Revisiting the Fisher and Statman Study on Market Timing

Pfau, Wade Donald

National Graduate Institute for Policy Studies (GRIPS)

9 March 2011

Revisiting the Fisher and Statman Study on Market Timing

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

Abstract

Valuation-based market timing demonstrates greater potential to improve risk-adjusted returns for conservative long-term investors than given credit by Fisher and Statman (2006). On a risk-adjusted basis, market-timing strategies 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, defining market timing as either 100 percent stocks or 100 percent Treasury bills does not provide a hedge against the possibility that valuations may depart from their historical averages for extended periods.

***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: The author is extremely indebted to suggestions and reading recommendations provided in the thread 'Any Studies on Long-term Market Timing?' at the Bogleheads Forum (www.bogleheads.org). Because market-timing strategies are specifically not part of John Bogle's investment philosophy, the author wishes to thank without implicating users including Adrian Nenu, afan, alec, Alex Frakt, bob90245, cjking, crl848, dmcmahon, DP, grayfox, Les, lostcowboy, market timer, matt, Mel Lindauer, Norbert Schlenker, peter71, pkcrafter, Rodc, SP-diceman, tadamsmar, wearethefall, and yobria. I am also extremely grateful to Rob Bennett for motivating this topic and contributing his experience and encouragement. I thank the financial support from the Japan Society for the Promotion of Science Grants-in-Aid for Young Scientists (B) #20730179.

Introduction

Using measures of market valuations and market sentiments, Fisher and Statman (2006) is one of the few extant studies that investigates 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 investigate from a U.S. investor's perspective the relative performance of valuation-based market-timing strategies over horizons of at least 10 years.

Fisher and Statman review the literature on this topic and provide evidence against the notion that existing valuation measures can guide long-term investors to improved risk-return outcomes. Arriving at opposite conclusions, 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.

Fisher and Statman 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 suggest markets are undervalued, and switches to 100 percent Treasury bills when markets are overvalued. Market timing is compared to a buy-and-hold strategy which maintains a 100 percent

stocks 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.

Fisher and Statman 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, they find that 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 Treasury bills 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 accumulates $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 accumulates $13,513. Finally, for the PE10 decision rule based on the historical median of 16.4, the market timer accumulates $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 without data mining, someone could not choose a reliable rule for outside of the test period. Most rules result in accumulations much lower than that 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.

But 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 three 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 making adjustments for risk. The 100 percent stocks buy-and-hold strategy is a rather risky benchmark to compare with the market-timing strategies, and it is generally too risky for

consideration by 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. As well, the U.S. enjoyed large equity premiums in the twentieth century that lead to more impressive results for 100 percent stocks than may be reliably expected in the future. When comparing the absolute returns for different strategies, the 50/50 fixed allocation strategy provides a more suitable benchmark, as it allows for comparisons of strategies with similar risk.

Second, Fisher and Statman test strategies only over the whole period 1871-2002, and 1964-2002. What is important is that they only consider cases ending with 2002, which occurs shortly after the most prolonged and unprecedentedly steep bull market in the historical period. Had a different ending date been chosen, the case for market timing is stronger.

Third, their treatment of market timing as an all-or-nothing strategy in which 100 percent stocks are used when valuations are below their historical median and 100 percent Treasury bills are used when valuations are above their historical median seems a bit nonsensical. Would someone really be comfortable with 100 percent stocks when the PE10 is 16.3, but switch to 100 percent Treasury bills when the 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 an extended period of time. Making this adjustment further improves the performance of market timing.

Methodology and Data

As with Fisher and Statman, I 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 (S&P 500). I consider a fixed allocation strategy of 50 percent stocks and 50 percent

Treasury bills (one-year yields) as well when discussing risk and the appropriate benchmark for market timing. The 50/50 strategy is not strictly "buy-and-hold," as I 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 Treasury bills 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 Treasury bills. 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. Taxes 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. More research is needed about this.

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 and forthcoming, 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). As the Wilson / Jones dataset is no longer updated, I use data for 1871-2009 from Robert Shiller's website (http://www.econ.yale.edu/~shiller/data.htm). 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, though there is no particular theoretical reason to pick precisely 10 years. The concept stems 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, and as it would imply further data mining, there is no particular need to test whether other measures would produce even better results.

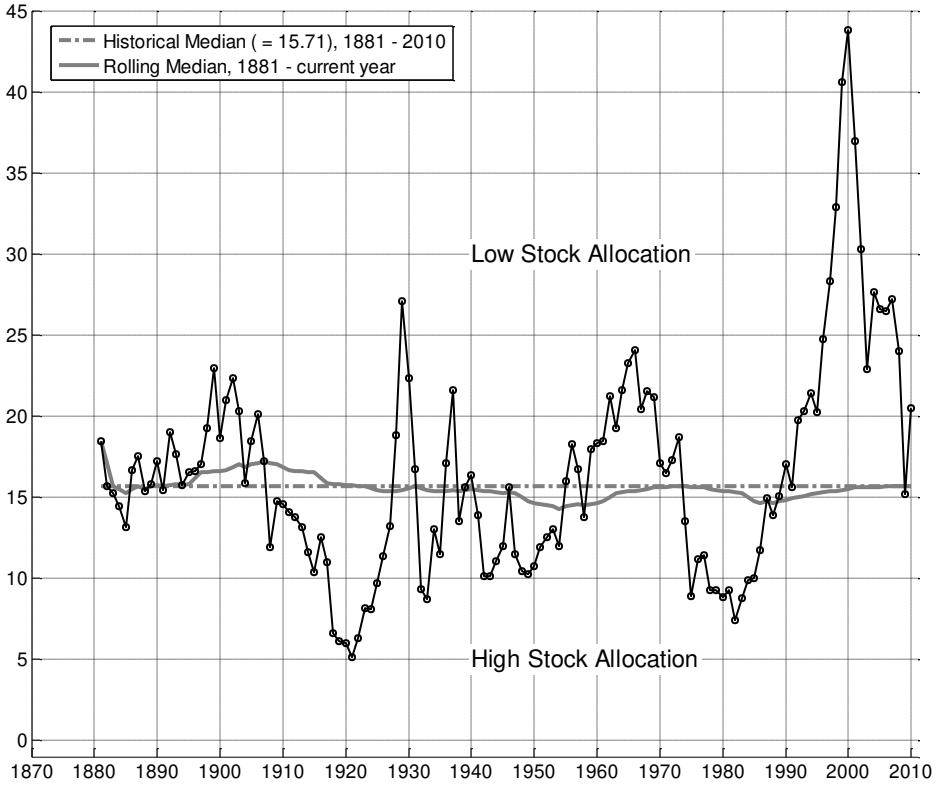

Figure I. Historical data for PE10. This data is from Robert Shiller's webpage. The PE10 value is the cyclically-adjusted price-earnings ratio at the beginning of the year. The historical median value is for the entire time period in which PE10 data is available, 1881 to 2010. This corresponds to the decision rule used for market timing in Fisher and Statman (2006). For years in which PE10 rises above 15.71, asset allocation is set to 100 percent Treasury bills, while 100 percent stocks is used for years when PE10 falls below 15.71. The rolling median is more realistic to have actually been used, and it also produces the worst results for market timing. In order to help avoid data mining, the rolling median measure will be used in subsequent comparisons.

As a point of comparison about the two 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 market-timing strategy with PE10, Fisher and Statman report a wealth accumulation of $72,750, while the Shiller data produces a value of $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. Because of the extreme asset allocation choices of Fisher and Statman's market-timing strategy, it is very sensitive to the returns experienced in the region where different decision rules would cause certain returns to fall on opposite sides for asset allocation purposes. Nevertheless, the two datasets appear to allow broadly similar results. Figure I provides more detail about the historical path of PE10 and an example of how different decision rules will have different implications for which way the returns in certain years are captured. Any points between these two decision rule curves would reflect different asset allocations for the market timer.

Results

Risk-Adjusted Returns

The test for market timing used by Fisher and Statman (2006) is to compare the wealth accumulation 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, Figure II 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, 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.

	Fixed 100/0	MT 0 - 100 historical median	MT 0 - 100 historical mean	MT 0 - 100 rolling mean	MT 0 - 100 rolling median	Fixed 50/50
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 $(R_P - R_{TB}) / \sigma_P$	0.297	0.353	0.371	0.337	0.336	0.298
Sharpe Ratio $(R_P - R_{TB}) / \sigma_{(P-TB)}$	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, ρ=2	2.04	3.06	3.31	2.84	2.82	1.82
GISW Performance Measure, ρ=4	-1.44	1.64	1.83	1.31	1.29	1.04
GISW Performance Measure, ρ=5	-3.38	0.96	1.11	0.54	0.53	0.64

Figure II. Return and Risk Measures for Fixed Allocation and Market-Timing Strategies. 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 bills, and market-timing strategies which alternate between 100 percent stocks and 100 percent bills based on four different decision rules. The "historical median" decision rule corresponds to that used by Fisher and Statman. It and the "historical mean" are calculated over the entire historical period. More realistically, the "rolling mean" and "rolling median" values are calculated from 1881 to each subsequent year across the range of historical data. Abbreviated terms include: R_P is the mean portfolio return, R_{TB} is the mean return on Treasury bills, σ_P is the standard deviation of portfolio returns, $\sigma_{(P-TB)}$ is the standard deviation of the portfolio returns in excess of Treasury bill returns, the information ratio compares the active return and active risk of the market-timing strategy relative to a benchmark 50/50 portfolio that shares the same ex-ante average stock allocation, and MAR is the minimum acceptable risk which only penalizes returns falling below zero. Finally, the GISW performance measure is a manipulation-proof utility-based measure developed by Goetzmann, Ingersoll, Spiegel, and Welch (2007), and ρ is investor risk aversion.

However, a problem with accepting this finding as evidence for market 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.

Figure II also shows the wealth accumulations for three other decision rules of this nature. 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. Because data mining is an important concern, the rest of the paper will use the rolling median method to calculate the PE10 decision rule, as it is a very reasonable criterion to have been chosen, and as it provides the worst results for market timing.

For the market-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 risk measure considered, the market-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 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. It is also worthwhile to note that the market-timing strategy results in changes to asset allocation only once every 5 to 6 years on average.

Finally, Fisher and Statman indicate that the Sharpe ratio is biased upward for market timing, and 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 Treasury bills provided by the strategy after accounting for risk aversion. A positive number indicates that an investor prefers that portfolio to Treasury bills. 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 market-timing strategies indicate that they provide superior risk-adjusted returns compared to the 100 percent stocks buy-and-hold strategy.

Market-timing strategies provide comparable returns as the 100 percent stocks strategy, but with substantially less risk. This happens in spite of the fact that market timing is out of stocks almost half of the time. The market timer could not fully enjoy the generous average equity premium for past U.S. stock investors. Given the large discrepancy in risk, perhaps a more appropriate benchmark to compare the market-timing strategy 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, Figure II 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. As a caveat, Bogle (2009) cautions that investors should not count on such a large equity premium in the future.

One might expect that the extreme market-timing strategy considered by Fisher and Statman would be much riskier than a corresponding fixed 50/50 asset allocation. Actually, the market-timing strategies do result in a comparable level of risk. Market timing does create larger standard deviations for portfolio returns, as well as larger downside deviations and lower Sortino ratios. At the same time, market 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 indicator that market timing provides superior risk-adjusted returns to the fixed 50/50 asset allocation. Finally, for the GISW performance measures, compared to the 50/50 fixed asset 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 market-timing strategies are slightly less but very close for risk aversion of 5. As the worst case market-timing strategy provides 7.07 times as much wealth as the fixed 50/50 strategy, it seems reasonable to conclude that the market-timing strategies provide significantly higher returns at a broadly comparable level of risk.

Market timing provides comparable returns to a 100 percent stock strategy, but with much less risk, while it provides much higher returns than a 50/50 asset allocation strategy for comparable risk. As well, an asset allocation of 100 percent stocks is too risky for a conservative investor. Figure II 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 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 market timing should use as a benchmark for comparison a strategy that has 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-case "rolling medians" market-timing strategy, Figure III 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 when conducting their research. 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 I. This naturally triggered the market-timing strategies to adopt 100 percent Treasury bills for an extended period of time while stock prices skyrocketed. With the market timer in Treasury bills since the start of 1989 in the rolling median scenario, the buy-and-hold strategy could catch-up to and surpass the market timer by 1996, though by the end of 2008 and 2009 they were once again close.

In this horserace of Figure III, 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. For long periods of the historical record, the market timer was well ahead.

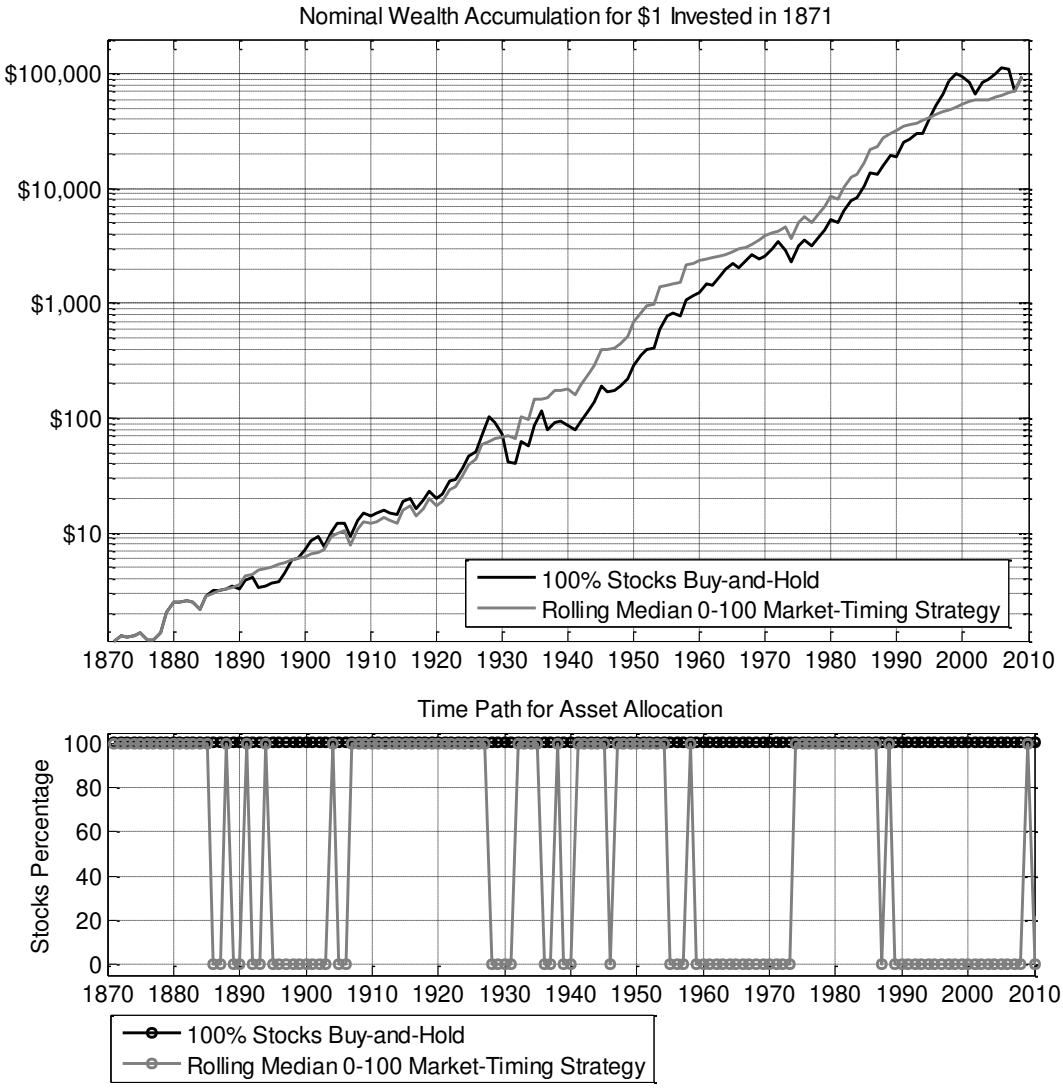

Figure III. Nominal Wealth Accumulation for $1 Invested in 1871. The y-axis has a logarithmic scale. Own calculations from data provided on Robert Shiller's webpage.

Hedging Against the Possible

Figure IV and Figure V illustrate an important flaw in the baseline market-timing strategy adopted by Fisher and Statman. For any market-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 market-timing strategy that shifts to 100 percent Treasury bills 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 of time.

	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 $(R_P - R_{TB}) / \sigma_P$	0.297	0.343	0.340	0.335	0.329	0.314
Sharpe Ratio $(R_P - R_{TB}) / \sigma_{(P-TB)}$	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, ρ=2	2.04	2.91	2.89	2.83	2.75	2.47
GISW Performance Measure, ρ=4	-1.44	1.32	1.13	0.93	0.65	-0.16
GISW Performance Measure, ρ=5	-3.38	0.53	0.24	-0.03	-0.40	-1.53

Figure IV. Return and Risks Measures for Fixed Allocation and Market-Timing Strategies. See Figure II 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 market-timing strategy.

Figure IV provides 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 market-timing strategies produce more wealth with less risk.

Among the 5 alternative market-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 other 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 Figure IV provides larger returns and lower risk than the 100 percent stocks buy-and-hold strategy.

Figure V compares the fixed 50/50 strategy against market-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 that provide lower risk and also hedge against the possibility that valuations do not immediately adjust to their medians. The only factor in favor of the fixed allocation strategy is that it provides the lowest standard deviation of 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.

	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 $(R_P - R_{TB}) / \sigma_P$	0.298	0.336	0.341	0.344	0.340	0.326
Sharpe Ratio $(R_P - R_{TB}) / \sigma_{(P-TB)}$	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, ρ=2	1.82	2.82	2.74	2.59	2.40	2.14
GISW Performance Measure, ρ=4	1.04	1.29	1.48	1.55	1.50	1.34
GISW Performance Measure, ρ=5	0.64	0.53	0.86	1.04	1.07	0.94

Figure V. Return and Risks Measures for Fixed Allocation and Market-Timing Strategies. See Figure II 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.

Conclusions

This study extends Fisher and Statman (2006) by suggesting that market timing as guided by use of PE10 decision rules actually has demonstrated strong potential to improve long-term investment returns. I offered three substantive critiques of Fisher and Statman's methodology. 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 making adjustments for risk. On a risk-adjusted basis, market-timing strategies provide very favorable results compared to a 100 percent stocks buy-and-hold strategy (comparable returns but with substantially less risk) and

compared to a 50/50 fixed asset allocation strategy (comparable risks but with much higher returns). Second, Fisher and Statman test strategies for periods ending in 2002, which occurs shortly after an unprecedented bull market that had sent valuations and stock prices soaring. Had a different ending date been chosen, the case for market timing is stronger. Third, their treatment of market timing as an all-or-nothing strategy in which the allocation is either 100 percent stocks or 100 percent Treasury bills does focus on what is most likely to happen, but fails to hedge against the possibility that valuations may depart from their historical averages for extended periods. Valuation-based market timing with PE10 has the potential to improve risk-adjusted returns for conservative long-term investors.

References

Azar, S. A., 2006, Measuring relative risk aversion. *Applied Financial Economics Letters* 2, 341-345.

Bogle, J., 2009, *Enough: True Measures of Money, Business, and Life* (John Wiley and Sons, Hoboken, NJ).

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).

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).

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.