources tabulation earlier in this Appendix.

The inflation measures are also extant, see Oliver and Williamson (2020) and the Bureau of Labor Statistics. Prior to 1913, the annual inflation is computed as [mean (year n+2, year n+1) / mean (year n+1, year n)]. This was done to bring the all-year averages of Oliver and Williamson into alignment with the January prices of the new stock and bond data. Year 0 is 1792. See Macaulay (1938) for use of a similar device in the calculation of yields.

A complete data file with individual stock and bond prices and source notes will ultimately be posted at edwardfmcquarrie.com.

Notes

[1] Siegel (2014) is the most recent book edition as this is written. The academic foundations were laid in Siegel (1992a, 1992b). Siegel (2005) contains some historical assertions not developed in those papers which will become relevant when dividends are discussed.

[2] Arnott (2009) and Zweig (2009) are notable exceptions.

[3] The data are discussed in Sylla, Wilson and Wright (2006) and can also be downloaded from EH.net.

[4] The origins of the NYSE and its path to dominance do not fit a linear progression from the Buttonwood agreement in 1792. Important transitions occurred in 1817 and in the late 1860s, as described in Werner and Smith (1991) and Sobel (1975). The new market history differs from that of Siegel in part because it includes stocks trading on exchanges outside of New York during the decades when the NYSE was large but not yet dominant.

[5] For instance, the higher stock returns in Siegel (2014) versus Siegel (1992a) came about because Siegel updated his dividend estimate to a higher value. Goetzmann et al. (2001), used in the most recent editions of Siegel's book, attempted to observe dividends but I found many omissions and not a few errors of commission.

[6] See Ibbotson and Sinquefield (1976) for details of the procedure.

[7] Durand (1942) gives an account of how corporate bonds once dominated the fixed income markets. But the inspiration for Siegel's work in the early 1990s was the equity premium puzzle. That literature demands a comparison between a risk instrument and a risk-free security. Stocks and government bonds function in the theory as proxies used to investigate the price of risk, hence corporate bonds have no place. For the tax privileges enjoyed by municipal bonds even before 1913, see Macaulay (1938).

[8] The history of the US bond market is complex. A detailed account is provided in the bond-oriented working papers tabled in the Appendix. Suffice to say that for the decades between the Civil War and WW I, long corporate bonds were *the* fixed income investment in the US, dwarfing all others in size and trading volume. The modern structure, in which a vast supply of Treasuries available at every maturity dominates and anchors the fixed income space, was not fully developed until well into the 1970s.

[9] Early compilers such as Smith and Cole (1935), Macaulay (1938) and Cowles (1939) anchored their series in January, and I chose to do the same when first extending market history backwards from 1871. Later compilers, such as Fisher and Lorie (1964), anchored in December. Everyone who links the two sorts of series needs to use an 11-month or 13-month year at some point, or stick with a January ending point, as I chose.

[10] Or as Macaulay (1938), one of Siegel's key underlying sources, put it: "The common stock is a blood brother of the preferred stock and the bond. Investments as a class constitute one family" (p.129).

[11] See, e.g., his Table 6-1, p. 96. He stopped with thirty-year rolls.

[12] The reader can verify this point by calculating a weighted average of the first two sub-periods in Table 2 to get the odds for the 1793 to 1942 super-period.

[13] I had to label this the stock advantage rather than the equity premium because, as noted earlier, the theoretical literature on the equity premium requires a government bond or other proxy for a risk-free security, whereas I use aggregate bond returns from the 1830s forward and corporate bonds exclusively from 1897.

[14] Some longer rolls are charted in subsequent figures. I do not believe that asset performance regimes have any fixed length or periodicity. Nor do I think longer rolls are per se superior. As a reductio ad absurdum, I also calculated hundred-year rolls for the second and third periods in Table 2. Stocks won only 48% of the time for century rolls ending 1893 to 1942, continuing the upside-down U pattern seen in the table, and 100% in century rolls ending after 1942, continuing the 30- and 50-year pattern for that sub-period. What matters is not holding period length but the ratio of bad years to good years for stocks in a particular period, and the magnitude of those good and bad years.

[15] See chapter 10 of Siegel (2005).

[16] This number is based on the monthly returns for the SBBI large company index (the S&P 90 at that point). Estimates using annual or daily data will differ, as will values for other indexes (e.g., the Dow Jones Industrials), and for nominal versus real calculations. Most of the values I have seen estimate a decline of between 75% and 90%.

[17] However, real returns suffered from the postwar inflation, so that it was only after 1949 that stock wealth fully recovered from the Crash, in the sense of not relapsing again.

[18] See McQuarrie (2009) for other examples of international declines that put 1929 into perspective. See also Part Three of Kaplan (2012) for a different perspective on worst case scenarios in the US (but limited to post-1871 data). Readers of Napier (2005) will recognize several of the turning points that recur in this discussion of worst-case scenarios; but he was limited to post-1896 data.

[19] In the latest edition of Stocks for the Long Run, Siegel (2014) makes several claims about stock returns over twenty years (pp. 94-95). But I do not believe Siegel is the only proponent of the twenty-year threshold for "can't lose in stocks." See, e.g., Exhibit 2.9 in the 2020

SBBI, which supports a similar inference.

[20] Note. Excepting the one-year results, these standard deviations are computed over a set of overlapping geometric means and may not have the same statistical properties or interpretation as the standard deviation of a series of separate returns. However, Siegel (2014) includes a similar chart, see his Figure 6-2.

[21] These rolls were not included in Table 1 because none exist for the first sub-period (covering only sixty-nine years) and the purpose of Table 1 was to show how roll results vary by sub-period. See note 14 for more on the hundred-year rolls.

[22] Siegel (2014) also presents this result (Figure 6-2 on p. 98 and the discussion following), but does not find it problematic, as I do.

[23] Recall that standard deviation is the square root of variance, and that variance is the average of the squared differences from the mean. Note that under the regime thesis, it is appropriate to use N rather than (N – 1) in the denominator for these calculations.

[24] This point is developed in Raghubir and Das (2009).

[25] Recently titled *Global Investment Returns Yearbook 2020*. I am grateful to Mike Staunton for providing me yearbook copies. New yearbook editions appear on or about February; the 2021 edition had not appeared when this paper was drafted.

[26] The next revision of this paper will consider the data of Jorda et al. (2019). A preliminary examination indicates concordance with the Yearbooks (comparatively weak returns) and not with the findings in *Triumph* or *Stocks for the Long Run*.

[27] Dimson et al. have the US return since 1900 as 6.5% real annualized. If instead we compared final wealth for Siegel's exact 6.6% estimate to the new World ex-USA estimate of 4.4%, and used a 226-year time frame as in this paper, then the ratio of final wealth under the two returns would be over 100-to-1. Put another way, at these rates the achievable wealth from a long-term stock investment, for the international investor, would be *two orders of magnitude* less than Siegel estimated.

[28] Siegel's (2014) rejoinder to critics who cite the case of Japan can be found on pp. 199-200. He concludes that a longer time frame dissolves the counterexample: "from 1970 onward Japanese stocks have matched the returns in other countries." However, the 2020 Credit Suisse yearbook does not appear any longer to support that claim, showing a fifty-year return for Japanese stocks of 3.6%, well below the 5.1% recorded for the World ex-USA.

[29] Country by country annual returns were computed by Bryan and shared with me in April of 2021. A description of some relevant GFD findings can be found at globalfinancialdata.com

[30] However, it does not hold for the Credit Suisse World ex-USA index, which includes Russian and Chinese markets that went to zero in 1917 and 1949. For it, the 20-, 30-, and 50-year worst case returns were -0.9, 1.6, and 0.9 percent, respectively.

[31] Personal communication from Mike Staunton, October 2017. The measurement was made with annual returns from their database.

[32] Reasons my correlation value is higher than in the SBBI include: 1) I use annual data keyed to January rather than annualized monthly data; 2) I use a different stock index, the CRSP total market index rather than the S&P 90 and S&P 500; and 3) my new corporate index corrects errors in the SBBI corporate index.

[33] Although inflation is not an asset per se (at least, not before the advent of TIPS), it is crucial to understanding the real return on assets, which in turn is crucial to understanding some of the regimes teased out in this paper.

[34] The quote is from Siegel (2014), pp. 79-80.

[35] To be clear, that is not what Siegel said, but only my estimation of what investors tend to infer from reading his account.

[36] And again: as seen in Figure 4, Siegel appropriately uses a log scale for inflation. But there is so much going on in his chart, and the scale needed to accommodate stock returns is so large, that when Siegel's chart is viewed, the record of 19th century inflation easily falls prey to the "ignore the lower left corner" heuristic.

[37] See Friedman (1952).

[38] Also, Friedman used wholesale prices, which gave more of a semblance of peaking in 1948 than seen in my annual chart of the consumer price level keyed to January.

[39] All these significance tests are t-tests that assume heterogeneous variance, except the earlier test of correlations, which used Fisher's *z* transformation.

[40] It might be thought that the inflation difference drives the real bond difference. But the difference across the first and second centuries for real stock returns is not significant ($t < 1$).

[41] All premia in the Ibbotson SBBI are calculated by ratio rather than subtraction. The ratio calculation that many investors will have previously seen concerns real return, given by [(1 + nominal return %) / (1 + inflation %)] – 1. Other sources discussing other kinds of premia may use subtraction, as in the Homer and Sylla (2005) treatment of corporate yield spreads, Table 56, pp. 412ff. But I thought it best to follow the SBBI approach.

[42] Available at: https://www.etf.com/publications/journalofindexes/joi-articles/3869-the-importance-of-investment-income.html. For a similar estimate and procedure, see Siegel (2005, p. 126). I do not mean to denigrate Mr. Bogle for picking the more

dramatic metric; that choice was appropriate to his rhetorical purpose. There also appears to be some rounding of either amounts or rates or holding period in the Bogle example. In my probes I used the exact compounding rates stated, 10.4% and 6.1%.

[43] See Shiller (2015) and Siegel (2005).

[44] Also, the dip in the years after WW I appears muted in Siegel's (2005, p. 126) chart, perhaps because of the inflation measure he used. The sharp dip to the low in the early 1920s in Shiller's (2015) chart (his figure 11.2, p. 210) looks more like mine.

[45] For the 120 years prior to the 1970s, banks mostly traded off-exchange, obscuring this distinctive feature of bank equity. In addition, none of the banks that failed in the 1930s were included in the CRSP database, in fact almost no banks at all, further obscuring from contemporary investors the bad things that can happen to bank stocks during times of crisis.

[46] It is also possible that the publication of the Modigliani - Miller theorem around 1960 created *de novo* the phenomenon that they had derived as timeless theory: that investors should be indifferent to whether they received their gains from dividends or some other source. Once most CFOs had been schooled in that theory, it is easy to see how the dividend regime might change. I am grateful to Meir Statman for pointing me to the Modigliani – Miller work as a potential transition point between regimes.

[47] I use this phrasing because I was able to produce the same finding when I compared corporate bond returns to those on government bonds: sometimes the riskier asset delivered the promised premium, sometimes it did not. See SSRN paper #3740190.