Yes, Stock Data do go back 200 years

By

Jeremy J. Siegel

Recently Jason Zweig (July 11) has questioned the quality of my early 19th century stock data that I use in my book, *Stocks for the Long Run*. Although he has no problems with the data or my analysis of the stock market since 1871, he oddly concludes that “the emperor [of the bull market] has no clothes.”

But it is Zweig who is running naked as he shows his ignorance of the recent research on historical US stock data. My first studies were indeed based on the path-breaking research of Prof. William Schwert of the University of Rochester, who published a paper in 1991 entitled, “Index of U.S. stocks prices from 1802 to 1897.”

Admittedly Schwert’s data, as Zweig indicates, has been questioned before. That is until two of the top researchers in the field of US stock return, Bill Goetzmann and Roger Ibbotson, have published a more recent article entitled “A New Historical Database for the NYSE 1815 to 1925: Performance and Predictability”. This work is by far the most thoroughly documented research on early US Stock returns, collecting monthly price and dividend data on more than 600 individual securities over more than a century of data.

This was a prodigious effort. They reported that it took their research team more than a decade of effort to track down individual share prices and dividends, mostly from original publications found in Yale’s Beinecke Rare Book Library. The data that they collected is free from the survivorship bias and other problems that Zweig cites in his critique of Schwert’s data.

Ibbotson and Goetzmann determined that the biggest source of uncertainty in these early data is the dividend yield, since many of the sources from which they obtained stock prices did not report dividends. As a result they formed two series of dividend yields, one assuming that those stocks for which they could not find dividends had zero dividends (their “low income” estimate), and another which uses the dividend yield of those stocks for which they could find dividends (their “high income estimate”).

They write, “The low income returns from the pre-1871 period is 3.77% per year. …. When we consider only the dividend paying stock during that era, however, we estimate much higher income returns – 9.27% per year. This higher income return estimate is consistent with the practice of paying out profits to keep stock prices in the early period trading near par values. The true dividend return to a capital-weighted investment in all NYSE stocks is undoubtedly somewhere in between these two extremes.”

This midpoint of their high and low estimates is 6.52%. The dividend yield for the early data that I published in *Stocks for the Long Run*, and which Zweig criticizes is 6.4%. Furthermore, their estimate of the capital gains of stocks during the period is actually slightly higher than the 0.3% per year estimate that I used.

I state in *Stocks for the Long Run* that the last 30-year period in which bonds beat stocks was from 1831 through 1861. No data have ever contradicted that. Furthermore, stocks, in sharp contrast to bonds, have never over any 20 year period or longer given negative after-inflation returns.

That is quite a record, and Zweig does not disagree with either of these statements. Nor does he disagree with any of my analysis of the data over the past 130 years. And although long-term Treasury bonds have just edged out stocks over the last 25-years (10.56% to 10.47% per year), I will be most happy to bet with Jason that over the next five years stocks will outperform Treasury bonds and keep their 150-year record intact.