Patience in investing is the capacity to endure periods of underperformance in hopes of achieving an investment objective. For many investors, outperforming the broad market through traditional active management or factor-based investing is a key goal—and one that requires significant patience to achieve.

In this paper, we quantify the wide range of frequencies, durations, and magnitudes of underperformance that both equity factor tilts and outperforming traditional active managers experience. Our findings show that almost all outperforming traditional active equity managers and equity factors have frequent periods of underperformance relative to the equity market, some of which are long in duration and large in magnitude.

We find that investors who have selected an outperforming manager or use factor-based investing can expect to experience a drawdown between 40% and 60% of all one-year evaluation periods. Additionally, an outperforming manager can expect to experience a continuous drawdown lasting two years or more every ten years and a drawdown of at least 20% over time.

Acknowledgment: The authors would like to thank Samantha Wetzel for her significant contribution to this research.
Introduction

As with most things in life, success in investing requires patience; investing, in fact, probably requires more patience than most endeavors. You need patience when what you are invested in is performing poorly—and you need it when what you haven’t invested in is performing well. At any given time, you might need to have extra patience with the market, a sub-asset class, a particular region, a strategy, an individual manager, or even the cash you have on the sidelines. As Benjamin Graham once said, “The investor’s chief problem—even his worst enemy—is likely to be himself.”

Unfortunately, patience is something that can easily wear thin if you aren’t benefiting from the current trend of an investment. A lack of patience can ultimately lead to bad decision-making if you give in to anxiety or deviate from the course that you initially intended. The truth is, patience is different for every investor. It’s ultimately rooted in your investment philosophy, conviction, and emotional fortitude. Investors aligned with a strategic asset allocation may need patience when an asset class is performing poorly. Other investors may need patience when a tilt toward a style factor or strategy is underperforming. Those in an actively managed fund will need patience with both the manager and the strategy.

However, there is good news: Our research shows that patience can be bolstered by education, data, a full understanding of the risks and opportunities involved, and, for many investors, the counsel of an advisor. Understanding what is reasonable to expect and how often to expect it can improve one’s ability to have patience when investments are not performing well. An investor must have a firm grasp of their current path and why it will help them reach their goals. One of the major benefits of using professional financial advice is behavioral coaching, which includes helping investors manage their expectations and thereby increase their patience with their investment choices. Kinniry et al. (2019) quantified the potential value of behavioral coaching for the typical client as 1.5% in net returns.

Most prior research around this topic addresses investors’ patience with an investment manager who is overseeing the overall portfolio (Goyal and Wahal, 2008) or when investing in the equity market.

We look to build upon this research because we know that patience is a key factor in active management success (Wallick, Wimmer, and Balsamo, 2015) and that active investors have tended to sell active investments when they are underperforming over relatively short periods. Investors’ levels of conviction in a manager and their level of active risk tolerance will directly affect their level of patience with that manager. The higher their conviction and risk tolerance, the more patience they are likely to have. Additionally, investors themselves may have a difficult time defining their level of conviction or active risk tolerance, and, therefore, the amount of patience that they might have.

We first look at the type of patience an investor historically needed when actively targeting common long-only equity style risk factors. We evaluate patience across five long-only equity factors and combinations of factors: size, value, momentum, low volatility, and a multifactor proxy. These factors have “worked in the past” and historically generated positive risk-adjusted excess returns relative to the market.

Next, we evaluate the patience investors would have needed if they had invested in a historically outperforming traditional active manager. There is no guarantee, of course, that a historically outperforming traditional active manager will be successful in the future—but if an investor does believe the manager will outperform, our analysis shows the levels of patience that is necessary for success.

Our research builds on the existing literature by quantifying the levels of “pain” investors might experience and providing insights investors can use when making time-dependent decisions in their portfolios across various active strategies and managers. In our analysis, we intentionally avoid a discussion of when to stop having patience and exit an investment; rather, we look to help investors and those that advise them understand what they might expect when investing in active strategies. We consider the necessary levels of patience across three dimensions of underperformance: frequency, magnitude, and duration.

---

1 Aside from parenting, marriage, and family holidays.
3 Although the two terms are similar, conviction is more about someone’s certainty of their belief in a specific investment to achieve an outcome, while active risk tolerance is the degree to which an investor can tolerate active performance uncertainty.
4 For further details on the empirical efficacy behind these factors, see Fama and French (1993), Carhart (1997), and Baker, Bradley, and Wurgler (2011).
Data and methodology

We begin by defining a few key terms. The period of underperformance, or drawdown period, is the length of time (usually measured in months) that a portfolio declines in value relative to a benchmark. Measured from a peak until the value recovers to the peak level at which the decline began, each drawdown period consists of a peak-to-trough phase and a recovery phase. The peak-to-trough phase is the number of months it takes for the drawdown magnitude to be realized (that is, the number of months from the peak to the trough). The drawdown magnitude is the cumulative peak-to-trough loss in portfolio value relative to a benchmark that occurs during the drawdown period. The recovery phase is the number of months it takes to offset the drawdown magnitude. Figure 1 illustrates these relationships. Factor tilts and traditional active funds can have multiple drawdowns over time.

Factor analysis data and methodology

Our analysis of equity factors uses data from the Kenneth R. French Data Library for the long-only size, value, momentum, and low-volatility (beta) factor returns and the risk-free (1-month Treasury bill) rate; see the Appendix for more information. For the multifactor return—and understanding that there are many ways to define multifactor—we used a portfolio equally weighted across size, value, momentum, and low volatility minus the market.

Figure 1. Key drawdown terms

Notes: Figure is for illustrative purposes only. Data are the month-end returns of the Standard and Poor’s 500 Index from January 1, 1995, through December 31, 2019. Past performance is not a guarantee of future results. Note that hypothetical illustrations are not exact representations of any particular investment. Index is unmanaged; therefore, direct investment is not possible. Net asset value indicates the cumulative value of one U.S. dollar invested at the beginning of the period. Sources: Vanguard calculations as of May 2020, based on data from Morningstar, Inc.
We calculate the factor excess returns for overlapping one-, three-, and five-year periods over the 25 years ended December 31, 2019, using month-end returns.\(^5\) We calculate the annualized excess return for the factors to show the additional returns an investor could have received for having patience. In addition, we calculate and analyze various drawdown metrics for each factor:

**Frequency**—the frequency of the periods of underperformance

**Magnitude**—the worst drawdown magnitudes (peak to trough) over various time periods

**Duration**—the longest period of underperformance, as measured by the length of time between a factor’s peak-to-trough phase and its subsequent return to that peak.

Traditional active equity data and methodology

Our traditional active analysis begins with the oldest share class for all U.S. domiciled actively managed open-ended equity funds in the Morningstar Direct database (including surviving and obsolete funds) during the 25 years ended December 31, 2019.\(^6\) We then remove any fund with fewer than ten holdings or less than ten years of returns, along with any fund not categorized in the traditional U.S. Morningstar nine-box style factors, emerging markets, foreign large-capitalization, or foreign small-/mid-cap categories during our sample period.\(^7,8\) This produces a sample of 2,593 funds.

Next, using each fund’s relevant Morningstar style benchmark, we calculate that fund’s cumulative net excess return over our sample period. Funds with a positive cumulative return net of expense ratio are labeled outperformers, giving us our final sample of 1,173 outperforming funds. This sample has 22,382 total years of performance and an average life of 19 years.

Looking at the sample in terms of fund category, we calculate the monthly cross-sectional net excess returns for each of the 300 months, then annualize the results to show what an investor might have received for having patience. We also calculate each fund’s net excess returns for each overlapping one-, three-, and five-year period and use this data to analyze various drawdown metrics for each outperforming fund. Most of these metrics fall under one of the three categories mentioned earlier, in the following ways:

**Frequency**—how often there were periods of underperformance relative to the fund’s style benchmarks, median peer, and 25th-percentile peer.

**Magnitude**—the worst underperformance for each fund over various time periods, and whether funds experienced drawdowns of various magnitudes.

**Duration**—the rate of drawdowns of various durations and the longest period of underperformance as measured by the length of time between a fund’s peak and its subsequent return to that peak.

---

\(^5\) Although we acknowledge that using a longer time series or a different starting point for the factor analysis may lead to differing results, we find that a period of 25 years gives us a fair representation of the characteristics of U.S. long-only equity factors and allows comparisons to our 25-year outperforming fund sample.

\(^6\) Henceforth referred to as “funds.”

\(^7\) The Morningstar Style Box is a nine-square grid that provides a graphical representation of the “investment style” of stocks and mutual funds. For stocks and stock funds, it classifies securities according to market capitalization and growth and value factors.

\(^8\) We define our sample’s categories as follows: emerging markets as diversified emerging markets; foreign large-capitalization as foreign large-cap blend, large-cap growth, and large-cap value; and foreign small-/mid-capitalization as foreign small-/mid-cap blend, small-/mid-cap growth, and small-/mid-cap value.
Historical patience results

What are patient investors getting for what they might endure? As Figure 2 shows, the median annualized net excess return generated by the funds in our sample is 0.9%; across style, size, and region, the median ranges between 0.4% and 2.0%. The table portion of the figure also shows that the average annualized excess returns of the long-only equity factors ranged from 0.9% to 2.9%, similar in magnitude to the outperforming active funds.

Figure 2. Annualized net excess returns of outperforming funds and factors (1995–2019)

<table>
<thead>
<tr>
<th>Long-only equity factors</th>
<th>Annualized excess returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>1.1%</td>
</tr>
<tr>
<td>Momentum</td>
<td>2.9</td>
</tr>
<tr>
<td>Low volatility</td>
<td>0.9</td>
</tr>
<tr>
<td>Size</td>
<td>0.9</td>
</tr>
<tr>
<td>Multifactor</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Notes: Data are for the period from January 1, 1995, to December 31, 2019. We calculated the annualized net excess return of each outperforming fund relative to its style benchmark as determined by Morningstar; we calculated the annualized excess returns for the long-only equity factors using the factors as described in the Appendix. The sample of outperforming funds totals 1,173. By category, the sample is made up of 83 U.S. large-cap blend, 192 U.S. large-cap growth, 88 U.S. large-cap value, 27 U.S. mid-cap blend, 85 U.S. mid-cap growth, 24 U.S. mid-cap value, 97 U.S. small-cap blend, 156 U.S. small-cap growth, 61 U.S. small-cap value, 235 foreign large-cap, 48 foreign small-/mid-cap, and 77 diversified emerging market funds.


9 The average annualized net excess return generated by the funds in our sample is 1.3%, across style, size, and region, the average ranges between 0.7% and 2.2%.
10 These results include neither implicit nor explicit implementation costs of the factors; for more information on implementation costs, see Novy-Marx and Velikov (2016) and citations therein.
Patience with U.S. long-only equity style factors

Well-known equity factors have historically generated positive risk-adjusted excess returns relative to the market, with extensive academic literature and empirical evidence supporting their potential long-term persistence (Grim et al., 2017). Successful long-term factor-based investing requires both the conviction that a positive risk-adjusted excess return will be realized going forward and the patience to stay invested through the inevitable drawdown periods. As we will show, all factor-based investing experiences extended stretches of relatively weak performance; what varies are the degrees of frequency, magnitude, and duration across the factors.

Our factor analysis uses excess returns, because we wanted to understand relative performance and the patience necessary when deviating from the market. As with other forms of active management, the performance of factor tilts relative to the broad markets is difficult to predict.

We first address the frequency with which the factors underperformed. In our analysis, frequency is defined as the percentage of overlapping time periods where a factor underperforms its benchmark. As Figure 3 shows, most factors underperform for about half of the overlapping one-year periods, with the frequency generally decreasing as the evaluation period increases.

Figure 3. Frequency of factor underperformance, as measured by percentage of overlapping time periods for which factors underperformed the market (1995–2019)

<table>
<thead>
<tr>
<th></th>
<th>1 year</th>
<th>3 year</th>
<th>5 year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>48%</td>
<td>49%</td>
<td>47%</td>
</tr>
<tr>
<td>Momentum</td>
<td>52%</td>
<td>47%</td>
<td>41%</td>
</tr>
<tr>
<td>Low volatility</td>
<td>46%</td>
<td>45%</td>
<td>32%</td>
</tr>
<tr>
<td>Size</td>
<td>44%</td>
<td>45%</td>
<td>36%</td>
</tr>
<tr>
<td>Multifactor</td>
<td>35%</td>
<td>31%</td>
<td>25%</td>
</tr>
</tbody>
</table>

Notes: Data are for the period from January 1, 1995, to December 31, 2019. Using monthly excess returns over the market, we calculated overlapping one-, five-, and ten-year performance of each long-only equity factor before costs. The market is defined as the value-weighted return of all CRSP firms incorporated in the U.S. and listed on NYSE, AMEX, or NASDAQ. The data presented is the percentage of the overlapping time periods that were negative. See Appendix for factor descriptions.

Sources: Vanguard calculations as of May 2020, based on data from Morningstar, Inc., and the Kenneth R. French Data Library.

11 We chose to evaluate the factors by using excess returns for the factors relative to the returns of the U.S. equity market and do not analyze the factors on a risk-adjusted basis. This is especially relevant for the low-volatility factor, which focuses on stocks that have historically exhibited lower volatility.
To determine if there was a difference in the frequency of drawdowns between factors, we test the difference of the means across all five equity factors and find a statistically significant difference to 5% between the frequency of momentum underperformance and the frequency of underperformance of most of the other factors across the three different time periods. This translates into momentum underperforming less frequently than other factors (between 7% and 22%, depending on the factor and the length of the evaluation period). Although we do not consider whether this will repeat itself in the future, our result is in line with Barraso and Santa-Clara (2015), who find that momentum is subject to severe, if less frequent, “momentum crashes.”

Many investments are evaluated on their performance over the most recent one-, three-, and five-year windows. It makes sense, then, for our next question to be, What is the worst drawdown for the factors over these time frames? Following similar methodology by Grim et al. (2017), we determine the magnitude of the underperformance that each equity factor has had during different periods (see Figure 4). All factors had drawdowns across all evaluation periods, and the magnitudes of underperformance were similar over the worst one-year period—but we see a wide range of magnitudes at five years.\(^2\)

Figure 4. Investors must be able to withstand difficult stretches of underperformance

Notes: We calculated performance of five long-only equity factors using monthly excess returns over the market for the period from January 1, 1995, to December 31, 2020. The market is defined as the value-weighted return of all CRSP firms incorporated in the U.S. and listed on NYSE, AMEX, or NASDAQ. For a description of the factors, see the Appendix. The data presented are each factor’s worst performance over the various aggregated time periods. Past performance is not a guarantee of future results. Note that hypothetical illustrations are not exact representations of any particular investment. Index is unmanaged; therefore, direct investment is not possible.

Sources: Vanguard calculations as of May 2020, based on data from Morningstar, Inc., and the Kenneth R. French Data Library.

\(^2\) We note that using different definitions for factors can lead to significantly different results. For example, Grim et al. (2017) used investable global indexes and found that volatility’s worst five-year cumulative return compared with the broad market was approximately –45%; when we used the U.S. only data from the Kenneth R. French Data Library, however, the result was –124%.
The future will always be different and could be better or worse. Even so, it helps to have a sense of what the magnitude and length of the maximum drawdown periods and longest drawdown periods have been historically. Some investors may not have the patience to withstand how large the drawdown magnitude might come to be, regardless of how long the drawdown period. Other investors may instead struggle with the length of the drawdown period. Therefore, we evaluate both. Figure 5 shows the maximum drawdown by length of time split into the two phases, peak-to-trough and recovery, along with maximum drawdown magnitude by percentage. We find that there are significant differences across the factors. Of special note is that the magnitude and length of the largest drawdowns for most of the factors have not been over the same time periods. The largest-magnitude drawdowns were in the late 1990s during the technology bubble, but the longest length of underperformance is much more recent (and is continuing as of this writing). Another key point is that multifactor doesn’t display the same magnitude or length of drawdowns as most single factors. This difference, in fact, is linked to one of the prime attractions of a multifactor product: its diversification benefit, which

Figure 5. Across equity factors, the lengths and magnitudes of the worst periods of underperformance can be quite large

<table>
<thead>
<tr>
<th>Factor</th>
<th>Maximum length of underperformance</th>
<th>Maximum drawdown magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Magnitude</td>
<td>Time period</td>
</tr>
</tbody>
</table>

Notes: We calculated each equity factor’s maximum (largest) cumulative decline relative to the equity market over the 25 years from January 1, 1995, through December 31, 2019. In the chart, the drawdown magnitude of each factor during its maximum drawdown period is shown as a blue diamond. The chart also shows each factor’s maximum length of underperformance. The table gives additional information, including the drawdown magnitudes for each factor’s maximum length of underperformance. Magnitude is the cumulative peak-to-trough loss in portfolio value relative to a benchmark that occurs during the drawdown period, length of underperformance is the total time from the start of the drawdown period until the cumulative magnitude recovers to zero. The drawdown period has two phases: the peak-to-trough phase and the recovery phase.

Sources: Vanguard calculations as of May 2020, based on data Morningstar, Inc., and the Kenneth R. French Data Library.
lets the investor take advantage of the performance cyclicalcy of the single factors. An investor holding such a product may therefore need less patience.

Patience with traditional active management

As noted earlier in this paper, success in traditional active investing requires both a conviction that the manager will outperform in the future and the active risk tolerance to stay invested through the underperforming or drawdown periods. Both outperforming active managers and outperforming factors, by definition, have historically generated positive net excess returns relative to the market—but an investor may have a different level of conviction for an active manager than they do for factors, either generally or during periods of underperformance.

The factors that we evaluated are those that have historically outperformed by generating higher returns, reduced volatility, or both. In other words, they have “worked” in the past.13 While there is no guarantee the factor outperformance will persist, there is a large body of evidence supporting some of the more common factors discussed here.14 That said, when a factor goes through extended drawdown periods (as we are currently seeing with a number of the factors), the efficacy of the factor starts to come into question. Similarly, an investment with an active manager requires the investor to not only have patience but to believe that the manager has the talent, skill, or edge to outperform after a period of underperformance.

We analyze historically outperforming funds to illustrate, using several different metrics, the kind of patience that has been necessary to succeed with active strategies. We first test to see whether the funds in our sample experienced a one-, three-, or five-year drawdown relative to their style benchmark, their median peer group, or the bottom quartile of their peer group. We find that close to 100% of outperforming funds have experienced a drawdown relative to their style and median peer benchmarks over one-, three-, and five-year evaluation periods—and that that 80% of outperforming funds had at least one five-year period where they were in the bottom quartile. This is especially important to understand given the results of a fairly recent survey of senior executives with asset allocation responsibilities for large institutional investors (State Street Global Advisors, 2016). The survey found that 89% of these executives would not tolerate underperformance for more than two years before seeking a replacement.

13 We focus on excess return in this analysis, but outperformance can also be measured on risk-adjusted basis.
14 See Footnote 4 for research on the efficacy of factors; see the Kenneth R. French Data Library for factor data dating back to 1926.
After finding that most outperforming active managers experienced drawdowns, we look at the frequency which the managers underperformed. Defining frequency as the percentage of time periods where the manager underperformed one of the three defined benchmarks, we first test the difference of the means across 12 previously defined style categories for each of the three evaluation periods and three different benchmarks—and find that while a small proportion of the differences between the styles were statistically significant to 1%, none were economically significant. Although there may not be significant differences across styles, we do find that individual managers within each style box had a broad range of levels of frequency of underperformance (see Figure 6).\textsuperscript{15}

As shown in the first bar of the figure, 837 out of 1,173 funds, or about 70% of outperforming funds, underperformed their style benchmarks between 40% and 60% of all one-year evaluation periods. Unsurprisingly, the percentage of the time periods for which outperforming funds underperform their style and peer benchmarks decreases as the evaluation period lengthens. We also find, as shown in the last bar, that outperforming funds are significantly less likely to frequently be in the bottom quartile of their peers, with more than 65% (771 out of 1,167 funds) of the outperforming funds doing so in fewer than 20% of the five-year periods.

Figure 6. Even funds with very good records of performance are likely to frequently underperform compared to various benchmarks

Notes: We evaluated all U.S.-domiciled, Morningstar nine-style-box U.S. active equity, emerging markets, and developed market foreign funds with a minimum of ten years of performance data over the period from January 1, 1995, to December 31, 2019, relative to their style benchmark and identified all net outperforming funds. For each group, we calculated overlapping one-, three-, and five-year performance for each year of the period and measured it relative to the relevant style benchmark, median peer, and 25th-percentile peer returns over the same period. The data presented are the percentage of the time periods that were negative, separated into five quintiles. Although there were statistically significant differences between the means of the categories, these differences were economically insignificant. Therefore, all style boxes were aggregated.

Sources: Vanguard calculations as of May 2020, based on data from Morningstar, Inc.

\textsuperscript{15} Although our analysis did not find economically significant differences across the styles of traditional active managers, we did find a wide cross-sectional dispersion across the individual managers. This suggests that a future body of research could address individual fund characteristics and how they might relate to the various drawdown characteristics.
After evaluating whether outperforming active managers underperform and how frequently they underperform, the next question we test is: When they do underperform, how bad does it get? Figure 7 presents the worst one-, three-, and five-year returns for each fund relative to style and peer benchmarks. We find that as we increase the time period the median manager’s worst performance increases and the dispersion of individual manager drawdowns increases dramatically, especially relative to style benchmarks.

Interestingly, when we compare the worst one-year performances of active funds with the worst one-year performances of factors, we find active funds averaging 19% (median 15%) and factors averaging 24%.

Additionally, the funds’ worst drawdowns for one, three, and five years have a similarly large range of outcomes across the outperforming managers and factors, with the range widening with time.

Some investors may lose patience if a manager underperforms by specific amounts. As shown in Figure 8 on page 12, we determine the percentage of outperforming funds that had drawdowns greater than various thresholds and find the likelihood of drawdowns breaching various performance levels to be similar for outperforming funds relative to style and median-peer-group benchmarks.

Figure 7. The magnitudes of the worst one-, three-, and five-year excess returns across the individual funds show a wide range of underperformance

Notes: We evaluated all U.S.-domiciled, Morningstar nine-style-box U.S. active equity, emerging markets, and developed market foreign funds with a minimum of ten years of performance data over the period from January 1, 1995, to December 31, 2019, relative to their style benchmark and identified all net outperforming funds. For each group, we calculated overlapping one-, three-, and five-year performance for each year of the period and measured it relative to the relevant style benchmark and median peer returns over the same period. The data presented are each outperforming fund’s worst performance over the various aggregated time periods, with the 5th-, 25th-, 50th-, 75th- and 95th-percentile fund returns shown. Although there were statistically significant differences between the means of the categories, these differences were economically insignificant. Therefore, all style boxes were aggregated.

Sources: Vanguard calculations as of May 2020, based on data from Morningstar, Inc.
The figure shows that 50% to 60% of outperforming active equity funds have underperformed their style and median peer benchmark by 20% or more. Across all return thresholds, we see a linear relationship between the magnitude of underperformance and the percentage of outperforming funds breaching those magnitude thresholds, with a 2.5–3.0% decrease in the percentage of funds per 1% drop in magnitude of underperformance. Although we found earlier that almost all outperforming funds end up in the bottom quartile peer group over various time periods, very few underperform the 25th percentile return by any significant amount.

All funds, including outperforming funds, will go through multiple drawdowns over their lives, some extremely small in magnitude and some, as we saw in Figure 8, fairly large. In terms of how often to expect continuous drawdowns of various durations (see Figure 9), we find that over a ten-year period investors should expect their outperforming manager to have, on average, one continuous drawdown relative to their style and median peer benchmark lasting two years or more.16

Figure 8. The majority of outperforming funds had drawdowns greater than 20%

![Figure 8](image)

Notes: We evaluated all U.S.-domiciled, Morningstar nine-style-box U.S. active equity, emerging markets, and developed market foreign funds with a minimum of ten years of performance data over the period from January 1, 1995, to December 31, 2019, relative to their style benchmark and identified all net outperforming funds. We calculated the magnitude of every drawdown of each fund over the sample period relative to its style benchmark, median peer, and 25th-percentile peer and used each outperforming fund’s worst drawdown in magnitude.

Sources: Vanguard calculations as of May 2020, based on data from Morningstar, Inc.

Figure 9. During any given ten-year period, outperforming funds should expect to have one continuous drawdown that lasts more than two years

![Figure 9](image)

Notes: We evaluated all U.S.-domiciled, Morningstar nine-style-box U.S. active equity, emerging markets, and developed market foreign funds with a minimum of ten years of performance data over the period from January 1, 1995, to December 31, 2019, relative to their style benchmark and identified all net outperforming funds. We calculated the length of every cumulative drawdown of each fund over the sample period for the funds that outperformed their style benchmark. The number of continuous drawdowns (defined as a length of time where the cumulative monthly return of the fund was never positive) relative to style, and median peer were 12,821, and 12,791, respectively.

Sources: Vanguard calculations as of May 2020, based on data from Morningstar, Inc.

In this analysis, continuous drawdown means that at no point over the time frame noted did the fund have a positive cumulative monthly return.
Our final analysis, shown in Figure 10, looks at the magnitude and length of the maximum drawdown period. Unlike our previous analyses, where the fund sample includes outperforming funds that may have still been recovering from their maximum drawdown, here we evaluate only the funds relative to their style benchmark that have fully recovered from their maximum drawdown. The average maximum drawdown magnitude and period of the outperforming funds that recovered were 24% and 5.5 years respectively, with the peak-to-trough phase averaging 3.5 years and the recovery phase averaging 2 years. Surprisingly, when we test the relationship between the magnitude and the length of the maximum drawdown, we did not find a statistically significant relationship between the two at the 5% level.

We also split our sample into those funds whose recovery phase was shorter than the peak-to-trough phase and those whose recovery phase was longer than the peak-to-trough phase. We did this to evaluate the potential relationship between the peak-to-trough phase and the recovery phase to the overall length and the magnitude of the maximum drawdown. Our analysis finds that the average maximum drawdown magnitude is the same whether it takes more or less time to recover than it did to reach the bottom; however, if the recovery phase is longer than the peak-to-trough phase, the overall underperformance period is longer by almost two years on average.

Interestingly, of the funds that recovered, more than a quarter did so after seven or more years of underperformance; another quarter recovered after less than three years of underperformance.

In terms of drawdown magnitude, a quarter of these funds recovered from maximum drawdowns that were greater than 30%, while half recovered from maximum drawdowns of less than 20%.

Figure 10. Similar maximum drawdown magnitudes are seen whether the recovery phase is shorter or longer than the peak-to-trough phase

Note: We evaluated all U.S.-domiciled, Morningstar nine-style-box U.S. active equity, emerging markets, and developed market foreign funds with a minimum of ten years of performance data over the period from January 1, 1995, to December 31, 2019, relative to their style benchmark and identified all net outperforming funds. We calculated the length and magnitude of every drawdown of each fund over the sample period for the 1,173 funds that outperformed their style benchmark and identified the largest drawdown by magnitude. From that sample, we identified the 478 funds that have recovered from the maximum drawdown.

Sources: Vanguard calculations as of May 2020, based on data from Morningstar, Inc.
Glimpse into the future: Simulating future patience

Second in difficulty to the initial decision to invest with an active manager is the decision of when to stop having patience with a manager that has been underperforming. Although our analysis here and in this body of research does not answer that important question, we do want to touch on the difficulty of this decision. Kaplan and Kowara (2019) and Donoho, Crenian, and Scanlan (2010) both attempt, using Monte Carlo simulations, to quantify how long an investor should have to wait before deciding that the manager is skilled or unskilled with the goal of confidently making the decision to fire or continue having patience when the fund is underperforming.

Our goal here is to provide an understanding, using a simulated excess return series, of the probability that a manager who has been underperforming is actually an outperforming manager (and therefore warrants continued patience). We assumed that the investor had chosen an active manager who would generate long-term excess returns. The long-term risk-adjusted returns that we use are similar to history, but evaluated as what the historical return path could have been or may be in the future.

Data and methodology\(^\text{17}\)

Our process was as follows:

- For each fund of the 25-year outperforming sample of historical U.S. active equity funds, define risk-adjusted performance by using the historical information ratio.\(^\text{18}\)
- Based on ranked information ratios, divide the outperforming funds into risk-adjusted performance groups (top decile, top quartile, top half, all outperformers).
- For each performance group, run 10,000 ten-year simulations to construct hypothetical monthly excess return paths that represent a randomly sampled “fund.” The simulation is modeled below using geometric Brownian motion.\(^\text{19}\)

\[
\ln \left( \frac{P_t}{P_{t-1}} \right) = \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma W_t
\]

- \(P_t\) = Price in month \(t\)
- \(P_{t-1}\) = Price in month \(t - 1\)
- \(\mu\) = Expected return
- \(\sigma\) = Standard deviation of returns
- \(W_t\) = Random normally distributed variable

Figure 11 on page 15 shows the results of the 10,000 simulations for the “all outperformers” group, expressed as excess returns. Results above a 1.0 net asset value indicate positive excess returns; results below 1.0 indicate negative excess returns. Repeating the analysis using the information ratios for the top decile, top quartile, and top half of the outperformers, we determine the percentage of the outperforming funds that will actually be outperforming at various time intervals (see Figure 12 on page 15). At five years, only 60% of the “true” outperformers will actually have generated positive excess returns, and the remaining 40% of the true outperformers will be underperforming. With 40% of the true outperformers actually underperforming at five years, the process used to evaluate funds cannot rely solely on a simple quantitative performance screen, but must include other evaluation criteria.

---

17 Our methodology and analytical techniques are most similar to Donoho, Crenian, and Scanlan (2010).
18 The inputs into information ratio are the excess net returns and standard deviation of the excess net returns versus the Morningstar style benchmark. Simulations are run based on the group of equity outperformers aggregated across the Morningstar nine-style box.
19 Geometric Brownian motion is a continuous, time-stochastic process in which the logarithm of the randomly varying quantity follows a Brownian motion (also known as Wiener process) with drift. The Wiener process, defined as \(W_t\), is a series of normally distributed random variables whose variances, as time goes on, increase to reflect more uncertainty in predicting the value of the process. For each of the 10,000 simulations, we obtain mean and standard deviation inputs in one of two ways: (1) by taking the median excess return and standard deviation of the performance group, or (2) by randomly selecting the median excess return and standard deviation for a fund from the performance group. Differences are statistically significant, but small in magnitude.
IMPORTANT: The projections and other information generated by the simulations referenced regarding the likelihood of various investments outcomes are hypothetical in nature, do not reflect actual investment performance, and are not guarantees of future results.
Conclusion

Patience is an investor’s capacity to endure periods of underperformance in hopes of achieving an investment objective. The possible benefits of using outperforming managers and factor tilts can be significantly eroded, however, for an investor who fails to maintain a long-term patient perspective. This is because the inconsistency inherent in returns wears on patience. In theory, most investors are fine with drawdowns. They have been told to expect them. Evidence shows that investors struggle when the frequency, duration, and magnitude of drawdowns mount.

All risky assets and strategies—whether associated with the equity market, style factors, or an outperforming traditional active manager—are prone to periods of underperformance. Investors who have selected an outperforming manager can realistically expect to be in a drawdown between 40% to 60% of one-year evaluation periods, to experience a continuous drawdown lasting two years or more every ten years, and to experience a drawdown of greater than 20% or more over time. Those with high conviction and the appropriate risk tolerance for active risk are more likely to have the patience necessary for success in active investing, as they will be able to prepare for and tolerate the frequency, magnitude, and length of these drawdowns.
References


Appendix

U.S. long-only equity factor definitions

The table below shows the criteria that were used to define and construct U.S. equity factors for the analysis reported in this paper.

Figure A-1. Factors used in analysis

<table>
<thead>
<tr>
<th>Equity factors</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market</td>
<td>Value-weighted return of all CRSP stocks incorporated in the U.S. and listed on NYSE, AMEX, or NASDAQ</td>
</tr>
<tr>
<td>Size</td>
<td>Value-weighted return of portfolio formed on small-capitalization CRSP stocks incorporated in the U.S. and listed on NYSE, AMEX, or NASDAQ minus market</td>
</tr>
<tr>
<td>Value</td>
<td>Value-weighted return of portfolio formed on highest book-to-market CRSP stocks incorporated in the U.S. and listed on NYSE, AMEX, or NASDAQ minus market</td>
</tr>
<tr>
<td>Momentum</td>
<td>Value-weighted return of portfolio formed on highest-momentum CRSP stocks incorporated in the U.S. and listed on NYSE, AMEX, or NASDAQ minus market</td>
</tr>
<tr>
<td>Low volatility</td>
<td>Value-weighted return of portfolio formed on lowest-beta CRSP stocks incorporated in the U.S. and listed on NYSE, AMEX, or NASDAQ minus market</td>
</tr>
<tr>
<td>Multifactor</td>
<td>Return of portfolio equally weighted across size, value, momentum, and volatility minus market</td>
</tr>
</tbody>
</table>

Sources: Vanguard calculations as of May 2020, based on data from Morningstar, Inc., and the Kenneth R. French Data Library.