



Munich Personal RePEc Archive

Long-term investors and valuation-based asset allocation

Pfau, Wade Donald

National Graduate Institute for Policy Studies (GRIPS)

09. March 2011

Online at <http://mpa.ub.uni-muenchen.de/35006/>

MPRA Paper No. 35006, posted 25. November 2011 / 15:44

Long-Term Investors and Valuation-Based Asset Allocation

by

Wade D. Pfau

Associate Professor

National Graduate Institute for Policy Studies (GRIPS)

7-22-1 Roppongi, Minato-ku, Tokyo 106-8677 Japan

Email: wpfau@grips.ac.jp

phone: 81-3-6439-6225

note, a previous version of this article circulated with the title,

“Revisiting the Fisher and Statman Study on Market Timing”

Abstract

Valuation-based market timing demonstrates strong potential to improve risk-adjusted returns for conservative long-term investors. Such timing strategies based on the cyclically-adjusted price-earnings ratio provide comparable returns as a 100 percent stocks buy-and-hold strategy but with substantially less risk. Meanwhile, market timing provides comparable risks and the same average asset allocation as a 50/50 fixed allocation strategy, but with much higher returns. Also, it is important to consider less extreme timing strategies as well, as defining market timing as either all stocks or all cash does not provide a hedge against the possibility that valuations may depart from their historical averages for extended periods. Finally, comparing the strategies over shorter rolling sub-periods reveals that a valuation-based market timing approach fairly consistently provides risk-adjusted returns superior to a fixed asset allocation strategy.

JEL Codes: C15, D14, G11, G14, G17, N21, N22

Keywords: market valuations, cyclically-adjusted price-earnings ratio, PE10, stock returns, market timing, long term, tactical asset allocation, buy and hold

Acknowledgements: Though market-timing strategies are specifically contrary to John Bogle's investment philosophy, the author is extremely indebted to suggestions and reading recommendations provided by countless users in the threads 'Any Studies on Long-term Market Timing?' and 'Valuation-based market timing with PE10 can improve returns?' at the Bogleheads Forum. I am also grateful to Rob Bennett for motivating this investigation, and to financial support from the Japan Society for the Promotion of Science Grants-in-Aid for Young Scientists (B) #23730272.

I. Introduction

Using measures of market valuations and market sentiments, Fisher and Statman (2006) is one of the few extant studies investigating whether a market-timing strategy with relatively few trades made over a long period of time can improve investment returns. There are numerous studies about the short-term performance of market timing, as well as about the predictability of long-term returns. But less common are studies which specifically investigate from a U.S. investor's perspective the relative performance of valuation-based asset allocation strategies over horizons of at least 10 years.

Fisher and Statman argue against the notion that existing valuation measures can guide long-term investors to improved risk-return outcomes. They derive market-timing strategies using price-to-earnings ratios (PE), dividend yields (DY), a cyclically-adjusted price-to-earnings ratio that is price divided by the average earnings over the previous ten years (PE10), and the Investors Intelligence Sentiment Index. The market-timing strategy invests 100 percent in stocks when the valuation measure suggests markets are undervalued, and switches to 100 percent cash when markets are overvalued. Market timing is compared to a buy-and-hold strategy which maintains a 100 percent stock allocation throughout the entire period. For PE, DY, and PE10, they have data between 1871 and 2002, and they test the Investors Intelligence Sentiment Index for 1963-2002.

They consider the portfolio balance in nominal terms at the end of 2002 for \$1 invested at the start of 1871. Over the 132-year period, keeping the asset allocation fixed at 100 percent stocks allowed a dollar to grow to \$67,672. For market timing, when PE was above its 132-year median value of 14.4 at the end of the year (overvaluation), the allocation is 100 percent cash for the next year. PE below 14.4 at the end of the year (undervaluation) results in a 100 percent stock allocation in the following year. Using this PE decision rule, the market timer only accumulated \$8,513 by the end of 2002. Applying this same approach to dividend yields centered around the historical median of 4.35 percent, the market timer accumulated \$13,513. Finally, for the PE10 decision rule based on the historical median of 16.4, the market timer accumulated \$72,750, which is more than buy-and-hold. They rightly indicate that basing the decision rule on the ex-post historical median is unrealistic. They consider other decision rules as well, and while a few occasionally work, their general consensus is that one could not choose a reliable strategy ex ante. Most rules result in lower accumulations than provided by a simple buy-and-hold strategy of 100 percent stocks. Fisher and Statman do not argue that market timing is impossible, as valuations and sentiment should be important guides to market returns. But they do conclude that existing measures are inadequate for determining true valuations and sentiments.

Both Smithers and Wright (2000) and Stein and DeMuth (2003) consider the same general framework of comparing a 100 percent stocks buy-and-hold strategy to a market-timing strategy that switches between stocks and fixed income assets based on valuations. They find strong evidence in favor of valuation-based timing approaches. To determine whether the market is over- or undervalued, Smithers and Wright specifically focus on Tobin's q , which is the ratio of stock market capitalization to the replacement cost of capital. Stein and Muth consider 15-year moving averages for a variety of valuation measures including stock price, price-earnings ratio, dividend yield, price-to-book ratio, and others. Solow, Kitces, and Locatelli (2011) also find support for valuation-based tactical asset allocation choices to improve risk-adjusted returns.

As well, I aim to demonstrate that valuation-based market timing as guided by use of PE10 decision rules actually has worked more effectively for conservative long-term investors than Fisher and Statman previously concluded for four fundamental reasons. First, Fisher and Statman only compare strategies on the basis of which provides the largest wealth accumulation at the end of a long historical period without adjusting for return volatility. The 100 percent stocks buy-and-hold strategy is a rather volatile benchmark to compare with the market-timing strategies, and it is generally too volatile for conservative household investors. Ex-ante, the market-timing strategies have an average asset allocation of 50 percent stocks. For a wide variety of risk measures, the 100 percent stocks strategy is noticeably more risky than market timing, while an alternative 50/50 fixed allocation strategy provides broadly comparable risks as market timing. When comparing the absolute returns for different strategies, the 50/50 fixed allocation strategy provides a more suitable benchmark, as it allows for comparisons between strategies with similar risk.

Second, Fisher and Statman test strategies only over the whole period 1871-2002, and 1964-2002. Importantly, they only consider cases ending with 2002, which occurs shortly after the most prolonged and unprecedentedly steep bull market in the historical period. The case for market timing is stronger with different end dates.

Third, their treatment of market timing as an all-or-nothing strategy in which all stocks or all fixed income are used based on where valuations stand with respect to their historical median seems a bit nonsensical. Would someone really be comfortable with 100 percent stocks when the PE10 is 16.3, but use 100 percent cash if PE10 is 16.5? While their market-timing strategies focus on what is most likely to happen, they fail to hedge against the possibility that valuations may depart from their historical averages for extended periods. Making this adjustment further improves the performance of valuation-based timing.

Finally, most household investors will find little relevance in results derived from extreme strategies over a 130+ year period. For a household investor with a long-term horizon who is considering valuation-based strategies, it is more relevant to compare the strategies over shorter rolling periods from the historical data. A final section considers wealth accumulations for long-term conservative investors who contribute new funds to a savings portfolio over a 30-year period. Rather than using extreme market-timing strategies, this investor may consider an approach such as that suggested in Graham and Dodd (1940), in which asset allocation is changed only when valuations move far away from the median, rather than switching between extremes as valuations fluctuate at levels close to the median. This approach provides further evidence supporting valuation-based timing, as such strategies consistently provide the same or more wealth over long portions of the historical record.

Valuation-based timing strategies as guided by use of PE10 decision rules actually have demonstrated the potential to improve long-term investment returns. On a risk-adjusted basis, valuation-based strategies provide comparable returns but with substantially less risk than a 100 percent stocks buy-and-hold strategy, and comparable risks but with much higher returns than a 50/50 fixed asset allocation strategy. As well, treating market timing as an all-or-nothing strategy in which the allocation is either 100 percent stocks or 100 percent cash does focus on what is most likely to happen, but less extreme asset allocations would allow hedging and reduce any transaction costs or capital gains taxes. These strategies tend to provide more wealth with less risk. Finally, when investigating shorter rolling periods from the historical data, we find evidence that valuation-based timing strategies with PE10 can fairly consistently improve risk-adjusted returns for conservative long-term investors.

II. Methodology and Data

As in Fisher and Statman, I first chart the nominal wealth accumulation of \$1 invested at the start of 1871. The buy-and-hold strategy is represented by 100 percent large-capitalization stocks (Standard and Poor Composite Stock Price Index). I consider a fixed allocation strategy of 50 percent stocks and 50 percent short-term fixed income (annual yield from six-month commercial paper rates, herein referred to as “cash”) as well when discussing risk and the appropriate benchmark for market timing.^a The 50/50 strategy is not strictly “buy-and-hold,” as I

^a I also replicated the analysis using the returns on 10-year Treasury Bonds for the fixed income category, but the differences were not material.

assume the investor rebalances to meet this target asset allocation at the start of each year. The baseline market-timing strategy chooses either 100 percent stocks or 100 percent cash at the start of each year, depending on whether the value of PE10 is below or above its “historical average” at that time. I consider four ways to define this average. When PE10 is above average, this suggests market overvaluation, and the investor chooses cash. When PE10 is below average, the investor chooses stocks. Following Fisher and Statman, I assume that 100 percent stocks is used (or the more aggressive allocation in later comparisons) for the years 1871-1880 when PE10 values could not yet be calculated.

Portfolio administrative and planning fees are not charged, and I do not attempt to account for taxes or transaction costs. Taxes and transaction costs could potentially be important, but as I will describe how asset allocation changes are relatively infrequent and as taxable dividends have otherwise played an important role in past returns, I suspect that the differing tax implications for the market timing and fixed allocation strategies will not be enough to overturn the results. As well, over the historical period, 27 changes to the asset allocation would imply about 30 percent less accumulation at the end with a one percent transaction cost, which would still place market timing ahead of a 50/50 strategy, for instance. Also, less extreme asset allocation strategies would have lower taxes and transaction costs than the baseline scenario described by Fisher and Statman. More research is needed about these aspects though.

I use data for 1871-2009 from Robert Shiller’s website (<http://www.econ.yale.edu/~shiller/data.htm>).^b The PE10 measure is the stock price in January divided by the average real earnings on a monthly basis over the previous 10 years. Campbell and Shiller (1998) justify this measure as a way to remove cyclical factors from earnings. The concept derives from Graham and Dodd (1940), who said “the period for averaging earnings would ordinarily be seven to ten years” (page 686). PE10 has become a widely accepted valuation measure despite lacking a precise theoretical underpinning for the choice of 10 years, and there is no particular need to test whether other measures might provide even stronger results.

^b Fisher and Statman’s data for stock returns and valuation measures for the years 1871-1999 are from the appendix of Wilson and Jones (2002). Fisher and Statman received subsequent data for 2000-2002 directly from Jack Wilson, who has passed away. Regarding the 2000-2002 data, and any more recent data as well, his co-author Charles Jones did not keep a copy of their spreadsheet. Meir Statman was very helpful, but unfortunately he also no longer has the later values either, making it impossible to precisely replicate the original results in Fisher and Statman (2006). The Wilson / Jones dataset is no longer updated, justifying the switch in data sources.

// Figure 1 About Here //

As a point of comparison for the datasets, Fisher and Statman indicate that the buy-and-hold 100 percent stocks strategy provided \$67,672 by the end of 2002. With Shiller's data, the corresponding value is \$66,512, which is 1.71 percent less. As for the valuation-based timing strategy with PE10, Fisher and Statman report a wealth accumulation of \$72,750, while Shiller's data leads to \$76,587, which is 5.27 percent more. A possible explanation for the discrepancy is that Shiller calculates PE10 using real earnings, while Wilson and Jones (2002) use nominal earnings. The extreme asset allocation choices in Fisher and Statman's market-timing strategy make it quite sensitive to the returns experienced in the area where different decision rules call for wildly different asset allocations. Nevertheless, the two datasets appear to provide broadly similar results.

Figure 1 shows the historical path for PE10 and an example for how different decision rules will have different implications for which way the returns in certain years are captured. Any points (there are six in total) between these two decision rule curves would reflect different asset allocations for the market timer. The historical median decision rule is the median PE10 value over the entire time period in which PE10 data is available, 1881 to 2010. This corresponds to the decision rule used for market timing in Fisher and Statman (2006). The rolling median, however, is more realistic to have actually been used as it is the median up to that historical point.

III. Results

Risk-Adjusted Returns

The test for market timing in Fisher and Statman (2006) compares wealth accumulations in nominal terms at the end of the historical period for \$1 invested at the beginning of 1871. For a fixed allocation of 100 percent stocks, Table 1 shows that the wealth accumulation over the 139 years from the beginning of 1871 to the beginning of 2010 is \$95,404. Using a decision rule based on whether PE10 is above or below its median value over the entire historical period ("historical median"), as Fisher and Statman do, provides a total wealth accumulation of \$124,147. This is consistent with Fisher and Statman's finding that market timing "worked" with PE10 by the end of 2002.

// Table 1 About Here //

However, a problem with accepting this finding as evidence for valuation-based timing is that past investors would not have known the median value of PE10 through 2010, and the results are quite sensitive to the specific breakpoint of the decision rule for these extreme market-timing

strategies. Table 1 also shows the wealth accumulations for three other decision rules. If the historical mean for the entire period is used, then the wealth accumulation is \$183,618, which is almost double the fixed stock allocation amount. But more realistically, if the decision rule is defined on an evolving basis using the mean PE10 value between 1881 and each subsequent year ("rolling mean"), the wealth accumulation is \$97,023. Finally, if the "rolling median" is used instead, the wealth accumulation is the lowest at \$94,866. To avoid look-back bias, the rest of the article will use the rolling median method to calculate the PE10 decision rule, as it is a reasonable criterion to have guided past decisions, and as it provides the worst results for valuation-based timing strategies.

For the timing strategy based on the rolling median PE10 value, the slightly lower wealth accumulation results from a geometric return of 8.59 percent, compared to 8.6 percent for the buy-and-hold strategy. The two strategies provide, essentially, the same returns. However, for every ex-post risk measure considered, the timing strategies result in less risk and higher risk-adjusted returns than the 100 percent stocks buy-and-hold strategy. The highest standard deviation for portfolio returns from market timing is 13.93 percent, compared to 18.02 percent for buy-and-hold. The Sharpe ratios are also larger using two different definitions, showing that market timing provides higher returns on a risk-adjusted basis. The information ratio provides a comparison to a 50/50 fixed allocation benchmark and will be discussed later. Meanwhile, the maximum drawdown, which is the maximum percentage drop in wealth between high points and any subsequent low points in the historical period, is also significantly less for market timing. The maximum drawdown was only 24.16 percent, compared to 60.96 percent for buy-and-hold. This, in turn, results in much larger ratios of geometric returns to maximum drawdowns for market timing. For risk measures which allow an investor to be more sensitive to losses than to gains, the measure of downside deviation using a minimum acceptable return (MAR) of zero is at most 6.89 percent with market timing, compared to 9.78 percent for buy-and-hold. This translates into a higher Sortino ratio of 1.367 for the worst-case market timer, compared to 1.037 for buy-and-hold. The timing strategies result in asset allocation changes only once every 5 to 6 years on average.

Finally, Fisher and Statman indicate that the Sharpe ratio is biased upward for market timing. The GISW performance measure developed by Goetzmann, Ingersoll, Spiegel, and Welch (2007) is not susceptible to manipulation or bias from active trading strategies. It is based on a power utility model which can directly incorporate the well-being and risk aversion (ρ) of an investor into its calculation. The computed statistics are the constant continuously-compounded premiums over cash provided by the strategy after accounting for risk aversion. A positive

number indicates that an investor prefers that portfolio to cash. Whichever strategy offers the highest premium for a given risk aversion coefficient maximizes the utility for that investor among the available choices, and therefore provides the highest risk-adjusted returns. For risk aversion, a value of zero represents risk neutrality, while increasingly positive values indicate increasing risk aversion. In surveying the literature, Azar (2006) finds general agreement that the realistic range for risk aversion is between one and five. The majority of studies use a value in this range, and a conservative investor may typically have a risk aversion coefficient of about 4 or 5. For the displayed risk aversion coefficients of 2, 4, and 5, the larger numbers for valuation-based strategies indicate that they provide superior risk-adjusted returns compared to the 100 percent stocks buy-and-hold strategy.

Valuation-based timing strategies provide comparable returns as the 100 percent stocks strategy, but with substantially less risk. This happens despite market timing being out of stocks almost half of the time and missing opportunities to earn the average equity premium for past U.S. stock investors. Given the large discrepancy in risk, perhaps a more appropriate benchmark for market timing is a fixed allocation strategy which provides the same average stock allocation. Ex-ante, this is a fixed 50/50 asset allocation rebalanced each year. With this strategy, Table 1 shows that the wealth accumulation is only \$13,426. The 100 percent stocks investor accumulated 7.11 times as much wealth as the investor using a 50/50 asset allocation.

One might expect that an extreme timing strategy would be much riskier than a corresponding fixed 50/50 asset allocation. Actually, the valuation-based timing strategies do result in a comparable level of risk. Timing does create larger standard deviations for portfolio returns, as well as larger downside deviations and lower Sortino ratios. At the same time, timing produces higher Sharpe ratios, lower maximum drawdowns, and higher returns over maximum drawdowns. As well, the information ratios compare the active returns from market timing to those from a benchmark 50/50 fixed allocation portfolio ($R_{MT} - R_{50/50}$), divided by the active risk from market timing compared to the same benchmark portfolio $\sigma_{(MT-50/50)}$. The positive values for the information ratio indicate that the active returns exceed the active risk relative to the benchmark, an indication that valuation-based timing provides superior risk-adjusted returns to the fixed 50/50 allocation. Finally, for the GISW performance measures, compared to the 50/50 allocation, market timing provides superior risk-adjusted returns for risk aversion of 2 or 4, while the values for the rolling mean and rolling median timing strategies are slightly less but very close for risk aversion of 5. As the worst-case timing strategy provides 7.07 times as much wealth as

the fixed 50/50 strategy, it seems reasonable to conclude that the timing strategies provide significantly higher returns at a broadly comparable level of risk.

Valuation-based timing provides comparable returns to a 100 percent stock strategy, but with less risk, while it provides higher returns than a 50/50 allocation strategy for comparable risk. As well, an asset allocation of 100 percent stocks is too volatile for a conservative investor. Table 1 confirms that the fixed 50/50 strategy provides higher risk-adjusted returns than 100 percent stocks for risk aversion of 4 or 5. As well, with Monte Carlo simulations, Pfau (2010) uses a similar constant relative risk-aversion utility function for the distribution of accumulated wealth to conclude that among the class of fixed asset allocation strategies followed over a 40-year career, conservative investors will maximize their expected utility with something closer to 50 or 60 percent stocks, rather than 100 percent stocks. A proper test of valuation-based timing should compare it to a benchmark with approximately the same risk and the same average asset allocation.

The Choice of Ending Date

Returning to the comparisons between 100 percent stocks and the worst-performing "rolling medians" market timing strategy, Figure 2 shows the path of wealth accumulations over the whole historical period. Fisher and Statman naturally choose 2002 as the most recent year for which they had sufficient data, but the choice of ending year is more important than they may have realized. Their ending year occurred shortly after the end of a major bull market that sent valuations to previously unseen levels, as was shown in Figure 1. This naturally triggered cash holdings for the timer over an extended period while stock prices skyrocketed. With the timer in cash since the start of 1989 in the rolling median scenario, the buy-and-hold strategy surpassed the market timer by 1996, though by the end of 2008 and 2009 they were once again quite close. More generally, market timing is ahead in terms of providing greater wealth in 51.8 percent of the historical end points, while buy-and-hold provides more wealth in 32.4 percent of these cases. The two strategies are essentially tied for another 15.8 percent of cases, which I define as years when the difference in wealth accumulations between the two strategies is less than one percent of the wealth accumulated by buy-and-hold.

// Figure 2 About Here //

Hedging Against the Possible

Table 2 and Table 3 illustrate an important flaw in the baseline market timing strategy adopted by Fisher and Statman. For any timing strategy to be useful, Jenkins (1961), who was describing stock formula plans that serve as important historical predecessors to this type of market timing, argues, "One of the characteristics of formulas is that they do not aim for one hundred percent accuracy, and always make allowances for the probable, while hedging against the possible" (page 18). The timing strategy that shifts to all cash when market valuations rise above their median is making an allowance for the probable, but it is completely failing to hedge against the possibility that PE10 may deviate from its historical average for an extended period.

// Table 2 About Here //

Table 2 demonstrates one simple way to hedge against the possible: do not reduce the stock allocation all the way to zero during times of overvaluation. The figure shows alternative stock allocations for overvaluation ranging from 20 percent to 80 percent stocks for the otherwise worst-performing rolling medians decision rule. In every case, these modified timing strategies produce more wealth with less risk than a fixed 100 percent stock allocation. Potential transaction costs and taxes would also be lower with these approaches. Among the five alternative timing choices, the highest wealth (\$122,140) is provided by the strategy which switches to 50 percent stocks when the market is overvalued. However, by many risk measures, the 20-100 strategy performs best. It has the highest Sharpe ratio, lowest maximum drawdown, highest returns over maximum drawdown, and the highest risk-adjusted returns for the GISW performance measures for all three risk aversion coefficients. By making this small hedging adjustment to use 20 percent stocks for market overvaluation, the risk-adjusted performance of market timing further improves in comparison to buy-and-hold. Indeed, each strategy in Table 2 provides larger returns and lower risk than the 100 percent stocks buy-and-hold strategy.

// Table 3 About Here //

Table 3 compares the fixed 50/50 strategy against valuation-based timing strategies whose lower and upper asset allocation bounds range from 0-100 to 40-60. This provides a way to compare less extreme forms of market timing. The only factor in favor of the fixed allocation strategy is that it produces the lowest standard deviation for portfolio returns. The 0-100 strategy provides the largest risk-adjusted returns for risk aversion of 2. For conservative investors, the 20-80 strategy provides the maximum risk-adjusted returns for risk aversion of 4, and 30-70 does best for risk aversion of 5. The 20-80 strategy also provides the highest Sharpe ratios, while the 30-70 strategy provides the smallest maximum drawdown and downside deviation, and the largest ratio of returns to maximum drawdown and Sortino ratio. Conservative investors could

have earned between 2.5 and 4 times as much wealth for less risk than a 50/50 asset allocation by adopting a less extreme market-timing strategy with bounds of 20-80 or 30-70.

Real Wealth Accumulations for Shorter Rolling Historical Periods

A final matter is to construct comparisons that will provide some applicability to actual long-term investors who will invest over shorter time horizons, who will contribute new savings over time, who will also hedge against the possible, and who will be unwilling to accept 100 percent stock allocations. For a more realistic valuation-based timing strategy, we defer to Graham and Dodd (1940). They suggest that a neutral strategic allocation be maintained as long as PE10 falls within a band between $2/3$ and $4/3$ of its historical average. More extreme allocations are used only when PE10 moves outside these bounds, with lower stock allocations when PE10 is high and higher stock allocations when PE10 is low. For the Graham and Dodd case, an investor treats 50/50 as the neutral portfolio, but is willing to adjust the stock allocation between 25 and 75 percent depending on PE10 values. Asset allocation changes can be made by directing new contributions and by rebalancing.

// Figure 3 About Here //

Figure 3 presents results for the rolling periods analysis, comparing the 50/50 fixed allocation strategy to the Graham and Dodd valuation-based strategy with a stock allocation of 25, 50, or 75 percent determined by the value of PE10 with respect to its rolling median value. The first part of the figure compares the wealth accumulations for the 110 rolling 30-year periods for careers beginning between 1871 and 1980, and ending between 1901 and 2010. The valuation-based strategy provides as much or more (defined as providing at least 99 percent of the wealth of a fixed 50/50 strategy) in 99 of the 110 rolling periods. The extreme valuations experienced in the late 1990s led to outperformance for the fixed allocation only for ending dates between 1997 and 2007.

The second part of Figure 3 shows the time path for the asset allocation with each strategy. Between 1871 and 1914, both strategies shared the same allocation in every year except 1898. This explains why the wealth accumulations were so close during the early part of the historical period. Between 1915 and 1944, the asset allocation does change rather frequently, often to the benefit of market timer as careers ending between the mid-1920s and 1960 show strong outperformance for the Graham and Dodd strategy. For 1944 through 1961, both strategies share the same allocation. For 6 years in the 1960s, the market timer uses a lower stock allocation, and from the mid-1970s to mid-1980s, the market timer tended toward a higher stock allocation.

From 1985 through 1992, the allocations were again the same, and then as the market boomed in the 1990s the timer used a lower stock allocation for all years between 1993 and 2008, except for 1995.

The third part of Figure 3 shows the sorted distribution of outcomes from the fixed allocation strategy with the corresponding wealth provided by the timing strategy in each case. In describing more generally about value stocks and growth stocks, Lakonishok, Schleifer, and Vishny (1994) argue, "To be fundamentally riskier, value stocks must underperform glamour stocks with some frequency, and particularly in the states of the world when the marginal utility of wealth is high" (page 1543). The figure shows that timing meets these criteria for being a less risky strategy. In recent years when the fixed allocation strategy outperformed timing, the fixed strategy was providing as much wealth as it ever had in the past and timing still provided more wealth than the fixed strategy had in at least half of the historical cases. The underperformance of timing in these cases is the price for insurance against the more likely outcome of lower stock returns when the market continues to soar in spite of high valuations. Importantly, timing did not underperform for the periods ending when both strategies were providing the lowest wealth accumulations, such as around 1920 and 1981.

Finally, the bottom part of the figure provides more detail about outperformance in rolling periods of lengths between one and 60 years. The two strategies are considered as "tied" when the difference in wealth accumulations between them is less than one percent of the wealth accumulation for the fixed strategy. As for strict outperformance, the fixed strategy outperforms in about 10 percent of the rolling periods regardless of their length. Quite strikingly, the valuation-based strategy does as well or better in about 90 percent of cases regardless of the rolling period length. Removing the "ties" and considering cases of strict outperformance, for 12-year rolling periods the timing strategy strictly outperforms in about 50 percent of cases, and it strictly outperforms in about 75 percent of cases for 30-year rolling periods and in about 85 percent of cases for 40-year rolling periods. Increasing the rolling period length leads to an increased probability for timing outperformance.

IV. Conclusions

This article provides favorable evidence based on the historical record for long-term conservative investors to obtain improved risk-adjusted returns using valuation-based asset allocation strategies. However, it is still important to emphasize a variety of caveats about these findings. First, since index funds did not exist until the 1970s, it would have been very costly to

replicate these strategies in the earlier historical period. This criticism applies to many studies using historical financial market returns, but it is especially relevant here since the historical transaction costs to replicate a valuation-based strategy would have been high. The question remains as to whether it was these costs or the accompanying taxes that prevented investors from using valuation-based strategies in the past. The answer is likely no, since implementing these strategies requires strong nerves to maintain a contrarian strategy that may only pay off in the long term. Another caveat relates to the fact that all risk measures considered here are ex post in nature. In hindsight, matters worked out fine for the valuation-based strategies, but historically it might have been quite unnerving to increase stock allocations after significant market drops, such as the call to increase the stock allocation at the start of 2009. As well, an important risk of valuation-based strategies is that the consequences of behavioral mistakes (abandoning the strategy to either increase stock allocations near a peak or reduce stock allocations near a trough) will be amplified due to the mistake being made from a more unfortunate initial level. Finally, as always, past performance does not guarantee future performance.

References

- Azar, S. A., 2006, Measuring relative risk aversion. *Applied Financial Economics Letters* 2, 341-345.
- Campbell, J. Y., and R. J. Shiller, 1998, Valuation ratios and the long-run stock market outlook, *Journal of Portfolio Management* 24, 11-26.
- Fisher, K. L., and M. Statman, 2006, Market timing in regression and reality, *Journal of Financial Research* 29, 293-304.
- Goetzmann, W., J. Ingersoll, M. Spiegel, and I. Welch, 2007, Portfolio performance manipulation and manipulation-proof performance measures, *Review of Financial Studies* 20, 1503-46.
- Graham, B., and D. Dodd, 1940, *Security Analysis (The Classic 1940 Second Edition)* (McGraw-Hill, New York, NY).
- Jenkins, D., 1961, *How to Profit From Formula Plans in the Stock Market* (American Research Council, Larchmont, NY).
- Lakonishok, J., A. Schleifer, and R. W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *The Journal of Finance* 49, 1541-1578.
- Pfau, W. D., 2010, Lifecycle funds and wealth accumulation for retirement: Evidence for a more conservative asset allocation as retirement approaches, *Financial Services Review* 19, 59-74.
- Smithers, A., and S. Wright, 2000, *Valuing Wall Street: Protecting Wealth in Turbulent Markets* (McGraw-Hill, New York, NY).

- Solow, K. R., M. E. Kitces, and S. Locatelli, 2011, Improving risk-adjusted returns using market-valuation-based tactical asset allocation strategies, *Journal of Financial Planning* 24, 38-49.
- Stein, B., and P. DeMuth, 2003, *Yes, You Can Time the Market* (John Wiley and Sons, Hoboken, NJ).
- Wilson, J. W. and C. P. Jones, An analysis of the S&P 500 index and Cowle's extensions: Price indexes and stock returns, 1870-1999, *Journal of Business* 75, 505-31.

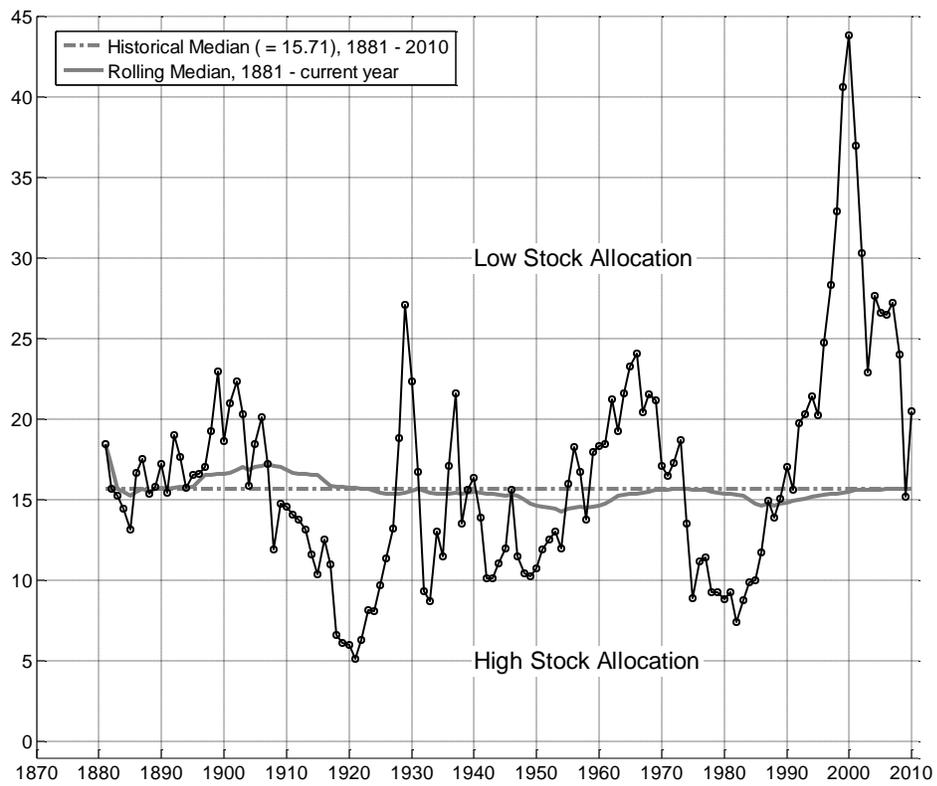


Fig. 1. Historical data for PE10

Note: See text for data source and description.

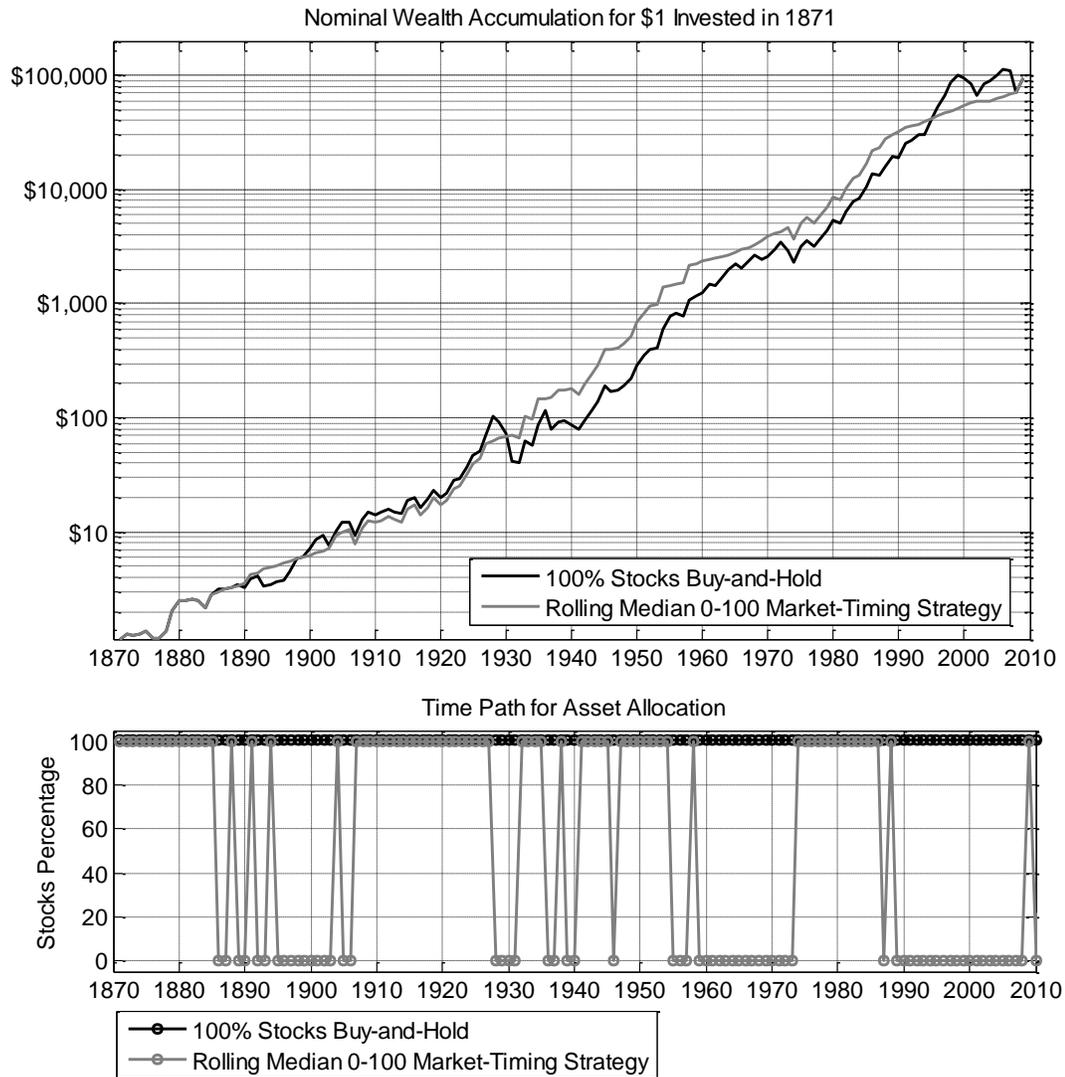


Fig. 2. Nominal wealth accumulation for \$1 invested in 1871

Note: The y-axis has a logarithmic scale.

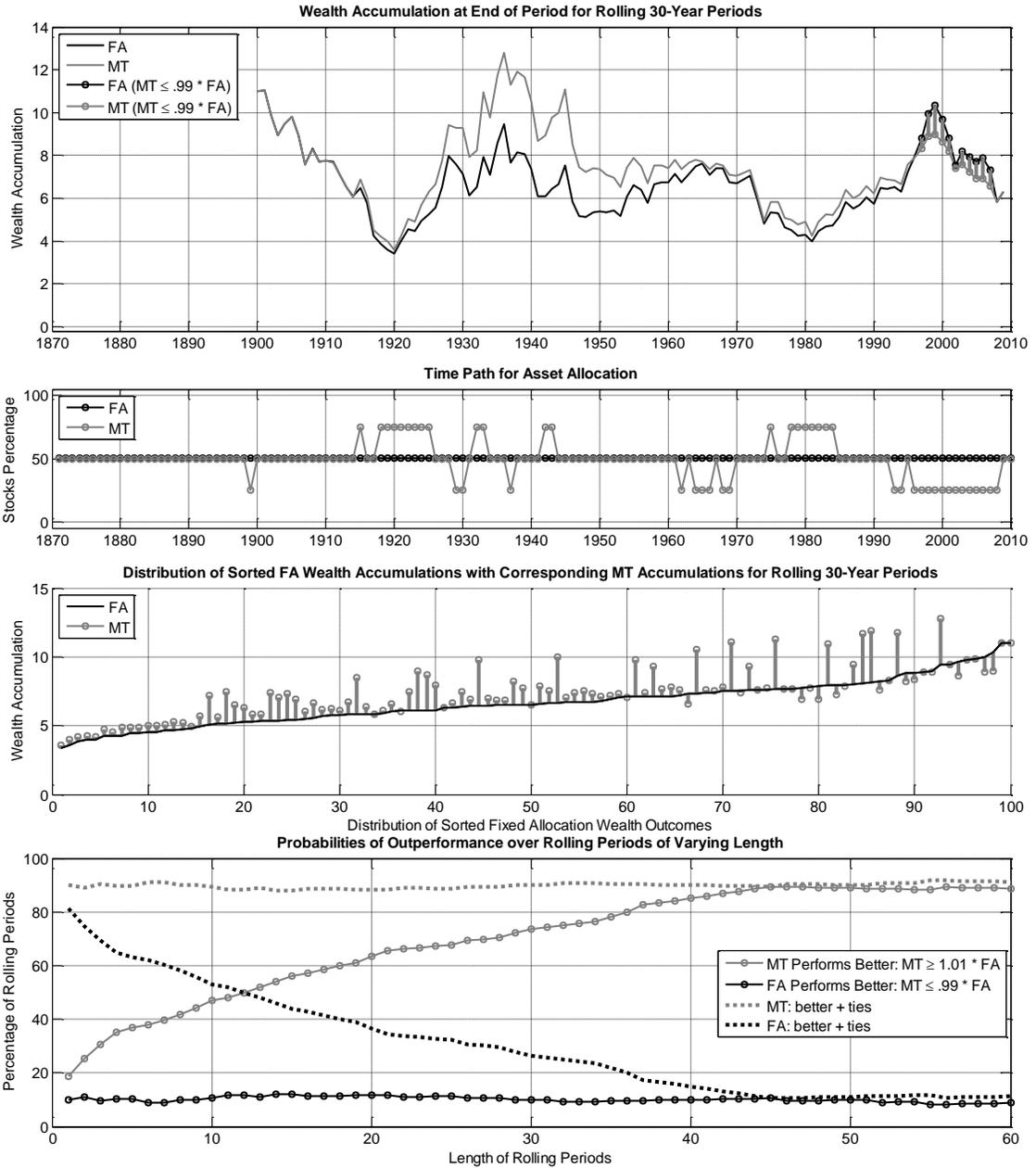


Fig. 3. Comparing a fixed allocation (FA) 50/50 strategy with a valuation-based market timing (MT) strategy

Note: MT strategy uses Graham and Dodd decision rule with 25-50-75 stock allocation, calculated using rolling medians for PE10, real wealth accumulation as a multiple of a constant real salary, and a 10 percent annual savings rate.

Table 1. Return and risk measures for fixed allocation and valuation-based market-timing

| | Fixed 100/0 | MT 0 - 100 | MT 0 - 100 | MT 0 - 100 | MT 0 - 100 | Fixed 50/50 |
|---|----------------|----------------------|--------------------|-----------------|-------------------|----------------|
| | | historical median | historical mean | rolling mean | rolling median | |
| Summary Statistics for Whole Period, January 1871 to January 2010 | | | | | | |
| Arithmetic Return | 10.13 | 9.60 | 9.94 | 9.43 | 9.41 | 7.46 |
| Geometric Return | 8.60 | 8.80 | 9.11 | 8.61 | 8.59 | 7.08 |
| Standard Deviation | 18.02 | 13.67 | 13.93 | 13.80 | 13.81 | 8.98 |
| Sharpe Ratio $(R_P - R_F) / \sigma_P$ | 0.297 | 0.353 | 0.371 | 0.337 | 0.336 | 0.298 |
| Sharpe Ratio $(R_P - RF) / \sigma_{(P-F)}$ | 0.289 | 0.342 | 0.359 | 0.324 | 0.323 | 0.289 |
| Information Ratio | --- | 0.208 | 0.229 | 0.189 | 0.191 | --- |
| Maximum Drawdown | 60.96 | 20.97 | 20.97 | 24.16 | 24.16 | 28.69 |
| Returns / Maximum Drawdown | 0.141 | 0.420 | 0.435 | 0.356 | 0.356 | 0.247 |
| Downside Deviation (MAR = 0) | 9.78 | 5.67 | 5.51 | 6.89 | 6.89 | 5.00 |
| Sortino Ratio (MAR = 0) | 1.037 | 1.694 | 1.804 | 1.369 | 1.367 | 1.490 |
| Average Stock Allocation | 100 | 53.9 | 57.4 | 53.9 | 53.2 | 50 |
| Average # Years bet. Allocation Changes | --- | 5.79 | 5.35 | 4.96 | 4.96 | --- |
| Value in 2010 of \$1 invested in 1871 | \$95,404 | \$124,147 | \$183,618 | \$97,023 | \$94,866 | \$13,426 |
| GISW Performance Measure, $\rho=2$ | 2.04 | 3.06 | 3.31 | 2.84 | 2.82 | 1.82 |
| GISW Performance Measure, $\rho=4$ | -1.44 | 1.64 | 1.83 | 1.31 | 1.29 | 1.04 |
| GISW Performance Measure, $\rho=5$ | -3.38 | 0.96 | 1.11 | 0.54 | 0.53 | 0.64 |

Note: Results are shown for the 100 percent stocks buy-and-hold strategy, a strategy that rebalances annually to a 50/50 portfolio of stocks and cash, and valuation-based timing (MT) strategies which alternate between 100 percent stocks and 100 percent cash based on four different decision rules. R_P : mean portfolio return; R_F : mean return on cash; σ_P : standard deviation of portfolio returns; $\sigma_{(P-F)}$: standard deviation of the portfolio returns in excess of cash returns; MAR: minimum acceptable risk which only penalizes returns falling below zero; GISW performance measure: manipulation-proof utility-based measure developed by Goetzmann, Ingersoll, Spiegel, and Welch (2007); ρ : investor risk aversion.

Table 2. Return and risk measures for fixed allocation and valuation-based market-timing

| | Fixed 100/0 | MT 20 - 100 | MT 40 - 100 | MT 50 - 100 | MT 60 - 100 | MT 80 - 100 |
|---|----------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | | rolling median | rolling median | rolling median | rolling median | rolling median |
| Summary Statistics for Whole Period, January 1871 to January 2010 | | | | | | |
| Arithmetic Return | 10.13 | 9.56 | 9.70 | 9.77 | 9.85 | 9.99 |
| Geometric Return | 8.60 | 8.72 | 8.78 | 8.79 | 8.79 | 8.73 |
| Standard Deviation | 18.02 | 13.95 | 14.49 | 14.90 | 15.39 | 16.59 |
| Sharpe Ratio $(R_P - R_F) / \sigma_P$ | 0.297 | 0.343 | 0.340 | 0.335 | 0.329 | 0.314 |
| Sharpe Ratio $(R_P - RF) / \sigma_{(P-F)}$ | 0.289 | 0.330 | 0.327 | 0.324 | 0.318 | 0.305 |
| Information Ratio | --- | 0.207 | 0.207 | 0.207 | 0.207 | 0.207 |
| Maximum Drawdown | 60.96 | 24.16 | 24.73 | 31.72 | 38.32 | 50.37 |
| Returns / Maximum Drawdown | 0.141 | 0.361 | 0.355 | 0.277 | 0.229 | 0.173 |
| Downside Deviation (MAR = 0) | 9.78 | 6.46 | 6.26 | 6.54 | 6.90 | 8.14 |
| Sortino Ratio (MAR = 0) | 1.037 | 1.480 | 1.550 | 1.495 | 1.426 | 1.228 |
| Average Stock Allocation | 100 | 62.6 | 71.9 | 76.6 | 81.3 | 90.6 |
| Average # Years bet. Allocation Changes | NaN | 4.96 | 4.96 | 4.96 | 4.96 | 4.96 |
| Value in 2010 of \$1 invested in 1871 | \$95,404 | \$111,011 | \$120,592 | \$122,140 | \$121,294 | \$112,449 |
| GISW Performance Measure, $\rho=2$ | 2.04 | 2.91 | 2.89 | 2.83 | 2.75 | 2.47 |
| GISW Performance Measure, $\rho=4$ | -1.44 | 1.32 | 1.13 | 0.93 | 0.65 | -0.16 |
| GISW Performance Measure, $\rho=5$ | -3.38 | 0.53 | 0.24 | -0.03 | -0.40 | -1.53 |

Note: See Table 1 for further explanation of the terminology. The only differences in this table are, for example, that "MT 20-100" means the market timer uses a 20 percent stock allocation when the market is overvalued relative to the rolling median for PE10, and uses a 100 percent stock allocation when the market is undervalued. The information ratios use benchmarks with fixed allocation strategies equal to the ex-ante average stock allocation of the timing strategy.

Table 3. Return and risk measures for fixed allocation and valuation-based market-timing

| | Fixed 50/50 | MT 0-100 | MT 10-90 | MT 20-80 | MT 30-70 | MT 40-60 |
|---|----------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | | rolling median | rolling median | rolling median | rolling median | rolling median |
| Summary Statistics for Whole Period, January 1871 to January 2010 | | | | | | |
| Arithmetic Return | 7.46 | 9.41 | 9.02 | 8.63 | 8.24 | 7.85 |
| Geometric Return | 7.08 | 8.59 | 8.35 | 8.08 | 7.78 | 7.44 |
| Standard Deviation | 8.98 | 13.81 | 12.43 | 11.19 | 10.17 | 9.41 |
| Sharpe Ratio $(R_P - R_F) / \sigma_P$ | 0.298 | 0.336 | 0.341 | 0.344 | 0.340 | 0.326 |
| Sharpe Ratio $(R_P - R_F) / \sigma_{(P-F)}$ | 0.289 | 0.323 | 0.328 | 0.330 | 0.327 | 0.314 |
| Information Ratio | --- | 0.191 | 0.207 | 0.207 | 0.207 | 0.207 |
| Maximum Drawdown | 28.69 | 24.16 | 21.12 | 18.08 | 16.34 | 22.06 |
| Returns / Maximum Drawdown | 0.247 | 0.356 | 0.395 | 0.447 | 0.476 | 0.337 |
| Downside Deviation (MAR = 0) | 5.00 | 6.89 | 6.00 | 4.72 | 4.00 | 4.17 |
| Sortino Ratio (MAR = 0) | 1.490 | 1.367 | 1.504 | 1.829 | 2.062 | 1.884 |
| Average Stock Allocation | 50 | 53.2 | 52.55 | 51.91 | 51.28 | 50.64 |
| Average # Years bet. Allocation Changes | NaN | 4.96 | 4.96 | 4.96 | 4.96 | 4.96 |
| Value in 2010 of \$1 invested in 1871 | \$13,426 | \$94,866 | \$69,608 | \$49,068 | \$33,211 | \$21,567 |
| GISW Performance Measure, $\rho=2$ | 1.82 | 2.82 | 2.74 | 2.59 | 2.40 | 2.14 |
| GISW Performance Measure, $\rho=4$ | 1.04 | 1.29 | 1.48 | 1.55 | 1.50 | 1.34 |
| GISW Performance Measure, $\rho=5$ | 0.64 | 0.53 | 0.86 | 1.04 | 1.07 | 0.94 |

Note: See Table 1 for further explanation of the terminology. The only difference in this table is, for example, that "MT 20-80" means the market timer uses a 20 percent stock allocation when the market is overvalued relative to the rolling median for PE10 and uses an 80 percent stock allocation when the market is undervalued.