Underperformance of Concentrated Stock Positions

Antti Petajisto*†

First draft: March 31, 2023

This draft: June 30, 2023

Abstract

Contrary to standard investment advice, many high-net-worth investors hold concentrated positions in single stocks, which may constitute 10-20% or more of their total portfolio assets. While most investors recognize that lack of diversification increases the volatility of portfolio returns, they may not understand that concentrated stock positions usually contribute negatively to portfolio returns. Since 1926, the median ten-year return on individual U.S. stocks relative to the broad equity market is –7.9%, underperforming by 0.82% per year. For stocks that have been among the top 20% performers over the previous five years, the median ten-year market-adjusted return falls to –17.8%, underperforming by 1.94% per year. Since the end of World War II, the median ten-year market-adjusted return of recent winners has been negative for 93% of the time. The case for diversifying concentrated positions in individual stocks, particularly in recent market winners, is even stronger than most investors realize.

JEL classification: G10, G11

Keywords: Diversification, skewness, long-term reversal, concentrated stock position, high-net-worth investor

---

* http://www.petajisto.net/, antti.petajisto@sloan.mit.edu.

† I wish to thank Tom Faust for discussions on the topic. I gratefully acknowledge financial support for this research from Morgan Stanley Investment Management.

Diversification is often called the only free lunch in investing: it reduces portfolio risk without reducing expected returns. Nevertheless, a large number of wealthy investors maintain stock portfolios that are not well diversified, having single-stock positions that account for 10-20% or more of their entire portfolio. In fact, many high-net-worth investors have become wealthy precisely *because* they have owned a concentrated equity position that has appreciated significantly in value – some have been company founders or senior executives with an equity stake, while others have simply bought a stock that subsequently performed very well. Given their personal experience, they may naturally question the value of diversification. Similar thinking sometimes applies even among institutional investors.

One notable and public case of a concentrated equity bet by an institutional investor is the Emory University endowment, which held an outsized position in Coca-Cola stock for decades, following a transformational grant to the university in 1979 funded in shares of Coca-Cola stock. In 2000, Emory had over half of its then roughly $5 billion endowment invested in Coca-Cola stock.[1] At the time, the university's Chief Operating Officer (COO) said they were "not uncomfortable holding what we hold now," they had "great confidence" in the future of Coca-Cola, and their "target allocation" to Coca-Cola stock was 48%. This extreme concentration was controversial, facing criticism from Emory's own finance faculty. In early 2000, the value of the university's Coca-Cola shares had dropped about 30%, or $1 billion, from the position's peak value in June 1998, even as the overall market had simultaneously risen about 30% (Figure 1). Despite the underperformance, the university held on to its Coca-Cola position, with the COO pointing out that they were still better off than if they had instead owned a diversified equity portfolio over the previous 15 years.

How did the Emory endowment debate turn out in the end? Subsequently, the stock did bounce back somewhat, although in the next five years it lost another cumulative 25% relative to the market. After that, the university finally pared its Coca-Cola position

[1] "Emory University Gets Hard Lesson as Coke's Stock Fails to Make Grade," *Wall Street Journal*, 1/28/2000.

to 10% of the endowment portfolio,[2] still concentrated but much less so than before. While the concentrated position worked out well for Emory until mid-1997, after that the university would have been significantly better off investing in a diversified equity market portfolio (including for the next 25 years until the end of our sample in 12/2022).

More generally, what is the cost to returns of investing a large part of a portfolio in a single stock? While most investors understand that concentrated stock positions exhibit significantly more volatility than the equity market as a whole, many view the return prospects of a single stock as essentially a coin toss, with an equal chance of outperforming or underperforming the market. Some holders of stocks that have outperformed in the past may extrapolate from past performance to expect the same in the future. However, this paper points out that concentrated stock positions are significantly more likely to underperform than to outperform the stock market as a whole over the long term. We find that the median ten-year single-stock return is –7.9% relative to the capitalization-weighted market portfolio, underperforming by 0.82% per year, which certainly is economically significant. What about investors holding concentrated stock positions that have made them a lot of money in the past? Actually, most past long-term winners underperform even more: stocks with top 20% performance over the past five years have a median ten-year return of –17.8% relative to the market, underperforming by 1.94% per year.

We find a high level of consistency in these results. In the time series of monthly returns from 1/1926 to 12/2022, the median performance of stocks with top 20% performance over the past five years lagged the market 86% of the time, and this share rises to 93% in the post-World War II period. This effect applies across industry groups, although more volatile and more rapidly changing industries exhibit a larger effect. It is non-monotonic across firm size: the biggest effect exists for the smallest firms, the effect is somewhat reduced for midcaps, and then it rises again for the largest firms.

[2] "Venture-Capital Bets Swell Stanford's Endowment," *Wall Street Journal*, 1/23/2006. Some numbers are from the author's calculations.

The fact that most stocks underperform the market as a whole over the long term arises from positive skewness in the pattern of long-term stock returns: while stock declines are limited to –100%, a few stocks can have extremely large positive returns, such as Dell beating the market in the 1990s by almost +20,000%. Positive skewness raises the mean stock return above the median and leaves a majority of stocks underperforming the market as a whole. We show that even if stocks follow a well-behaved lognormal distribution, skewness drastically increases over long horizons, and we document this empirically as well.

Skewness and its implications for equity investors have certainly been discussed in the literature before. Arditti (1967) shows empirical evidence that investors prefer stocks with more positive skewness. Simkowitz and Beedles (1978) document significant positive skewness at the individual stock level, but zero to slightly negative skewness for diversified portfolios. The closest paper to our study is by Bessembinder (2018). He presents the striking result that more than half of all U.S. stocks have underperformed Treasury bills over the long term, with the favorable overall equity market returns observed over most long-term periods driven by a relatively small number of extremely successful stocks. However, Bessembinder (2018) uses every single stock publicly listed in the U.S., so his results end up being heavily influenced by microcaps: in the last 50 years, the number of publicly listed U.S. stocks has averaged about 5,000 (Figure 3), yet most practitioners would not pay much attention to stocks outside of the top 3,000 due to their small capitalizations and illiquidity.

In contrast to Bessembinder (2018), our analysis seeks to examine stock returns from a perspective that is most relevant to U.S. high-net-worth investors. We start by defining a concentrated stock position. We exclude microcaps and concentrate on a universe similar to the Russell 3000, because this is where most high-net-worth investors' public equity holdings are focused (our results would be significantly more extreme with microcaps included). We choose a capitalization-weighted U.S. equity market portfolio, rather than Treasury bills, as our benchmark because a cap-weighted equity index is more similar in return expectation and a more likely alternative to the concentrated stock positions held by high-net-worth investors. Importantly, our analysis includes tests of the long-term persistence of the return patterns we identify. Because wealthy investors holding concentrated stock positions are, almost by definition, sitting on large capital gains, they tend to be long-term investors. While Bessembinder (2018) focuses more on the handful of top-performing stocks, we dive deeper into median returns, including time-series analysis to test whether the effects observed are still likely to prevail today, nonparametric dependence on market capitalization and prior return, and a breakdown into industry groups to see how broad-based the effects are. Furthermore, our analysis has some methodological differences relative to Bessembinder (2018), such as using overlapping time periods starting each month (as opposed to a few fixed non-overlapping ten-year periods), which eliminates any impact from an arbitrary choice of a time period start date and ensures we use all data as efficiently as possible.

But why do high-net-worth investors continue to hold concentrated stock positions? We briefly discuss both rational reasons, which include taxes (to avoid realizing large capital gains) and retaining corporate control, as well as behavioral reasons rooted in investor psychology.

In the aforementioned case of Emory's concentrated Coca-Cola position, taxes were not a factor (because Emory, as a non-profit, is tax-exempt), and control rights were probably also not important to the university. Instead, the tendency to extrapolate past trends into the future, a preference for status quo and loyalty to the donors who contributed the Coca-Cola stock were more likely to have an impact; indeed, the university sold the biggest part of its position only after a period of significant underperformance, which struck at the heart of the justification about this one stock being special.

In summary, when taxes and control rights are not significant factors, investors holding concentrated stock positions should seek to overcome the behavioral biases that often interfere with diversification plans. Diversification away from a concentrated stock position not only reduces portfolio risk but also boosts long-term returns in most cases, particularly when the concentrated position has a history of strong long-term performance. If the risk-reduction benefit of portfolio diversification is considered a free lunch, then the combined benefit of higher returns *and* lower risk from diversification is tantamount to *getting paid* to eat lunch.

This paper proceeds as follows. Section I starts by defining a concentrated position. Section II describes the data and the selection of a relevant universe of U.S. stocks. Section III reports the key findings on the median stock return, including its behavior in different time periods. Section IV further investigates the median and mean returns across different levels of market capitalization and prior return, as well as across industry groups. Section V briefly discusses both rational and behavioral reasons for the observed lack of diversification. Section VI concludes.

I. Defining a Concentrated Position

When does an individual stock position become "concentrated?" We care about concentration because, as the opposite of diversification, concentration increases portfolio volatility. We can therefore define a concentrated position based on where the volatility-reducing benefits of diversification start to disappear.

The top panel of Figure 2 shows the volatility of a portfolio that combines a single concentrated position with a position in the cap-weighted market portfolio. We assume the market portfolio has an annual volatility of 15%, and we show the results for three levels of idiosyncratic volatility of the concentrated position: 20%, 30%, and 40%.[3]

We see that portfolio volatility is essentially unaffected by holding a single position at weights between zero and 10%, with a maximum increase in portfolio volatility from 15% to 15.5% (3% increase). Depending on the concentrated stock position's idiosyncratic volatility, the effect on portfolio volatility starts to become meaningful at single-stock position weights of 10-20%, where the portfolio volatility can rise from 15% to 17% (13%

[3] These numbers are calibrated from our data: The time-series median rolling five-year volatility of the market portfolio is 15%. The quartiles of the market-adjusted return volatility for our universe of individual stocks are roughly 20%, 30%, and 40%. The market-adjusted return volatility also varies from a median of around 20% for the largest firms to a median of around 40% for the smallest firms in our universe. The time-series behavior of both market and stock-level volatility is relatively stable, with the main exceptions being a general volatility spike in the 1930s and a stock-level volatility spike in the tech bubble and crash around 2000.

increase). At about a 30% single-stock position weight, the function becomes approximately linear, so any further increase in portfolio concentration has a constant first-order effect on portfolio volatility, and diversification no longer helps. Diversification is the reason for the nonlinearity up to 20-30% portfolio weight: that is why the curve starts with zero slope, i.e., a small single position has negligible impact on portfolio volatility.

The bottom panel shows portfolio volatility when the portfolio consists of single stocks with equal weights on each (and no weight on the market portfolio). Now 100% of the portfolio is subject to idiosyncratic volatility. With single-stock position weights of 1% or less, portfolio volatility essentially matches the volatility of the market portfolio. As the single-stock weights increase to 10%, volatility rises from 15% to a range of 16.3%–19.6%, depending on the idiosyncratic volatility of the constituent stocks, which starts to be economically meaningful.

Based on these numbers, we suggest defining 10% as the cutoff for a concentrated stock position. It also implies that an investor needs to hold more than ten stocks to avoid concentrated positions. For additional context, the largest index weights in the popular S&P 500 index have typically been around 5%, which is still well below our suggested threshold for a concentrated stock position. The precise cutoff is of course somewhat arbitrary, as portfolio volatility is a smooth function of portfolio weights, but roughly this magnitude of portfolio weight seems an appropriate and economically meaningful definition of a concentrated stock position.

II. Data and Universe Selection

Because we are interested in long-term stock returns, we select a sample that consists of all U.S.-incorporated publicly traded common stocks (share codes 10 and 11) in the Center for Research in Security Prices (CRSP) database[4] over almost a century,

[4] Bessembinder (2018) also includes firms incorporated outside the U.S. if their primary exchange listing is in the U.S. (share code 12), but based on the results from Cremers, Petajisto and Zitzewitz (2013), we exclude such non-U.S. firms in order to stay close to the universe of the popular Russell and S&P indices.

from 1/1926 to 12/2022. If a stock is delisted, we add in the delisting return reported by CRSP.[5]

We are most interested in results for stock positions typically held by high-net-worth investors, so we exclude microcap stocks from virtually all our analysis. We do this by creating a monthly universe of the largest 3,000 U.S. stocks based on market capitalization, roughly corresponding to the Russell 3000 index of large, midcap, and small stocks. However, because of the smaller total number of stocks before the 1980s, we need a different approach for the earlier period.[6] As a secondary screen, we choose the largest stocks comprising 99.5% of total U.S. equity market capitalization. Stocks in our selected universe need to satisfy both criteria, so it is always the smaller of the two. To mitigate the impact of temporary price spikes or potentially erroneous data points, we form the universe using the median market capitalization over the prior three months, rather than always the latest month-end value.[7]

Figure 3 shows the total number of U.S. stocks both in our selected universe and in the entire CRSP database. The CRSP database originally included only stocks listed on the New York Stock Exchange, with stocks listed on the American Stock Exchange added in 1962 and Nasdaq stocks added in 1972. Since late 2008, our selected universe has actually had fewer stocks than the Russell 3000. For example, in the 2022 annual reconstitution of the Russell 3000, the market cap cutoff was $240 million (based on end-of-May data), versus a cutoff of $310 million for our universe.

[5] For the small number of cases where the CRSP delisting return is missing, mostly in older data, we follow the common approach of assuming a –30% delisting return.

[6] The Russell 3000 index was not created until 1984, and the total number of stocks in the CRSP database does not reach 3,000 until 1972.

[7] We also exclude GameStop from the sample due to its outsized effect on short-term skewness and kurtosis, all based on a single 1,625% return in January 2021. This exclusion has no effect on other numbers is our tables.

III. The Median Return across Investment Horizons

A. Methodology

To understand the differences between the investment returns of concentrated stock positions and a diversified equity portfolio, we need a benchmark. As our benchmark, we choose the capitalization-weighted portfolio of stocks in our selected universe,[8] which in recent decades roughly approximates the Russell 3000 index. For each stock and each month, we compute the market-adjusted return by subtracting the market return from the individual stock return.

For a given investment horizon, e.g., ten years, we start with all stocks currently in the universe and then compute the compounded "buy and hold" market-adjusted return on each stock for that horizon going forward. If a stock is delisted during the investment horizon, we set its market-adjusted return to zero for the remaining months after delisting; in effect, the capital invested in a delisted stock is treated as moving to the market portfolio upon delisting. A stock near the universe cutoff may also drop out of the universe during the investment horizon if its relative market-cap ranking declines, but as long as the stock was initially in the universe, we still compute its future return for the full investment horizon.

Concentrated stock positions are typically long-term holdings, so we focus on investment horizons of five, ten and 20 years. However, we also compute results for one-month and one-year horizons to contrast short-term and long-term holding period results.

B. Return Distribution: Skewness

Figure 4 shows the shape of the return distribution on stocks in our selected universe for different investment horizons. The numbers shown are cumulative market-

[8] In spite of excluding microcaps, this benchmark has a return correlation of 99.997% with the Fama-French market portfolio. However, we preferred to make sure that the individual stocks and their benchmark are all chosen from the exact same universe of stocks.

adjusted returns, but the distributions would have the same shape for cumulative absolute returns as well.

In the short run, such as over a month, stock returns are relatively symmetric and moderately well described by a normal distribution. However, the shape of the return distribution changes completely for longer horizons. Stock returns are naturally truncated at –100%, while the upside is essentially unlimited. As the time horizon increases, the distributions of stock returns become more positively skewed. Furthermore, the distributions start to look more lognormal, which is a common model for long-term equity returns.

In fact, if stock returns follow a lognormal distribution, the distribution is always positively skewed, and positive skewness increases with investment horizon. Figure 5 shows how the skewness of the lognormal distribution increases for three different levels of annualized volatility. At 40% annualized volatility, skewness rises from 0.3 and 1.3 (near symmetry) at the one-month and one-year horizons to 13.8 for a ten-year horizon, and 129 for a 20-year horizon.

High skewness arises from those rare growth stocks that experience returns in the hundreds or even thousands of percent over time. Prominent examples include ten-year market-adjusted returns of almost +20,000% for Cisco and Dell in the 1990s, and around +5,000% for Apple in the 2000s and Tesla in the 2010s.

While the performance of these and other high-flying stocks of past years has been spectacular, all stocks eventually face limits in their ability to sustain above-market returns. Once a firm becomes one of the largest in the economy, it attracts greater competitive and regulatory pressures, suggesting an upper bound on the percentage of the U.S. equity market accounted for by a single stock. Volatility also tends to decline as a stock rises in market capitalization, reflecting lower risk for the largest stocks. The actual skewness estimated from data, shown as dots on the figure, is initially closest to the 40%

volatility curve, although for the 20-year horizon it is much closer to the 30% volatility curve, which is consistent with these effects.[9]

C. Return Distribution: Median

Panel A in Table I shows the key statistics for cumulative market-adjusted returns on individual stocks in our selected universe over forward-looking investment horizons from one month to 20 years. Our focus here is on the median return, which is a representative outcome for an undiversified investor.

The one-month market-adjusted return distribution is relatively symmetric, with a median of –0.4%, mean of 0.1%, skewness of +1.66, and 52% of observations negative. However, the ten-year median market-adjusted return is significantly negative at –7.9% (or –0.82% per year), with 54.6% of values negative. Most investors would presumably view these as rather unattractive returns.

Why have the low median market-adjusted returns on individual stocks not attracted more attention in prior studies? Most attention has been on mean returns instead. The mean is indeed probably the most important metric for a well-diversified portfolio: if you buy and hold all the stocks in the investment universe, you will earn the cross-sectional (weighted) mean return on those stocks. The equal-weighted mean stock return has in fact been strongly positive, outperforming the cap-weighted benchmark by an average of a cumulative 21% over rolling ten-year horizons. In practice, investing $10 million across 3,000 stocks consistent with this approach would mean buying $3,333 worth of each stock and then passively holding each position for ten years. Such a portfolio would have a few spectacular winners that would be responsible for most of the total return and would grow to dominate the portfolio over time. However, trying to gamble on identifying those few stocks with outsized returns would be a bad idea for any investor with standard

[9] Excluding GameStop from our sample has no impact on any numbers in the table except one-month and one-year skewness and kurtosis. GameStop's 1,625% return in January 2021 alone increases one-month skewness from 1.66 to 2.76 and one-year skewness from 3.81 to 5.65; it also raises one-month kurtosis from 28 to 180 and one-year kurtosis from 68 to 371.

preferences (a declining marginal utility of wealth, which leads to risk aversion), because it would drastically increase portfolio volatility. For an investor seeking to evaluate the merits of holding a large position in a single stock, the median return is arguably the more relevant metric than the mean.

The ten-year investment horizon produces an "effective horizon" of 95 months – due to delistings, the average lifespan of a stock over a ten-year horizon is only about eight years. This makes it harder to interpret results for very long horizons: e.g., going from a ten-year horizon to a 20-year horizon increases the effective horizon by only five years, not ten years. However, we can adjust for the shorter effective horizon as follows: Notice that a $100 investment for ten years starts at $100, but as firms are delisted, the invested share of the capital goes down over the years, with an average of $79.2 (as seen in the table) and a final value below that. If we scale up our initial investment by dividing by 0.792, then we begin with a $126.3 investment that starts to decline with delistings and averages to exactly $100 over ten years. We could therefore say that the adjusted ten-year return is –7.9% / 0.792 = –10.0%. This adjustment makes our case about a disappointing median market-adjusted return on concentrated stock positions even stronger.

While our focus here is on stocks in our selected universe, which excludes microcaps, Panel B in Table I shows the same set of results for a broader sample that includes all U.S. stocks in CRSP. These results are significantly more extreme: the ten-year median cumulative market-adjusted return falls to –19.1%, compared with –7.9% for our selected universe. Skewness and kurtosis are rather extreme as well, mostly because a few individual extreme observations have a disproportionate effect on these measures. Survivorship rates are also meaningfully lower: over 5% of stocks are delisted within the first year and the average stock survives only seven out of the next ten years. Of course, this broader sample still includes the moderating impact of the largest firms, indicating that microcaps exhibit extreme volatility and extreme positive skewness.

D. Conditioning on Prior Winners

Most investors holding a concentrated stock position are in that position because they have experienced past success owning the stock at issue (e.g., Tesla's 743% surge in 2020 created many "Teslanaires"). Is the forward-looking return distribution different for stocks that have outperformed in recent years?

Panel C in Table I shows the same return statistics across multiple investment horizons as Panels A and B, but only for stocks that have been recent winners. Here we define "recent winners" as the top 20% of stocks in the selected universe based on their cumulative five-year past performance. This makes a significant difference in forward-looking long-term returns: the ten-year median cumulative market-adjusted return falls to –17.8% (or –1.94% per year) for recent winners, compared with –7.9% (or –0.82% per year) across all stocks. Furthermore, fully 60% of recent winners underperformed the cap-weighted market over a ten-year horizon, meaning that 50% more of such stocks underperformed as outperformed. The odds are rather stacked against a concentrated position in a recently top-performing stock beating the equity market as a whole over the next ten years, yet this is a common situation faced by investors who hold concentrated positions.

E. Consistency over Time

To evaluate long-term returns, we need a long sample period. The downside of a long sample period is that the observed effects could potentially apply only in the distant past and no longer be relevant today. So how consistent is our finding about significantly negative median market-adjusted long-term returns on individual stocks?

The top panel in Figure 6 shows the cross-sectional median of the forward ten-year market-adjusted return as a monthly time series starting in 1/1926. Because the effect is most relevant for recent winners, we include only the top 20% of stocks in the selected universe based on their prior five-year performance, as in Panel C of Table I. Because our sample ends in 12/2022, the last ten-year forward period starts in 1/2013.

The median market-adjusted ten-year return of recent winners has been negative in 86% of all the months, so if we randomly select a time period, it is likely that the median market-adjusted ten-year return will be negative. The main exceptions are the periods that start in the late 1930s and include World War II, which was a heavily regulated and unusual economic period. More growth stocks were added to the sample in the 1960s and 1970s, providing a universe that is more reflective of today's market. Since the end of World War II, which is probably more relevant for today's investors, the result has been strikingly consistent, with 93% of periods showing a negative median market-adjusted ten-year return for recent winners.

The bottom panel in Figure 6 shows the median market-adjusted return of recent winners for investment horizons other than ten years. Naturally the shortest horizons show the greatest volatility, but an overwhelming percentage of the median market-adjusted return numbers are still negative. The most recent years (computed for shorter horizons of one or five years) do not suggest any weakening of the observed underperformance of most recent winners. Thus, it seems likely that today's investors still live in a regime where this effect applies.

IV. Robustness of Results

We have already seen that our results are consistent over time. But how consistent are they across cross-sectional stock characteristics, specifically market capitalization and industry group? And how sensitive are the results to the choice of the top 20% prior return cutoff?

A. Prior Return

To measure the effect for a specific prior five-year return range, each month we assign each stock to its prior five-year return percentile in the cross-section of all stocks in the selected universe as of that month. We then compute the median ten-year forward market-adjusted return as a function of the prior five-year return percentile (from 0 to 1) in the pooled sample. We compute the median as a nonlinear kernel estimate (essentially a local median value), using a Gaussian kernel with a bandwidth of 0.05.

The top panel of Figure 7 shows the median market-adjusted return estimate. The median return exhibits a humped shape, with a maximum at the 20th percentile of prior five-year return and declining values above that. Above the 70th percentile, all conditional medians are below the unconditional median. The negative slope becomes steeper around the 80th percentile, so picking, for example, the top 10% based on prior five-year performance would generate more extreme results than our top 20% cutoff.

The bottom panel of Figure 7 shows the conditional mean market-adjusted return. Here we see the forward ten-year return as a monotonically decreasing function of the prior five-year return. This is consistent with the long-term return reversal pattern documented by De Bondt and Thaler (1985) and others. However, our methodology sheds more light on this previously observed effect: we find that the biggest prior five-year losers do *not* have the highest future median returns, even though they have the highest future mean returns, so the high future mean comes from just a small number of stocks with extremely high returns. In other words, positive skewness is exceptionally high for prior five-year losers.

B. Market Capitalization

To measure the effect for different levels of market capitalization, we follow a similar approach as with prior returns. Each month we assign each stock in the selected universe to its cross-sectional market cap percentile, and we compute the median ten-year forward market-adjusted return as a function of the market cap percentile in the pooled sample. The median is again a nonlinear Gaussian kernel estimate with a bandwidth of 0.05.

The top panel of Figure 8 shows that the median market-adjusted return has a humped shape, with a maximum slightly above the 60th percentile of market cap. The smallest 20% of stocks have the lowest median return, underperforming the cap-weighted market by a cumulative 14% over ten years. Among the very largest stocks, the median stock also significantly underperforms both the market and the unconditional median stock, losing to the market by around 12% over ten years.

The bottom panel of Figure 8 shows the mean ten-year market-adjusted return as a function of market cap. This relationship is essentially flat for the smallest half of stocks in our selected universe, but it declines steeply for the biggest stocks. The mean return therefore shows a significant size effect, although the size effect is driven primarily by the underperformance of the biggest stocks rather than the outperformance of the smallest stocks. Furthermore, our methodology again shows different patterns for the median and the mean: because the smallest stocks have by far the lowest median but the highest mean, they must have a significantly more positively skewed return distribution, where a handful of stocks with spectacular returns make up for many hundreds of stocks with weak returns.

C. Industry Groups

The core findings of this paper that most stocks underperform and that market returns are driven by a relatively small universe of exceptional performers may seem intuitive for firms with potentially disruptive technologies or business models, such as Google and Amazon. It is far less clear that stocks in highly regulated industries such as utilities should exhibit a similar effect. To investigate the effect across industry groups, we use the SIC codes to sort firms first into 48 industries defined by Fama and French,[10] and then aggregate these into ten more-comparably sized industry groups as in Kacperczyk, Sialm, and Zheng (2005).[11]

Table II shows the median ten-year returns on individual stocks in the selected universe relative to the cap-weighted market portfolio, broken down by industry group.

---

[10] The industry definitions were obtained from Kenneth French's data library, publicly available on the web at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[11] NAICS is another more modern approach to create industries, but NAICS codes only first appear in CRSP in 2004. In contrast, SIC codes have full coverage for our time period from 1926 to 2022 (with the exception of some newly added Nasdaq stocks in the 1970s), so the SIC codes are a more appropriate choice for our long historical sample.

The median market-adjusted stock return is negative for all industries except finance, where it exactly matches the market return, so this is indeed a broad-based effect.

The largest effect can be seen in the business equipment and services group, which includes many of today's technology-driven stocks, such as the aforementioned Google and Amazon: the median ten-year market-adjusted return here is –18.4%, even though the equal-weighted mean return is positive at +20.6%. Many other industry groups similarly have over 30% spread between their median and mean ten-year returns; as one example, the spread for health care stocks is essentially the same at 37.2%. Even utility stocks exhibit a negative median return at –11.8%, although utilities are still the least volatile and have the lowest spread between the mean and the median ten-year returns.

V. Why Do Concentrated Positions Still Exist?

Why do many high-net-worth investors continue to hold concentrated stock positions? While this paper has provided new evidence on their low median returns going forward, most investors are already aware of the risk-reduction benefit of diversification, yet they willingly forgo that.

One reason is taxes: if a stock position has appreciated by many multiples, a rebalancing sale transaction would require realizing the gain and paying substantial capital gain taxes. Another reason, likely relevant only to larger shareholders, is to retain voting control over the firm.

Behavioral reasons are likely to have an impact as well. The representativeness heuristic (e.g., Kahneman and Tversky (1972)) suggests that people tend to ignore base-rate probabilities and assume that a small sample is more representative of the population than it actually is. In financial markets, this implies that investors tend to extrapolate trends: if a stock has beaten the market by a large margin in past years, many would view it as a "good stock" and assume that the stock is also more likely to outperform going forward (even though the exact opposite is true in the data). The endowment effect says that people tend to prefer things they already have; similarly, the status quo bias is about people's preference to stay with the default option (Kahneman, Knetsch, and Thaler

(1991)). In addition, simple loyalty to one's own company has been shown to lead to under-diversification (Cohen (2009)).

Anchoring (e.g., Sherif, Taub, and Hovland (1958)) refers to the tendency of people to pick an initial reference point, sometimes completely unrelated, and insufficiently adjust their estimate away from the reference point (e.g., a recently observed outcome of a roulette wheel, fixed to show either 10 or 65, impacted people's estimate of the number of African counties in the United Nations). In our context, investors can anchor their thinking around their baseline of current portfolio holdings, and a large reduction even in a concentrated position can thus be viewed as a radical move. People also exhibit ambiguity aversion (e.g., Ellsberg (1961)) and generally prefer gambles that they feel they understand better; for example, surveys show people considering their own-company stock as less risky than a diversified stock portfolio.

These behavioral reasons are, by definition, not rational, and investors should try to resist such influences on their portfolio choices. This is, of course, easier said than done, but a reasonable first step to inoculate oneself against such biases is to be aware of them. In contrast, taxes and voting rights may sometimes be valid reasons not to liquidate an appreciated concentrated position; however, not diversifying does come with a nontrivial cost, as this paper has shown, so concentrated stock investors should evaluate the costs and benefits of taking at least incremental steps toward diversification.

VI. Conclusions

Unlike most papers in the literature, this one focuses on median stock returns, not mean returns. Because individual stock returns exhibit significant positive skewness, which naturally increases over longer investment horizons, we find that the median stock has underperformed the cap-weighted market portfolio by 7.9% over rolling ten-year investment periods (or 0.82% per year) since 1926. The relative underperformance over rolling ten-year periods increases to 17.8% (or 1.94% per year) when considering only stocks whose performance ranked in the top 20% over the prior five years. This is not just about small growth firms in the technology sector: the observed underperformance of the median stock applies across all industry groups and among both the smallest and largest stocks. The effect is even more consistent in the post-World War II period and shows no sign of weakening in recent years.

These results are particularly relevant for investors holding concentrated stock positions. Sometimes they hold onto concentrated positions for tax reasons or to have corporate control, but there are also well-documented psychological impulses that can interfere with achieving prudent levels of portfolio diversification. For most investors, the benefits of diversifying concentrated stock positions may be even greater than they realize: increasing diversification not only reduces portfolio volatility but usually also increases portfolio return.

References

Arditti, Fred D., 1967, Risk and the required return on equity, *Journal of Finance* 22, 19-36.

Bessembinder, Hendrik, 2018, Do stocks outperform Treasury bills? *Journal of Financial Economics* 129: 440-457.

Cohen, Lauren, 2009, Loyalty-based portfolio choice, *Review of Financial Studies* 22, 1213-1245.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz, 2012, Should benchmark indices have alpha? Revisiting performance evaluation, *Critical Finance Review* 2, 1-48.

De Bondt, Werner F.M., and Richard Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793-805.

Ellsberg, Daniel, 1961, Risk, ambiguity, and the Savage axioms, *Quarterly Journal of Economics* 75, 643-669.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983-2011.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler, 1991, Anomalies: The endowment effect, loss aversion, and status quo bias, *Journal of Economic Perspectives* 5, 193-206.

Kahneman, Daniel, and Amos Tversky, 1972, Subjective probability: A judgment of representativeness, *Cognitive Psychology* 3, 430-454.

Sherif, Muzafer, Daniel Taub, and Carl I. Hovland, 1958, Assimilation and contrast effects of anchoring stimuli on judgments, *Journal of Experimental Psychology* 55, 150-155.

Simkowitz, Michael A. and William L. Beedles, 1978, Diversification in a three-moment world, *Journal of Financial and Quantitative Analysis* 13, 927-941.

Table I. The Distribution of Market-Adjusted Stock Returns across Investment Horizons. This table shows the cumulative forward returns of individual stocks relative to the cap-weighted market portfolio. The first column is the investment horizon in months. If a stock is delisted during the investment horizon, the delisting return is included but subsequent market-adjusted returns are assumed to be zero (i.e., post-delisting returns are assumed to match the market return). The effective horizon is the mean lifetime of a stock for each investment horizon, shown both in months and as a percentage of the target investment horizon. Panels A and C include all U.S. common stocks except microcaps; Panel B extends the sample to include microcaps. All statistics are computed for the pooled sample from 1/1926 to 12/2022.

Panel A: All Stocks in Selected Universe

Horizon (Months)	Median	Negative Values	Mean	Stdev	Skewness	Kurtosis	Effective Horizon	
							Months	Percent
1	-0.4%	52.2%	0.1%	11.9%	1.66	28	1.00	99.9%
12	-2.4%	53.7%	1.8%	43.4%	3.81	68	11.74	97.8%
60	-6.0%	54.4%	12.0%	104.3%	6.73	131	53.28	88.8%
120	-7.9%	54.6%	21.0%	160.9%	19.88	1,349	95.05	79.2%
240	-10.2%	55.0%	30.9%	208.3%	15.13	627	156.61	65.3%

Panel B: All Stocks in CRSP

Horizon (Months)	Median	Negative Values	Mean	Stdev	Skewness	Kurtosis	Effective Horizon	
							Months	Percent
1	-0.5%	52.7%	0.1%	16.2%	8.38	488	0.98	97.7%
12	-4.4%	56.3%	1.1%	66.4%	108.34	61,031	11.35	94.6%
60	-13.6%	59.0%	9.7%	150.0%	41.30	10,230	49.37	82.3%
120	-19.1%	59.9%	16.2%	199.9%	18.98	911	84.63	70.5%
240	-24.6%	60.4%	24.0%	299.2%	52.06	6,438	131.41	54.8%

Panel C: Top 20% Winners in Prior Five Years in Selected Universe

Horizon (Months)	Median	Negative Values	Mean	Stdev	Skewness	Kurtosis	Effective Horizon	
							Months	Percent
1	-0.3%	51.7%	0.1%	11.0%	0.91	11	1.00	100.0%
12	-2.8%	54.2%	1.1%	40.0%	3.37	58	11.71	97.6%
60	-12.4%	58.8%	3.1%	90.2%	7.33	200	54.30	90.5%
120	-17.8%	60.2%	5.3%	119.3%	7.86	159	98.99	82.5%
240	-20.6%	60.4%	12.9%	162.5%	13.85	566	166.24	69.3%

Table II. The Distribution of Ten-Year Market-Adjusted Stock Returns by Industry Group. This table shows the cumulative ten-year forward return on individual stocks relative to the cap-weighted market portfolio, separately for each of ten industry groups. The first column is the investment horizon in months. If a stock is delisted during the investment horizon, the delisting return is included but subsequent market-adjusted returns are assumed to be zero (i.e., post-delisting returns are assumed to match the market return). The effective horizon is the mean lifetime of a stock for each investment horizon, shown both in months and as a percentage of the target investment horizon. The selected universe includes all U.S. common stocks except microcaps. All statistics are computed for the pooled sample from 1/1926 to 12/2022.

Industry	Median	Negative Values	Mean	Stdev	Skewness	Kurtosis	Effective Horizon	
							Months	Percent
Consumer non-durables	-6.7%	53.7%	21.0%	130.8%	5.66	73	97.93	81.6%
Consumer durables	-10.4%	55.6%	21.5%	213.9%	35.05	2,031	100.76	84.0%
Healthcare	-6.5%	53.5%	30.7%	175.4%	5.69	58	87.22	72.7%
Manufacturing	-6.3%	53.6%	21.6%	130.4%	6.53	111	99.86	83.2%
Energy	-13.2%	56.3%	21.3%	168.1%	7.62	120	94.06	78.4%
Utilities	-11.8%	61.8%	-1.8%	62.1%	4.73	56	108.28	90.2%
Telecom	-3.2%	52.2%	26.0%	183.4%	8.52	115	80.31	66.9%
Business equipment and services	-18.4%	58.7%	20.6%	231.9%	27.78	1,586	89.45	74.5%
Wholesale and retail	-10.0%	54.9%	24.2%	159.8%	6.37	78	95.12	79.3%
Finance	0.0%	50.0%	19.1%	117.0%	6.67	112	88.11	73.4%
All	-7.9%	54.6%	21.0%	160.9%	19.88	1,349	95.05	79.2%

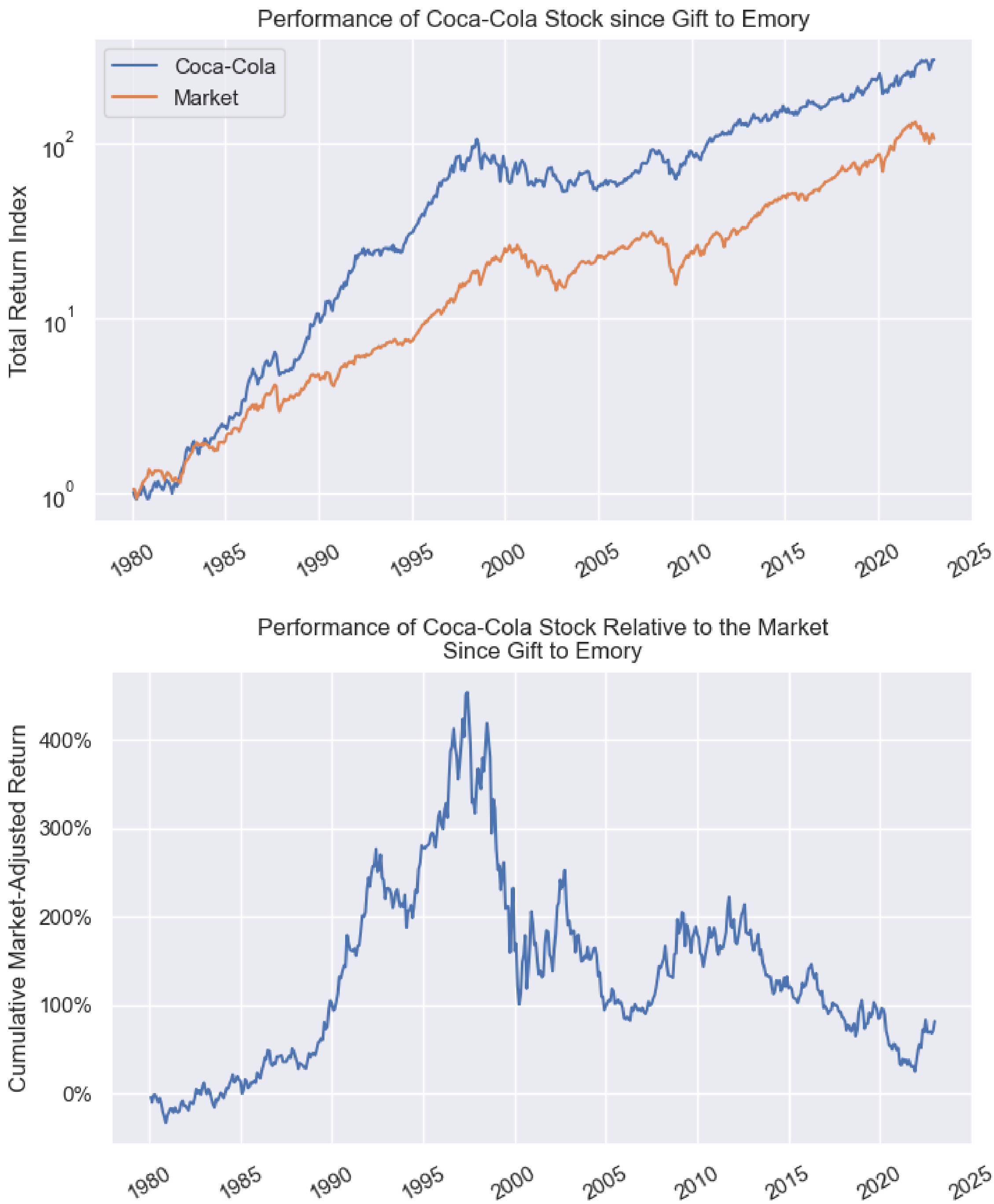

Figure 1. Performance of Coca-Cola Stock.

The figure shows the cumulative total return on Coca-Cola stock relative to the U.S. equity market portfolio, starting in 1/1980 around the time of transfer of a large gift of Coca-Cola shares to Emory University. The top panel shows the cumulative total nominal return on Coca-Cola and the market portfolio separately. The bottom panel shows the cumulative market-adjusted return on Coca-Cola, computed from monthly returns. The sample period ends in 12/2022.

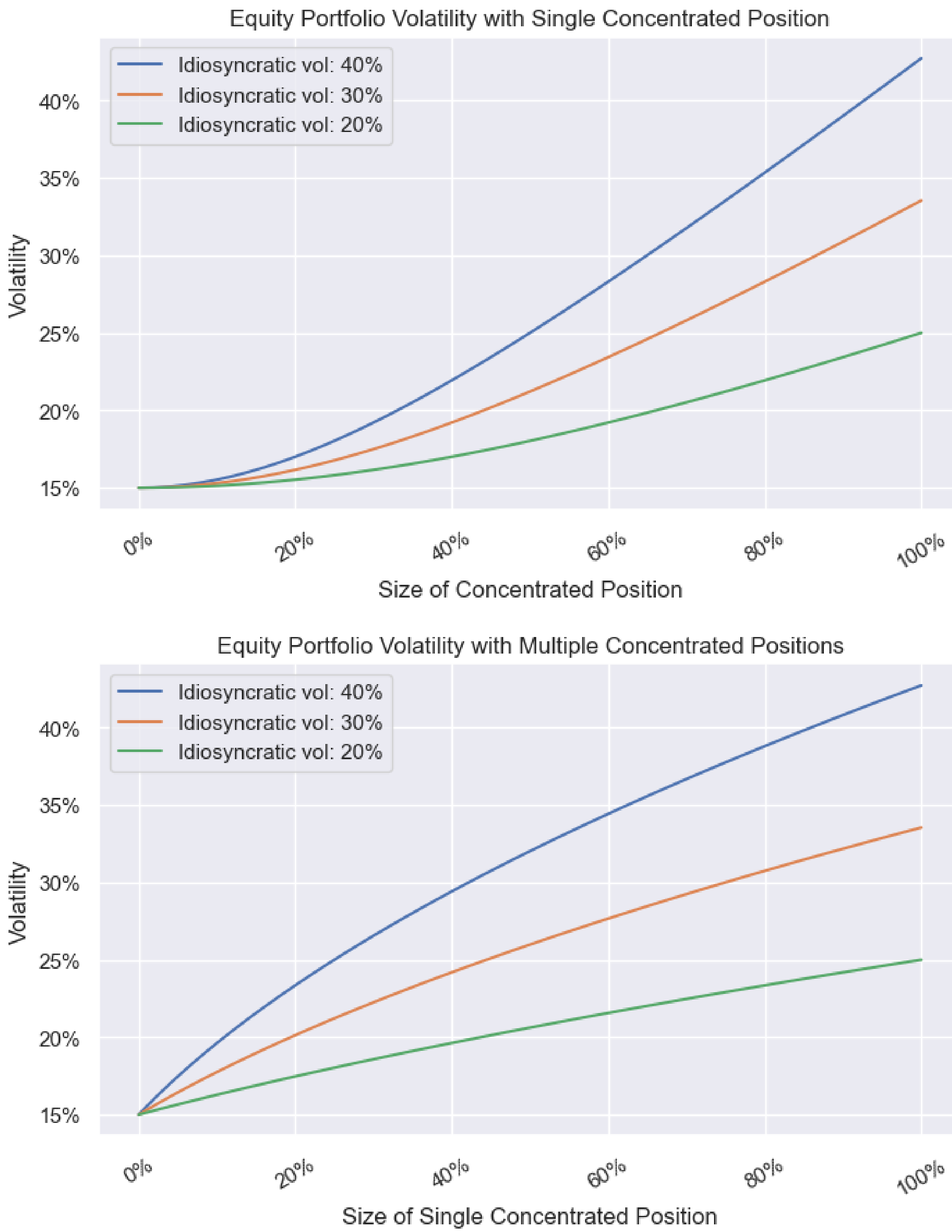

Figure 2. Equity Portfolio Volatility with Concentration Positions.
The figure shows the annual total portfolio volatility for different levels of idiosyncratic volatility. The top panel shows the volatility of a portfolio consisting of a single concentrated position and the rest invested in the cap-weighted U.S. equity market portfolio. The bottom panel shows the volatility of a portfolio consisting of multiple concentrated positions, each accounting for the same size (e.g., five positions of 20% each). The volatility of the market portfolio is assumed to be 15%.

Figure 3. Universe of U.S. Stocks.

The figure shows the number of stocks in the selected universe compared with the total number of all U.S. common stocks in the CRSP database. The selected universe includes the largest 3,000 U.S. stocks by market cap, or the largest stocks comprising 99.5% of the total U.S. market cap, whichever produces a smaller number, with the objective to exclude microcaps. The sample period is from 1/1926 to 12/2022.

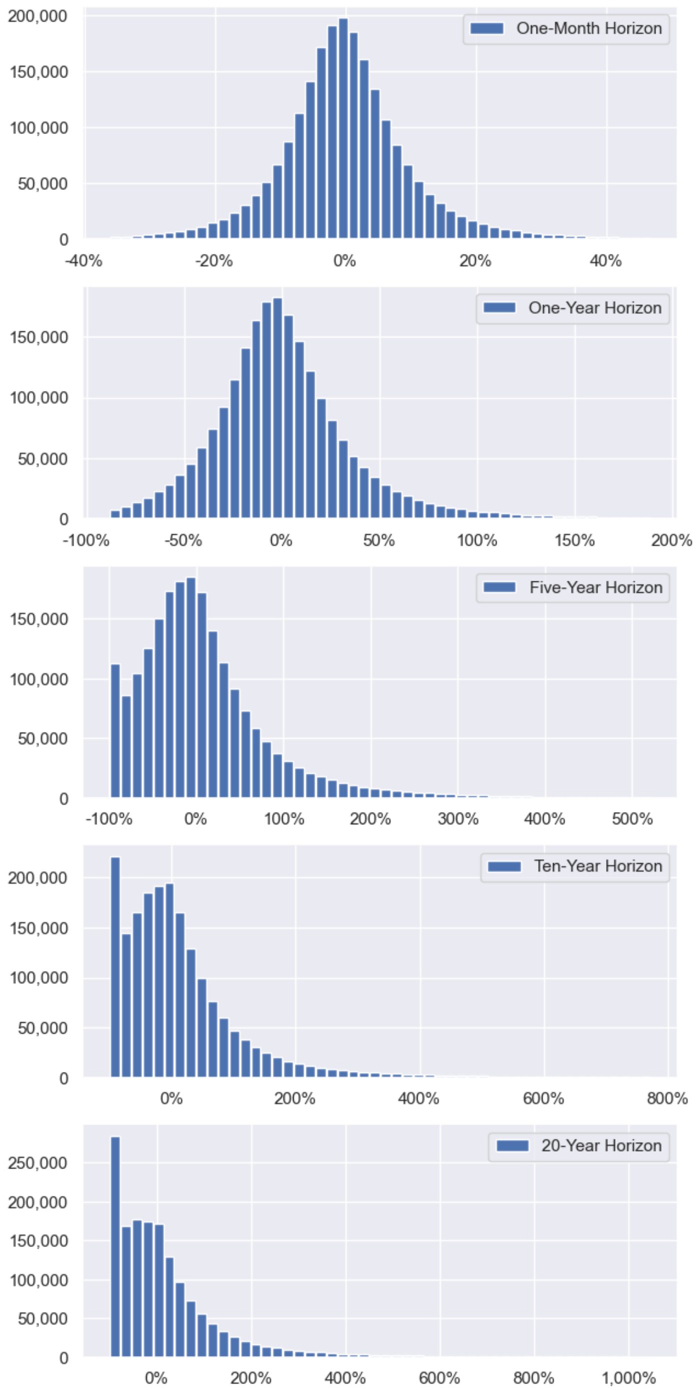

Figure 4. The Distribution of Market-Adjusted Stock Returns across Investment Horizons. The figure shows the distribution of cumulative stock returns relative to the cap-weighted market portfolio, for five different investment horizons. If a stock is delisted during the investment horizon, the delisting return is included but subsequent market-adjusted returns are assumed to be zero (i.e., post-delisting returns are assumed to match the market return). The selected universe includes all U.S. common stocks except microcaps from 1/1926 to 12/2022.

Figure 5. Skewness of Lognormally Distributed Return as a Function of Investment Horizon. Assuming a stock return follows the lognormal distribution, this figure shows how the skewness of the final return distribution increases with investment horizon, for two different levels of annualized return volatility. The dots indicate actual skewness of U.S. stock returns estimated from our sample as reported in Table I.

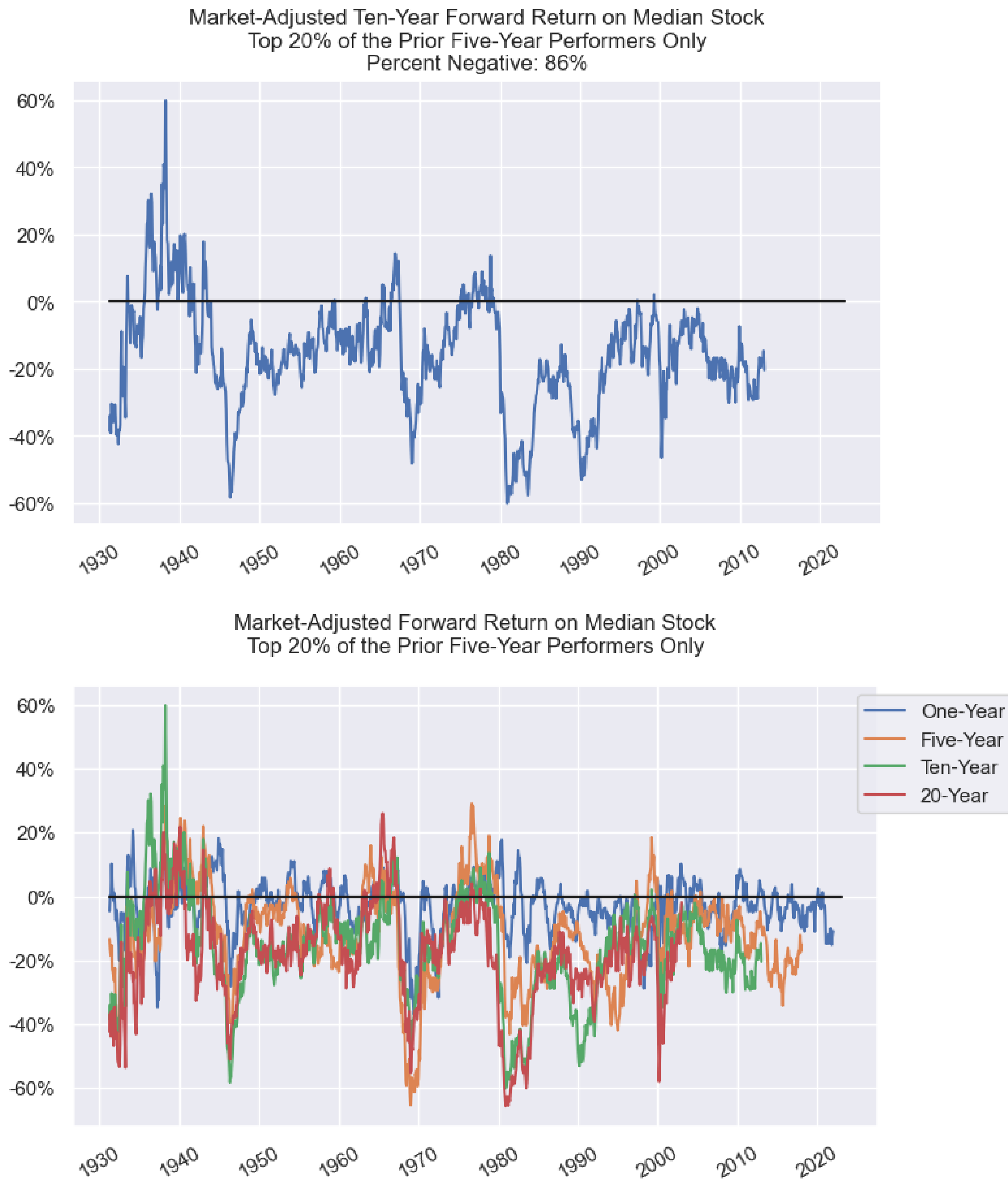

Figure 6. Market-Adjusted Forward Return on the Median Stock over Time.
This figure shows the cumulative forward return on the median stock relative to the cap-weighted market portfolio, computed each calendar month. Only the 20% of stocks with the highest prior five-year return are included. The top panel shows ten-year returns. The bottom panel shows returns for horizons of one, five, ten and 20 years. The selected universe includes all U.S. common stocks except microcaps from 1/1926 to 12/2022.

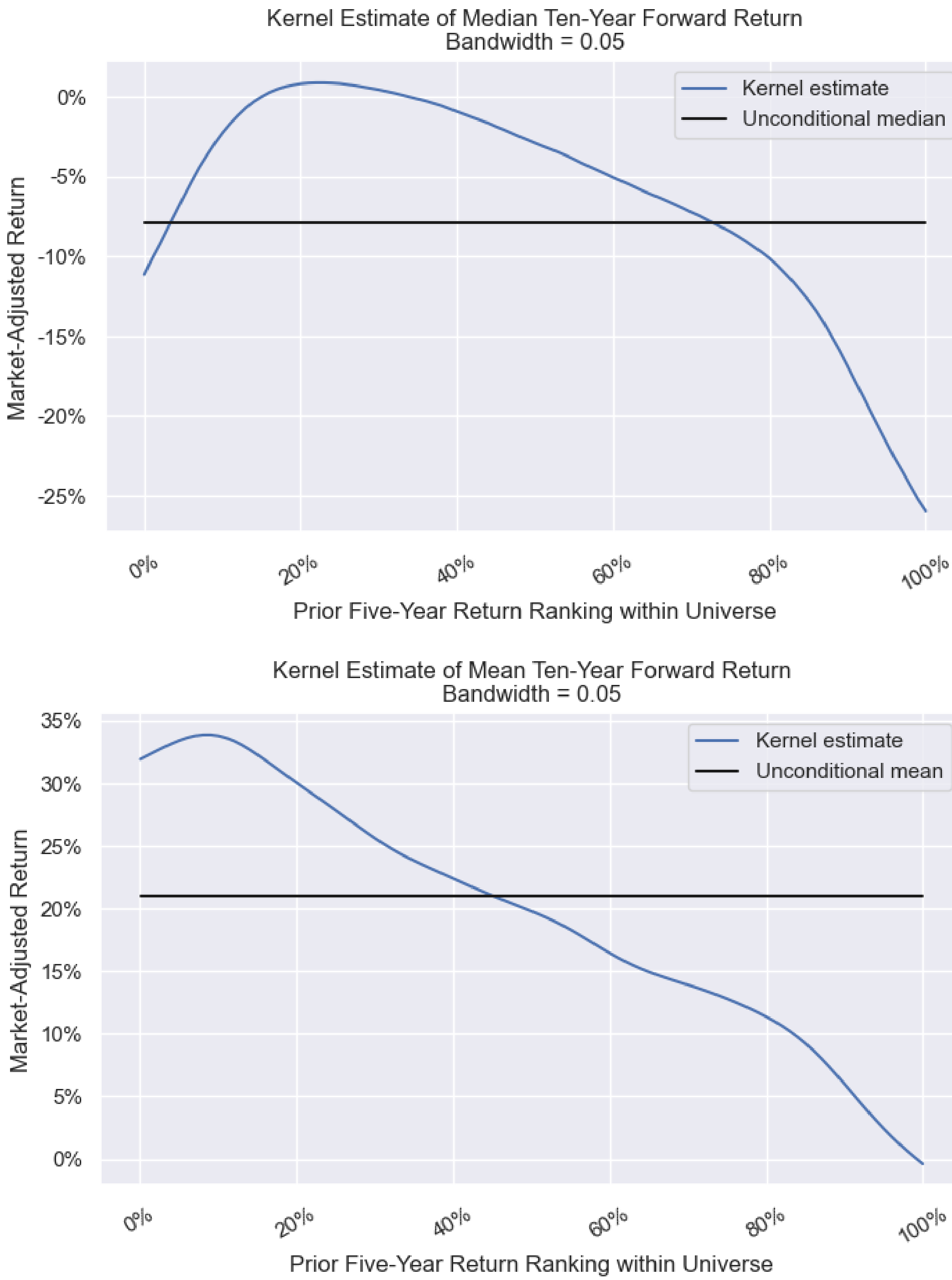

Figure 7. Median and Mean of Ten-Year Forward Return as a Function of Prior Return. This figure shows the cumulative ten-year forward return on individual stocks relative to the cap-weighted market portfolio. The top panel shows the median, and the bottom panel shows the mean. The x-axis is a stock's cross-sectional percentile ranking by five-year cumulative market-adjusted performance within the universe. Estimates are local constants based on a Gaussian kernel with bandwidth=0.05. The selected universe includes all U.S. common stocks except microcaps. The statistics are computed for the pooled sample from 1/1926 to 12/2022.

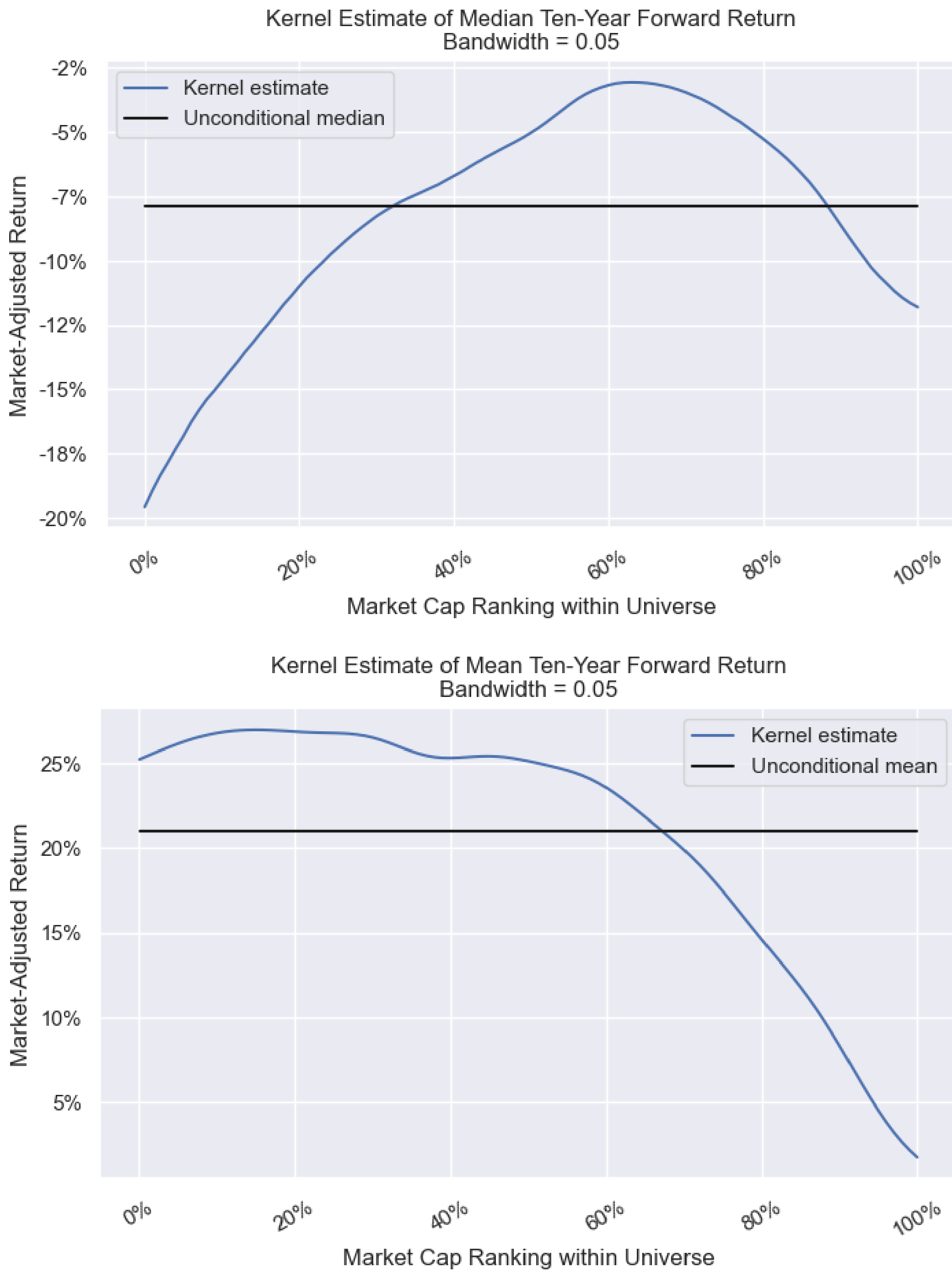

Figure 8. Median and Mean of Ten-Year Forward Return as a Function of Market Cap. This figure shows the cumulative ten-year forward return on individual stocks relative to the cap-weighted market portfolio. The top panel shows the median, and the bottom panel shows the mean. The x-axis is a stock's cross-sectional percentile ranking by market capitalization within the universe. Estimates are local constants based on a Gaussian kernel with bandwidth=0.05. The selected universe includes all U.S. common stocks except microcaps. The statistics are computed for the pooled sample from 1/1926 to 12/2022.