ANNUAL REVIEWS

Annual Review of Financial Economics

Long-Run Asset Returns

David Chambers,[1,2] Elroy Dimson,[1,3] Antti Ilmanen,[4] and Paul Rintamäki[5]

[1]Cambridge Judge Business School, University of Cambridge, Cambridge, United Kingdom; email: e.dimson@jbs.cam.ac.uk

[2]Centre for Economic Policy Research, London, United Kingdom

[3]European Corporate Governance Institute, Brussels, Belgium

[4]AQR Capital Management, Greenwich, Connecticut, USA

[5]Aalto University School of Business, Aalto University, Espoo, Finland

Annu. Rev. Financ. Econ. 2024. 16:435–58

The *Annual Review of Financial Economics* is online at financial.annualreviews.org

https://doi.org/10.1146/annurev-financial-082123-105515

JEL codes: G11, G12, G32, G38, N20

Keywords

asset pricing, historical returns, stock market index, investment management, risk premium, stocks, bonds, real estate, commodities

Abstract

The literature on long-run asset returns has continued to grow steadily, particularly since the start of the new millennium. We survey this expanding body of evidence on historical return premia across the major asset classes—stocks, bonds, and real assets—over the very long run. In addition, we discuss the benefits and pitfalls of these long-run data sets and make suggestions on best practice in compiling and using such data. We report the magnitude of these risk premia over the current and previous two centuries, and we compare estimates from alternative data compilers. We conclude by proposing some promising directions for future research.

1. INTRODUCTION

In this article, we survey the growing body of evidence on the magnitude and consistency of historical returns across the major asset classes of stocks, bonds, and real assets. Most financial research uses data extending back to the beginning of the last century; we try to look even further into the past.

There are two main benefits from taking a long-horizon perspective on asset returns (Chambers & Dimson 2016). The first is simply that it gives us more data at a stroke. Using as much historical data as possible helps investors form better return expectations. Statistically, a longer sample period gives more reliable average return estimates thanks to the law of large numbers and the resultant smaller standard errors. Second, we can test the robustness of in-sample results with the out-of-sample evidence that distant historical data provide. This is especially useful given that accumulating out-of-sample evidence as we move forward is painfully slow in a situation where the in-sample period has been thoroughly data mined.

At the same time, there are two important caveats to bear in mind when using long-run data. First, the longer the run of financial data, the more likely it is that regime shifts occur that complicate the interpretation of long-run average return estimates. Second, the quality of financial data deteriorates the further we retreat into the past. For example, when stocks and bonds are delisted, it can be impossible for scholars to discover whether the firm was acquired for value or whether it went bankrupt and investors lost all their money.

Pástor & Stambaugh (2001), Lettau, Ludvigson & Wachter (2008), and others have sought to identify and predict regime shifts. The latter—particularly in the case of more gradual structural changes—are obscured by the evolution of the investment landscape (expanding asset universes, falling trading costs) and by the changing nature of the investor community (the shift from individual to institutional investors, or enhanced investor protection).

Paying adequate attention to the quality of long-run historical data is often difficult. In this review, we highlight four common challenges in seeking to uphold data quality, namely, easy-data bias, macro-consistency, replicability, and total-return estimation. At the same time, we make suggestions for best practices when compiling, interpreting, and using such data.

The growth in the empirical literature on long-run asset returns has provided a considerable number of long-run data sets that are now at our disposal. Increasingly, data collection will be aided by artificial intelligence–based machine reading tools, so the number and variety of historical data sets can be expected to grow. These long-run data sets enable us to address questions using broader and deeper evidence than was possible until recently. The rest of this study deals with five important research questions.

First, we address one of the central questions in empirical asset pricing, namely, whether stocks consistently beat bonds in the long run. Much of this evidence is based on studies of data series starting in 1900, or notably, 1926, in the case of US stocks and bonds—the start date of the CRSP (Center for Research in Security Prices) data set. The choice of start date matters. The evidence regarding average returns drawn from post-1900/1926 studies delivered the conventional wisdom that the long-run equity risk premium over bonds was substantially positive in the United States and in every other country with a long history. However, in returning to this question, we examine the pre-1900/1926 history of stock and bond returns—primarily in the United States and the United Kingdom. The evidence suggests that the gap between the average returns on stocks and on bonds was low in the United Kingdom and seemingly negligible in the United States.

Second, we examine just how abnormally low bond yields were in the two opening decades of the twenty-first century. When viewing the yields of government bonds (or other bonds with a low credit risk) in a much longer-run perspective, we conclude that recent yield experience was indeed unprecedented.

The third question concerns corporate bonds and the credit risk premium. The existing evidence on the extent of the excess returns or premia attributable to credit risk is based on a relatively short history stretching back to the 1970s. We therefore ask whether (or not) long-horizon corporate bond data histories can provide insights on credit excess returns. In brief, the evidence is at best mixed due to estimation problems and a paucity of data.

Our two remaining questions deal with real assets—real estate and commodities—and how compelling their long-run returns have been compared to equities. When reviewing the evidence on the investment performance of real estate, we discuss the challenges presented by the heterogeneity and immovability of this asset class. The evidence has largely focused on housing and suggests that total returns are below those from equities. We then turn to commodities and critically examine the claims of the recent studies pointing to surprisingly high long-run returns. Here, we conclude that the long-run historical returns on a diversified portfolio of futures have come close to approximating those on equities.

While we focus on historical asset class premia, we only fleetingly cover their drivers or higher moments and the resulting asset pricing implications. The literature on using long data sets to relate macroeconomic and firm-specific factors to asset returns as well as to test various asset pricing models is beyond the scope of this survey.

We begin with a discussion of the challenges in handling long-run financial data in Section 2. In Section 3, we then deal with each of our important research questions that can be addressed with historical data. In Section 4, we conclude with some potentially promising directions for future research.

2. CHALLENGES OF LONG-RUN FINANCIAL DATA

While we are proponents of the use of long-run data sets, it is important to exercise care when using such data. The danger of not understanding the limitations of long-horizon data increases the longer the history. There may, for example, be structural changes in the characteristics of the data such as the number of stocks in the universe and the degree of sector diversification. It is therefore important for historical data sets to include as full a description as possible of the key characteristics of the underlying records and how they were processed.

Below, we briefly outline four common challenges when handling long-run historical data. They are various biases, macro-consistency, replicability, and total-return estimation. We focus on stock indexes, but similar issues arise in other asset classes.

2.1. Biases: Easy-Data, Survival, and Success

Researchers can be misled into assuming that long-term historical financial series are of a similar high quality to contemporary indexes. Dimson, Marsh & Staunton (2002) noted that severe data weakness can result from reliance on easy-to-collect data. Seemingly innocuous decisions, such as selecting a base date that is subsequent to a period of market trauma, can distort estimates of expected return. They highlight the fact that, in 2002, the standard publications on long-term stock market returns in the United Kingdom started at the end of 1918, in the Netherlands in 1947, in Germany in 1952, and in Japan in 1971. They also point toward index compilers' inclusion of markets that had grown, and omission of those that had been laggards. Looking at 16 standard sources, 1 for each country, the equity returns over the varying periods covered in prior studies exceeded the returns over the full period starting in 1900 by an average of 3 percentage points per year. Easy-data bias alone gave rise to estimated long-run equity premia that were almost double the unbiased estimate.

We need to be particularly vigilant about possible biases that could make a data set unrepresentative of the true long-term market experience we are seeking to document. That is, not just adding noise but adding bias. Easy data may expose the researcher to a variety of selection biases in terms of both the securities included and the countries studied. One cause of this bias is that the financial returns of stocks or countries are correlated with decisions about asset inclusion or exclusion.

Survivorship bias may be illustrated through comparisons of equity market country returns. Dimson, Marsh & Staunton (2021), and previously Brown, Goetzmann & Ross (1995) as well as Jorion & Goetzmann (1999), stressed that US equities have been among the best-performing stock markets over the long term, outpacing the rest of the world by 2–3% per annum. Not coincidentally, the United States has until recently been the primary focus of most academic research. Meanwhile, markets that lacked full capital market histories during the past century and lagged in performance comparisons—including Russia, China, Argentina, and Zimbabwe—have often been excluded from historical studies of global equity market performance. As a result, researchers may have reported an unrealistically favorable picture of long-run equity returns.

Of course, the fact that the United States and United Kingdom are recognized as markets with relatively accessible historical data going back to the 1800s and 1900s also reflects success bias. Easy-data bias points in the same direction as success bias, which often means that historical samples begin after some difficult periods when asset valuations may have been low. Some studies omit observations that were outliers or uninvestable: Periods that were chosen with hindsight. The German hyperinflation of 1923 is one example. Such issues are hard enough to address with recent data; they may be unresolvable with nineteenth-century data.

2.2. Macro-Consistency

Early indexes were rarely weighted by the market value of each constituent. They were typically unweighted—computed on an equally weighted basis—or alternatively, like the Dow Jones Industrial Average in New York, were sometimes weighted by the price of each constituent stock. Many other indexes followed investment approaches that could not be replicated by investors, including the *Financial Times* 30-share index (FT Index; a UK index launched in 1935) and the Value Line Index (a US index launched in 1961).

In the UK, the FT Index is geometric and by construction it must underperform an arithmetic index with the same constituents and weights. Consistent with this, Marks & Stuart (1971) calculated an arithmetic version of the FT Index from 1935 to 1970 and estimated a level for the index that was 38% higher than the actual geometric index. In 1971, a constituent declared bankruptcy and the level of the FT Index fell to zero, which was corrected by reconstituting the index retrospectively. The Value Line Index was also geometric, and replicating its performance was impossible, as shown by Brennan & Schwartz (1985), Eytan & Harpaz (1986), and Thomas (2002).

Differing philosophies have been followed in assembling consolidated research databases. The Dimson, Marsh & Staunton (2024) approach has been to adopt broad commercial indexes supplemented by peer-reviewed back histories for each country. Taylor's (2023) approach on behalf of Global Financial Data (GFD) has been to follow a sampling procedure based, where possible, on the largest feasible number of stocks in each market—but capped at a maximum of 100 constituents. These two approaches are illustrated in the sidebar titled Example of Differing Index Design Philosophies, which reproduces for these two index compilers their descriptions of the respective US equity series. Similar details are provided by Dimson, Marsh & Staunton (2024) and Taylor (2023) for other markets.

Indexes can be constructed in many ways. Most useful are indexes that measure the investment performance of a portfolio that contains every security in the investment universe, with each one held in proportion to its outstanding value. Such market capitalization–weighted indexes are macro-consistent (Sharpe 2010). They reflect the fact that the average investor must hold the market portfolio (Cochrane 2011).

EXAMPLE OF DIFFERING INDEX DESIGN PHILOSOPHIES

The Dimson, Marsh & Staunton (2024) Approach

"The broadest index of early US stock market returns is the one presented in Wilson & Jones (2002), which we use for this study. Our data set commences with the Wilson-Jones index data over 1900–25. For 1926–61, we use the University of Chicago's Center for Research in Security Prices (CRSP) capitalization-weighted index of all New York Stock Exchange stocks. From 1962 onward, we use the CRSP capitalization-weighted index which includes stocks quoted on all major exchanges including the NYSE, the American Stock Exchange (now NYSE AMEX Equities), and since 1972 Nasdaq."

—Dimson, Marsh & Staunton (2024, p. 270)

The Global Financial Data Approach

"We use the GFD-100 Index to calculate returns to stock markets in the United States. . . . The GFD-100 index is revised in January of each year. It uses the largest stocks in the United States from 1792 until 1825, the 50 largest stocks from 1825 until 1850 and the 100 largest stocks from 1850 to date. The index is capitalization-weighted and only includes liquid stocks that trade on a regular basis. Stocks from all regional exchanges and the over-the-counter market are included."

—Taylor (2023, p. 244)

When we use a historical index, we therefore need to know how the index is defined. Does it represent a broad all-stock universe or a narrower subset? What has been included, which stocks have been omitted, and why? When there is a choice between broader and more focused coverage, we favor using an all-share index, which provides the broadest coverage in every country. Indexes should wherever possible be value-weighted, not equal-weighted, and index composition should reflect the experience of real-world investors. This preference is challenged by the difficulty of creating stock market indexes for early periods.

2.3. Replicability

Stock market indexes should be designed so that, without any foresight about future performance, investors can replicate index composition with real money. We have already discussed index design and the importance of having a value-weighted arithmetic index rather than, say, a geometric index. We now turn to the challenge of stocks that disappear from databases—such securities often have low returns toward the end of their lives. What should a financial historian do to estimate stock market performance when stocks are deleted?

The crudest approach is to exclude stocks that, with the benefit of hindsight, no longer exist. This is equivalent to assuming the investor is clairvoyant—that she exits from a holding before a disaster occurs. When delistings occur, researchers may assign over-optimistic estimates of pre-delisting returns. Grossman (2002) showed a wide range of UK equity return estimates for the period 1870–1913, depending on the delisting return assumptions used. Shumway (1997) showed that, because of delisting bias, estimates of the long-run US equity return, and especially the small-cap premium, were significantly overstated. Foreman-Peck & Hannah (2012) and Campbell, Grossman & Turner (2021) reported that a standard equity index used in the United Kingdom had serious omissions that made the index "inadequate" (Gregory 2024, p. 5).

Even transitory or occasional omissions matter. Pioneers will frequently see their early work as financial archeologists criticized and their foundational research subsequently improved upon by

THE DIFFERENCE BETWEEN MCQUARRIE'S INDEX AND SIEGEL'S

"Stock returns before 1871 look different now because of the reduction in survivorship bias. The old record omitted the largest 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 US market capitalization. It failed as the Panic proceeded, with shares dropping in price from $120 to $1.50, and never recovered."

—McQuarrie (2024, p. 14)

subsequent scholars. In the sidebar titled The Difference Between McQuarrie's Index and Siegel's, McQuarrie (2024) compared his new index for US common stocks in the nineteenth century with Siegel's (2014) earlier computations. He reveals that even one missing stock can have a devastating impact on an estimate of long-run returns. Taylor (2018) provides additional detail.

Omissions can be numerous. Zweig (2009) reported that "Siegel exclude[d] 97% of all the stocks that existed in the earliest years of the US market." He noted that "more than 300 companies in various industries had been launched in the US by 1801. Most ended up fading into oblivion, leaving no record of their hapless investors' losses. Yet the earliest of the data in *Stocks for the Long Run* tracks only seven bank stocks—all of which were selected because they did not go bust" (Graham & Zweig 2024, chapter 3). The large number of omissions from the index meant that it failed the test of macro-consistency. Moreover, unless an investor was clairvoyant—knowing which firms were destined to disappear—the investment strategy of the index could not have been replicated in real life.

Many real-world indexes apply liquidity screens that take account of constituent information from before an index is rebalanced. Historical reconstructions, such as those estimated by GFD, can require a prospective constituent to have a minimum number of observations over the following year or to have other information for the year after rebalancing (Taylor 2018). Taken together with rebalancing bias (e.g., Roll 1981; Lyon, Barber & Tsai 1999), there may be a significant gap between estimated index returns and true buy-and-hold returns. We defer this concern to future research.

The expanded interest in research on long-term investment returns, supported by extensive digitization of historical data sources, has given rise to multiple approaches to creating and analyzing security price data. There are single-country studies, such as those mentioned above for the United States and United Kingdom. More recent initiatives include the work of Annaert, Buelens & Riva (2016), which covered three centuries of the Belgian Stock Exchange history, and Hautcœur & Riva's (2018) comprehensive database on the French stock markets since 1795.[1]

2.4. Total-Return Estimation

The real return on a portfolio held over the long haul is driven by the income it generates. The importance of income was recognized in the first careful studies of returns on the New York Stock Exchange by Fisher & Lorie (1964, 1968), who computed not only capital appreciation but also total returns including dividends. Ibbotson & Sinquefield (1976) also noted the importance of dividend income in their landmark study of US stock returns. In general, as we go back in time, asset price data are more easily available than dividend data, and so price-only return estimates tend to

[1]For more information, see the Study Center for Companies and Exchanges (SCOB) at **https://www.uantwerpen.be/scob** and the Data for Financial History (DFIH) database at **https://dfih.fr**, respectively.

be more readily compiled than total-return estimates. Good estimates of dividend income are difficult to assemble, especially for intervals from long ago and for markets with incomplete records.

Jorion & Goetzmann's (1999) analysis of investment returns from stock markets around the world was based on long-term capital appreciation indexes in 39 countries. They noted that the inclusion of dividend income would elevate return estimates in both the United States and other countries. Moreover, small adjustments to estimated dividend yields have a sizeable impact on long-term returns. Zweig (2009) highlighted inaccuracy in dividend estimates in Siegel's (1992) early data and the resulting upward bias in computed returns.

When we consider pre-1900 data, income return dominated price appreciation in the total real returns of stocks, bonds, and real estate. However, income or payout data are harder to source. Collecting data on realized as opposed to contractual income payments becomes particularly challenging in the case of coupons on bonds and rents on properties. In the case of bond defaults and of vacated properties, identifying recovery rates and the length of rental voids may not be possible. Plugging in the contractual cash flows or interpolating data will overstate realized returns.[2]

Finally, we close this section with some remarks on best practice, in the spirit of Annaert, Buelens & Riva (2016) and Baltussen, van Vliet & van Vliet (2023). Scholars should do the best they can with the available historical data, with all its inherent problems and shortcomings. At the same time, there should be a clear disclosure of data sources and design and measurement choices, while highlighting challenges, potential biases, and interpolation or back-casting of data. Good practice involves robustness checks of reported results reflecting the different possible design choices. Last and not least, we should give full credit to earlier researchers, especially those who have undertaken the painful effort to create and clean data sets and to make them available to others.

3. THE BIG QUESTIONS

With long and broad data sets, we can take a fresh look at many issues. Our choice of questions discussed below represent those that the authors believe are of first-order importance to academics and practitioners alike and that have been directly addressed in the prior literature using data from the past century.

3.1. Equities: Was the Twentieth Century an Anomaly?

An abundance of evidence has shown that the total returns of equities have exceeded those of (government) bonds over various long-run periods since the beginning of the twentieth century—both in the United States (Ibbotson & Sinquefield 1976; Siegel 1994, 2022) and in the rest of the world (Jorion & Goetzmann 1999; Dimson, Marsh & Staunton 2002). First, we address one of the central questions in empirical asset pricing, namely, whether stocks consistently beat bonds over the long run. As we noted above, this evidence relies primarily on data series starting in 1900 or, in the case of the United States, 1926, when the University of Chicago's CRSP data starts. In this section, we first discuss the history of stock and bond returns before 1900/1926 and critically review these out-of-sample estimates of the equity risk premium before making comparisons with twentieth-century estimates. As well as looking at the United States, we consider the United Kingdom—the home of the premier stock and government bond markets in the nineteenth century.

[2]Apart from total versus price-only returns—or nominal versus real returns, or arithmetic versus geometric means—many return measurement issues have received increasing research attention but are beyond the scope of this article. Bessembinder et al. (2023) summarized issues related to the measurement horizon, the rebalancing strategy, the reinvestment rate assumptions of interim cash flows, as well as the gap between investment and investor return.

Below, we review the long-run evidence on the total returns achieved by stocks. The ideal measure of stock returns is the compound average total return, including price changes and any payouts to existing shareholders, for a market capitalization–weighted universe of all listed stocks, adjusted for delistings and corporate actions.

Researchers have undertaken admirable detective work in physical and digital archives to gather data and construct estimates of nineteenth-century stock returns. In the United States, the first pre-CRSP evidence on stock returns drew on the efforts of Smith & Cole (1935), Macaulay (1938), and Cowles (1938). These estimates were later used by Schwert (1990) and Siegel (1992) among others as their pre-CRSP source and subsequently modified by Goetzmann, Ibbotson & Peng (2001). Most recently, Baltussen, van Vliet & van Vliet (2023) have made a detailed and valuable contribution to our understanding of US stock returns from 1866 to 1926 based on a careful data collection exercise. McQuarrie (2024) has made a similar heroic effort going back to 1800 for the United States (with much help from Sylla, Wilson & Wright 2005).

Similarly, in the United Kingdom, considerable efforts have been made to extend stock return histories back to 1829 (Acheson et al. 2009; Campbell, Grossman & Turner 2021) and to preceding periods (Campbell et al. 2018). An even earlier equity series (Golez & Koudijs 2018) can be used to extend the Campbell, Grossman & Turner (2021) indexes back to 1800. Thomas & Dimsdale (2018) provided the corresponding bond returns and inflation rates, reflecting the adjustments proposed by Klovland (1994).

Table 1 summarizes the annualized nominal returns of equities and bonds, the equity risk premium (over bonds), and the annualized real returns of equities over 1800–1899, 1900–1999, and 2000–2023. It covers the US returns reported by Siegel (2014) and later roughly reconstructed by McQuarrie (2024), accompanied by estimates from Taylor (2023) and Dimson, Marsh & Staunton

Table 1 Annualized percentage returns on US stocks and bonds from 1800

Period	Siegel (2014)	McQuarrie (2024)[a]	Taylor (2023)	Dimson, Marsh & Staunton (2024) + extension[b]	Siegel (2014)	McQuarrie (2024)[a]	Taylor (2023)	Dimson, Marsh & Staunton (2024) + extension[b]
	US equities nominal return				US bonds nominal return			
1800–1899[c]	6.8	5.5	6.8	5.5	5.1	6.2	5.5	6.2
1900–1999	10.5	10.2	11.2	10.3	4.8	4.9	4.4	4.5
2000–2023[d]	7.5	6.9	6.5	7.0	5.1	5.2	4.0	5.2
All	8.5	7.7	8.7	7.8	5.0	5.5	4.9	5.3
	US equity-bond premium				US equities real return			
1800–1899[c]	1.6	–0.6	1.2	–0.6	7.2	5.9	6.7	5.9
1900–1999	5.5	5.1	6.5	5.5	7.3	7.0	7.8	7.0
2000–2023[d]	2.2	1.6	2.5	1.7	4.8	4.3	3.9	4.4
All	3.4	2.1	3.7	2.3	7.0	6.2	6.9	6.3

Sources: McQuarrie (2024), including his estimates of Siegel (2014); Taylor (2023); Dimson, Marsh & Staunton (2024); and the authors of this article. All returns are annualized geometric means in local currency.

[a]McQuarrie (2024) roughly reconstructs Siegel (2014) data from original Smith & Cole (1935), Schwert (1990), Goetzmann, Ibbotson & Peng (2001), Cowles (1938), and the Center for Research in Security Prices (post-1926) data.

[b]The Dimson, Marsh & Staunton (2024) + extension series merges the McQuarrie (2024) data up to 1899 with Dimson, Marsh & Staunton (2024) estimates for 1900–2023.

[c]For 1800–1801, we use Siegel (2014) data to extend McQuarrie (2024) data back to 1800.

[d]For 2020–2023, we use Dimson, Marsh & Staunton (2024) data to extend the McQuarrie (2024) and Siegel (2014) series forward.

Table 2 Annualized percentage returns on UK stocks and bonds since 1800

Period	Three UK studies (see sources footnote)	Taylor (2023)	Dimson, Marsh & Staunton (2024)	Dimson, Marsh & Staunton (2024) + extension[a]	Golez & Koudijs (2018) and Campbell, Grossman & Turner (2021)	Taylor (2023)	Dimson, Marsh & Staunton (2024)	Dimson, Marsh & Staunton (2024) + extension[a]
	UK equities nominal return				UK bonds nominal return			
1800–1899	5.4	5.6	N/A	5.4	4.0	4.0	N/A	4.0
1900–1999	N/A	10.2	10.2	10.2	N/A	4.9	5.4	5.4
2000–2023	N/A	4.6	4.5	4.5	N/A	3.7	3.9	3.9
All	5.4	7.5	9.1	7.4	4.0	4.4	5.1	4.6
	UK equity-bond premium				UK equities real return			
1800–1899	1.4	1.5	N/A	1.4	5.4	5.8	N/A	5.4
1900–1999	N/A	5.0	4.6	4.6	N/A	6.0	6.1	6.1
2000–2023	N/A	0.9	0.6	0.6	N/A	1.1	2.0	2.0
All	1.4	3.0	3.8	2.7	5.4	5.4	5.3	5.5

Sources: Taylor (2023), Dimson, Marsh & Staunton (2024), and three UK studies: Golez & Koudijs (2018) over 1800–1829, Campbell, Grossman & Turner (2021) over 1830–1899, and Thomas & Dimsdale (2018) for 1800–1899 bond return and inflation data.

All returns are annualized geometric means in local currency.

Abbreviation: N/A, not available.

[a]The Dimson, Marsh & Staunton (2024) + extension series merges the Golez & Koudijs (2018) and Campbell, Grossman & Turner (2021) 1800–1899 equity returns, as well as Thomas & Dimsdale (2018) 1800–1899 bond and inflation data, with Dimson, Marsh & Staunton (2024) estimates for 1900–2023.

(2024). All the quoted figures are geometric means (compound returns) and thus lower than arithmetic means (simple returns), especially for more volatile series.

The main message is that the US equity premium over (government) bonds at and above 5% in the 1900s was considerably higher than estimates for the 1800s ranging between +1.6% and –0.6%. The disagreement between these two figures is due to differing estimates of equity and bond returns in roughly equal measure. First, Siegel (1992, 2014) overestimated equity returns and (as noted above) his estimate suffers from both survivorship bias and a crucial stock omission. Second, there are problems with Siegel's (1992, 2014) attempts to proxy risk-free bonds in the nineteenth century.

McQuarrie (2024) eschewed such a risk-free proxy in favor of the return on a bond index, drawing on a comprehensive data set of municipal, federal, and corporate bond returns. This will inevitably capture some credit premium in municipal and corporate bonds, and not a pure riskless return, partly explaining McQuarrie's (2024) higher bond return and lower equity premium in the early data.[3] Hence, the 1800–1899 US equity premium over truly riskless bonds would probably not be a negative figure but most likely a small positive number and still a lot smaller than the premium for the 1900s.[4] Given that government bonds did not exist during the nineteenth century in the United States, we look next at evidence for the United Kingdom, where default-free government bonds did exist.

Table 2 provides geometric mean returns for the United Kingdom over 1800–1899 for equities (Golez & Koudijs 2018; Campbell, Grossman & Turner 2021) and for bonds (Thomas & Dimsdale 2018), accompanied by estimates up to 2023 (Taylor 2023; Dimson, Marsh & Staunton 2024). The Dimson, Marsh & Staunton (2024) + extension series links the Golez & Koudijs (2018) and the

[3]All bond sources use long-dated government bonds since the 1920s, though with different maturities.

[4]Difference-in-means tests show that the equity-bond premium was significantly lower in the 1800s than the 1900s, with t-statistics ranging from –2.1 to –3.6.

Campbell, Grossman & Turner (2021) equity series and the Thomas & Dimsdale (2018) bond series for 1800–1899 with Dimson, Marsh & Staunton (2024) estimates for 1900–2023.

The UK equity-bond premium for the 1800s is estimated near 1.5% (Golez & Koudijs 2018; Thomas & Dimsdale 2018; Campbell, Grossman & Turner 2021). We consider the UK estimate for the nineteenth century to be a more reliable measure of the equity premium than the US estimate, given that the bond returns are based on obligations of the British government (so-called Consols)—the prime credit available to investors throughout the 1800s. The 224-year estimate of the annualized UK equity-bond premium is in the range of 2–3%, which is again well below the twentieth-century premium of 4–5%.[5]

Recent empirical studies of equity and bond returns in the nineteenth century for Belgium and France—both were among the most important capital markets in the world at this time—also indicate a low premium compared to the twentieth century. Annualized nominal returns for Belgian equities and government bonds are estimated as 4.0% and 4.4%, respectively, over 1838–1900 (Van Mencxel, Annaert & Deloof 2024, table 5). This slightly negative equity-bond premium compares with 3% for 1900–1999 (Dimson, Marsh & Staunton 2024, table 22). In France, annualized nominal returns over 1854–1914 for equities were estimated at 5.3% and for government bonds at 4.0%, implying an equity-bond premium of 1.3% (Le Bris & Hautcœur 2010, table 10). This compares with 4.3% for 1900–1999 (Dimson, Marsh & Staunton 2024, table 35). In short, the robust and substantially positive equity premium of the 1900s—almost pervasive across countries in the Dimson, Marsh & Staunton (2002, 2024) data set—was much smaller and quite close to zero in the previous century.

A full discussion of the reasons for the lower equity returns in the nineteenth versus the twentieth century is beyond the scope of this review, but here are some thoughts. Mechanically, it could reflect lower growth—and, indeed, the 1800s experienced lower growth in per capita income, lower productivity growth, and lower dividend growth than the 1900s. The payoff from the first industrial revolution (steam) of the nineteenth century was modest compared to that of the second (electricity and the internal combustion engine) in the early twentieth century and that of the third (information and communication technology) around the turn of the last century.

A decomposition analysis of total equity returns supports such claims about the relatively poorer fundamentals of the 1800s. From a return source perspective, the components of total real equity returns are the dividend yield, the real growth in dividend per share, and the change in equity valuation. The decompositions for the United States and the United Kingdom in the 1800s and 1900s show that dividend yield is overwhelmingly the most important contributor to returns in both centuries.[6] The real growth in dividends and the upward trend in equity valuations of the twentieth century were nowhere as apparent in the earlier period.

From a rational asset pricing perspective, these growth arguments are beside the point. Instead, lower average equity premia over a century suggest less risk in the 1800s than in the 1900s. This

[5]There are fewer sources of short-term bill returns than long-term bond returns before 1900, so we do not include them in **Table 1** and **Table 2**. GFD estimates a 1.2% lower return for US bills than bonds in the 1800s, reflecting a healthy term premium (compared to a term premium of 0.7% in the 1900s and 2.3% in the 2000s). The term premium was lower in the United Kingdom: 0.2% in the 1800s, near 0% in the 1900s, and near 2% in the 2000s. Chen et al. (2023) discussed the shifting location over several centuries of the world's dominant safe-haven asset supplier—from Holland to the United Kingdom to the United States.

[6]Total equity return can be decomposed in increasingly granular ways. The most common decomposition is in the spirit of the dividend discount model, splitting the return into dividend yield (d/p), dividend-per-share growth, and valuation (p/d) change components (e.g., Campbell 2018). Dimson, Marsh & Staunton (2024, table 12) presented this decomposition for 21 national markets over the period 1900–2023. Kuvshinov & Zimmermann (2022) documented the declining role of dividend income in stock returns across centuries.

seems hard to justify, especially in the United States, given political instability and immature markets. That said, measured equity market volatility was lower, and bonds were hardly riskless in the 1800s (with their high real return raising the bar for stocks).

Moreover, realized average stock returns may not reflect expected returns even over a century-long history, especially if there was significant repricing. The twentieth century saw many improvements in corporate governance and disclosure necessary to induce outside equity ownership as well as improved equity investability (via pooling and diversification) and higher investor risk tolerance. All of these developments likely contributed to the general rerating of equities and to stocks outpacing bonds over the century. Realized stock returns were boosted by falling required returns (rising valuation multiples) and thus likely exceeded the returns investors had been expecting.[7]

To sum up, the equity risk premium was exceptionally high in the twentieth century, while being low in the nineteenth century—notably in the United States. The reasons for these differences are worthy of further research. Both centuries may have been abnormal. Consistent with this, leading experts at a recent equity premium symposium estimated a prospective equity-bond premium of between 0% and 5% (Siegel & McCaffrey 2023, appendix A). The best guidance on future premia may be between the experience of the lucky 1900s and that of the disappointing 1800s.

3.2. Bonds: How Abnormal Were the Recent Low Bond Yields?

Having looked at bond returns in the previous section, here we consider bond yields in a long-run perspective. The finance literature has anchored on the fact that recent bond yields appeared abnormally low compared to the past and that the common investor response was to reach for yield.[8] Certainly, bond yields in the past decade were as low as they had been since 1900 (reaching negative real yields in major bond markets, even negative nominal yields in Japan, Germany, and the United Kingdom). Given that bond markets have existed for much longer than stock markets, what does the pre-1900 data tell us about just how unusually low real yields became?

The longest data sources relating to loan or bond yields are by Homer & Sylla (2005) and Schmelzing (2020). The former dates back to ancient times and the latter to the early fourteenth century. However, interpreting these estimates of historical bond yields is problematic. Bond contracts in the nineteenth century incorporated many features that make estimating a consistent series of bond yields very difficult. Beyond maturity and coupon, there are complications such as sinking funds and embedded options (mostly for the issuer's benefit), tax features, and even lotteries. Furthermore, bond ratings for either government or corporate bonds were not introduced until the 1900s, making credit risk virtually impossible to estimate before then.

Notwithstanding such challenges, Schmelzing (2020)—through his own efforts as well as those of others—assembled a global history of yields on loans or bonds since the early 1300s. We focus on GDP-weighted average yields across eight countries once each became available: Italy (1310), Great Britain (1310), Germany (1326), France (1387), Spain (1400), the Netherlands (1400), United States (1786), and Japan (1870). Among the six European states, only France

[7]The cyclically adjusted price/earnings ratio of US equities rose from 18.5 at end-1899 to 44.2 at end-1999, contributing to roughly 1% annual repricing gains. Moreover, both average growth and average inflation likely surprised on the upside (cf. **Table 3**).

[8]Historical analysis shows that there was nothing new in the 2010s reach-for-yield activities. This is what investors have done repeatedly when yields fall to abnormally low levels. Korevaar (2023) covered Dutch investors in the 1600–1700s, and Turnbull (2017) discussed British investors in the 1800–1900s. Chancellor (2022) quoted Walter Bagehot most memorably from the 1850s: "John Bull can stand many things, but he can't stand two percent" (p. xxii).

Table 3 Centuries of declining nominal and real bond yields (%)

	1314–1399	1400–1499	1500–1599	1600–1699	1700–1799	1800–1899	1900–1999	2000–2018
Nominal rate	11.5	12.2	9.2	7.4	5.1	5.4	5.4	3.0
Real rate	10.4	11.5	7.1	6.8	4.2	4.9	1.0	1.3
Inflation (7-year, centered)	1.2	0.8	2.1	0.6	0.9	0.5	4.4	1.7
Real growth (GDP %)	0.4	0.1	0.2	0.2	0.2	0.9	2.2	1.1

Sources: Schmelzing (2020) and Maddison-Groningen Database (2020).

All series are GDP-weighted averages across countries and arithmetic means across years of annualized rates from (up to) eight countries since data are available, in most cases before the birth of the nation. Nominal rate refers to loans or bonds, and real rate is the difference between the nominal rate and the 7-year centered average inflation rate. The last series, real GDP growth per capita, is from the Maddison-Groningen Database (2020) and where possible reflects the same GDP-weights as the Schmelzing (2020) series.

existed when the time series began, so the loans in the early sample were to local rulers, princes, or cities. As capital markets evolved over time, loans were replaced by tradeable bonds, and for the past century, the data set covers 10-year government bonds. Earlier, maturity and credit quality varied over time and across countries.

Table 3 tracks the evolution of nominal and real (inflation-adjusted) yields as well as inflation and real GDP growth. [The last series is from the Maddison-Groningen Database (2020), to be discussed below.] The table shows that nominal and real yields were in double digits in the Middle Ages but fell to approximately 5% near the Industrial Revolution and ever lower thereafter.[9] Global inflation was mildly positive but near zero in most centuries before the 1900s. Centennial average inflation near zero did not mean stable inflation near zero, as there were oscillating waves of inflation and deflation.

Looking at Schmelzing's (2020) data in **Table 3** reveals that, before the twentieth century, inflation was very low and there was little difference between nominal and real yields, which both fell to 3–4% around 1900. The twentieth century saw more persistent price rises, particularly the Great Inflation of the 1970s, which pushed nominal rates up to double digits—a level last seen in the 1500s. The post-1982 Great Moderation witnessed global convergence in inflation rates toward 2%—back to the low levels that characterized the pre-1900 economy.

Even after inflation expectations stabilized at low levels in the 2000s, real yields and policy rates kept falling to negative territory. Eventually, even nominal yields in many European countries and Japan turned negative, or reached record low levels in places where they didn't (the United States and the United Kingdom). The global rise in inflation in 2021–2022 finally turned bond yields higher after a 40-year downtrend. The yield evidence in **Table 3** stretching back over centuries provides confirmation that the recent low (even negative) bond yields were truly exceptional.

Alongside real bond yield data, we have very-long-run estimates of income per capita growth across countries assembled by the Groningen Growth and Development Centre (see Maddison-Groningen Database 2020). Classic economic theories would suggest that equilibrium real yields and real growth should be similar. However, very-long-run empirical evidence suggests the opposite.[10] Real economic growth was 0.1–0.4% per year before the Industrial Revolution, whereas the past two centuries have seen approximately 1–2% real annual growth in advanced economies (last

[9]Looking even further back, Homer & Sylla (2005) documented yields near 20% in Mesopotamia 4 or 5 millennia ago. A local trough occurred amid Pax Romana near Common Era 1 with maybe 4% yields, before double-digit yields again became commonplace.

[10]From Wicksell (1898) to Woodford (2003), the natural (equilibrium) real rate is tied to the productivity of capital, serving as one key anchor to monetary policy rules. In neo-Keynesian dynamic stochastic general

row of **Table 3**). Thus, the recent two centuries of relatively fast economic growth and higher inflation coincided with low real yields, while the earlier centuries of slow growth and low inflation witnessed much higher required yields.[11]

This apparent puzzle of high real interest rates in a slow-growth world may be explained by capital scarcity, the impatience of savers at a time when most people lived near subsistence level, as well as high financial intermediation costs (Bernstein 2022). Most economists studying the fundamental determinants of real interest rates have focused on the expected consumption growth rate rather than time preference. While this is appropriate in the modern world, it seems less so in preindustrial times. Rogoff, Rossi & Schmelzing (2024) explored the Schmelzing data for structural breaks but find none after the Industrial Revolution; instead, real yields are trend stationary. Like us, the authors highlight opposing trends in real yields and real growth across centuries.[12]

3.3. Do Long Histories Give Us Clear Evidence on Excess Returns to Credit Risk?

We next turn to corporate credits. The credit premium is defined as the (expected or realized) excess return of corporate bonds over comparable government bonds. We discuss evidence on both realized excess returns and forward-looking credit spreads, and since the latter need to be adjusted for defaults, we also comment on the realized default experience. For much of the long histories we study there are no available credit ratings, so we cannot split the results by rating class or between investment-grade and speculative-grade bonds. The same data issues around accurately capturing the various features of bonds discussed above in the context of bond yields make it nearly impossible to measure the credit premium in early corporate bond markets.

The financing needs of the railroads stimulated the growth of large corporate bond markets from the mid-nineteenth century onward in the United States and the United Kingdom. Consequently, there are a few studies that add to our knowledge in this area. Giesecke et al. (2011) estimated annual value-weighted percentage default rates for US non-financial corporate bonds over 1866–2008.[13] The authors reported the annual time series of default rates and observed a decline over time with rates averaging 4%, 1.4%, and 0.3% for 1866–1899, 1900–1945, and 1946–2008, respectively. To estimate the credit spread, they utilize Homer & Sylla's (2005) yield data on high-grade railroad bonds and high-grade New England municipal bonds for 1866–1914 and prime corporate bonds and long-term government bonds after 1914. These category definitions can only crudely take account of credit differences across firms in the era before bond rating. We also again note the problems of proxying US government bond yields pre-1914. With this data, the authors reported an average credit spread of 1.5% for the entire sample period 1866–2008 (without disclosing the annual time series). The average default rate is also 1.5%. Given their broad-brush assumption of a 50% average recovery rate of defaulted losses, they estimated the

---

equilibrium models, this rate reflects productivity growth and investor time preference. Consumption-based asset pricing links the riskless real rate to consumption growth (as well as time preference and Jensen's inequality terms; see Campbell 2018).

[11]This evidence contradicts Piketty's (2013) claim that the rate of interest (or broader investment return) always exceeds the rate of economic growth, $R > G$. We bypass this topic because we could debate endlessly about which interest rate and which growth rate to use.

[12]The study also discusses estimation of inflation expectations (needed for the computation of real yields). We cannot do justice to the related literature here beyond noting that the challenge is made harder by structural changes in inflation, growth, and real rate processes. Farmer, Nakamura & Steinsson (2024) stressed how pernicious gradual structural changes are since we learn about them very slowly.

[13]Van Mencxel (2023) estimated default rates for Belgian corporate bonds in the late nineteenth and early twentieth centuries.

average credit risk premium to be approximately 0.75%.[14] This figure compares with an average credit premium of 1% for investment-grade corporate bonds from 1973 to 2020 (Ilmanen 2022).

Coyle & Turner (2013) reported annualized real returns on UK corporate bonds over the period 1861–2002 of 2.4% and a credit spread of approximately 1% versus government bonds. The spread was higher at 1.7% for 1861–1913 than afterward. Since ex post default losses were negligible, the credit premium is of a similar magnitude. As with the United States, however, there is no available information on corporate credit quality in the nineteenth and early twentieth centuries. Given the sample comprised all corporate bonds traded on the London stock market for which listing requirements were relatively strict, we might infer that credit quality of this bond sample was generally high.

A further data point on corporate bonds was provided in the recent study by Van Mencxel, Annaert & Deloof (2024) on Belgium from 1838 to 1939. As mentioned above, the Brussels stock market was relatively large in the 1800s, and there is an abundance of data. This study reports overall a small positive corporate credit spread over government bonds of approximately 0.4%. There is no discussion of default losses and recovery rates to enable us to estimate the credit premium.

Overall, we can only measure corporate credit spreads and credit premia quite crudely. In the absence of bond rating data for corporate bonds before the 1920s, it seems difficult for scholars to be able to provide more precise estimates of the corporate credit premium than the few we have. The existing sparse evidence suggests that there was a modest credit premium in the nineteenth as well as the twentieth century. Given that Kizer, Grover & Hendershot (2019) and McQuarrie (2020) found problems even with the US corporate bond return data in the early decades after 1926, it is likely to remain the case that the most reliable evidence on the credit premium begins only in 1973.

Finally, we do have an estimate of the sovereign credit premium. Meyer, Reinhart & Trebesch (2022) have estimated a credit spread before default losses of 4.3% on foreign-currency government bonds between 1815 and 2016 over UK and US government bonds. This spread was approximately 1% lower for the first 100 years, 1815–1914. In credit events (sovereign debt restructurings), the typical realized haircut was less than half, implying an average annual default loss near 1%. The resulting average ex post sovereign credit premium of approximately 3% is considerably above that for corporate bonds.

3.4. Real Estate: Do Housing Returns Match Equities in the Long Run?

Much of the academic literature regarding real estate performance has focused on housing, given its importance within household wealth. There are numerous historical studies of house price changes over time.[15] However, very few attempts have been made to estimate total returns, including rental yield net of costs, to housing over long time periods largely due to the data demands of making such estimates. Total return is the focus of the discussion in this section.

There are two main types of historical data that are employed to analyze long-run housing returns. Macro studies seek to exploit the aggregated national statistical index series of house prices and rents. Jordà et al. (2019) collected much of this macro data in their impressive multicountry study of housing returns (alongside stocks and bonds) from 1870 to 2015.[16]

[14]Giesecke et al. (2011, p. 243) cited research by Hickman (1960) that implies an average recovery rate of 62.5% of par value for 1900–1944.

[15]The historical US house price series of Shiller (2005) are the best known but have subsequently been improved upon by, for example, Lyons et al. (2023).

[16]In the MacroFinance and MacroHistory Lab, **https://www.macrohistory.net**, the authors provide a free multicountry data set of their collected asset class returns and select macroeconomic series since 1870.

In contrast, micro studies seek to exploit detailed, property-level transaction prices, realized rents, and the associated costs to the landlord typically in a specific location. Such archival evidence of the required quality and without discontinuities is harder to source for properties than for stocks or bonds. As a result, there have been only a handful of micro studies of total returns to date—Paris over 1809–1943 and Amsterdam over 1900–1979 by Eichholtz et al. (2021) and England over 1901–1983 by Chambers, Spaenjers & Steiner (2021).

There are five important differences when estimating total returns on housing compared to publicly traded stocks and bonds. First, the price formation and trading process of real estate are different. Trading is infrequent, and market values are much more difficult to establish consistently. Hence, investment returns can only be estimated with noise.

Second, (residential) real estate prices reflect private value, namely, a use value derived from ownership that is nonmonetary and nontradable.[17] It is private in the sense that it depends on the identity of the owner. For private-value assets, any two potential buyers will be willing to pay different amounts—reflecting differences in preferences and relative wealth—even when they have identical resale strategies and agree on future monetary cash flows. Because of the illiquidity of the markets in which these assets are traded, variation in private values can translate into systematic differences in transaction prices and thus financial returns between market participants.

Third, quality improvements: A new three-bedroom house today is built to a higher standard than one built 100 years ago in the same location. Hence, we need to adjust for such quality improvements when estimating returns—both price changes and income. It is noteworthy that Eichholtz, Korevaar & Lindenthal (2023), who addressed the issue of quality adjustments head on by using new historical data, concluded that "most of the surge in housing expenditure that did occur over time is due to increasing housing quality rather than rising rent" (p. 1).

Fourth, investors must bear the substantial costs of maintaining their properties to realize capital gains in line with a quality-adjusted price index. Yet, prior studies do not always adequately take such costs into account.

Last, we need to be aware of the so-called superstar city bias in the case of housing (Gyourko, Mayer & Sinai 2013). Prior historical studies—even of national housing prices—with a focus on capitals and other large cities are thought to exhibit an upward bias in their average rate of house price appreciation. On the other hand, a very recent and extensive long-run study claims that, when we estimate total returns, this bias reverses due to lower city rents (Amaral et al. 2023). The higher total return in rural housing is perhaps compensation for the increased risk in the form of higher covariance with income growth and lower liquidity.

When estimating housing returns, therefore, scholars need to be particularly mindful of the last three of these real estate characteristics. Both prices and rents need adequate adjustment for housing-quality improvements, and maintenance and other costs need to be fully accounted for. The superstar city debate highlights the importance of assembling income and capital appreciation with particular care.

Quality adjustment is usually addressed by estimating price appreciation using repeat sales and rental changes of the same set of properties. This obviously requires a very large sample of properties to generate sufficient transacted price observations given the infrequency of trading. It is possible to use smaller samples along with hedonic regressions, provided that sufficiently detailed data are available to control for the heterogeneity in property characteristics.

When estimating income returns, there are two further considerations to be kept in mind. We require data on realized (as opposed to contractual) rents that take account of any discounts or voids. Furthermore, income yields need to be estimated from prices and rents observed

[17]Goetzmann, Spaenjers & Van Nieuwerburgh (2021) reviewed the literature on other private value assets including real estate.

Table 4 Comparison of annualized percentage real returns to housing

	France, 1871–1943		Holland, 1900–1979		United Kingdom, 1901–1983	
	Eichholtz et al. (2021)	Jordà et al. (2019)	Eichholtz et al. (2021)	Jordà et al. (2019)[a]	Chambers, Spaenjers & Steiner (2021)	Jordà et al. (2019)[a]
Capital gains	−0.6	−0.7	−0.6	0.7	−0.7	0.7
Total returns	2.8	4.2	4.8	7.2	2.3	4.7

All figures are geometric means.
[a]Note that the capital gains estimate for Jordà et al. (2019) comes from Knoll, Schularick & Steger (2017). See Chambers, Spaenjers & Steiner (2021, p. 3600) for details on estimating the UK returns from the Jordà et al. (2019) data.

concurrently on the same set of individual properties. Where such a time series of realized income yields on a set of properties cannot be estimated, scholars have resorted to starting with a contemporary rent-to-price ratio and then extrapolating backward using historical price and rent index changes. This methodological approach risks compounding measurement errors over any long-run period. According to Eichholtz et al. (2021, p. 3629), "it is difficult if not impossible to accurately derive the evolution of rental yields over time from series of house prices and rents that relate to different sets of dwellings or use different methodologies."

Turning to the empirical evidence, the Jordà et al. (2019) study asserted that on average housing delivered a total return that at least matched that of equities.[18] Housing performed better than equities in 9 (not including the United States or United Kingdom) out of the 16 countries in the Jordà et al. (2019) sample over 1870–2015. Across all countries, the equal-weighted real return was 7.3% on housing compared to 6.7% on stocks (see Jordà et al. 2019, table VII). When switching to GDP-weighting, housing (6.7%) performs slightly worse than equities (7.1%).

Dimson, Marsh & Staunton (2018) were among the first to challenge this surprising conclusion in their study of broad house price indexes over the period 1900–2017 in 11 national markets. Taking account of population size, supercity bias, maintenance and insurance costs, and quality improvements, they estimated annualized real house price appreciation considerably below the Jordà et al. (2019) estimate.

Turning to total returns, the micro housing studies of France, Holland, and England report lower estimates than Jordà et al. (2019) (see Eichholtz et al. 2021; Chambers, Spaenjers & Steiner 2021). **Table 4** summarizes the estimates of capital gains and of total returns in real terms for all three studies. While Jordà et al. (2019) estimated country-wide returns, the micro studies estimated returns for Paris, Amsterdam, and England, respectively. The Paris and Amsterdam samples are sufficiently large to yield enough repeated price observations for the estimation of capital gains by using repeat-sales regressions.[19]

Each of the three micro studies report total-return estimates substantially lower than those of Jordà et al. (2019). This most likely reflects the difficulties in a macro study arising from not fully adjusting for quality improvements and fully accounting for all costs, from the misestimation of realized income yields by extrapolating backward using different national price and rent index series, as well as from smoothing and splicing issues inherent to using historic price indexes.[20] Our

[18]Jordà et al. (2019) also reported lower volatilities for housing compared to equities and hence higher Sharpe ratios. Given that most retail investors in housing would be unable to diversify their housing exposure beyond a single property or outside their home market, such low levels of risk would be unattainable.
[19]In the absence of sufficient repeat sales, the Chambers, Spaenjers & Steiner (2021) study adopted a different approach to estimating capital gains. The data set does, however, allow them to comprehensively account for the impact of quality changes on rents and for costs when estimating income returns.
[20]Eichholtz et al. (2021, pp. 3626–28) noted the upward biases to Jordà et al. (2019) capital gains estimates due to smoothing and splicing problems for both France and Holland.

conclusion is that the evidence from the small number of single-country studies is nonetheless sufficient to suggest that the total returns to housing are lower than previously claimed.

The other important conclusion upon which all four empirical studies agree is that income yield accounts for the bulk of total return. Net rental yield most likely represents the ceiling on expected total returns—and not the floor to be augmented by long-term real growth and repricing.

Our final word on real estate is an appeal for more long-run empirical studies of total returns to housing using the best available data and avoiding the measurement problems highlighted above. It is also worth emphasizing that little has been done on the long-run performance of commercial real estate, farmland, and infrastructure.[21] This is particularly important given their greater relevance to institutional investor portfolios than housing.

3.5. Commodities: Are Commodities Rewarded in the Long Run?

Many empirical studies suggest that commodities are a poor investment over the very long run. For example, GFD reports an annual excess return over cash for an equal-weighted portfolio of diversified spot commodities of –0.4% over 1791–2020.[22] This can be rationalized by the fact that commodities are not income-generating assets like stocks and bonds, by theories pointing to commodity-price-hedging and general-inflation-hedging services, or by the Prebisch-Singer hypothesis predicting a negative relative price trend for commodities versus manufactured goods (Harvey et al. 2010).

However, when we move away from spot returns, empirical evidence based on a well-diversified portfolio of commodity futures suggests that there is a positive commodity premium (long-term excess return over cash). Early studies found positive rewards for equal-weighted composites of commodity futures: Bodie & Rosansky (1980) used 1950–1976 data; Erb & Harvey (2006) studied 1970–2004; and Gorton & Rouwenhorst (2006) studied 1959–2004. Erb & Harvey (2006, p. 84) contrasted such portfolio evidence with the much more modest returns from individual commodities, explaining the gap by the so-called diversification or rebalancing return (see Booth & Fama 1992), dubbing this phenomenon as "turning water into wine."

More recent studies have extended the history of commodity futures returns to the nineteenth century: Levine et al. (2018) back to 1877; Bhardwaj, Janardanan & Rouwenhorst (2019) back to 1871; and Geczy & Samonov (2019) back to 1877. All three studies confirmed a positive premium over cash for an equal-weighted portfolio of commodity futures from the 1870s up to the present. While the data sets vary and always involve narrower universes in the early decades, the main takeaway remains the same whether one starts in the 1870s, 1900, 1950, or 1970. Indeed, **Figure 1** shows that commodity futures almost matched the long-run returns of US equities over a 152-year window (and in fact exceeded the investment return from equities since 1900).

To better understand the nature of the commodity premium (focusing on the long-run compound return over cash), we examine its three possible sources. First, the higher long-run returns for commodity futures over spot commodities suggest a positive long-run roll return (carry) in commodity futures. However, the only long-run futures study that splits commodity futures returns into spot and roll returns, Levine et al. (2018), reached mixed conclusions on their relative long-run contribution. Unlike spot commodity investments, futures investments naturally delay

[21] Chambers, Spaenjers & Steiner (2021) reported farmland total returns in England over 1901–1983.

[22] For comparison, Jacks (2019) implied an annual excess return over cash of –0.2% and slightly lower real return for an equal-weighted composite of spot commodity returns between 1851 and 2015, while *The Economist*'s (2020) index of industrial spot commodities had a flat cumulative real return between 1860 and 2020. Using an even longer GFD data set, Zaremba, Bianchi & Mikutowski (2021) documented –0.1% annual real return over multiple centuries. Erb & Harvey (2013) found little change in the real price of gold over two millennia.

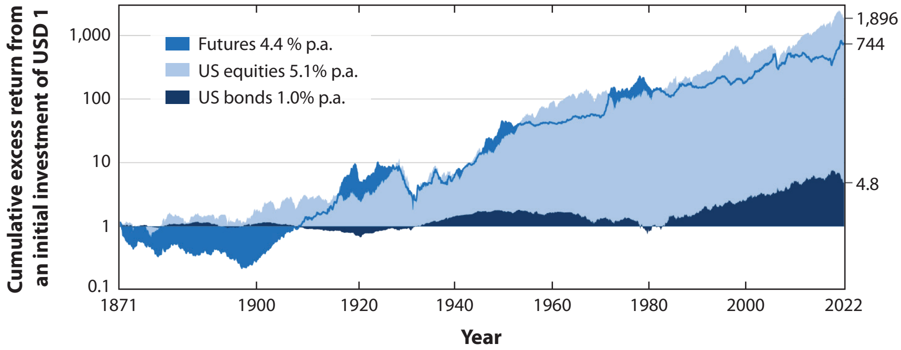

Figure 1

Commodity futures, US stocks, and US bonds, 1871–2022. Sources: Dimson, Marsh, & Staunton (2023, chapter 8), using the equally weighted commodity futures index created by Bhardwaj, Janardanan & Rouwenhorst (2019) and updated by SummerHaven Investment Management, linking into the AQR Capital Management equally weighted commodity futures index after November 2021. The US inflation rate and equity series are from Dimson, Marsh & Staunton (2023). Figure adapted with permission from Dimson, Marsh & Staunton (2023).

the cash outlay, so futures effectively earn excess-of-cash returns. The relative roles of spot and roll returns in commodity futures investing depend mechanically on whether the cash return is subtracted from the roll or spot return in performance attribution.[23]

Second, it is possible that there is a selection bias in futures contracts that are created or that survive. We find little evidence that commodities for which futures contracts have been created have higher average spot returns than those without futures. However, in their broad universe of 230 contracts, Bhardwaj, Janardanan & Rouwenhorst (2019) found clear evidence of survivorship bias, in that commodity futures with shorter contract histories have lower average returns.

Last, the diversification return, while small for most asset classes, is a major feature for commodity portfolios, thanks to the combination of high volatilities and low correlations among constituents. The greater the volatility drag or variance drain, the more it hurts compounding and thus the geometric mean return (for a given arithmetic mean). Reducing portfolio volatility by diversification mitigates this variance drain. Rebalancing boosts the portfolio geometric mean compared to a typical single-commodity geometric mean.

Table 5 illustrates the beneficial impact of the diversification return in commodities utilizing data on an equal-weighted futures portfolio.[24] This diversified portfolio generated over 1877–2020 a geometric mean excess return over cash of 4.3%, compared to a mean of 0.4% across single-commodity portfolios. This additional return reflects the volatility reduction from an average of 37% across individual commodities to 22% for the diversified portfolio. Similarly, the equal-weighted spot commodity portfolio (GFD data) has a 1.1% geometric mean excess return over cash, compared to −1.5% single-commodity mean experience, reflecting the volatility reduction from 32% to 15%. The AQR spot experience is shown to disentangle different effects.

---

[23]The 1877–2020 compound excess return over cash of a commodity futures portfolio is 3.0% (p.a.) with monthly rebalancing (used by Levine et al. 2018) and 4.3% with annual rebalancing (used below for consistency with GFD data). Slower annual rebalancing benefited from commodities' trending tendencies by allowing the momentum to play out.

[24]Equal weighting is an especially simple active strategy, as each futures contract is regularly rolled into a new contract in the same size irrespective of performance. More realistic weighting schemes that incorporate asset volatility or liquidity can achieve comparable volatility reduction if their rebalancing rules ensure broad diversification.

Table 5 The commodities diversification return: excess return over cash 1877–2020 (% p.a.)

	AQR futures		AQR spot		GFD spot	
	EW portfolio	Commodity	EW portfolio	Commodity	EW portfolio	Commodity
GM	4.3	0.4	0.3	−1.2	1.1	−1.5
AM	6.5	6.0	2.3	3.4	2.1	2.7
Vol	22.2	37.0	20.9	32.3	14.5	32.0
SR	0.29	0.16	0.11	0.10	0.14	0.08

Sources: AQR refers to the data set used in Levine et al. (2018) extended by AQR Capital Management to 2020 using annual rebalancing.
EW portfolio is the equal weighted portfolio return across commodities. Commodity is the mean statistic across commodities when each commodity is held individually. The reported statistics for all data sets include the whole available universe of commodities at each point in time [rising from single digits to 30 (AQR) or 46 (GFD), which explains different AMs for EW portfolios and single commodities].
Abbreviations: AM, arithmetic mean; GFD, Global Financial Data; GM, geometric mean; SR, Sharpe ratio; Vol, volatility or standard deviation.

Summarizing, the historical evidence suggests that while single commodities may not be expected to outperform cash in compound returns, a diversified commodity futures portfolio can offer an expected premium of 3–4% per annum. Furthermore, it has achieved this over long periods—whether the past 150, 100, or 50 years. Overall, we conclude that this diversification return is the most important and empirically robust contributor to the positive long-term commodity premium.[25]

4. CONCLUSION

The differing approaches to research on long-run asset returns summarized above deserve both credit and encouragement. We especially salute Alfred Cowles, Lawrence Fisher, James Lorie, and Roger Ibbotson, who inspired many subsequent empirical researchers to create and analyze new series of raw historical data. Partly thanks to technological advances making historical data extraction easier, we are arguably experiencing a golden age in the research of financial history.

Also, a cohort of applied researchers are addressing interesting questions using very long data histories created by others—not least, GFD. An example is the Anarkulova, Cederburg & O'Doherty (2022) team, who deployed raw data from GFD in imaginative ways, showing that the frequency of negative long-horizon returns over almost 200 years is an order of magnitude higher when studying a global equity data set compared to US-only equity returns, or Arnott & Bernstein (2002), who studied the components and determinants of the US equity premium.

Some researchers combine both aspects—compiling new data and undertaking novel empirical tests. For example, Baltussen, van Vliet & van Vliet (2023) have extended the CRSP data set of US stocks back to 1866–1926 to study the robustness of cross-sectional return predictability (factor premia), and elsewhere, Baltussen, Swinkels & van Vliet (2023) analyzed asset class performance in different inflation environments.

It is increasingly important to evaluate best practices for research on long-term financial history. In keeping with others, we wish to emphasize the importance of good documentation of data sources and choices/assumptions as well as more focus on possible biases. The most obvious empirical research on US and UK equity and bond markets before 1900 has already been made. It is desirable to refine these return series by improved estimates of delisting bias, default rates, broad payout measures, etc.

[25]As to why there is a positive commodity premium, the long history enabled Levine et al. (2018) to study the macroeconomic sensitivities of commodities and the related asset pricing implications and suggest that positive growth beta is at least a partial answer.

Another natural path is to extend empirical research beyond these core stock and bond markets to other countries and to other asset classes. Any global comparisons will require data on foreign exchange rates (if unhedged), relevant short rates (if hedged), or purchasing power parity estimates. Some equity analyses could be extended before 1800 following the pioneering steps of Le Bris, Goetzmann & Pouget (2019) and Golez & Koudijs (2018), among others. In the case of real estate, further country micro studies seem warranted and feasible, given the even larger role of this asset class in investor wealth before 1900. We can hope for more histories of house prices and rental income that engage with the data challenges, including the need for quality adjustments as houses age.

We chose to focus our empirical survey on the average returns of major asset classes. Further analysis of forward-looking expected returns (time series predictability), or volatility and other higher moments, can supplement analysis of historical average returns. Most importantly, relating asset returns to macro developments and various risk factors is essential for asset pricing models. Longer and better-quality data histories of all these series will allow both larger samples and the evaluation of long-horizon relationships.

Finally, very long historical data sets face methodological challenges. Their potential usefulness for long-horizon predictability studies is balanced by the statistical problems highlighted by Boudoukh, Israel & Richardson (2019, 2022). Similar econometric research may help deal effectively with structural changes and missing data as well as gradually improve asset breadth and data quality.

DISCLOSURE STATEMENT

A.I. is employed by AQR Capital Management and affirms that he coauthored this article in a personal capacity. The authors are not aware of any other affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Allison Adams, Andrew Ang, Rob Arnott, Jan Annaert, William Bernstein, Hank Bessembinder, Gareth Campbell, John Campbell, Mark Carhart, Frank Fabozzi, Rik Frehen, William Goetzmann, Amit Goyal, Campbell Harvey, Roger Ibbotson, Matti Ilmanen, Sara Jonsson, Peter Koudjis, Colin Lizieri, Paul Marsh, Edward McQuarrie, Mikko Niskanen, Peter Nyberg, Elias Ohneberg, Jarkko Peltomäki, Zhenkai Ran, Raghu Rau, Geert Rouwenhorst, Lucio Sarno, Stephen Schaefer, Moritz Schularick, Jeremy Siegel, Laurence Siegel, Christophe Spaenjers, Mike Staunton, Laurens Swinkels, Bryan Taylor, Annalisa Tonetto, John Turner, Kevin van Mencxel, Stijn van Nieuwerburgh, Josef Zechner, and *Annual Review of Financial Economics* coeditor, Matthew Richardson, for data provision and helpful conversations. We similarly thank our colleagues at the Centre for Endowment Asset Management (CEAM) at Cambridge Judge Business School and seminar participants at the Centre for Economic Policy Research; the Institute for Quantitative Investment Research, London; the London School of Economics; Queen's University Belfast; and Stockholm University. We also received valuable comments from conference participants at the AQR Capital Management London Forum at the London Stock Exchange; the Centre for Economic Policy Research Economic History Symposium in Dublin; the ETF Ecosystem Unwrapped Conference in London; the Financial Management Association Consortium on Asset Management in Cambridge, United Kingdom; the Neue Zürcher Zeitung Real Estate conference at Davos, Switzerland; and the Quantitative Endowment Investing Symposium at Potsdam, Germany.

LITERATURE CITED

Acheson GG, Hickson CR, Turner JD, Ye Q. 2009. Rule Britannia! British stock market returns, 1825–1870. *J. Econ. Hist.* 69:1107–37

Amaral F, Dohmen M, Kohl S, Schularick M. 2023. *Superstar returns? Spatial heterogeneity in returns to housing*. CEPR Discuss. Pap. 16806

Anarkulova A, Cederburg S, O'Doherty MS. 2022. Stocks for the long run? Evidence from a broad sample of developed markets. *J. Financ. Econ.* 143(1):409–33

Annaert J, Buelens F, Riva A. 2016. Financial history databases: old data, old issues, new insights. In *Financial Market History: Reflections on the Past for Investors Today*, ed. D Chambers, E Dimson, pp. 44–58. Charlottesville, VA: CFA Inst. Res. Found.

Arnott R, Bernstein P. 2002. What risk premium is normal? *Financ. Anal. J.* 58(2):64–85

Baltussen G, Swinkels L, van Vliet P. 2023. Investing in deflation, inflation, and stagflation regimes. *Financ. Anal. J.* 79(3):5–32

Baltussen G, van Vliet B, van Vliet P. 2023. *The cross-section of stock returns before CRSP*. SSRN Work. Pap. 3969743

Bernstein W. 2022. Review of Antti Ilmanen: *Investing Amid Low Expected Returns*. *Advisor Perspectives*, April 11

Bessembinder H, Chen T-F, Choi G, Wei KCJ. 2023. *How should returns to long term investors be measured?* SSRN Work. Pap. 4528681

Bhardwaj G, Janardanan R, Rouwenhorst KG. 2019. *The commodity futures risk premium: 1871–2018*. SSRN Work. Pap. 3452255

Bodie Z, Rosansky V. 1980. Risk and return in commodity futures. *Financ. Anal. J.* 36(3):27–39

Booth DG, Fama EF. 1992. Diversification returns and asset contributions *Financ. Anal. J.* 48(3):26–32

Boudoukh J, Israel R, Richardson M. 2019. Long-horizon predictability: a cautionary tale. *Financ. Anal. J.* 75(1):17–30

Boudoukh J, Israel R, Richardson M. 2022. Biases in long-horizon predictive regressions. *J. Financ. Econ.* 145(3):937–69

Brennan MJ, Schwartz ES. 1985. On the geometric mean index: a note. *J. Financ. Quant. Anal.* 20(1):119–22

Brown SJ, Goetzmann WN, Ross SA. 1995. Survival. *J. Finance* 50(3):853–73

Campbell G, Grossman R, Turner JD. 2021. Before the cult of equity: the British stock market, 1829–1929. *Eur. Rev. Econ. Hist.* 25(4):645–79

Campbell G, Quinn W, Turner JD, Ye Q. 2018. What moved share prices in the nineteenth-century London stock market? *Econ. Hist. Rev.* 71(1):157–89

Campbell JY. 2018. *Financial Decisions: A Course in Asset Pricing*. Princeton, NJ: Princeton Univ. Press

Chambers D, Dimson E, eds. 2016. *Financial Market History: Reflections on the Past for Investors Today*. Charlottesville, VA: CFA Inst. Res. Found.

Chambers D, Spaenjers C, Steiner E. 2021. The rate of return on real estate: long-run micro-level evidence. *Rev. Financ. Stud.* 34(8):3572–607

Chancellor E. 2022. *The Price of Time. The Real Story of Interest*. London: Allen Lane

Chen Z, Zhengyang J, Lustig H, Van Nieuwerburgh S, Xiaolan M. 2023. *Exorbitant privilege gained and lost: fiscal implications*. NBER Work. Pap. 30059

Cochrane JH. 2011. Presidential address: discount rates. *J. Finance* 66:1047–108

Cowles III A. 1938. *Common Stock Indices, 1871–1937*. Cowles Comm. Res. Econ. Monogr. 3. Bloomington, IN: Principia Press

Coyle C, Turner JD. 2013. Law, politics, and financial development: the great reversal of the U.K. corporate debt market. *J. Econ. Hist.* 73:810–46

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

Dimson E, Marsh P, Staunton M. 2018. Private wealth investments. In *Global Investment Returns Yearbook 2018*, pp. 64–82. Zurich: Credit Suisse Res. Inst.

Dimson E, Marsh P, Staunton M. 2021. American exceptionalism. *J. Portf. Manag.* 47(7):14–26

Dimson E, Marsh P, Staunton M. 2023. *Global Investment Returns Yearbook 2023*. Zurich: Credit Suisse Res. Inst.

Dimson E, Marsh P, Staunton M. 2024. *Global Investment Returns Yearbook 2024*. Zurich: UBS Invest. Bank
Eichholtz P, Korevaar M, Lindenthal T. 2023. *The housing affordability revolution*. Work. Pap., Maastricht Univ., Maastricht, Neth. **https://www.aeaweb.org/conference/2023/program/paper/9fDzbzen**
Eichholtz P, Korevaar M, Lindenthal T, Tallec R. 2021. The total return and risk to residential real estate. *Rev. Financ. Stud.* 34(8):3608–46
Erb CB, Harvey CR. 2006. The tactical and strategic value of commodity futures. *Financ. Anal. J.* 62(2):69–97
Erb CB, Harvey CR. 2013. The golden dilemma. *Financ. Anal. J.* 69(4):10–42
Eytan H, Harpaz G. 1986. The pricing of futures and options contracts on the Value Line Index. *J. Finance* 43(3):843–55
Farmer LE, Nakamura E, Steinsson J. 2024. Learning about the long run. *J. Political Econ.* In press
Fisher L, Lorie JH. 1964. Rates of return on investments in common stocks. *J. Bus.* 37(1):1–21
Fisher L, Lorie JH. 1968. Rates of return on investments in common stock: the year-by-year record, 1926–65. *J. Bus.* 41(3):291–316
Foreman-Peck J, Hannah L. 2012. Extreme divorce: the managerial revolution in UK companies before 1914. *Econ. Hist. Rev.* 65:1217–38
Geczy CC, Samonov M. 2019. *Two centuries of commodity futures premia: momentum, value and basis*. SSRN Work. Pap. 3336406
Giesecke K, Longstaff FA, Schaefer S, Strebulaev I. 2011. Corporate bond default risk: a 150-year perspective. *J. Financ. Econ.* 102:233–50
Goetzmann WN, Ibbotson RG, Peng L. 2001. A new historical database for the NYSE 1815 to 1925: performance and predictability. *J. Financ. Mark.* 4(1):1–32
Goetzmann WN, Spaenjers C, Van Nieuwerburgh S. 2021. Real and private-value assets. *Rev. Financ. Stud.* 34(8):3497–526
Golez B, Koudijs P. 2018. Four centuries of return predictability. *J. Financ. Econ.* 127:248–63
Gorton GB, Rouwenhorst KG. 2006. Facts and fantasies about commodity futures. *Financ. Anal. J.* 62:86–93
Graham B, Zweig J. 2024. *The Intelligent Investor*. New York: Harper Collins. Rev. ed.
Gregory A. 2024. *The expected cost of equity in the UK revisited*. SSRN Work. Pap. 4516096
Grossman R. 2002. New indices of British equity prices, 1870–1913. *J. Econ. Hist.* 62(1):121–46
Gyourko J, Mayer C, Sinai T. 2013. Superstar cities. *Am. Econ. J. Econ. Policy* 5(4):167–99
Harvey D, Kellard N, Madsen J, Wohar M. 2010. The Prebisch–Singer hypothesis: four centuries of evidence. *Rev. Econ. Stat.* 92(2):367–77
Hautcœur P-C, Riva A. 2018. *The Data for Financial History (DFIH) Database*. Paris: Paris Sch. Econ.
Hickman WB. 1960. *Statistical Measures of Corporate Bond Financing Since 1900, National Bureau of Economic Research, Studies in Corporate Bond Financing*. Princeton, NJ: Princeton Univ. Press
Homer S, Sylla R. 2005. *A History of Interest Rates*. New Brunswick, NJ: Rutgers Univ. Press
Ibbotson RG, Sinquefield RA. 1976. Stocks, bonds, bills and inflation: year-by-year historical returns. *J. Bus.* 49:11–47
Ilmanen A. 2022. *Investing Amid Low Expected Returns.* Hoboken, NJ: Wiley
Jacks DS. 2019. From boom to bust: a typology of real commodity prices in the long run. *Cliometrica* 13(2):202–20
Jordà Ò, Knoll K, Kuvshinov D, Schularick M, Taylor AM. 2019. The rate of return on everything, 1870–2015. *Q. J. Econ.* 134(3):1225–98
Jorion P, Goetzmann WN. 1999. Global stock markets in the twentieth century. *J. Finance* 54(3):953–80
Kizer J, Grover S, Hendershot C. 2019. *Re-examining the credit premium*. SSRN Work. Pap. 3147005
Klovland JT. 1994. Pitfalls in the estimation of the yield on British consols, 1850–1914. *J. Econ. Hist.* 54(1):164–87
Knoll K, Schularick M, Steger T. 2017. No price like home: global house prices, 1870–2012. *Am. Econ. Rev.* 107(2):331–53
Korevaar M. 2023. Reaching for yield and the housing market: evidence from 18th-century Amsterdam. *J. Financ. Econ.* 148(3):273–96
Kuvshinov D, Zimmermann K. 2022. The big bang: stock market capitalization in the long run. *J. Financ. Econ.* 145(2):527–52

Le Bris D, Goetzmann WN, Pouget S. 2019. The present value relation over six centuries: the case of the Bazacle company. *J. Financ. Econ.* 132(1):248–65
Le Bris D, Hautcœur P-C. 2010. A challenge to triumphant optimists? A blue chips index for the Paris stock exchange, 1854–2007. *Financ. Hist. Rev.* 17(2):141–83
Lettau M, Ludvigson SC, Wachter JA. 2008. The declining equity premium: What role does macroeconomic risk play? *Rev. Financ. Stud.* 21(4):1653–87
Levine A, Ooi YH, Richardson M, Sasseville C. 2018. Commodities for the long run. *Financ. Anal. J.* 74(2):55–68
Lyon J, Barber B, Tsai C-L. 1999. Improved methods for tests of long-run abnormal stock returns. *J. Finance* 54(1):165–201
Lyons R, Shertzer A, Gray R, Agorastos D. 2023. *The price of housing in the United States, 1890–2006*. Work. Pap., Yale Univ., New Haven, CT
Macaulay FR. 1938. Some theoretical problems suggested by the movements of interest rates, bond yields and stock prices in the United States since 1856. *J. R. Stat. Soc.* 101(3):620–23
Maddison-Groningen Database. 2020. *Maddison Project database 2020*. Groningen Growth Dev. Cent., Univ. Groningen, Groningen, Neth. **https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2020**
Marks P, Stuart A. 1971. An arithmetic version of the *Financial Times* Industrial Ordinary Share Index. *J. Inst. Actuar.* 97(2–3):297–324
McQuarrie EF. 2020. *When do corporate bond investors earn a premium for bearing risk? A test spanning the Great Depression of the 1930s*. SSRN Work. Pap. 3740190
McQuarrie EF. 2024. Stocks for the long run? Sometimes yes, sometimes no. *Financ. Anal. J.* 80(1):12–28
Meyer J, Reinhart CM, Trebesch C. 2022. Sovereign bonds since Waterloo. *Q. J. Econ.* 137(3):1615–80
Pástor Ľ, Stambaugh RF. 2001. The equity premium and structural breaks. *J. Finance* 56(4):1207–39
Piketty T. 2013. *Capital in the Twenty First Century*. Cambridge, MA: Harvard Univ. Press
Rogoff KS, Rossi B, Schmelzing P. 2024. Long-run trends in long-maturity real rates, 1311–2022. *Am. Econ. Rev.* 114:2271–307
Roll R. 1981. A possible explanation of the small firm effect. *J. Finance* 36(4):879–88
Schmelzing P. 2020. *Eight centuries of global real interest rates, R-G, and the 'suprasecular' decline, 1311–2018*. Work. Pap. 845, Bank Engl., London
Schwert GW. 1990. Indexes of US stock prices from 1802 to 1987. *J. Bus.* 63(3):399–426
Sharpe WF. 2010. Adaptive asset allocation policies. *Financ. Anal. J.* 66:45–59
Shiller R. 2005. *Irrational Exuberance*. Princeton, NJ: Princeton Univ. Press. 2nd ed.
Shumway T. 1997. The delisting bias in CRSP data. *J. Finance* 52(1):327–40
Siegel JJ. 1992. The equity premium: stock and bond returns since 1802. *Financ. Anal. J.* 48(1):28–46
Siegel JJ. 1994. *Stocks for the Long Run*. New York: McGraw Hill. 1st ed.
Siegel JJ. 2014. *Stocks for the Long Run*. New York: McGraw Hill. 5th ed.
Siegel JJ. 2022. *Stocks for the Long Run*. New York: McGraw Hill. 6th ed.
Siegel L, McCaffrey P, eds. 2023. *Revisiting the Equity Risk Premium*. Charlottesville, VA: CFA Inst. Res. Found.
Smith WB, Cole AH. 1935. *Fluctuations in American Business 1790–1860*. Cambridge, MA: Harvard Univ. Press
Sylla RE, Wilson JW, Wright RE. 2005. *Price quotations in early United States securities markets, 1790–1860*. Inter-Univ. Consort. Political Soc. Res., Inst. Soc. Res., Univ. Michigan, Ann Arbor. **https://www.icpsr.umich.edu/web/ICPSR/studies/4053/datadocumentation**
Taylor B. 2018. *Global Financial Data's 100-share United States Stock Index. Global Financial Data*. Work. Pap., Global Financ. Data, San Juan Capistrano, CA. **https://globalfinancialdata.com/global-financial-datas-100-share-united-states-stock-index**
Taylor B. 2023. *GFD guide to global stock markets, 2023 edition*. Work. Pap., Global Financ. Data, San Juan Capistrano, CA
The Economist. 2020. Oil and commodity prices are where they were 160 years ago. *The Economist* April 27. **https://www.economist.com/graphic-detail/2020/04/27/oil-and-commodity-prices-are-where-they-were-160-years-ago**
Thomas R, Dimsdale N. 2018. *A millennium of macroeconomic data for the UK*. Data set, Bank Engl., London. **https://www.bankofengland.co.uk/statistics/research-datasets**

Thomas S. 2002. The saga of the first stock index futures contract: benchmarks, models, and learning. *J. Money Credit Bank.* 34(3):767–808

Turnbull C. 2017. *A History of British Actuarial Thought*. Cham, Switz.: Palgrave Macmillan

Van Mencxel K. 2023. *The cross-section of corporate bond returns: the pre-OTC evidence*. PhD Thesis, Univ. Antwerp, Belgium

Van Mencxel K, Annaert J, Deloof M. 2024. Bonds for the long run? The rate of return on corporate bonds in Belgium, 1838–1939. *Econ. Hist. Rev*. In press. **https://doi.org/10.1111/ehr.13330**

Wicksell K. 1898 (1936). *Interest and Prices*, transl. RF Kahn. New York: Sentry Press

Wilson J, Jones C. 2002. An analysis of the S&P 500 Index and Cowles's extensions: price indexes and stock returns, 1870–1999. *J. Business* 75(3):505–34

Woodford M. 2003. *Interest and Prices*. Princeton, NJ: Princeton Univ. Press

Zaremba A, Bianchi RJ, Mikutowski M. 2021. Long-run reversal in commodity returns: insights from seven centuries of evidence. *J. Bank. Finance* 133(7):106238

Zweig J. 2009. Does stock-market data really go back 200 years? *Wall Street Journal*, July 11