Momentum in a Funds Context— Critically Different Choices

A WORKING PAPER— FEBRUARY 2018

by Mike O'Connor

mike@thetradedportfolio.com

www.thetradedportfolio.com

Abstract

The strategy of simply holding stocks of high momentum is amazing for the amount of support that it has gotten from normally skeptical academics. But in practice there have been flies in the ointment. For example, it has been acknowledged by the investment advisor and management firms Dorsey Wright Money Management and AQR Capital Management that their own stock selection strategies that were based on momentum failed, crucially, in 2007-08 (DWMM, 2016; Asness, 2017).

Herein *four* things are accomplished, with the intended application being to a portfolio of *funds,* not stocks. One is to ratify a simple cure for momentum failures during panics and to specify it well. The other three are: (1) providing a systematic way of finding an optimized momentum strategy that *adapts* to secular changes in the marketplace; (2) using it to discover that the optimal form of the strategy for a portfolio of funds is critically different from the form that works with stocks; and (3), doing so while presenting a simple hypothesis testing procedure that is more of a simulation than an abstract exercise in mathematical statistics.

It is shown, as some advisors have concluded, that with a portfolio of funds volatility could have been reduced and events such as the 2007-08 panic could have been avoided by having refrained from shorting funds and by having opted to go into cash whenever the trailing returns of funds were exceeded by the returns on cash.

The provided means for optimizing the momentum strategy also answers the following vital question: Out of a given list of candidates, how choosy should the manager be at

picking the funds of highest momentum? It was found that on at least three-quarters of the trials the policy of resorting to cash when trailing returns falter would have performed optimally when the maximum number of funds to be held out of a candidates list of 10 was greater than 1 (tantamount to not restricting holdings to anything like only the top decile); if instead staying 100% invested were the policy then that fraction would have dropped to at least as far as one-third, meaning that at least two-thirds of the time it would generally have paid to be very picky and to have held only one fund at a time notwithstanding the minimal amount of diversification.

At the heart of a momentum strategy there is necessarily a lookback period L, here the number of trailing months over which returns are tallied as a measure of momentum. Certainly we want to use a lookback period over which returns are positively correlated with future returns. But if L is too large the strategy will ignore, fail to avoid, not only the potholes but also chasms into which we would definitely prefer not to fall. And if L is too small there will be responsiveness not only to major sustained moves but also to every little zig or zag, possibly leading to excessive trading and lots of embarrassing little whipsaw losses. The procedures of this paper specify a momentum strategy via a "walk-forward" procedure that adapts to changing market conditions by selecting for use the L value that has been producing the best returns over a *trailing* interval of time of a fixed duration of some years; the L value that would have worked best over the entire period of record is not used.

"Out-of-sample" testing, buttressed by Monte Carlo permutation, was used to test the hypothesis that momentum is efficacious — using Professor Kenneth R. French's online monthly dividend-adjusted returns data for US stocks from 1926-2017. The data are compiled as groups of portfolios, formed out of decile sorts of the CRSP US stock market database broken out by market capitalization (Size) and book-to-market value (Value) and out of industrial sectors (Industry). As such the portfolios can be regarded as mimicking a value fund, a big-cap fund, a small-cap fund, an industrial sector fund, etc.

Given that the system for specifying a momentum strategy and using it that is presented in this paper is purposefully adaptive it cannot be called upon to produce an announceable fixed value for the lookback period L, certainly not one proffered as usable for decades hence. That said, the following may be of interest: It was found that lookback periods as long as 12 months would have *seldom* been optimal with funds as securities if the policy were to have been to resort to cash when fund momentum drops below that of cash; if however the policy had instead been to stay 100% invested, which is conventional, then a lookback period L of 12 months or thereabouts would fairly often have been optimal for the Value and Size groups in recent decades but not at all for the Industry group, for which an L of one month was by far and away most often optimal; and, relatedly, deleting the most recent month's return from the lookback-period returns would have lead to a substantially less successful momentum strategy, with or without resorting to cash when fund returns falter.

Among the decile portfolios of the Size and Value groups the average allocations of the strategy respectively went predominately to the low-cap and high book-to-market portfolios, which is as we would have anticipated given the famous Fama-French findings about market capitalization and book-to-market as factors.

The basic finding of this study is that momentum works. It works if the policy is to remain 100% invested at all times and it works with the option of substituting cash for funds whose trailing returns falter which is helpful with regard to quelling volatility and achieving of the goal of avoiding debacles such as 2007-08. Also notable is the finding that momentum works especially well with the Industry group, which is consonant with the results of an earlier analysis by Moskowitz and Grinblatt (1999) which was conducted differently.

Introduction

With appropriate adjustments for dividends, a suitable measure of the momentum of a security on a given date would simply be the price on that date divided by the price L months prior to that date, where L is the “lookback” period — a simple ratio. Research has generally shown that stocks that show the highest momentum with L fixed somewhere in the range of a few months to a year produce the best near-term returns.

The most substantial early academic account of momentum (1990) appears to have been one by Jegadeesh. He found that monthly returns of stocks yielded predictability. Consistent with findings of prior researchers, the returns on his portfolios formed monthly out of decile rankings based on a 12-month lookback period showed a high positive correlation with lookback-period returns as represented by the decile ranking; the correlation for the one-month lookback period was also large but negative.

Asness later produced (1995) a working paper that sorted out findings such as those of Jegadeesh and other predecessors and focused on the high positive momentum correlations of year-long lookback periods (L=12), while also investigating the influences of firm size and book-to-market. And to avoid the negatively-correlated most recent month, that month was simply deleted from the lookback period so that average returns over trailing months 2-12 were used — still L=12, but with that proviso concerning the most-recent month. That established a viable momentum strategy, for *stocks*.

This paper instead examines versions of the momentum strategy that might be suitable for *funds* — for efficiently-tradable ETFs, which are now abundant, and for mutual funds that can feasibly be traded on occasions as often as once a month. Given that such events as earnings surprises, drug approvals from the FDA and patents granted or overthrown may substantially affect the prices of some stocks held by a fund yet hardly affect the fund's price we certainly have to pause to consider that the utter diversifying away of specific security risk that a fund provides might be quite consequential with regard to how we should react to signals that we get from its recent price history — as we go about applying a momentum strategy to *funds*, not stocks. Is a particular rule for momentum investing that works best with a portfolio of stocks, such as the now-popular one involving the use of returns over trailing months 2-12, also the best rule for active management of a portfolio of ETFs or other tradable funds?

In the very least we certainly want a scheme for actively managing a portfolio of ETFs or other funds that avoids crashes of all kinds. But whereas we could avoid certain peculiar failures of momentum stratagems that in part involve selling lowest-momentum securities short so as to effect a hedge (Daniel and Moscowitz, 2016), by neither engaging in short sales nor buying “inverse” funds, that would leave us fully exposed to losses during panics. In lieu of incorporating short sales and inverse funds some practitioners have resorted to the commendably simple and seemingly sane solution of bailing out of long positions in funds when few if any funds are to be found with momentum exceeding that of cumulative returns on cash. Ahh... but how exactly is that to be done? And does it really work?

A particular concern arises: If, with resort to cash being permitted, the policy were to be to hold at most only a very small percentage of all of the funds on a candidates list, just those of the very highest momentum, might that mean that at the start of a bear market and for months thereafter it could still be possible to find such a small percentage of funds on the candidates list with momentum still exceeding that of cash? In that extreme of being so very choosy we should worry that the eventual resort to cash could be too tardy or weak. We would like to know what hold-at-most percentage of the candidates list would be optimal.

Other things about momentum and panics concern us. If we consider what occurs with a momentum strategy during a sudden steep decline in the stock market with a lookback period L of 12 months then we can quickly start to wonder why we hadn't chosen a *shorter* lookback period, one that would be more responsive to sudden changes. One answer might be that a shorter lookback period might cause to us to get whipsawed on the occasion of every little dip or pullback, posting a small loss each time.

The idea that the lookback period L should be treated as if it were something like a constant of nature, such as the speed of light which varies neither with the time nor the frame of reference, can reasonably be questioned. Hasn't the character of the marketplace for stocks and other securities changed radically over just the last two decades — quantitative easing, increased globalization, increased indexing and passive investing, high-frequency trading, etc.? And there is a related question: Must the lookback period really be the same for every asset class, or could it be that allowing it to be different for substantially different asset classes works better? The choice made in this paper is to examine the efficacy of determining the optimal lookback period using only relatively recent historical data, such as of the last decade.

US stock market data from the online library of Professor Kenneth R. French that extend back to 1926 were used. Different versions of the basic momentum strategy were studied, such as with or without resort to cash being allowed, with annual returns maximized and with risk-adjusted returns maximized, and with and without the Asness step of omitting the return of the most recent trailing month in the lookback period return calculation (which was developed for stocks). Statistical significance was evaluated using a combination of “out-of-sample” testing and a Monte Carlo permutation (MCP) method, both of which are easy to understand because they answer “what-if” questions that come easily to mind by actually doing the implied experiment — over and over again. Out-of-sample testing is an easily understandable and irrefutably unbiased means of “hypothesis testing” that utterly avoids “data-mining bias”, with the main drawback being that substantial quantities of data are needed. For the purposes of this article sufficient data *are* available, thanks to French's data series being so extensive and having a uniform starting date of July 1926.

Beyond its basic structure, a strategy is specified by a choice of its parameters. Academic studies of momentum have usually concluded with determinations of parameters that worked over some particular past time period. Investment managers may want to follow suit, but what if ten years goes by and the strategy isn't working? In this paper we provide a way to discover what happens, through the decades, subsequent to any use of prior data to re-optimize the strategy.

Synopsis of the Momentum Strategy with Variations

- Monthly data are used. There are no short sales.
- Optionally, the portfolio does not hold a long position in a security if the security's actual trailing lookback-period price ratio (dividend-adjusted) does not exceed the return ratio on cash.
- Out of a candidates list of N securities, the size of the allocation to any one security is fixed at 1/M where $1 \le M \le N$. No leverage is used so the integer M is also the maximum number of securities held.
- The Asness step of deleting the return of the most recent month within the lookback period is also optionally considered (though it was intended for application to stocks, not funds).
- The lookback period L has an allowed range: $1 \le L \le 18$ months.
- With N given, L and M values are found that are in-sample optimal over a multi-year data window that trails a given month in the time-domain. Performance with

the thus-chosen values is then assessed over a subsequent out-of-sample period that could also be chosen to be of any duration (but one-month and 120-month durations are featured). The given month is then advanced to the next month and the process is repeated, hence "walk-forward."

- For referencing strategy performance a benchmark portfolio is constructed that consists of the N securities on the candidates list allocated as 1/N and rebalanced monthly. Sharpe ratios are provided for various versions of the strategy and also for the benchmark and are based on returns in excess of those on cash.
- Interest is paid on the cash when trailing returns are less than the return on cash, at the FRED Effective Federal Funds rate (approximated before July 1954).

N can actually be of any magnitude and may simply represent all of the securities of an asset class or classes that are deemed to be potentially suitable as investments. For example, N could be 100 and it could ultimately be determined through optimization that M should then be 15. With M set at 15 a maximum of 15 securities would be held, each position of size $1/15^{th}$. With a long-cash portfolio that implements the option of going into cash when an insufficient number of securities on the candidates list have lookback-period returns in excess of the return on cash, if say only 8 securities out of the N were to have lookback-period returns exceeding those of cash then there would only be 8 securities in the portfolio with $7/15^{th's}$ held as cash.

Some Procedural Details

Professor French's online data library features CRSP data that he has sorted and used to form portfolios for which he has computed dividend-adjusted monthly returns. The portfolios include some based on market capitalization and book-to-market criteria that he and Professor Fama famously developed. In this paper only the US data are used.

— Portfolios of This Paper & Relevance to Currently-Active Funds

In particular, it is French's decile sorts on market capitalization and book-to-market values that are used in this paper. At the end of each June he uses his CRSP decile sorts to form a stock portfolio out of the 10% of the stocks that have the highest market capitalization, another out of the 10% that have the smallest market capitalization, and eight others in between. And likewise for book-to-market. Also used in this paper are his portfolios that are formed from 10 industrial sectors (in his notation): NoDur, Durbl, Manuf, Enrgy, HiTec, Telcm, Shops, Hlth, Utils, and Other. Within these three categories of Industry, Value and Size “Equal Weight Returns” were used (potentially important!).

The asserted relevance of this paper's findings to current-day ETFs is based on the idea that the French portfolios each serve to simulate the performance of a stock fund, such as an ETF having the corresponding stock selection criteria — a small-cap fund, a big-cap fund, a value fund, an industrial sector fund, etc.

— Number-Crunching

The computational means for performing the analyses consists of original code written by the author in the computer language Python. The program runs on a desktop computer and is done in a functional programming style using the NumPy library which greatly eases computations involving large multi-dimensional arrays via the use of “broadcasting”, “advanced indexing” and a number of built-in functions. The algorithm is however conceptually very simple and is fully specified by this paper so that any programmer could implement it in virtually any programming language. French's decile sorts and Industry group each give us 10 cumulative return records that we are treating as simulating the performance of 10 funds out of which we are forming a portfolio of funds — so that M, the maximum number of securities held, could range from one to 10. So, in all, for any particular version of the momentum strategy the number of possible combinations of the two parameters L and M that specify it is $18 \times 10 = 180$. That is indeed not a number that challenges todays desktop computers. The task is to find a combination that works optimally with in-sample testing, for use out-of-sample.

— General Features of the Out-of-Sample Hypothesis Testing Scheme

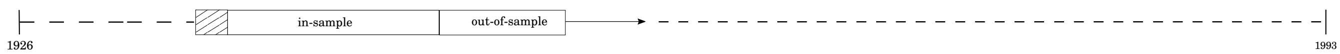
Schemes for conducting out-of-sample hypothesis testing can assume various forms. The one employed in this paper can be feasibly applied to French's monthly data from 1926 onward. We could not convincingly conduct the very same sort of procedure using, say, only a decade or so of the histories of some currently-active ETFs.

Simply put, the out-of-sample hypothesis testing procedure determines the parameters L and M through optimization of a figure of merit that measures performance, namely either the Sharpe ratio or the annualized return, using data from an in-sample period that extends back in time prior to a given past month. But then, using the thus-optimized parameter values the procedure applies the strategy to out-of-sample years thereafter and assesses the performance over those years, using both the Sharpe ratio and the annualized return. The in- and out-of-sample periods are sometimes respectively called the “training” and “testing” periods. It is only the performance figures of the strategy with the optimized parameters over the *out-of-sample* period that interest us, that we will consider to be indicative of efficacy.

This scheme generally corresponds, to a considerable degree, to what naturally happens when an investment manager reads an academic article on momentum that reports on research on choosing values of L and M (e.g., a decile) and on price data from some span of years gone by and then proceeds to implement the strategy using the article's proposed best values of those parameters — except that with the scheme of this paper we simulate that process for every available month of the historical record, thereby accumulating statistics that tell us a great deal about the odds of success should we adopt the strategy. In the main, the out-of-sample testing procedure is a faithful simulation of what can happen in the real world. And that is a good part of its charm.

But the other part of its charm stems from the fact that in an alternative mode the out-of-sample procedure can also be put into regular use monthly. It's just a matter of setting the out-of-sample period to one month rather than many months. Indeed, runs done in that way are reported on below on page 19 [cf. (a) and (b) of Figure 3].

— Details of the Out-of-Sample Hypothesis Testing Scheme



To implement the out-of-sample testing scheme, co-moving in-sample and out-of-sample "data windows" were superimposed over the historical record, as illustrated in the sketch above. The hatched area butts up to the labeled adjacent in-sample period and brackets initial data that must be compiled in order to start up the position selection rule for application by the first day of the in-sample period. Given that L was permitted to have a maximum value of 18 months the hatched area needed to be 18 months long.

As is suggested by the arrow on the figure, after one in-sample determination of the best parameters to use is completed and tested in the out-of-sample period of the window that is forward of it in time, the two windows are both moved forward in time together by one month and the entire procedure is repeated. Thus, upon completion of the entire scheme we have a time series of Sharpe ratios and one of annualized returns, one entry in each for each monthly registration of the abutted windows in the history. There is no special need for the in-sample and out-of-sample periods to be of the same duration.

Given that French's data are monthly and start in 1926, the two time series of out-of-sample Sharpe ratios and annualized returns that we compile with the testing scheme will have hundreds of entries each (the total number of months of the history minus the sum of the durations in months of the in-sample, out-of-sample and 18-month startup periods). However, it's essential to realize and to deal with the consequences of the fact that since the out-of-sample periods overlap to the n^{th} degree, the entries are hardly "independent." If we had spaced a few such in-and-out-of-sample window pairs across the history without overlap then, yes, we would have independence. But the drawback of that arrangement would be that since the outcome in each out-of-sample period would critically depend upon its placement relative to interludes of turbulent market conditions there would be volatility of the performance outcomes that would be problematical, especially so since we would have only a few such truly independent samples.

— Results at Face Value & Determining Statistical Significance Via MCP

Does the lack of independence of the hundreds of Sharpe ratios and annualized return values mean that the out-of-sample test procedure is not fully objective and unbiased? No. It is the good simulation that it appears to be. The lack of independence only has consequences for assessing the statistical significance of the findings, which concern was resolved by resorting to Monte Carlo permutation (MCP) rather than by the use of alternatives that require independence. Those two series of hundreds of Sharpe ratio and annualized return values that we get from the out-of-sample testing procedure provide a distribution of possible outcomes the like of which would befall us should we decide to henceforth adopt and use the studied momentum strategy in that way — fixing the L and M values at one point in time and thereafter using them for a substantial period of time. The found distributions being essentially ready to use as is, a first use could be to compare the mean values of the hundreds of Sharpe ratios and annualized returns with those of the benchmark portfolio over the same out-of-sample periods. Other statistics of particular interest have been compiled and are presented in Figure 1. (Figures follow the References at the end of this document.)

Given the utter objectivity of the out-of-sample testing procedure and the afforded opportunities for realistic comparisons of strategy performance with that of the benchmark, there would indeed seem to be little need to do more in the way of testing the strategy for efficacy. However, since there are nothing like hundreds of independent findings for the possible outcomes — it's as if there are only a few — we do not really have in our distribution of hundreds of out-of-sample outcomes a distribution that is fully representative of the probabilities of occurrence of all of the outcomes that could be experienced. In order to flesh out our distribution, making out of it a "sampling distribution" that is reliable in the tails well as in the core, we might want to somehow "pull ourselves up by our own bootstraps", as the saying goes, and somehow derive a good sampling distribution from the one that we have.

Indeed, there is a "resampling" procedure of statistics called "the bootstrap", which is for that very purpose. And some advanced versions of it, the "moving-block" and “stationary” bootstraps, are suitable for use with non-independent time series such as our hundreds of out-of-sample outcomes for the performance statistics. However those two procedures are rather unsettled recent developments that are complicated and are still undergoing revisions and improvements. In lieu of those, this paper makes use of MCP, which does not provide an improved distribution of possible outcomes of the strategy but which instead provides such a distribution for an antithetical “null hypothesis” and thereby deals nonetheless, in a different way, with the principal concern of whether or not the estimated good performance of the strategy is meaningful.

The MCP method is described in a working paper by Timothy Masters and in David Aronson's book (2007), in some generality. The combined out-of-sample/MCP approach taken in this paper is very close to a particular combination that is favored by Masters (2006):

> “My own favorite permutation test when training a complex model on a single dataset is to use either walkforward testing or cross validation to generate a set of out-of-sample positions, and then apply the ordinary permutation test to these results. In other words, hold out a small block of data and train the model on the remainder of the dataset. Apply the trained model to the hold-out set and record the position vector. Then replace the hold-out set and remove a different block. Repeat this test until all cases have been held out exactly once. (This is cross validation, my preference. You may wish to walk forward instead.) After the

complete set of positions is found, apply the single-sample permutation test in order to compute a probability value [p value] for the return... Of course, if you are simultaneously training several competing models..., you will need to use the best-of version of the test."

(Here, what he calls "walkforward" may indeed be taken to refer to the very scheme of co-moving in-sample and out-of-sample windows that is described above. The "several competing models" can in our case be taken to pertain to the 180 combinations of the parameters L and M that specify the strategy. And "probability value [p value] for the return" doesn't rule out instead computing the p value for the Sharpe ratio, Masters having made that clear elsewhere in his paper.)

The thus-computed "p value" is the probability that random rearrangements of the ordering of the strategy's selected positions would do as well or better than the strategy. To be a bit crude about it, the null hypothesis here is the claim that an investment advisor wielding a version of our momentum strategy is no better than a monkey pulling monthly position selections from a hat. So we fervently hope to refute the null hypothesis! A popular but entirely arbitrary choice is to declare the monkey beaten if $p < 0.05$ — if in more than 95% of many, many hat-emptying trials by the monkey he is beaten by the strategy.

Test Results— Eight Versions of Momentum

We have described the option to resort to cash should trailing price ratios fail to exceed the return ratio on cash, and, the optional Asness choice of disregarding the most recent month when calculating the trailing returns (which was intended for application to stocks, not funds, but which is tested here anyway). And a third option is to conduct the in-sample optimization using simply the return and not the Sharpe ratio (but we don't need to pick one or the other of these statistics when using them to assess *out-of-sample* performance... for that we will simply state both). These are three binary choices and so it's 2 ⨯ 2 ⨯ 2 = 8 possibilities in all.

— Performance Statistics for Industry, Value and Size Groups

Figure 1 provides a complete summary of out-of-sample results:

- Comparing the Avg Sharpe Ratios for the several strategy versions with those of the benchmarks quickly tells us the story: The strategies of columns 1, 2, 5 & 6 seem to be effective, essentially confirming that those four flavors of momentum work, more so for the Industry group than for the Value group and more for the Value group than for the Size group; and, the step of disregarding the return of the most recent trailing month, which has been found to be somewhat helpful with stocks, is shown to not be comparatively effective with any of the studied groups.
- Optimizing using the Sharpe ratio in-sample does indeed produce a somewhat better out-of-sample Sharpe ratio than optimizing with the annual return. And with the option of resorting to cash (column 6) it seems, for the most part, to help

ensure that not only Min Sharpe Ratios but also Min Ann Returns are better than those of the benchmark.

- Comparing the resort to cash option with the no-cash option while optimizing the in-sample Sharpe ratio, the former (column 6) produces substantially better out-of-sample Avg Sharpe Ratios than the latter (column 2) but that advantage comes at the expense of somewhat lower Avg Ann Returns and substantially higher Total Trades/Yr.
- The p values for the strategy versions of columns 2 and 6 are satisfactory for the Industry group but appear to either be not quite good enough or bad elsewhere on the figure. However, with the resort to cash option (column 6) the p values of the last 25 years have generally been much better — 0.01 and 0.01 for the Value and Size groups with the resort to cash option (but respectively 0.40 and 0.43 for the same groups with the 100%-invested policy as in column 2).

— Strategy Internals for Figure-1 Runs With and Without the Resort to Cash Option

Figure 2 is focused on comparisons of the out-of-sample results and in-sample parameter choices of the momentum strategy under the conditions of columns 2 and 6 of Figure 1 — for the Industry group. We now go row by row through Figure 2.

- Some of the Sharpe-ratio outcomes of the strategy fail to be better than those of the benchmark — 7% with the resort to cash option and 4% with no cash allowed — as is indicated by the extent to which the scatter-plotted points of the

Sharpe Ratio Improvements charts lie almost entirely above the drawn diagonals. But note especially that with the resort to cash option all of the untoward outcomes happened prior to 1961; with the 100%-invested policy only one of the very few that occurred happened after 1976. Also noteworthy is the failure of any Sharpe ratio to fall below or even approach zero with the resort to cash option. A Sharpe ratio of zero would mean that the portfolio's returns were on balance the same as the returns on cash. To put it another way, the resort to cash option fails at a high level — when the benchmark Sharpe ratio is high. That is almost certainly due to the momentum strategy occasionally getting “faked out” of long positions during minor pullbacks in a strong bull market phase.

- The lookback period L is very often chosen to be 1 month (but that is not a frequent choice for the Value and Size groups with the 100%-invested policy).
- If remaining 100% invested is the policy then the favored choice of M, the maximum number of securities held, is indeed 1, 1 out of 10, analogous to the top decile; with the resort to cash option much higher values of M are fairly often optimal.
- The p value charts show the values for each possible position of a 10-year wide out-of-sample window in the historical record, with the plotted date being the date of the end of the window. The mean values of these charts are the p values as stated in Figure 1. Notably, the p values of the last 25 years or so are much

improved over the worst values of prior years, especially with the resort to cash option. Somewhat high p values do not surely indicate that the strategy was outperformed by the benchmark. Of the several upward spikes on the resort to cash option p-value chart only the tall ones before 1962 were associated with strategy Sharpe ratios falling a bit below those of the benchmark.

— Supplementary Information

Figure 3 begins with some output from a the walk-forward procedure of the program being used with the out-of-sample window simply being the month that follows the trailing in-sample period. Every month we optimize the values of the lookback period L and of the maximum number of funds held M and use the optimized values to determine the positions that are to be held during the next month. The in-sample data window is moved through the history just as when in the normal hypothesis-testing mode [page 10]. As with the normal hypothesis-testing mode, this procedure leaves the in-sample window width, here 10 years, and the upper limit of the L values (18 months) as the only analyst-chosen parameters. But, on page 22 we will see exemplified the fact that the sensitivity of the outcomes to the width of the trailing in-sample window is minimal over quite a range, and the second row of Figure 2 illustrates that the analyst-chosen 18-month upper limit of L is hardly ever hit. There is a minor drawback: In this mode the value of M is not constant, which would increase the required frequency of trading somewhat.

- So charts (a) and (b) show very favorable outcomes for the strategy with the Industry group, with and without the option of resorting to cash. Note especially that while the policy of remaining 100% invested produces a very significantly higher return the resultant cumulative return plot does dip substantially more at the 2007-08 crash. This is consistent with Figure 1 which was derived using the normal hypothesis-testing mode — column 2 of that figure has higher returns than column 6 but lower Sharpe ratios. We also see, on (b), that with the resort to cash option there was a very significant decline suffered by the strategy in the early going, ending in about 1942, which decline hardly happened with the 100%-invested policy or even with the benchmark. However in other tests it was found that the strategy with the 100%-invested policy took the plunge in 1929 along with the rest of the market, as it did in 2007-08, whereas with the resort to cash option both declines were largely avoided. Not illustrated is the fact that in this walk-forward mode with the out-of-sample period being just the next month the strategy does not fare nearly as well with Value and Size groups as it does with the Industry group. But that is entirely consistent with the corresponding Figure 1 results and there are nonetheless substantial improvements over the benchmark.
- Charts (c) and (d) show that the Fama-French choices, Hi Value and Lo Size, are not-surprisingly given the highest allocations by the momentum strategy. And so the question comes to mind: What if we had simply bought and held those two

portfolios and said "To hell with momentum"? Ahh... but another question comes to mind as well: What if we had applied momentum, with the policy of resorting to cash if the trailing returns falter, directly and separately to just the Hi Value and Lo Size portfolios? These questions are answered in the Conclusions section below.

- Charts (e) and (f) are extras that pertain to the discussion of page 3 about how being too choosy with momentum could cause the resort to cash option to not do as well as hoped. However with these charts *stocks* are the securities, not funds, and an entirely conventional approach of forcing L=12 and disregarding the return of the most recent month has been taken. Note the utter failure of the 100%-invested policy to avoid the 2007-08 debacle for any value of M; it is not thus for the resort to cash option, with which M=112 misses the benchmark's return by a small margin but substantially avoids the crisis.
- Finally, chart (g) shows the L values that were found to be optimal in the normal hypothesis-testing mode of the system of this paper. So each plotted L value is one that was found to be optimal within an in-sample window of 10 years duration, the plotted date being the date at the end of the window. Not illustrated is the fact that the apparent downward trends are also evident for the Value and Size groups with the resort to cash option, but not so with the policy of remaining

100% invested. With the latter policy L=12 has often been optimal in recent decades for both the Value and Size groups.

Discussion

The computational scheme of this paper is first and foremost neither a creator nor a promulgator of nuisance parameters such as the lookback period L and the maximum number of funds held M; it is a destroyer of them. The normal hypothesis testing mode which is introduced on page 10 and the walk-forward mode with a one-month out-of-sample period which is described on page 18 both conclude with only two parameters remaining in existence: the duration of the trailing in-sample window and the very maximum value that the lookback period L is permitted to have. The latter, 18 months, is seldom touched during the runs [cf. Figure 2, 2nd row] and the performance outcomes are insensitive to the former. For example the Avg Sharpe Ratios for the Industry group with trailing in-sample window widths of 10 years, 20 years and 30 years with the policy of resorting to cash when trailing returns are exceeded by those of cash are respectively 0.99, 1.01 and 0.97; the corresponding values for the Avg Annual Return % values are 23.4, 21.3 and 22.0 . For the Value group the corresponding figures are 0.82, 0.85, 0.85 and 18.2, 17.6, 17.4; for the Size group they are 0.76, 0.76, 0.73 and 16.1, 16.2, 15.7 . The runs were configured so that the values all pertain to the same range of out-of-sample window positions, covering 1958-2017. The variations are rather insubstantial, are not systematic and are of the same character in both of those regards when the policy is instead to remain 100% invested. The most helpful aspect of this is that it seems to mean that to apply the momentum strategy to funds we may only need about a dozenyear history (10 years for the in-sample window plus a year and a half to allow the lookbacks to happen and produce the first position sizes).

We now briefly address the matter of how momentum might possibly be implemented so as to allow us to also profit from the Fama-French findings about Value and Size as factors. The following table is of runs of the normal hypothesis-testing mode of the program, in the configuration that was used to produce column 6 of Figure 1.

	Sharpe Ratio	Return	Sharpe Ratio	Return
	Hi Value		Lo Size	
Buy-and-Hold	0.67	21.1	0.48	16.6
Strategy (With Resort to Cash)	0.97	22.0	0.85	18.8
	Entire Value Group		Entire Size Group	
Strategy (With Resort to Cash)	0.85	17.3	0.78	15.4

So clearly the momentum strategy with the policy of resorting to cash applied separately to just the Hi Value and Lo Size portfolios and not to either of the groups to which they belong produced the better outcomes. Trades per year were about three or four, with the average allocation being about 55-60%. The p values were 0.12 for Hi Value and 0.08 for Lo Size. However strategy Sharpe ratios respectively failed to exceed those of buy-and-hold 20% and 13% of the time.

We have raised the question, is it somehow necessary or advisable that every asset class be given the same universal lookback period? Trying to answer that definitively would be beyond the scope of this paper, but we can resolve one little thing: What happens to the strategy performance with the Industry group if, instead of allowing both

L and M to be optimized in the trailing in-sample training period, we fix L=12 (the popular choice for stocks, which was found to fairly frequently be optimal for both the Value and Size groups with the 100%-invested policy)? The answer, for the policy of remaining 100% invested, is that the Avg Sharpe Ratio and Avg Ann Return % drop substantially from 0.91 and 26.7 as in column 2 of Figure 1, to decidedly less-optimal values of 0.76 and 21.7 .

A general procedure for finding optimal and adaptive momentum strategies for portfolio management has been presented above, along with understandable methods for determining the odds of success and also the odds of the seeming success having been due to chance. The latter odds have been substantial at times, especially in the early decades of the period of record from 1926. But momentum seems to have been improving! The cumulative return plots for the strategy on (a) and (b) of Figure 3 show a distinct increase in the rate of growth starting at roughly 1975, which is more pronounced for the 100%-invested policy. And p values have also generally declined, for Industry, Value and Size groups with the resort to cash option and also for the Industry and Value groups with the 100%-invested policy (with that policy Size-group p values have no apparent trend).

Among the critically-different choices to be made are the following:

- With the procedures of this paper we have found that short lookback periods (predominately L=1) have been optimal for Industry, Value and Size groups if the resort to cash option is adopted, and for the Industry group even if it isn't. This lookback-period finding is generally contrary to academic research on stocks. Forcing L=1 on the Value and Size groups with the 100%-invested policy, which policy is usually assumed in academic research, produces higher Sharpe ratios and annual returns than the benchmark — meaning also that were we to use a longer lookback period we could hardly discard the returns of the most recent month, as has been the recommended practice for stocks.

- It is shown above that the resort to cash option successfully quells volatility and leads to the highest possible Sharpe ratio. It is especially helpful at containing losses during panics. But this comes at the expense of not maximizing returns and having to trade more frequently.
- It is imperative that with the resort to cash option the investment manager not be too choosy, not set M the maximum number of funds to be held at just a small fraction (e.g. 10%) of the funds on the candidates list. With a substantially larger fraction major declines in the market would cause a quicker resort to cash as it would become increasingly impossible to find that many funds on the candidates with lookback-period returns in excess of those on cash. The critical difference here is between the resort to cash option and being 100% invested. With the latter policy it *does* pay to be choosy.

References

Narasimham Jegadeesh (1990), “Evidence of Predictable Behavior of Security Returns”, The Journal of Finance 45(3), July, 881–898 [https://doi.org/10.1111/j.1540-6261.1990.tb05110.x]

Clifford S. Asness (1995), “The Power of Past Stock Returns to Explain Future Stock Returns”, Social Science Research Network (SSRN), January [https://ssrn.com/abstract=2865769]

Tobias J. Moskowitz and Mark Grinblatt (1999), “Do Industries Explain Momentum?”, Journal of Financial Economics 54(4), August, 1249-1290 [http://dx.doi.org/10.1111/0022-1082.00146]

Timothy Masters (2006), “Monte-Carlo Evaluation of Trading Systems”, Working Paper [http://evidencebasedta.com/montedoc12.15.06.pdf]

David R. Aronson (2007), *Evidence-Based Technical Analysis*, John Wiley & Sons, Inc. [ISBN: 978-0-470-00874-4]

Kent Daniel and Tobias J. Moskowitz (2016), “Momentum Crashes”, Journal of Financial Economics 122(2), November, 221-247 [https://doi.org/10.1016/j.jfineco.2015.12.002]

DWMM (2016), “Momentum Demystified”, Systematic Relative Strength– the Official Blog of Dorsey Wright Money Management, December 28 [http://business.nasdaq.com/marketinsite/2017/Momentum-Demystified.html This is a 5/9/2017 reprint by the acquirer of the firm, from which the part admitting to the 2008 momentum failure has been deleted. But a link to additional information from Morningstar that refers in part to the failure is provided.]

Clifford S. Asness (2017), “The August of our Discontent: Once More Unto the Breach?”, Cliff's Perspectives, August 7 [https://www.aqr.com/cliffs-perspective/the-august-of-our-discontent-once-more-unto-the-breach]

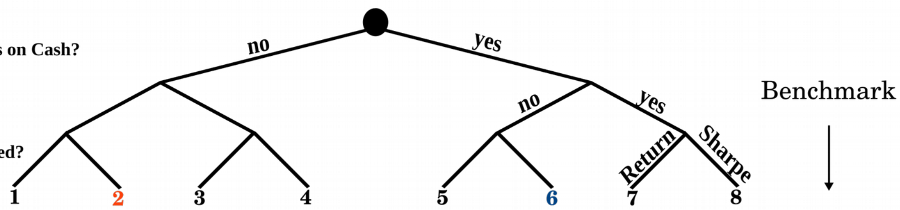
Figure 1: Out-of-Sample Results of In-Sample-Optimized Momentum Strategies

Derived from French Monthly Dividend-Corrected Returns for Decile & Industry Groups— June 1926 to October 2017

Require Trailing Returns to Exceed Returns on Cash?

Disregard Most-Recent Month's Return?

In-Sample Return or Sharpe Ratio Optimized?



Group		1	2	3	4	5	6	7	8	Benchmark
Industry	Avg Sharpe Ratio:	0.84	0.91	0.64	0.69	0.83	1.04	0.54	0.58	0.53
	Avg Ann Return %:	25.6	26.7	18.2	18.9	24.3	22.0	15.2	14.9	14.4
	Min Sharpe Ratio:	-0.16	-0.15	-0.27	-0.07	-0.30	0.30	-0.37	-0.17	-0.14
	Min Ann Return:	2.5	2.5	-1.5	1.0	0.8	7.8	-4.4	2.7	3.3
	p Value:	0.10	0.05	0.21	0.12	0.13	0.05	0.33	0.29	NA
	Avoid 2007-08 Debacle:	Partly	Partly	No	No	Yes	Yes	No	No	No
	Avg % Long:	100	100	100	100	84	58	89	84	100
	Total Trades/Yr:	13	16	11	15	12	30	12	16	120
Value	Avg Sharpe Ratio:	0.67	0.67	0.61	0.61	0.72	0.85	0.54	0.55	0.53
	Avg Ann Return %:	19.5	19.1	18.2	17.8	17.5	17.3	14.7	13.3	14.7
	Min Sharpe Ratio:	0.02	0.02	-0.05	-0.02	0.05	0.06	-0.25	-0.22	-0.15
	Min Ann Return:	5.4	6.3	5.3	5.5	7.9	5.1	2.2	2.9	3.0
	p Value:	0.18	0.13	0.25	0.23	0.18	0.13	0.38	0.36	NA
	Avoid 2007-08 Debacle:	No	No	No	No	Partly	Partly	Partly	Partly	No
	Avg % Long:	100	100	100	100	70	60	84	78	100
	Total Trades/Yr:	7	13	8	13	13	29	8	16	120
Size	Avg Sharpe Ratio:	0.61	0.62	0.55	0.53	0.68	0.78	0.49	0.47	0.47
	Avg Ann Return %:	17.9	17.4	16.2	15.3	16.0	15.4	13.5	11.6	13.0
	Min Sharpe Ratio:	0.04	0.05	-0.25	-0.25	0.03	0.19	-0.08	-0.17	-0.24
	Min Ann Return:	5.4	3.9	1.6	1.6	3.5	5.9	1.9	1.9	1.3
	p Value:	0.21	0.18	0.27	0.33	0.18	0.11	0.38	0.38	NA
	Avoid 2007-08 Debacle:	No	No	No	No	Partly	Partly	Partly	No	No
	Avg % Long:	100	100	100	100	71	68	82	82	100
	Total Trades/Yr:	10	15	8	11	14	33	8	16	120

- In- and out-of-sample windows are both of 10 years duration.
- The out-of-sample data windows occupy the range from June 1939 to October 2017.
- Sharpe ratios are derived from returns in excess of cash.
- The p values pertain to the Sharpe ratios.
- Annual returns are total returns.
- "Avg" and "Min" respectively refer to averaging and to taking the minimum over all of the registrations of the co-moving in- and out-of-sample windows with respect to the historical record.

Figure 2: Industry Group With & Without the Resort to Cash Option

With the In-Sample Sharpe Ratio Optimized as in Columns 2 and 6 of Figure 1

Corresponding to Column 2 of Figure 1 | Corresponding to Column 6 of Figure 1

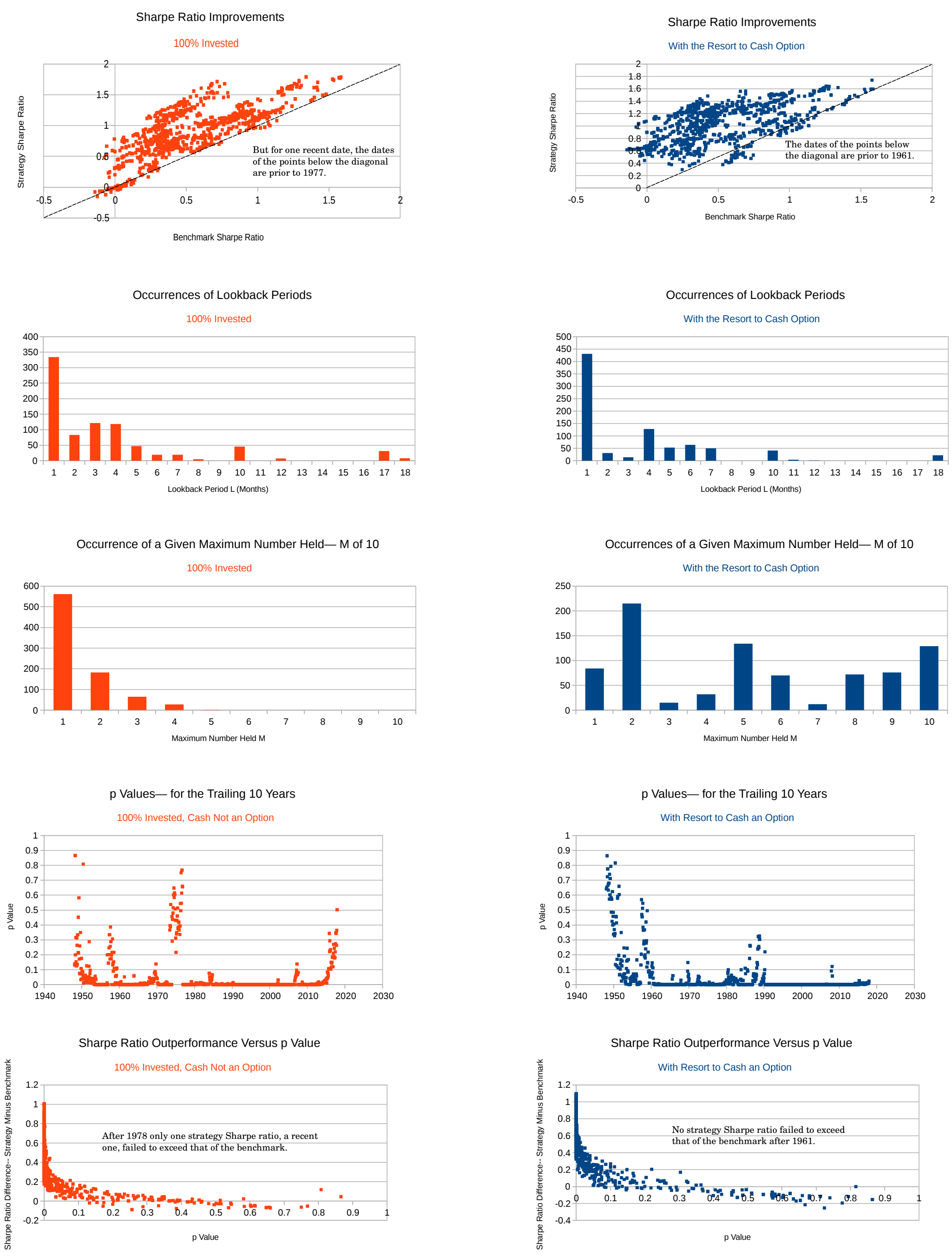
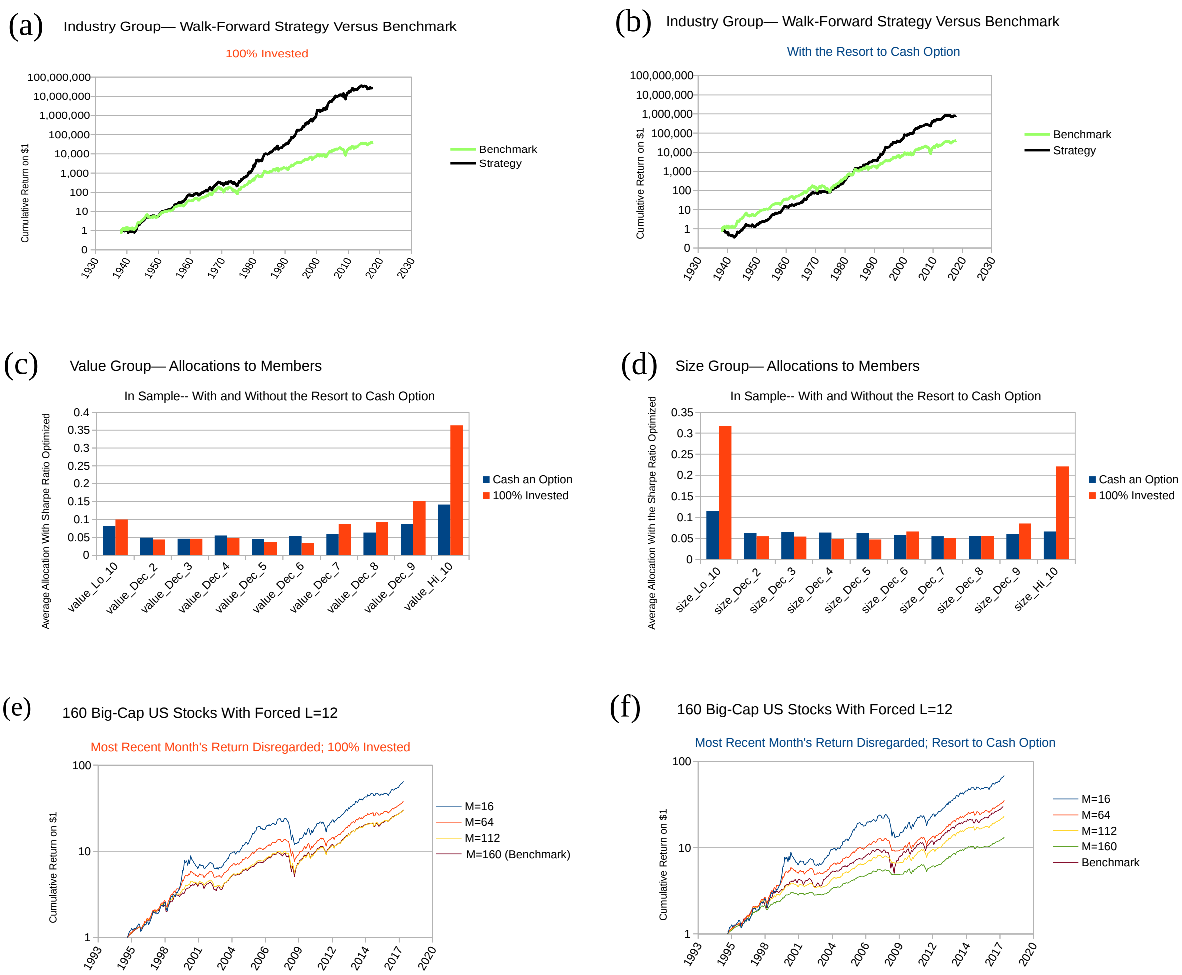
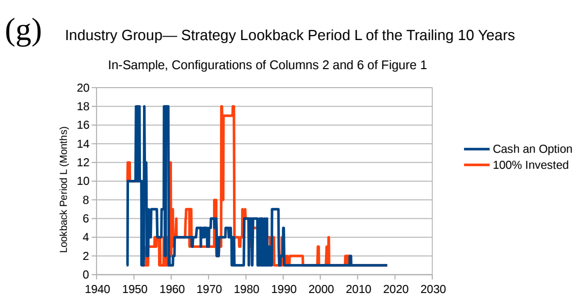


Figure 3: Special In-Sample and Out-of-Sample Results of Interest

All But (e) and (f) Derived from French's Monthly Dividend-Corrected Returns— June 1926 to October 2017





- ∗ In-sample windows were of 10 years duration.
- ∗ Sharpe ratios are derived from returns in excess of cash.
- ∗ In-sample Sharpe ratios were optimized, not returns.
- ∗ Charts (a) through (d) and (e) and (f) are of out-of-sample data.
- ∗ Data of chart (c), (d) and (g) are in-sample.