Compound Returns
Adam Farago and Erik Hjalmarsson*
March 26, 2019

Abstract
We provide a theoretical basis for understanding the properties of compound returns. At long horizons, multiplicative compounding induces extreme positive skewness into individual stock returns, an effect primarily driven by single-period volatility. As a consequence, most individual stocks perform very poorly. However, holding just a few stocks (instead of a single one) greatly improves the long-run prospects of an investment strategy, indicating that missing out on the “lucky few” winner stocks is not a great concern. We show analytically how this somewhat counterintuitive result arises from an interaction between compounding and diversification that has seemingly not been previously noted.

*JEL classification: C58, G1.

Keywords: Compound returns; Diversification; Long-run returns; Skewness.

---

*Both authors are at the Department of Economics and the Centre for Finance, University of Gothenburg. Contact information: P.O. Box 640, SE 405 30 Gothenburg, Sweden. Email: adam.farago@cff.gu.se and erik.hjalmarsson@economics.gu.se.
1 Introduction

To a long-run investor, the total compound returns over the investment horizon is the key quantity of interest. Despite this obvious fact, the properties of compound stock returns have been left relatively unexplored in most financial research. However, as shown in recent work by Bessembinder (2018), multiplicative compounding induces effects that are not evident when simply looking at the properties of single-period returns. Through simulation exercises, Bessembinder illustrates how compounding induces positive skewness into multiperiod returns—even if the single-period return is symmetric—and shows that over long horizons this skewness becomes a dominant feature of the distribution of individual compound stock returns. The extreme skewness at long horizons results in a majority of stocks performing very poorly, with a few exceptions that perform extremely well. In short, compounding induces a “few-winners-take-all” effect.

In this paper, we aim to provide a firm theoretical basis for the properties of compound returns. We first derive an expression for the higher order central moments (including skewness) of compound returns, which can be seen as a theoretical verification of the simulation-based findings in Bessembinder (2018). Our theoretical results show that the effects of compounding are actually considerably more extreme than is evident from simulations. These effects are primarily driven by the level of volatility in the single-period return – the higher the volatility, the more extreme the effects – and are not qualitatively affected by the specific distribution (or skewness) of the single-period returns. In the second part of our analysis, we therefore consider the most tractable case, where returns are log-normally distributed. In this setting, we derive some simple but informative results on the properties of long-run compound returns. The results highlight the key role of volatility and show that even a small amount of diversification can tremendously improve the long-run prospects for an investment strategy. In the final part of the analysis, we further analyze how to reconcile the clear long-run benefits of even small degrees of diversification, with the fact that extreme skewness concentrates all the (long-run) returns to just a small fraction of stocks and the apparent implication that failure to own these specific stocks would lead to very poor returns.
Our study is related to the recent work by Bessembinder (2018) and also to other recent papers that explicitly study skewness in individual stock returns (e.g., Neuberger and Payne, 2018, and Oh and Wachter, 2018). Fama and French (2018) establish some empirical facts regarding aggregate (market-wide) compound returns. In comparison to these studies, our clearest contribution is to provide an essentially model-free theoretical result for the higher-order moments in compound returns. This theoretical result turns out to be important, since we also show that direct empirical inference on the skewness in the compound returns of individual stocks is essentially impossible for horizons of 10 years and longer. This holds true also for the types of sample sizes available in simulation exercises, and an immediate corollary to this result is that simulation-based evidence on the skewness in long-horizon compound returns is unreliable and possibly misleading.

In effect, the theoretical results show that skewness in compound returns of individual stocks will tend to grow at a pace even faster than that suggested by the (bootstrap) simulations in Bessembinder (2018). Our results thus reinforce and sharpen the conclusions from Bessembinder’s study and show that the effects of compounding are, by all measures, extreme: 30-year compound returns, for a stock with a monthly volatility of 17%, have a skewness in excess of one million. These results hold irrespective of whether the single-period returns are symmetric or not. A (large) positive skewness in the single-period returns does reinforce the skew-inducing effect of compounding, but the qualitative effects of compounding are identical for symmetric single-period returns. We also analyze the impact of mean reversion on long-run skewness, but even large degrees of mean reversion in returns cannot affect the qualitative conclusions. The dominant factor in determining the skewness of long-run compound returns is the volatility of the single-period returns, and for sufficiently volatile assets, extreme skew-inducing effects from compounding seem inevitable. In practice, this implies that long-run compound individual stock returns will

---

1 There is a large literature on the implications of higher moments for portfolio choice and asset pricing. Early references include Arditti and Levy (1975), Krauss and Litzenberger (1976), Simkowitz and Beedles (1978), Scott and Horvath (1980), and Kane (1982). Examples of more recent studies include Brummermeier, Gollier, and Parker (2007) and Conrad, Dittmar, and Ghysels (2013).

2 Neuberger and Payne (2018) work with an alternative to the standard moment-based measure of skewness, which we use here. Under their measure, the log-normal distribution has zero skew, whereas we show here that for long horizons, log-normality can imply extreme levels of skewness in individual stock returns.
tend to exhibit extreme skewness, whereas compound market returns will be considerably less skewed; skewness for the 30-year return on the market portfolio would typically be below 10.

The extreme effects of compounding renders skewness and other higher-order moments rather meaningless as summary statistics of long-run returns. Not only is it next to impossible to interpret and compare skewnesses of these magnitudes but, as mentioned above, it is also next to impossible to estimate these moments. We instead argue that one should focus on the quantiles of the compound returns, which can both be reliably estimated and offer straightforward interpretations.

Analytical calculations of quantiles require, of course, knowledge of the entire distribution of the compound returns. For sufficiently long horizons, one would expect compound returns to be (almost) log-normally distributed per the central limit theorem. Empirically, we show that the log-normal approximation works reasonably well and the implied long-run performance of various strategies, using the single-period parameter values and the log-normal distribution assumption, is similar to the directly estimated long-run performance of these strategies. As a device for understanding the first order properties of long-run compound returns, the log-normal distribution therefore appears quite adequate.

The empirical results, using the CRSP sample of U.S. stocks, highlight the very strong benefits of diversification for long-run returns. During the 30-year period from January 1987 to December 2016, the total return from a single-stock investment underperforms the investment in one-month T-bills with 82.4% probability, and it underperforms the equal-weighted market portfolio with 94.5% probability. However, investing in a portfolio containing only 10 stocks during the same period provides a total return that outperforms the T-bill investment with 93.7% probability, and investing in a portfolio of 50 stocks brings the probability of beating the equal-weighted market portfolio close to 50%.

We end with an analysis aimed at more fundamentally understanding how diversification can so effectively eliminate the extreme properties of the individual compound stock returns and result in portfolio returns that exhibit little skewness. The extreme skewness in the individual long-run stock returns implies that just a few stocks will end
up generating most of the long-run returns. From a long-run investor perspective, this fact seems to imply that missing out on some, or many, of these top stocks would be devastating for portfolio performance and, absent very good stock-picking skills, one would need to hold a portfolio with extremely many stocks to ensure against such an outcome. In essence, unless one holds the market portfolio, one would run the risk of missing out on the top performers and suffer considerably worse portfolio performance. In contrast, our results show that even small portfolios (e.g., holding 50 stocks out of the several thousand available) get close to the performance of the market portfolio. It is well established in the case of single-period (monthly) returns that relatively small portfolios can attain a large fraction of the total benefits of diversification: Holding 50 stocks has proven to be adequate in recent samples. Our results are important, because they show that the same rule of thumb also applies to the case of compound returns.

How can we reconcile the clear power of diversification for long-run investors with the extreme skewness in individual stock returns and the few-winners-take-all empirical finding in Bessembinder (2018)? Theoretical analysis of compounding in a portfolio setting quickly becomes intractable since it combines a summation in the cross-section and a multiplication in the time-series dimension. We overcome this problem by analyzing the properties of compound portfolio returns in a simple, but still highly revealing, binomial model with just 2 or 3 stocks, where analytical results can be derived. We show how the simple intuition of viewing portfolio returns as (weighted) averages of individual returns no longer holds in a multi-period compound setting. Most strikingly, we find that the

---

3 Simple combinatorics quickly reveal how large a portfolio one would need. For instance, if there are 4,000 stocks (approximately the current number of unique listings in the CRSP data base) and an investor wants an ex ante probability of 90% to hold at least 50 (75) of the 100 top performers, she would have to hold a portfolio of 2,232 (3,186) stocks out of the 4,000.

4 The question of how many stocks are needed for a well-diversified portfolio has long been studied in the case of single-period returns. The conclusion has been changing through time as the number of listed stocks and firm-level idiosyncratic volatility has been rising (Campbell et al., 2001). Evans and Archer (1968) conclude that the benefits of diversification are exhausted when a portfolio contains approximately 10 stocks. Bloomfield et al. (1977) find that around 20 stocks are needed. Statman (1987) argues that a well-diversified portfolio should contain around 30 stocks. Campbell et al. (2001) and Campbell (2017) argue that almost 50 stocks are required in recent subsamples.

5 Bessembinder (2018) carefully documents how a tiny fraction of all stocks have generated the vast majority of wealth for investors: The top 90 U.S. stocks of all time (out of roughly 25,000) contributed more than 50 percent of all wealth accrued to investors. Just five firms generated ten percent of all wealth.
probability that the compound portfolio return beats the best individual compound stock return is non-zero. This is in sharp contrast to the single-period case where the return on the portfolio can never outperform the best individual component. In a specific example with 3 stocks, we find that there is a 72% probability that the equal-weighted portfolio beats all of its 3 individual constituents over a 30-year horizon. For larger numbers of stocks, we use simulations to document similar effects. If there are 50 stocks available, the equal-weighted portfolio outperforms the third best stock with a 44% probability and the 7th best stock with a 99% probability, at a 30-year horizon. Finally, we also show that these results extend to a setting where only a subset of stocks is chosen from a larger universe. If an equal-weighted portfolio is formed by randomly choosing 50 out of 1,000 stocks, there is a 50% probability that the portfolio outperforms the 60th best stock on the market, and a 97% probability that it outperforms the 100th best stock, over a 30-year horizon. These model results explain how diversified portfolios can so quickly approach the performance of the market portfolio and eliminate the extreme behavior of the individual stocks.

2 A motivating empirical exercise

To set the stage for our theoretical analysis, we start with some data-based summary statistics for U.S. stock returns. For short horizons (i.e., from one month to a year), the summary statistics can easily be obtained using monthly and annual returns of individual stocks. However, for longer horizons (e.g., 10 or 30 years), such direct measurement becomes more problematic since far from all stocks exist over such long periods. To get around this issue, we follow Bessembinder (2018) and focus on returns from single-stock strategies that randomly select one stock in each period from all the available stocks in that period. In a bootstrap-like manner, we construct returns for a great number of such random strategies and use these to calculate the return characteristics for different holding periods. The procedure is described in detail in Appendix A, and the number of simulated strategies is set to 200,000. It is worth observing that while the procedure is similar in spirit to a typical bootstrap exercise, the resulting portfolio returns represent
actual empirical returns to feasible strategies. That is, the procedure simply generates returns for the strategy that chooses a single new random stock in each period (month), and the temporal ordering of the underlying return data is maintained. The simulation is implemented using monthly CRSP data on individual stock returns for the 30-year period between January 1987 and December 2016. We restrict ourselves to a 30-year sample period, since we will later compare the directly estimated properties of long-run (30-year) returns, to inferred long-run properties based on short-run (1-month) parameters. Such an exercise only makes sense if the short- and long-run quantities are based on the same sample, as they are when the total sample is 30-year. In the main empirical analysis presented in Section 4, results for earlier sample periods are also shown.

Table 1 shows summary statistics for returns of such single-stock strategies. The first row corresponds to the one-month returns. The monthly average return is 1%, the monthly standard deviation is 19%, and the monthly skewness is close to 4. The remaining rows in Table 1 show summary statistics for compound returns at the 1, 5, 10, 20, and 30 year horizons. The mean and volatility increases with the horizon and, most importantly, so does the skewness. The estimated skewness of the 5-year and 30-year compound returns is 44 and 339, respectively. This result reiterates the message in Bessembinder (2018), namely that the distribution of compound returns over long horizons is highly asymmetric.

The aim of our paper is to provide a deeper understanding of the nature of this asymmetry; its determinants and consequences. The column labeled “Impl Skew” shows the implied skewness of compound returns calculated using the one-period moments (i.e., the one-month mean, variance, and skewness from the first row of the table) and an iid assumption; the explicit formulas for calculating the implied moments of the compound returns are derived in Section 3.1. It is immediately apparent that the implied skewness at longer horizons is vastly greater than the directly estimated skewness. We argue in the next section that the discrepancies between estimated and implied skewness values in Table 1 reflect the fact that skewness is not a suitable measure to understand the asymme-

---

6The numbers in the first row of Table 1 are close to those that one would obtain from a direct calculation of the same summary statistics using the entire pooled CRSP sample of 1-month returns. Essentially, we draw a random sample of 200,000 returns from the pooled sample and calculate the statistics on this random sub-sample.
try of compound returns of individual stocks: First, as we show in Section 3.3, estimated skewness values for long-horizon returns (i.e., those in column “Skew” in Table 1) are severely downward biased. Second, the theoretically implied skewness values are so high (e.g., in the order of billions for 30-year returns) that the values in column “Impl Skew” are impossible to interpret.

We argue instead for focusing on the mean and quantiles of the distribution. Table 1 reports the 10th, 50th, and 90th percentiles. The 30-year mean return is 20.9, whereas the 30-year median return is 0.12, and the 90th percentile of the 30-year returns is 7.88. The fact that the mean is considerably higher than the 90th percentile indicates the severe asymmetry of the distribution.

The final three columns in the table show the percent of realized strategy returns that end up beating either the returns on the risk-free asset (the rolled over 1-month Treasury Bill) or the market portfolio (equal- and value-weighted) over the same period. These probabilities are strictly decreasing in the length of the holding period. If one pursues a strategy of holding a single stock (possibly picking a new stock every month) for a 30-year horizon, the probability of beating the risk-free investment is only around 18%, and the probability of beating the market is a mere 6%. This is in line with the other important message of Bessembinder (2018): In the long-run, the typical stock (or single-stock strategy) tends to perform much worse than the risk-free asset or the market portfolio.

In Section 4 we argue that log-normality provides a convenient and reasonably well-working approximation to understand the above results regarding the quantiles and probabilities of compound returns. Our results also reveal how diversification can vastly improve upon the disappointing long-run performance of the single-stock strategy discussed above.
3 Skewness of compound returns

3.1 Implied higher-order moments

With some abuse of notation, we use lower-case letters to denote single period variables, and upper-case letters to denote compound multi-period variables. Let \( x \) represent the one-period gross return on a given asset or portfolio. Throughout the paper, we will denote the expected value, standard deviation, and skewness of the one-period return as

\[
\mu \equiv E[x], \quad \sigma \equiv Std(x) = \sqrt{E[(x - \mu)^2]}, \quad \gamma \equiv Skew(x) = \frac{E[(x - \mu)^3]}{\sigma^3} .
\] (1)

Define the product process \( X_T \) as

\[
X_T = x_1 \times x_2 \times ... \times x_T ,
\] (2)
where the \( x_t \)s are assumed to be independently and identically distributed (iid) and have the same distribution as \( x \). That is, \( X_T \) represents compound returns over \( T \) periods. Since \( x_t \) is iid for all \( t \), it is straightforward that the \( k \)-th order (non-central) moment of \( X_T \) can be calculated as

\[
E \left[ X_T^k \right] = E \left[ x_1^k \right] \times E \left[ x_2^k \right] \times ... \times E \left[ x_T^k \right] = E \left[ x_1^k \right]^T .
\] (3)

The mean and variance of \( X_T \) can easily be computed using (3) as

\[
E \left[ X_T \right] = \mu^T \quad \text{ and } \quad Var \left( X_T \right) = (\mu^2 + \sigma^2)^T - \mu^{2T} .
\] (4)

Proposition 1 provides a formula for calculating the \( k \)-th order standardized moment of \( X_T \) for any \( k > 2 \).

**Proposition 1** Let \( x \) and \( x_t, t = 1, ..., T \), be iid random variables, and denote

\[
\theta_j = \frac{E[x^j]}{E[x]^j} .
\] (5)
Define the compound process \( X_T = \prod_{t=1}^{T} x_t \). For \( k > 2 \), the \( k \)-th order standardized moment of the compound process is given by

\[
E \left[ \left( X_T - E[X_T] \right)^k \right] = \theta_T^k + \left( \sum_{j=1}^{k-2} \binom{k}{j} (-1)^j \theta_{k-j} T \right) + (-1)^k (1 - k) \left( \theta_2^T - 1 \right)^{k/2}.
\] (6)

**Proof.** See the proof in Appendix B. ■

With the help of Proposition 1, all the higher-order standardized moments of \( X_T \) can easily be obtained. Since we focus on the skewness of compound returns, it is useful to spell out the formula for skewness in a separate corollary.

**Corollary 1** Let \( x \) and \( x_t, t = 1, \ldots, T \), be iid random variables with mean \( \mu \), variance \( \sigma^2 \), and skewness \( \gamma \). The skewness of the compound process \( X_T = \prod_{t=1}^{T} x_t \) is

\[
Skew(X_T) = \frac{\theta_T^T - 3 \theta_2^T + 2}{\left( \theta_2^T - 1 \right)^{3/2}}.
\] (7)

where

\[
\theta_2 = \frac{\sigma^2}{\mu^2} + 1 \quad \text{and} \quad \theta_3 = -2 + 3 \theta_2 + (\theta_2 - 1)^{3/2} \gamma.
\] (8)

**Proof.** This is a straightforward application of Proposition 1 for \( k = 3 \). ■

Table 2 tabulates the skewness of \( X_T \) calculated via Corollary 1, when the single-period returns correspond to monthly returns with \( \mu = 1.01 \) (i.e., 1% per month) and volatility that varies across the columns. Compound returns corresponding to 1-, 5-, 10-, 20-, and 30-year horizons are presented.

Panel A shows the skewness of compound returns, when the single-period returns are symmetric (zero-skew). Several results are worth noting. First, compound returns are positively skewed, and their skewness increases non-linearly with the horizon. That is, compounding induces skewness in long-horizon returns even if single-period returns are symmetric. Second, skewness increases dramatically and highly non-linearly in \( \sigma \), for a given \( T \). In other words, the single-period volatility has a huge effect on the degree of

\footnote{Arditti and Levy (1975) also derive a related result on the third moment of compound returns, but their focus is on portfolio choice and they do not examine the long-run implications of compounding.}
skewness induced by compounding. If the volatility of the monthly returns is $\sigma = 0.05$, which corresponds to a well-diversified portfolio (annual volatility around 17%), then the effect of compounding is relatively modest: The skewness of the 30-year returns is 5.19. On the other hand, for $\sigma \geq 0.14$, which is more typical for individual stocks, the skewness induced by compounding increases very rapidly with the horizon. This leads to our third observation: For large $T$ and $\sigma$, the skewness values are extreme. For example, the skewness of 30-year returns when $\sigma \geq 0.17$ is in the order of millions. Arguably, it is hard to give an interpretation to any skewness level larger than 10, and even more difficult to (intuitively) compare distributions with very large but different skewness values. Finally, it is worth highlighting that the results in Panel A hold for any symmetric distribution. That is, they are equally valid if one-period returns are normally or uniformly distributed. Since the uniform distribution has a fully bounded support, the extreme skewness in long-horizon compound returns is therefore not due to the possibility of extremely large return realizations (i.e., it is not due to an infinite support of the one-period return distribution).\footnote{A given pair of $\mu$ and $\sigma$ pin down the two parameters of the uniform distribution, i.e., the bounds of its support. Specifically, the lower and upper bound can be calculated as $a = \mu - \sqrt{3}\sigma$ and $a = \mu + \sqrt{3}\sigma$, respectively. For example, $\mu = 1.01$ and $\sigma = 0.17$ lead to $a = 0.716$ and $b = 1.304$, i.e., the single-period return cannot be lower than -28.4% or higher than 30.4%}

The rest of Table 2 helps us understand the effect of single-period skewness. Panel B corresponds to the case where monthly returns have a skewness equal to that of a log-normal distribution.\footnote{The log-normal distribution does not have an explicit skewness parameter, but its skewness is a function of the mean and variance of the distribution. Specifically, $\gamma = \frac{\sigma}{\mu} \left( \frac{\sigma^2}{\mu^2} + 3 \right)$. As examples, for $\mu = 1.01$ and $\sigma = 0.05$ (0.17), the skewness is equal to 0.15 (0.51).} Panel C and D represent cases with more greatly skewed one-period returns, with $\gamma = 2$ and $\gamma = 4$, respectively. Our main observation is that the effect of single-period skewness depends on the level of the single-period volatility. When $\sigma$ is low (corresponding to well-diversified portfolios), single-period skewness does not have a large effect on the skewness of long-horizon returns (up to a 30-year horizon). Take the column with $\sigma = 0.05$; the skewness of the 30-year returns is 5.19 when $\gamma = 0$, and 6.77 when $\gamma = 4$. That is, the difference in skewness at the 30-year horizon is actually lower than at the monthly level. When single-period volatility is high, single-period skewness can have a large effect, in absolute terms, on the skewness of compound returns, especially
at long horizons. For example, if $\sigma = 0.17$, the skewness of 30-year returns is of the order of $10^6$ when $\gamma = 0$, and of the order of $10^8$ when $\gamma = 4$. However, large absolute differences between the corresponding cells of different Panels in Table 2 only occur when the values in Panel A (where $\gamma = 0$) are already extreme. In these cases, it is hard to give an interpretation to the differences in the extreme skewness levels. Coming back to the example of 30-year returns when $\sigma = 0.17$, it is difficult to interpret the difference between $\text{Skew}(X_T) = 10^6$ (Panel A) and $\text{Skew}(X_T) = 10^8$ (Panel D).

Figure 1 provides a graphical illustration of the results in Table 2 by plotting the skewness of compound returns as a function of horizon. Single-period volatility, $\sigma$, is varied across the panels, while differing single-period skewness, $\gamma$, is represented by different lines. Panel A clearly illustrates that for low single-period volatility, the skewness in long-horizon compound returns is almost identical regardless of inherent skewness in the single-period returns. As the volatility of the single-period returns increases (through Panels B-D), the skewness in compound returns can easily reach extreme values. However, for a given volatility, the cases with $\gamma = 0$ and $\gamma = 4$ result in qualitatively similar patterns. To that extent, it is the volatility of the single-period returns, and not their skewness, which is of first order importance for the skewness of compound returns. In other words, the patterns in $\text{Skew}(X_T)$ are more similar within the panels of Figure 1 (where single-period skewness is varied), than they are across the panels (where single-period volatility is varied).

The assumption of iid single-period returns was used to derive the above results. In the External Appendix, we relax the iid assumption and analyze the effects of serial dependence on the skewness of compound returns. We rely on a heuristic approximation based on the log-normal case, and arrive at a conclusion similar to the one obtained when looking at the effect of single-period skewness. When $\sigma$ is low, the effect of serial dependence on long-horizon skewness is small. When $\sigma$ is high, the effect of serial dependence can be sizable, but only in the range of extreme skewness levels, where interpretation of the different skewness values is not straightforward any more. To that extent, the effect of serial dependence is also of second order importance compared to the effect of single-period return volatility.
3.2 Intuition from compound binomial returns

The above analysis highlights the extreme effects of compounding on higher order moments, as long as the volatility of the single-period returns is sufficiently high. To get some intuition behind these results, we consider a simple binomial model. Assume that the single-period return, \( x \), can only take two values: there is an “up-tick” in the price with probability \( \pi \) that results in a gross return \( u \), and there is a “down-tick” with probability \( 1 - \pi \), resulting in a gross-return \( d \). Moreover, to isolate the effect of compounding from that of single-period skewness, let \( \pi = 0.5 \), which is equivalent to assuming that the distribution of \( x \) has zero skewness.\(^{10}\)

The mean and standard deviation of \( x \) are then

\[
\mu = \frac{u + d}{2} \quad \text{and} \quad \sigma = \frac{u - d}{2},
\]

so a given pair of mean and volatility can be matched by setting \( u = \mu + \sigma \) and \( d = \mu - \sigma \).

If the \( x_t \)s are \( iid \), then the total return evolves along a recombining binomial tree, and the compound return over \( T \) periods can take on \( T + 1 \) values:

\[
X_T = u^M d^{T-M} = \begin{cases} 
(ud)^M d^{T-2M} & \text{if } M \leq T/2 \\
(u)^{T-M} u^{2M-T} & \text{if } M > T/2
\end{cases},
\]

where \( M \in \{0, 1, ..., T\} \) denotes the number of up-ticks over the investment horizon and \( T - M \) is the number of down-ticks. The second formulation in equation (10) reveals that every possible value of \( X_T \) can be rewritten as a product of pairs of up- and down-ticks, \( ud \), and some remaining up-ticks (if \( M > T/2 \)) or down-ticks (if \( M < T/2 \)).

The probability of experiencing \( M \) up-ticks over \( T \) periods follows a binomial distribution with parameters \( \pi \) and \( T \). The first three columns of Table 3 tabulates the probability of \( M \) or less up-ticks during a 30-year horizon, with the second column indicating the value of \( X_T \) in each case, using the formulation in equation (10). As is seen, it is much more likely to observe a similar number of up- and down-ticks than it is to observe disproportionately more moves in one direction. In other words, we are likely to

\(^{10}\)The skewness of \( x \) in the general case is \( \gamma = \frac{1-2\pi}{\sqrt{\pi(1-\pi)}} \), which is equal to zero only if \( \pi = 0.5 \).
observe a relatively large number of \(ud\) pairs, which highlights the relevance of the second formulation in equation (10). For \(T = 360\), the maximum possible number of paired up- and down-ticks is 180, and there is a 97% chance that we observe at least 160 \(ud\) pairs (since \(P(160 \leq M \leq 200) \approx 0.97\)). Also, \(P(175 \leq M \leq 185) \approx 0.44\), so there is a 44% chance to have at least 175 \(ud\) pairs.\(^{11}\)

The value of \(ud\) will therefore have a major impact on the behavior of long-run compound returns. The fourth column of Table 3 provides actual values of \(X_T\) for a given \(M\) when \(\mu = 1.01\) and \(\sigma = 0.17\), which is used to represent individual stocks. As \(ud = 0.991\) in this case, the investment loses roughly 1% of its value after every \(ud\) pair. Since the number of \(ud\) pairs is likely to be large, the compound effect of these losses will be highly detrimental to the investment. There is a 72% chance that the total compound return over the 30-year period will not exceed 11% (i.e., \(X_T \leq 1.11\)) as seen from row \(M = 185\) of Table 3. On the other hand, since \(u = 1.18\) is relatively large, if the number of up-ticks happens to be disproportionately large (e.g., \(M \geq 210\)), \(X_T\) takes on extremely large values. However, the probability of this happening is very low as \(P(M \geq 210) = 0.001\). The fact that \(X_T\) takes on low values with high probability and exceedingly large values with very low probability creates the extremely asymmetric distribution of long-run compound returns that is typical in the case of individual stocks.

In the last column of Table 3, values of \(X_T\) are shown when \(\mu = 1.01\) and \(\sigma = 0.05\), which is used to represent well-diversified portfolios. This parameterization implies a completely different behavior. Since \(ud = 1.018\), the investment gains almost 2% after every \(ud\) pair, and the compound effect of these gains will be highly beneficial for the total return. Consequently, there is a 99.99% chance that the total compound return over the 30-year period will be higher than 18%, and with 68% probability the return will exceed 1300% (as seen from rows \(M = 150\) and \(M = 175\) of Table 3, respectively). On the other hand, since \(u = 1.06\), \(X_T\) does not take on such extreme values when \(M\) is large compared to the case with \(\sigma = 0.17\). Altogether, these imply that the distribution of \(X_T\) is much less asymmetric when the single-period volatility is low.

\(^{11}\)The high probability to observe a similar number of up- and down-ticks is generally true when \(\pi = 0.5\) and \(T\) is large.
It is straightforward to show from equation (9) that \( gd = \mu^2 - \sigma^2 \). As long as \( \mu > 1 \), low single-period volatility \( (\sigma < \sqrt{\mu^2 - 1}) \) implies \( g > 1 \), which leads to similar behavior as in column 5 of Table 3, while high volatility \( (\sigma > \sqrt{\mu^2 - 1}) \) implies \( g < 1 \), leading to similar behavior as in column 4.

### 3.3 Estimating skewness

It is often of natural interest to directly estimate the properties of compound returns, both in the strict empirical sense but also in Monte Carlo (or bootstrap) simulation exercises. Bessembinder (2018), Fama and French (2018), and Oh and Wachter (2018) all rely on simulations when studying the skewness of long-horizon compound returns. However, as we demonstrate below, relying on simulations in the case of individual stock returns can be highly misleading, because skewness estimators are generally biased, and the bias can be extremely severe in this context.

A natural estimator of skewness is

\[
g \equiv \frac{1}{n} \sum_{i=1}^{n} (z_i - \bar{z})^3 \left( \frac{1}{n} \sum_{i=1}^{n} (z_i - \bar{z})^2 \right)^{-\frac{3}{2}},
\]

where \( z \) denotes a general random variable, \( z_i, i = 1, \ldots, n \) denotes a sample of size \( n \), and \( \bar{z} \) is the sample average. For non-normal distributions, \( g \) is typically biased, but theoretical expressions for the bias are generally not available (e.g., Joanes and Gill, 1998). However, a very simple and often overlooked result implies that skewness estimates of long-horizon compound returns from individual stocks are severely downward biased. Wilkins (1944) shows that there is an upper limit to the absolute value of \( g \), which depends solely on the sample size:

\[
|g| \leq \frac{n - 2}{\sqrt{n - 1}}.
\]

For sample sizes of \( n = 20,000 \) and \( n = 200,000 \), the upper limits are 141.4 and 447.2, respectively.\(^{12}\) When estimating the skewness of long-horizon compound returns from individual stocks, these limits are highly restrictive. As discussed in Section 3.1 and

\(^{12}\)Another commonly used skewness estimator is based on the unbiased central moment estimates, and
illustrated in Table 2, the skewness of long-run individual stock returns can be extreme. If we take the example of log-normal single-period returns with a volatility of $\sigma = 0.17$, the skewness of the 30-year compound returns is $3.6 \times 10^6$. A sample size of $1.3 \times 10^{13}$ would be needed just for the upper limit in (12) not to be binding when estimating such a high level of skewness (and the estimate would still be downward biased).

In the External Appendix, we show in simulation exercises that the upper limit on $g$ is indeed binding for feasible sample sizes. We also show that the asymptotic standard errors on $g$ are extremely large, even when the upper limit in equation (12) is no longer binding. Orders of magnitudes larger sample sizes than those hinted at above would therefore be needed to obtain skewness estimates with any meaningful precision.

Direct estimation of the moment-based measure of skewness for long-horizon compound returns on individual stocks is therefore essentially impossible in practice. Instead, we argue that it is more meaningful to estimate the quantiles of the distribution of the compound returns. This is not to imply that the quantiles are an exact substitute for skewness, but absent the possibility of direct estimation of skewness, estimating different tail quantiles seems like the most reasonable thing to do.\(^\text{13}\)

Let $F_z(w) = \Pr(z \leq w)$ denote the cumulative distribution function of a general random variable $z$, and let $q_\alpha$ denote the $\alpha$-quantile of this distribution, with $0 < \alpha < 1$. That is, $q_\alpha$ is the number that solves $\alpha = F_z(q_\alpha)$, and the sample quantile is given by

$$\hat{q}_\alpha \equiv \inf \left\{ w : \frac{1}{n} \sum_{i=1}^{n} I\{z_i \leq w\} \geq \alpha \right\} . \quad (13)$$

We show in the External Appendix that quantiles of long-horizon compound returns can be estimated with much higher precision (compared to skewness). In particular, results can be written as

$$G = \frac{\sqrt{n(n-1)}}{n-2} g .$$

It is straightforward to see that (12) implies $|G| \leq \sqrt{n}$, which translates into essentially the same limits as for $g$ if the sample size is not too small (i.e., for $n \geq 100$). Samples sizes of $n = 20,000$ or $n = 200,000$ are available in the type of bootstrap exercises that we conduct in this paper (we use $n = 200,000$ draws in all cases). In a purely empirical analysis, the sample sizes would typically be considerably smaller.

\(^{13}\)There exists alternative (non moment-based) measures of skewness that can be constructed from the quantiles of the distribution. However, an analysis of the actual quantiles seems more informative from the perspective of learning about long-run compound returns, and we do not pursue such quantile-based measures of skewness here.
based on simulations and on the asymptotic distribution of \( \hat{q}_a \) both show that quantiles can be reasonably well estimated for relevant sample sizes. Moreover, quantiles offer straightforward interpretations, unlike the extreme values of skewness obtained for long-horizon individual stock returns. Therefore, we strongly advocate using quantiles when studying the distribution of long-horizon compound returns. This will be our focus in the following section.

4 Long-horizon returns in the log-normal case

4.1 Log-normality as an approximation

Characterizing the distribution of long-run compound returns with quantiles rather than moments is considerably more robust from an empirical perspective. However, in terms of deriving theoretical properties for the compound returns, the use of quantiles is more restrictive. The results in Section 3.1 on the (higher-order) moments of compound returns apply to all distributions in the \( iid \) case. In contrast, theoretical calculations of quantiles require knowledge of the full distribution of the compound returns, which is only available in specific cases. Most prominent of these is, of course, the log-normal distribution.

As previously, let \( x \) represent the one-period gross return on a given asset or investment strategy, and let the compound return corresponding to horizon \( T \) be \( X_T = \prod_{t=1}^{T} x_t \), where \( x_t \) are \( iid \) for all \( t \) and have the same distribution as \( x \). By the central limit theorem, (standardized) long-run compound returns will be asymptotically (i.e., as \( T \to \infty \)) log-normally distributed under very general assumptions on the distribution of \( x \), allowing for both serial dependence and heterogeneity (i.e., neither independence nor identical distribution would be required for asymptotic log-normality to hold; see White, 2001). For “large” \( T \), \( X_T \) should therefore be approximately log-normally distributed.

Without any specific assumptions on the distribution of \( x \), define the following quantities,

\[
\psi \equiv \log \left( \frac{\mu^2}{\sqrt{\sigma^2 + \mu^2}} \right) \quad \text{and} \quad \eta \equiv \sqrt{\log \left( \frac{\sigma^2}{\mu^2} + 1 \right)} .
\] (14)
Note that for typical $\mu$ and $\sigma$ values corresponding to monthly stock (or portfolio) returns, $\eta \approx \sigma$. Given the iid assumption in the definition of $X_T$, $\psi$ and $\eta$ scale up with the horizon according to

$$
\log \left( \frac{E[X_T]^2}{\sqrt{\text{Var}(X_T) + E[X_T]^2}} \right) = T\psi \quad \text{and} \quad \sqrt{\log \left( \frac{\text{Var}(X_T)}{E[X_T]^2} + 1 \right)} = \sqrt{T\eta} . \quad (15)
$$

Further, *if we assume* that $x$ is log-normally distributed, $\psi$ and $\eta$ are the parameters of the distribution, i.e.,

$$
E[\log(x)] = \psi \quad \text{and} \quad \text{Std}(\log(x)) = \eta . \quad (16)
$$

Coupling these observations with the implications of the central limit theorem, we have

$$
X_T^{\approx} \sim \text{LN} \left( T\psi, T\eta^2 \right) . \quad (17)
$$

The above distributional result on compound returns is exact when $x$ is log-normal, while it is an approximation (via the central limit theorem) when $x$ has a different distribution.

Given (17), standard results yield that the $\alpha$-quantile of compound returns can be calculated as

$$
q_{\alpha}(X_T) = e^{T\psi + \sqrt{T\eta}\Phi^{-1}(\alpha)} , \quad (18)
$$

where $\Phi^{-1}(\cdot)$ denotes the inverse cdf of the standard normal distribution. By comparing quantiles based on (18) with the “actual” (bootstrapped) quantiles, Panels A and B in Figure 2 provide fairly strong support for the practical applicability of the log-normal approximation. The lines in the graphs show the quantiles calculated via equation (18), using the estimated mean and standard deviation of the monthly returns reported in Table 1 (i.e., $\mu = 1.0102$ and $\sigma = 0.186$ are used, which imply $\psi = -0.0065$ and $\eta = 0.1826$). The round markers in Panels A and B of Figure 2 correspond to the quantiles estimated directly from the single-stock bootstrap exercise described in Section 2 (and reported in Table 1), and can be thought of as representing the “actual” quantiles of the
distribution as a function of the horizon $T$.\textsuperscript{14} This exercise, and similar ones below, is the main reason for focusing on a 30-year sample, where the data used to calculate the short-run parameters exactly correspond to the data used for forming the 30-year quantiles and other properties. We focus the discussion below on the evidence from the 1987-2016 sample, but we also provide results for the two preceding 30-year periods covering 1957-1986 and 1928-1958, which we briefly discuss towards the end of Section 4. Overall, the results from the different 30-year samples are qualitatively similar.

As is seen, the round markers line up quite well with the log-normal quantiles, suggesting that for the distribution of the bootstrapped returns log-normality provides a decent approximation. The correspondence between the lines and round markers is to some extent remarkable, given that the only input to the former is the mean and volatility of monthly returns, while the latter rely on a bootstrap exercise using all the monthly returns to capture the “actual” distribution of long-horizon compound returns.

This is not to say that we think the log-normal distribution provides a perfect characterization of long-horizon stock returns, but we would argue that it seems reasonable as a first order approximation.\textsuperscript{15} In the following sub-sections, we state theoretical results for long-run compound returns under the log-normal approximation, and based on some of these we further corroborate this claim.

\textsuperscript{14}More specifically, the values of the round markers for $\alpha = 0.1$, 0.5, and 0.9 in Panels A and B of Figure 2 are taken from columns “p10”, “Median”, and “p90” of Table 1, corresponding to the quantiles of bootstrapped compound returns over 1, 5, 10, 20, and 30 years. The round markers corresponding to $\alpha = 0.25$ and 0.75 are not presented in Table 1, but come from the same bootstrap exercise.

\textsuperscript{15}Oh and Wachter (2018) state what is effectively the opposite conclusion: The log-normal distribution implies much too extreme tail behavior at long horizons. Specifically, the extreme skewness of the distribution of long-run compound returns suggest that most wealth (i.e., stock value) will eventually be concentrated to just a few stocks (and in the limit only one stock). We do not disagree with their conclusion, but note that our results do not concern the extreme tail behavior of the distribution. In practice, there are likely “real-world” restrictions on the absolute size of firms (consider, for instance, the break-ups of Standard Oil and AT&T), and some of the theoretical long-run tail implications from a simple stylized model might therefore be too extreme.
4.2 Properties of long-run compound returns

4.2.1 Quantiles

We start with some further analysis of the quantiles under the log-normal approximation. The median of $X_T$ corresponds to $\alpha = 0.5$ in equation (18), and can thus be calculated as $\text{Median}(X_T) = e^{T\psi}$. If $\psi = 0$, the median is one at all horizons. If $\psi < 0$, the median decreases and approaches zero as the horizon increases, while $\psi > 0$ implies that the median increases with the horizon. The single-stock strategy of Section 2 has $\psi = -0.0065$, and correspondingly, the median of the compound return gets close to zero as the 30-year horizon is reached (Panel A of Figure 2).

Equation (18) has important implications for the other quantiles as well. To highlight these implications, it is useful to look at the derivative of the quantile with respect to horizon:

$$\frac{\partial q_\alpha(X_T)}{\partial T} = q_\alpha(X_T) \left( \psi + \frac{\eta \Phi^{-1}(\alpha)}{2\sqrt{T}} \right).$$

Consider first the case when $\psi < 0$. All the lower quantiles ($\alpha < 0.5$) are decreasing with the horizon (since both $\psi < 0$ and $\Phi^{-1}(\alpha) < 0$) and they approach zero as $T \to \infty$. Some upper quantiles may initially increase with the horizon, but for any fixed $\alpha$ there is a $T$ value where the second factor in equation (19) becomes negative, and all quantiles will eventually decrease and approach zero as $T$ becomes sufficiently large. This is well illustrated in Panels A and B of Figure 2: the 75th percentile of the compound return distribution decreases when $T \geq 90$ (i.e., after 7.5 years), while the 90th percentile decreases when $T \geq 322$ (i.e., after approximately 27 years).

Turning to the $\psi > 0$ case, it is clear that all the upper quantiles ($\alpha \geq 0.5$) increase with the horizon (since both $\psi > 0$ and $\Phi^{-1}(\alpha) > 0$). Following the same argument as in the previous case, there are lower quantiles that initially decrease with the horizon, but for any fixed $\alpha$ there is a $T$ value, where the corresponding quantile starts to increase and keeps on increasing as $T$ grows further.
4.2.2 Probability of beating the risk-free investment

When trying to determine the long-run success of an asset or investment strategy, it is natural to think about the probability that it beats a certain benchmark over a specific horizon. One popular benchmark is the return on the risk-free asset. Let $R_f$ denote the one-period (monthly in our examples) gross return on the risk-free asset. Empirically, we use the return on the 1-month T-Bill to proxy for the 1-month risk-free rate. The total return over $T$ periods is then $R_T^f$. It is straightforward to show that under the log-normality assumption in (17),

$$P(X_T \geq R_T^f) = \Phi\left(\sqrt{T}\frac{\psi - r_f}{\eta}\right),$$

(20)

where $r_f = \log(R_f)$ and $\Phi(\cdot)$ denotes the cdf of the standard normal distribution. Equation (20) provides a very clear-cut categorization. If $\psi = r_f$, then $P(X_T \geq R_T^f) = 0.5$ irrespective of the horizon. If $\psi > r_f$, the probability that the risky investment beats the risk-free asset is always larger than 0.5, and increases with the horizon (approaching one in the limit). If $\psi < r_f$, the probability that the risky investment beats the risk-free asset is always lower than 0.5, and decreases with the horizon (approaching zero in the limit).

While the value of $\psi$ dominates the asymptotic probability of beating the risk-free rate, the value of $\eta$ still plays a role for finite $T$. Specifically, for $\psi \neq r_f$, the value of $\eta$ will determine how quickly $P(X_T \geq R_T^f)$ converges to one or zero. A smaller $\eta$ implies less variable returns, and a quicker convergence to the asymptotic probability.\(^{16}\)

The average monthly risk-free rate during our sample period from 1987 to 2016 is 0.26%, implying $r_f = 0.0026$, which makes $\psi < r_f$ the relevant case for the single-stock strategy. Panel C of Figure 2 shows the probability that the compound return from the single-stock strategy is higher than the compound risk-free rate, as a function

\(^{16}\)The formula in equation (20) implicitly assumes that the risk-free rate is constant over time. In practice, the risk-free rate varies across periods, and a multi-period investment that rolls over the risk-free asset each period is not risk-free to a long-run investor, in the sense that the final return realization is not known at the time of the investment. However, in times when the 1-month risk-free rate is relatively stable, equation (20) should provide a good approximation when evaluating the risky asset against a “risk-free” benchmark. If the 1-month risk-free rate varies considerably, the formula in equation (23) likely provides a better approximation.
of the horizon. The line shows the probability calculated via equation (20), while the round markers correspond to the “actual” probabilities based on the bootstrap exercise of Section 2 (reported in column “%>Rf” of Table 1). The round markers line up very well with the probabilities implied by log-normality, providing further support for the practical applicability of the log-normal approximation.

4.2.3 Probability of beating a risky benchmark

Some other typical benchmarks are risky investments themselves. As before, let $x$ represent the single-period gross return on a given asset or investment strategy. Consider now another risky return, $x_m$, that represents the return on a benchmark investment. Let

$$
\rho \equiv \log \left( \frac{\text{Cov}(x,x_m)}{\text{E}[x|x_m] \text{E}[x_m]} + 1 \right). \tag{21}
$$

For typical values corresponding to monthly stock (or portfolio) returns, $\rho \approx \text{Corr}(x,x_m)$.

The compound returns on the benchmark strategy is defined as $X_{Tm} = \prod_{t=1}^{T} x_{tm}$. We assume, as a natural extension to the above analysis, that $(x_t,x_{tm})'$ for $t = 1, ..., T$ are iid and have the same joint distribution as $(x,x_m)'$. The log-normal approximation in the two-risky-asset case corresponds to assuming that the returns on the two strategies are jointly log-normally distributed according to

$$
\begin{pmatrix}
\log x \\
\log x_m
\end{pmatrix} \sim N \left( \begin{pmatrix}
\psi \\
\psi_m
\end{pmatrix}, \begin{pmatrix}
\eta^2 & \eta \eta_m \rho \\
\eta \eta_m \rho & \eta_m^2
\end{pmatrix} \right). \tag{22}
$$

In this case, $\rho = \text{Corr}(\log x, \log x_m)$.

Standard calculations show that

$$
P(X_T \geq X_{Tm}) = \Phi \left( \sqrt{T} \frac{\psi - \psi_m}{\sqrt{\eta^2 + \eta_m^2 - 2\eta \eta_m \rho}} \right). \tag{23}
$$

The probability crucially depends on the relation of the parameters $\psi$ and $\psi_m$. If $\psi = \psi_m$, then $P(X_T \geq X_{Tm}) = 0.5$ irrespective of the horizon. If $\psi < \psi_m$, the probability that
the risky investment beats the benchmark is always lower than 0.5, and decreases with
the horizon (approaching zero in the limit). If $\psi > \psi_m$, the probability that the risky
investment beats the benchmark is always larger than 0.5, and increases with the horizon
(approaching one in the limit).

Panel D of Figure 2 shows the probability of the single-stock strategy beating the equal-
weighted market portfolio, as a function of $T$. Similar to the other graphs in Figure 2, the
line shows the probability based on the log-normal approximation (in this case, calculated
via equation (23)), while the round markers represent the “actual” probabilities based on
the bootstrap exercise of Section 2 (reported in column “%>EW” of Table 1). The
log-normal probabilities line up almost exactly with the bootstrapped ones.

All the graphs in Figure 2 suggest that the log-normal distribution, at a minimum,
provides a decent approximation to the behavior of long-run compound returns, in line
with the predictions of the central limit theorem.

4.3 Long-run performance of strategies

In the previous subsections we established three simple rules that help us understand the
behavior of long-horizon compound returns, all of which are related to the parameter $\psi$
of the single-period return distribution. First, if $\psi < 0$, all quantiles of the compound
return distribution approach zero as the horizon goes to infinity, while for $\psi > 0$, all the
quantiles diverge as the horizon increases. Second, if $\psi < r_f$, the probability that the
risky investment beats the risk-free asset approaches zero as the horizon goes to infinity,
while for $\psi > r_f$, the same probability approaches one. Third, if $\psi < \psi_m$, the probability
that the risky investment (represented by $\psi$) beats the benchmark investment strategy
(represented by $\psi_m$) approaches zero as the horizon goes to infinity, while for $\psi > \psi_m$,
the same probability approaches one.

From the definition in equation (14), it is clear that $\psi$ is a non-linear function of $\mu$
and $\sigma$. Therefore, it is instructive to plot different investment strategies in the expected-
return/volatility space, together with curves corresponding to the three rules above. Panel

\footnote{Using the value-weighted market return as the benchmark produces similar (unreported) results.}
A of Figure 3 does so for the single-stock strategy discussed so far in the paper. The round marker represents the single-stock strategy, with $\mu = 1.0102$ and $\sigma = 0.186$. The three curves represent mean/volatility combinations for which $\psi = 0$ (the solid line), $\psi = r_f$ (the dashed line), and $\psi = \psi_m$ (the dotted line) where the risky benchmark is the equal-weighted market portfolio.\(^{18}\) Any point to the right (left) of one of these curves indicates a mean/volatility combination with a strictly smaller (larger) value of $\psi$ than the value represented by the given curve. The single-stock strategy is far to the right of all three curves, indicating $\psi < 0 < r_f < \psi_m$, as discussed in Section 4.2.

Investment strategies in the upper left corner on the graphs of Figure 3 represent those with the greatest long-run growth prospects.\(^{19}\) This area can be approached either by increasing the expected value of single-period returns (moving up) or decreasing their volatility (moving left). One straightforward way to achieve the latter is diversification.

Panel B of Figure 3 illustrates the effect of diversification. The panel shows the mean-volatility characteristics of bootstrapped portfolio strategies, where the equal-weighted portfolio of $N$ randomly selected stocks is created every month.\(^{20}\) These strategies have the same expected return (a very small variation in the actual mean is due to the bootstrap procedure), and the increase in number of stocks thus induces a strict left-ward shift of the round markers in the graph. The lowest variance (highest diversification) is achieved by the equal-weighted market portfolio (represented by the diamond marker), but the portfolio with 50 stocks is already very close. The positive effects of diversification are clearly seen in terms of the compound returns from these strategies. Going from the

\(^{18}\)In the case of $\psi_m$, the corresponding $x_m$ represents the monthly gross returns on the equal-weighted portfolio of all CRSP stocks over the sample period 01/1987-12/2016 with $E[x_m] = 1.0105$ and $\text{Std}(x_m) = 0.058$.

\(^{19}\)To be clear, our results are meant to illustrate the statistical properties of long-run compound returns as a function of their mean and variance, and highlight how both the mean and variance affect long-run returns. As argued forcefully by Samuelson (1969, 1979) and Merton and Samuelson (1974), convergence to the log-normal distribution for long-run compound returns does not imply that all investors should choose the portfolio with the highest $\psi$.

\(^{20}\)The same bootstrap procedure is carried out as in the case of the single-stock strategy in Section 2. The only difference is that instead of selecting a single stock in each month, multiple stocks are selected randomly ($N \in \{2, 5, 10, 25, 50, 100\}$), and the equal-weighted return of the selected stocks is calculated for the given month. A new portfolio is picked for each month. The universe of available stocks and the sample period is exactly the same as in Section 2. These strategies thus capture the returns on monthly rebalanced portfolios, where the stocks are picked randomly and the portfolio is equal-weighted at the beginning of each period.
single-stock strategy to picking two stocks already ensures that the compound returns will not eventually drop to zero (it is above the $\psi = 0$ curve), and including five stocks ensures that the strategy eventually beats the risk-free rate (it is above the $\psi = r_f$ curve).

Table 4 accompanies Figure 3 and elaborates on its findings. The first two columns give the expected return and volatility of the monthly returns for each strategy, simply tabulating what is shown graphically in Figure 3. The next two columns show the corresponding $\psi$ and $\eta$ values. The remainder of the table shows the probabilities that the 30-year compound returns from the strategies beat the return on the risk-free investment and the market portfolio (equal- or value-weighted) over the same period. The columns labeled “actual” use the distribution of 30-year bootstrapped returns for each strategy, while the columns labeled “implied” show the corresponding values implied by the log-normal approximation and the single-period parameters in the first columns. Panel A of Table 4 corresponds to the single-stock strategy, while Panels B and C present the portfolio strategies. In general, there is a close correspondence between the values in the “actual” and “implied” columns, and only in a few cases do the two probabilities differ by more than a few percentage points.\textsuperscript{21} Overall, the results in Table 4 support the previous notion that the log-normal distribution works well as an approximation for long-run compound returns.

Panel B of Table 4 reiterates the benefits of diversification through the example of equal-weighted portfolios (in which case the relevant risky benchmark is the equal-weighted market portfolio). While the probability of the single-stock strategy beating the risk-free investment on a 30-year horizon is only 17.6%, the same probability for portfolios containing as few as 10 stocks is 93.7%. The probability that the single-stock strategy beats the (equal-weighted) market on a 30-year horizon is a mere 5.5%, but the same probability for a portfolio containing only 50 stocks is 40.1%. However, it is essentially impossible to push the latter probability above 50% just via diversification, since it leaves

\textsuperscript{21}In some of these cases, the difference between the “actual” and “implied” probabilities differ by a somewhat substantial margin. To that extent, the initial results presented in Figure 2 might overstate the precision of the log-normal approximation. We stress that we do not view the log-normal distribution as a perfect representation for long-run returns, but the correspondence is close enough that the use of the log-normal distribution as a tool for understanding the main properties of long-run compound returns seems justified.
\[ \mu \text{ unchanged and decreases } \sigma \text{ to the level of the market at best (as shown in Panel B of Figure 3), and hence it cannot achieve } \psi > \psi_m. \] In order to achieve a probability of beating the market in excess of 50%, one needs to find strategies that deliver higher expected single-period returns than the market. There is an enormous literature trying to uncover factors that help to predict cross-sectional patterns in expected single-period returns (for recent overviews see, e.g., Harvey et al., 2016, and Kewei et al., 2019). While the long-run implications of the results in this literature are certainly of interest, they are outside the scope of the current paper.

Panel C of Figure 3 and Panel C of Table 4 present value-weighted portfolios. In this case, the relevant risky benchmark is the value-weighted market portfolio. The conclusions are essentially unchanged: the probability of a 10-stock portfolio beating the risk-free investment on a 30-year horizon is 86.7%. At the same time, the probability that a portfolio containing 50 stocks beats the (value-weighted) market over 30 years is 41.7%.

The empirical results above focus on the 30-year period from January 1987 to December 2016. Table 5 shows that the conclusions are qualitatively unchanged if previous non-overlapping 30-year periods are considered instead (namely, January 1957 to December 1986 or January 1927 to December 1956).

5 Diversification in the long run

The above analysis of compound returns highlights two conclusions. First, for individual stocks the distribution of long-run compound returns is extremely skewed, such that most stocks deliver very poor returns while a few deliver exceptionally large returns. Second, this extreme skewness disappears quickly through diversification (e.g., with 50 stocks in the portfolio). Purely mathematically, this finding is not surprising given the results in Section 3.1, which show that skewness-via-compounding is primarily induced by the volatility of the single-period returns. Diversification quickly brings down the volatility, and the skewness effect thus disappears.

Intuitively, however, this result is less obvious. The extreme skewness of long-horizon individual stock returns indicates that large long-run returns are concentrated to just
a few stocks. Simple intuition might suggest that the failure to own (most of) these stocks would severely negatively affect portfolio performance. But with long-run returns concentrated to just a tiny fraction of firms, one would need to hold a very large number of stocks to ensure that one does not miss out on these extreme performers. Holding just 10 or 50 stocks should not be enough.

In this section, we continue with the binomial model introduced in Section 3.2, in order to offer some intuition for how diversification impacts long-run compound returns. Whereas the results in Section 3.1 provide a “reduced form” explanation of the effects of diversification (through lowered volatility) the subsequent analysis is intended to provide a more “structural” description of the actual mechanics of diversification in compound portfolio returns.

5.1 Compound portfolio returns in the binomial model

Assume that there are \( N \) stocks, whose single-period gross returns, \( x_{ti} \), are iid both across stocks (indexed by \( i \)) and across time (indexed by \( t \)). We are interested in the properties of the compound return on the portfolio of these stocks. To this end, define the following objects:

- \( x_{tp} = \frac{1}{N} \sum_{i=1}^{N} x_{ti} \) is the single-period equal-weighted portfolio return,
- \( X_{Ti} = \prod_{t=1}^{T} x_{ti} \) is the compound return (over \( T \) periods) on stock \( i \),
- \( X_{Tp} = \prod_{t=1}^{T} x_{tp} \) is the compound return on the portfolio that is rebalanced every period to be equal-weighted,
- \( x_t(k) \) is the \( k \)-th largest element of \( \{x_{t1}, ..., x_{tN}\} \), i.e., the \( k \)-th largest return in period \( t \) from the returns on the \( N \) stocks, and
- \( X_T(k) \) is the \( k \)-th largest element of \( \{X_{T1}, ..., X_{TN}\} \), i.e., the \( k \)-th largest total compound return over \( T \) periods from the \( N \) stocks.

Note that \( x_t(1) \) denotes the largest return among all the stocks in period \( t \), while \( X_T(1) \) is the total compound return on the best performing stock.
In the case of one-period returns, it is straightforward to see that $P(x_{tp} > x_t(1)) = 0$ for any portfolio size $N$. That is, in any single period, the portfolio return cannot be higher than the return on the best performing constituent. The interesting question is whether the same simple intuition also holds for compound returns, i.e., whether $P(X_{Tp} > X_T(1))$ is also zero? We demonstrate below that this is far from true and that diversification interacts with compounding in a surprisingly strong way, especially over long horizons. As far as we are aware, these results have not been explored before.

To illustrate how compounding and diversification interacts, suppose there are two stocks and the investment horizon is three periods ($N = 2$ and $T = 3$). The compound return on the stocks are $X_{T1} = x_{11}x_{21}x_{31}$ and $X_{T2} = x_{12}x_{22}x_{32}$, while the total return on the equal-weighted portfolio (rebalanced every month) is

$$X_{Tp} = x_{1p}x_{2p}x_{3p} = \left(\frac{x_{11} + x_{12}}{2}\right) \left(\frac{x_{21} + x_{22}}{2}\right) \left(\frac{x_{31} + x_{32}}{2}\right)$$

$$= \frac{1}{2^3} (x_{11}x_{21}x_{31} + x_{11}x_{21}x_{32} + x_{11}x_{22}x_{31} + x_{11}x_{22}x_{32}$$

$$+ x_{12}x_{21}x_{31} + x_{12}x_{21}x_{32} + x_{12}x_{22}x_{31} + x_{12}x_{22}x_{32}) \quad (24)$$

The last formulation reveals that $X_{Tp}$ can be interpreted as the average of the compound returns on all possible single-stock strategies that can be formed from the underlying stocks. Recall from Section 2 that a single-stock strategy randomly selects one of the two (since $N = 2$) available stocks in each period. When $N = 2$ and $T = 3$, there are eight possible single-stock strategy strategies. In general, there are $N^T$ possible strategies. For long horizons, this number becomes extremely large: If there are still just two stocks but the horizon is 30 years ($T = 360$), the number of single-stock strategies is of the order of $10^{108}$. Some of these possible single-stock strategies might lead to extremely large total returns that can have a considerable impact on the overall value of $X_{Tp}$.

To capture these effects in the simplest possible example, consider two stocks and a two-period horizon ($N = T = 2$), and let us revisit the binomial model from Section 3.2. That is, $x_{ti}$ can take two values, $u$ and $d$, both with a probability of 0.5. In order to represent individual stocks, we use $u = 1.18$ and $d = 0.84$ throughout this section,
corresponding to $\mu = 1.01$ and $\sigma = 0.17$. Both stocks can realize four different return sequences, \{uu, ud, du, dd\}, with equal probability. Therefore, the two stocks can have 16 unique return sequence combinations, as illustrated in Figure 4. Within each cell of the matrix, the top sequence corresponds to stock 1, while the bottom sequence corresponds to stock 2, and each cell realizes with equal probability of $\frac{1}{16}$.

Consider now the case of observing the sequence \{ud\} for one of the stocks and the sequence \{du\} for the other (the cells highlighted in gray). Then,

$$X_{T1} = X_{T2} = X_T (1) = ud = 0.99.$$  \hspace{1cm} (25)

If we form all possible single-stock strategy paths, we get each of the sequences \{uu, ud, du, dd\} once, and taking the average of their total returns as in equation (24), we get

$$X_{Tp} = \frac{d^2 + 2 \times ud + u^2}{4} = \frac{0.71 + 2 \times 0.99 + 1.39}{4} = 1.02,$$  \hspace{1cm} (26)

which implies $X_{Tp} > X_T (1)$. That is, the portfolio outperforms both underlying stocks in this case. The key intuition is that while both constituent stocks experience one up- and one down-tick over the investment horizon, leading to a -1% total return, one of the four possible single-stock paths will experience two up-ticks, leading to a total return of 39%, which raises the overall average such that $X_{Tp} > X_T (1)$. Since this holds (only) in the two highlighted cells, $P (X_{Tp} > X_T (1)) = \frac{2}{16} = 0.125$. Clearly, the probability that the portfolio performs better than any of the constituent stocks is not zero, unlike in the case of single period returns.

In Appendix C.1, we analytically derive the probabilities that a portfolio will outperform any of the constituent stocks in the case with $N = 2$ or $N = 3$ stocks, at an arbitrary horizon $T$. Based on these analytical results, Panel A of Figure 5 extends the above simple example and shows the probability of the portfolio beating its constituents as a function of the investment horizon.\footnote{The results in Figure 5 are based on the binomial model where the single-period return, $x_{ti}$, can take only two values. However, the conclusions regarding the behavior of long-run compound returns do not hinge on this assumption. In the External Appendix we provide simulation evidence that if $x_{ti}$ is assumed to be normally distributed instead (having the same mean and variance), the corresponding...} As the horizon grows, there is a general in-
crease in $P(X_{Tp} > X_T(1))$. At the 30-year horizon, there is a 74% chance that the total compound return on the equal-weighted portfolio is higher than the total return on the better performing stock. This result clearly illustrates the benefits of diversification for long-horizon compound returns.

These benefits become even more obvious if the number of stocks in the portfolio increases. Panel B of Figure 5 corresponds to the case with three stocks in the portfolio. If we consider single-period returns ($T = 1$), the probability that the equal-weighted portfolio performs better than the second best constituent, $X_T(2)$, is 37.5% and the probability that the portfolio beats the best performing constituent, $X_T(1)$, is zero. If the investment horizon is 30 years instead, the probability of the portfolio beating its second best and best performing constituents is 99.8% and 72%, respectively. Based on simulations (see Appendix C.3 for details), Panels C and D of Figure 5 show results for $N = 10$ and $N = 50$ stocks, respectively. When there are 10 stocks (Panel C), the probability that the portfolio beats its third best, second best, and best constituents in any single period is 5.4%, 1%, and zero, respectively. However, the same probabilities on a 30-year horizon are 99.4%, 89.7%, and 47%, respectively. Finally, when the portfolio consists of 50 stocks (Panel D) the probability that it beats its 20th best performing constituent in any single period is merely 6% (and obviously, the probability of beating any of the better ranked constituents is lower than this). In contrast, on a 30-year horizon, the same portfolio beats its 7th best constituent with 99% probability and the 5th best stock with 88% probability.

Finally, consider the case when the market consists of a large number of stocks, and the portfolio contains only a small subset of all available stocks. A similar question to the one above can now be asked: What is the probability that the long-run return on the portfolio is higher than the long-run return on the $k$-th best performing stock on the market? That is, we are interested in $P(X_{Tp} > X_T^*(k))$, where we use the star superscript to emphasize that $X_T^*(k)$ is the $k$-th best performing stock on the market, and not only within the portfolio. Figure 6 shows the probability that an equal-weighted portfolio of 50 stocks results are practically identical to those in Figure 5 for horizons longer than 5 years ($T > 60$).
beats the $k$-th best from 1,000 stocks over a 30-year investment horizon in the binomial model. As is seen, the probability that a portfolio of 50 stocks has better total return than the 60th best stock (out of 1,000) is above 50%. The probability that the portfolio beats the 100th best stock is 97%. These results further highlight the surprisingly strong effect of diversification on long-run compound returns. The results in Figure 6 are based on an approximate analytical solution, with details provided in Appendix C.2.

To sum up, the above results explain how diversification can so effectively eliminate the extreme behavior of individual compound stock returns. In a multiperiod compound setting, the portfolio return is no longer a simple average of the constituent stocks’ returns, but rather an average of a very large number of single-stock strategies based on the constituent stocks. Some of these single-stock strategies are likely to have extremely large returns, even if the constituent stocks themselves are not among the extreme winners.

Coming back to the example in Figure 6, with a universe of 1,000 stocks. If an investor holds a portfolio of 50 stocks randomly selected from this universe, the probability that the investor’s portfolio contains at least 10 of the 100 top performers is approximately 2%; the probability follows from simple combinatorics. Yet, over a 30-year investment horizon, the portfolio beats the 100th best stock (i.e., the portfolio itself is “among” the 100 top performers) with 97% probability. That is, missing out on most of the extreme winners is not as problematic as it would initially seem. A moderate level of diversification (e.g., having 50 stocks in the portfolio) is enough to mostly eliminate the negative effects of the extreme skewness in long-run individual stock returns, explaining the close performance of moderately diversified portfolios and the market portfolio seen in Panel B of Figure 3.

6 Conclusion

We provide a theoretical analysis and characterization of the properties of compound returns of both individual stocks and portfolios. Our key theoretical results can be summarized as follows: (i) Compounding induces extreme skewness in the distribution of

\[\text{\textsuperscript{23}}\text{In the External Appendix we provide evidence that the two approximating assumptions that we make when deriving the analytical formula are fairly accurate.}\]
long-run individual stock returns; (ii) The skew-inducing effect of compounding is primarily driven by the level of single-period return volatility and diversification across stocks quickly eliminates most of the skewness effects of compounding, by bringing down the volatility; (iii) Diversification and compounding interact such that compound portfolio returns can outperform the best of the underlying stocks. The last result provides an explanation of why the concentration of large positive long-run returns to just a few stocks (as implied by the theory and as documented empirically by Bessembinder, 2018) does not imply that failure to hold these stocks is catastrophic for portfolio performance, provided an otherwise diversified portfolio is held.

We also argue that higher-order moments are not a useful way of characterizing the statistical properties of long-run compound returns. Skewness can easily be of the order of millions for individual stock returns at a 30-year horizon and cannot be given a meaningful interpretation. Instead, we suggest that one should study the quantiles of the distribution. We also show that the quantiles can be reliably estimated using feasible sample sizes, whereas skewness (and higher order moments) cannot be reliably estimated for compound returns of individual stocks at horizons greater than 10 years. The latter result is very strong and rules out meaningful inference on skewness of compound returns for any computationally feasible sample size, including those available in simulation and bootstrap studies. In the empirical analysis, we highlight that the log-normal distribution provides a surprisingly accurate tool for modelling some key properties of long-run returns.
Appendix

A Bootstrap exercise

We use monthly returns on all CRSP stocks from the thirty year period between January 1987 and December 2016; the same method applies to earlier sub-samples as well. For various investment horizons denoted by $H$ (e.g., $H = 12$ for a one-year horizon), we implement the following bootstrap procedure:

i. We randomly pick an $H$-month long sub-period within the full 30-year sample denoted by month $\tau$ to month $\tau + H - 1$. Obviously, when $H = 360$, this always corresponds to the full 30-year period from 1987 to 2016.

ii. At the start of month $\tau$, we pick a stock randomly (denoted by $i_\tau$) from all the stocks available in CRSP for the given month. Let $x_{\tau, i_\tau}$ represent the gross return on stock $i_\tau$ in month $\tau$.

iii. At the start of month $\tau + 1$, we pick a new stock randomly (denoted by $i_{\tau+1}$). Let $x_{\tau+1, i_{\tau+1}}$ represent the gross return on stock $i_{\tau+1}$ in month $\tau + 1$.

iv. We repeat the procedure in (iii) for months $\tau + 2$ to $\tau + H - 1$. The resulting return series $\{x_{\tau, i_\tau}, x_{\tau+1, i_{\tau+1}}, \ldots, x_{\tau+H-1, i_{\tau+H-1}}\}$ represents the monthly returns from a strategy of holding a single random stock in each month over the period chosen in (i). Let

$$X_H = \prod_{j=0}^{H-1} x_{j, i_j}$$

represent the total return (ignoring transaction costs) on this strategy.

v. We repeat (i) to (iv) a large number of times (we use 200,000 iterations) to obtain a bootstrap distribution of the total return after $H$ months.
B Proof of Proposition 1

Denoting $E[x] = \mu$ and variance $Var(x) = \sigma^2$, it is straightforward to show that

$$\theta_1 = \frac{E[x]}{E[x]} = 1$$

$$\theta_2 = \frac{E[x^2]}{E[x]^2} = 1 + \left( \frac{E[x^2] - E[x]^2}{E[x]^2} \right) = 1 + \frac{\sigma^2}{\mu^2}. \quad (A1)$$

To determine $\theta_k$ for $k > 2$, start with the binomial expansion of the $k$-th central moment,

$$\frac{E[(x - E[x])^k]}{Var(x)^{k/2}} = \frac{E\left[\sum_{j=0}^{k} \binom{k}{j}(-1)^j x^{k-j} E[x]^j\right]}{Var(x)^{k/2}} = \frac{\sum_{j=0}^{k} \binom{k}{j}(-1)^j E[x^{k-j}] E[x]^j}{Var(x)^{k/2}}$$

$$= \frac{\left(\sum_{j=0}^{k-2} \binom{k-2}{j}(-1)^j E[x^{k-j}] E[x]^j\right) + (-1)^k (1-k) E[x]^k}{Var(x)^{k/2}}$$

$$= \frac{\theta_k + \left(\sum_{j=1}^{k-2} \binom{k-2}{j}(-1)^j \theta_{k-j}\right) + (-1)^k (1-k)}{(\theta_2 - 1)^{k/2}}. \quad (A2)$$

To get to the second line above, spell out the terms with $j = k - 1$ and $j = k$. To get to the third line, divide both the numerator and the denominator by $E[x]^k$, apply the definition of $\theta_j$, and separate the term with $j = 0$ from the sum. Rearranging equation (A2) yields

$$\theta_k = (-1)^k (k - 1) - \left(\sum_{j=1}^{k-2} \binom{k-2}{j}(-1)^j \theta_{k-j}\right) + (\theta_2 - 1)^{k/2} \frac{E[(x - \mu)^k]}{\sigma^k}. \quad (A3)$$

Define the compound process $X_T = x_1 \times \ldots \times x_T$. Since $x_t$ are iid, we have $E[X_T^j] = E[x^j]^T$, which also implies

$$\frac{E[X_T^j]}{E[X_T]^j} = \left(\frac{E[x^j]}{E[x]^j}\right)^T = \theta_j^T. \quad (A4)$$

Using the binomial expansion of the $k$-th central moment of $X_T$ (for $k > 2$), the same steps as in equation (A2), and the equality in equation (A4), we get
\[
\frac{E \left[ (X_T - E[X_T])^k \right]}{\text{Var} \left( X_T \right)^{k/2}} = \frac{E \left[ \sum_{j=0}^{k} \binom{k}{j} (-1)^j X_T^{k-j} E[X_T]^j \right]}{\text{Var} \left( X_T \right)^{k/2}} = \frac{\sum_{j=0}^{k} \binom{k}{j} (-1)^j E \left[ X_T^{k-j} \right] E[X_T]^j}{\text{Var} \left( X_T \right)^{k/2}}
\]
\[
= \frac{\left( \sum_{j=0}^{k-2} \binom{k}{j} (-1)^j E \left[ X_T^{k-j} \right] E[X_T]^j \right) + (-1)^k (1 - k) E[X_T]^k}{\text{Var} \left( X_T \right)^{k/2}}
\]
\[
= \frac{\theta_T^k + \left( \sum_{j=1}^{k-2} \binom{k}{j} (-1)^j \theta_T^j \right) + (-1)^k (1 - k)}{(\theta_T - 1)^{k/2}}.
\]

(A5)

C  Results for the binomial model

Suppose the random variable \( x \) represents single-period gross returns and can take two values: \( u \) with probability \( \pi \), and \( d \) with probability \( 1 - \pi \). Without loss of generality, let \( u > d \). The moments of \( x \) are

\[
E [x] = d + \pi (u - d), \quad \text{Std} (x) = \sqrt{\pi (1 - \pi) (u - d)^2}, \quad \text{Skew} (x) = \frac{1 - 2\pi}{\sqrt{\pi (1 - \pi)}}. \quad (A6)
\]

All the results in this Appendix are valid for a general \( \pi \), but in the main text we focus on the case where \( \pi = 0.5 \).

Let \( x \) and \( x_t, t = 1, ..., T \) be iid random variables and define \( X_T = \prod_{t=1}^{T} x_t \), which represents compound returns over \( T \) periods. \( X_T \) can be written as

\[
X_T = u^M d^{T-M}, \quad (A7)
\]

where \( M \) is a random variable with the support \( \{0, 1, ..., T\} \), representing the number of periods when the single-period return is \( u \). The random variable \( M \) follows a binomial distribution with parameters \( \pi \) and \( T \), i.e., its probability mass function (pmf) and cumulative distribution function (cdf) are

\[
P(M = m) = b(m; T, \pi) = \binom{T}{m} \pi^m (1 - \pi)^{T-m}
\]
\[
P(M \leq m) = B(m; T, \pi) = \sum_{j=0}^{m} b(j; T, \pi). \quad (A8)
\]
Consequently, the pmf of $X_T$ is $P(X_T = u^md^{T-m}) = b(m; T, \pi)$.

C.1 A portfolio beating its $k$-th best constituent

Assume now that there are $N = 2$ stocks. The joint distribution of the single-period returns, $(x_{t1}, x_{t2})$, is described by the following table

<table>
<thead>
<tr>
<th>$x_{t1}$</th>
<th>$x_{t2}$</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u$</td>
<td>$u$</td>
<td>$\pi_{uu}$</td>
</tr>
<tr>
<td>$u$</td>
<td>$d$</td>
<td>$\pi - \pi_{uu}$</td>
</tr>
<tr>
<td>$d$</td>
<td>$u$</td>
<td>$\pi - \pi_{uu}$</td>
</tr>
<tr>
<td>$d$</td>
<td>$d$</td>
<td>$1 - 2\pi + \pi_{uu}$</td>
</tr>
</tbody>
</table>

The returns on the two stocks are identically distributed and have the same distribution as before ($u$ with probability $\pi$ and $d$ with probability $1 - \pi$). The parameter $\pi_{uu}$ denotes the probability that both stocks have a return $u$ in period $t$. To ensure that all probabilities in the above table are non-negative, we need to set $\max(2\pi - 1, 0) \leq \pi_{uu} \leq \pi$. It is straightforward to show that

$$\text{Corr}(x_{t1}, x_{t2}) = \frac{(\pi_{uu} - \pi^2)}{\pi(1 - \pi)}.$$  \hspace{1cm} (A9)

The two stocks are uncorrelated (independent) if $\pi_{uu} = \pi^2$, and have a positive (negative) correlation if $\pi_{uu} > \pi^2$ ($\pi_{uu} < \pi^2$). In the main text, we focus on the case with $\pi_{uu} = \pi^2$.

Consider a $T$ period setting and assume that the joint distribution of $(x_{t1}, x_{t2})$ is iid across time for $t = 1, ..., T$. We introduce the following random variables:

- $L_{uu}$ is the number of periods where $x_{t1} = x_{t2} = u$,
- $L_{ud}$ is the number of periods where $x_{t1} = u$ and $x_{t2} = d$,
- $L_{du}$ is the number of periods where $x_{t1} = d$ and $x_{t2} = u$,
- $L_{dd}$ is the number of periods where $x_{t1} = x_{t2} = d$. 
Note that $0 \leq L_{uu}, L_{ud}, L_{du}, L_{dd} \leq T$ and $L_{uu} + L_{ud} + L_{du} + L_{dd} = T$. The joint distribution of the above variables is a multinomial distribution and has the pmf

$$P(L_{uu} = l_{uu}, L_{ud} = l_{ud}, L_{du} = l_{du}, L_{dd} = T - l_{uu} - l_{ud} - l_{du}) = \frac{T!}{l_{uu}!l_{ud}!l_{du}!(T - l_{uu} - l_{ud} - l_{du})!} \pi_{uu}^{l_{uu}} \pi_{ud}^{l_{ud}} \pi_{du}^{l_{du}} (1 - 2\pi_{uu} + \pi_{uu})^{T - l_{uu} - l_{ud} - l_{du}} \quad (A10)$$

The compound return over $T$ periods on the two stocks and on the equal-weighted portfolio (rebalanced every period) are

$$X_{T1} = u^{l_{uu} + l_{ud}} d^{l_{du} + l_{dd}}$$
$$X_{T2} = u^{l_{uu} + l_{du}} d^{l_{ud} + l_{dd}}$$
$$X_{Tp} = u^{l_{uu}} \left(\frac{u + d}{2}\right)^{l_{ud} + l_{du}} d^{l_{dd}} \quad (A11)$$

For a set of values $(L_{dd} = l_{dd}, L_{ud} = l_{ud}, L_{du} = l_{du}, L_{uu} = l_{uu})$, $X_{Tp} > X_{T1}$ is satisfied when

$$u^{l_{uu}} \left(\frac{u + d}{2}\right)^{l_{ud} + l_{du}} d^{l_{dd}} > u^{l_{uu} + l_{ud}} d^{l_{du} + l_{dd}} \iff (l_{ud} + l_{du}) \left(\frac{\log \left(\frac{u + d}{2d}\right)}{\log \left(\frac{u}{d}\right)}\right) > l_{ud} \quad (A12)$$

Similarly, $X_{Tp} > X_{T2}$ is satisfied when

$$u^{l_{uu}} \left(\frac{u + d}{2}\right)^{l_{ud} + l_{du}} d^{l_{dd}} > u^{l_{uu} + l_{du}} d^{l_{ud} + l_{dd}} \iff (l_{ud} + l_{du}) \left(\frac{\log \left(\frac{u + d}{2u}\right)}{\log \left(\frac{u}{d}\right)}\right) < l_{ud} \quad (A13)$$

Observe that $A_u > A_d > 0$.

Let us first derive $P(X_{Tp} > X_T(1))$. It can be seen from (A12) and (A13) that the portfolio compound return will be higher than both of the individual compound returns, i.e., $X_{Tp} > X_T(1)$ if

$$(l_{ud} + l_{du}) A_u > l_{ud} > (l_{ud} + l_{du}) A_d \quad (A14)$$
The probability that the portfolio beats the better performing stock can be written as

\[ P (X_{Tp} > X_T (1)) = \sum_{l=0}^{T} P (X_{Tp} > X_T (1) \mid L_{ud} + L_{du} = l) P (L_{ud} + L_{du} = l) . \tag{A15} \]

We now derive the probabilities from the above equation. Using equation (A10), the following probabilities can be derived (details are in the External Appendix)

\[ P (L_{ud} + L_{du} = l) = b (l; T, 2 (\pi - \pi_{uu})) , \tag{A16} \]
\[ P (L_{ud} = l_{ud} \mid L_{ud} + L_{du} = l) = b (l_{ud}; l, 0.5) , \tag{A17} \]

where \( b (\cdot; \cdot) \) represents the pmf of the binomial distribution (see equation (A8)). Combining (A14) and (A17) we get

\[ P (X_{Tp} > X_T (1) \mid L_{ud} + L_{du} = l) = P (lA_u > l_{ud} > lA_d \mid L_{ud} + L_{du} = l) \]
\[ = \sum_{lA_u > j > lA_d, j \in \mathbb{N}} b (j; l, 0.5) . \tag{A18} \]

Finally, substituting (A16) and (A18) into (A15), we arrive at

\[ P (X_{Tp} > X_T (1)) = \sum_{l=0}^{T} \left( b (l; T, 2 (\pi - \pi_{uu})) \sum_{lA_u > j > lA_d, j \in \mathbb{N}} b (j; l, 0.5) \right) . \tag{A19} \]

Let us now derive \( P (X_{Tp} > X_T (2)) \). Since \( N = 2 \),

\[ P (X_{Tp} > X_T (2)) = 1 - P (X_{Tp} \leq X_{T1}, X_{Tp} \leq X_{T2}) . \tag{A20} \]

We can see from (A12) and (A13) that \( X_{Tp} \leq X_{T1} \) and \( X_{Tp} \leq X_{T2} \) are jointly satisfied if

\[ (l_{ud} + l_{du}) A_u \leq l_{ud} \quad \text{and} \quad (l_{ud} + l_{du}) A_d \geq l_{ud} . \tag{A21} \]

However, since \( A_u > A_d > 0 \), this is only satisfied if \( l_{ud} = l_{du} = 0 \), that is if the two stocks have the same returns (both have a return \( u \) or both have a return \( d \)) in all periods. In this case, \( X_{T1} = X_{T2} = X_{Tp} \). It can be shown that \( P (L_{ud} = 0, L_{du} = 0) = (1 - 2\pi + 2\pi_{uu})^T \)
(see the details in the External Appendix). Therefore,

\[ P(X_{T_p} \not\geq X_T(2)) = 1 - \left(1 - 2\pi + 2\pi_{uu}\right)^T. \] (A22)

For the \( N = 3 \) case, we refer the reader to the External Appendix.

C.2 A portfolio beating the \( k \)-th best stock on the market

Assume that \( X_{T_1}, X_{T_2}, ..., X_{T_N} \) are \( iid \) random variables and let \( X_T[k] \) be the \( k \)-th order statistic, i.e., the \( k \)-th smallest value if we take a sample of \( X_{T_1}, X_{T_2}, ..., X_{T_N} \). It is important to note that in the main text we use \( X_T(k) \) to denote the \( k \)-th largest element, and thus

\[ X_T(k) = X_T[N - k + 1]. \] (A23)

Let us also consider the \( iid \) random variables \( M_1, M_2, ..., M_N \), where \( M_i \) represents the number of periods with return \( u \) corresponding to \( X_{T_i} \), and let \( M[k] \) be the \( k \)-th order statistic of \( M_1, M_2, ..., M_N \). Clearly,

\[ X_T[k] = u^{M[k]}d^{T - M[k]}. \]

The cdf of \( M[k] \) is given by

\[ P(M[k] \leq m) = \sum_{j=0}^{N-k} \left(\begin{array}{c} N \\ j \end{array}\right) P(M > m)^j P(M \leq m)^{N-j} \] (A24)

where \( m \in \{0, 1, ..., T\} \), and \( M \) has the distribution described in (A8). Using the notations from (A8), we can rewrite the above as

\[ P(M[k] \leq m) = \sum_{j=0}^{N-k} \left(\begin{array}{c} N \\ j \end{array}\right) (1 - B(m; T, \pi))^j (1 - (1 - B(m; T, \pi)))^{N-j} \] (A25)

\[ = B(N - k; N, 1 - B(m; T, \pi)) \].
The corresponding pdf is then
\[
P(M[k] = m) = \begin{cases} 
B(N - k; N, 1 - B(0, T, \pi)) & \text{if } m = 0 \\
B(N - k; N, 1 - B(m, T, \pi)) - B(N - k; N, 1 - B(m - 1, T, \pi)) & \text{if } m > 0.
\end{cases}
\] (A26)

Consider now a portfolio with \(N\) stocks. It is straightforward to show that the portfolio’s single-period mean and standard deviation are
\[
\mu_p = \mu \quad \text{and} \quad \sigma_p = \frac{\sigma}{\sqrt{N}}. \quad (A27)
\]

Using these values, define
\[
\psi_p = \log \left( \frac{\mu_p^2}{\sigma_p^2 + \mu_p^2} \right) \quad \text{and} \quad \eta_p = \sqrt{\log \left( \frac{\sigma_p^2}{\mu_p^2} + 1 \right)}.
\] (A28)

Using the log-normal approximation, \(X_{Tp} \sim LN(T\psi_p, T\eta_p^2)\), and therefore
\[
P(X_{Tp} \leq y) \approx \Phi \left( \frac{\log y - T\psi_p}{\sqrt{T\eta_p}} \right). \quad (A29)
\]

Let us assume that the market has \(N^*\) stocks, and denote the order statistics of the compound returns on all the stocks on the market as \(X^*_T[k]\). The probability that the portfolio of \(N\) stocks beats the \(k\)-th worst performing stock from the market is
\[
P(X_{Tp} > X^*_T[k]) = \sum_{m=0}^{T} P(X_{Tp} > X^*_T[k], X^*_T[k] = u^m d^{T-m})
\]
\[
= \sum_{m=0}^{T} P(X_{Tp} > u^m d^{T-m}) P(X^*_T[k] = u^m d^{T-m}) \quad \text{(A30)}
\]
\[
= \sum_{m=0}^{T} \left[ 1 - P(X_{Tp} \leq u^m d^{T-m}) \right] P(X^*_T[k] = u^m d^{T-m})
\]
\[
= \sum_{m=0}^{T} \left[ 1 - \Phi \left( \frac{\log (u^m d^{T-m}) - T\psi_p}{\sqrt{T\eta_p}} \right) \right] P(M^*[k] = m),
\]

where we make the assumption that \(X_{Tp}\) and \(X^*_T[k]\) are independent when going from the fist line to the second. The formula for \(P(M^*[k] = m)\) is given in equation (A26).
with the only difference that $N^*$ have to be used instead of $N$, as the order statistic now refers to the market. Finally, using (A23),

$$P \left( X_{T_p} > X^*_T (k) \right) = P \left( X_{T_p} > X^*_T [N^* - k + 1] \right).$$  \hspace{1cm} (A31)

When deriving the above formula, we made the approximating assumptions that (i) $X_{T_p}$ is log-normal, and (ii) $X_{T_p}$ and $X^*_T [k]$ are independent. Assumption (i) is based on the central limit theorem and constitutes a good approximation if $T$ is large. Assumption (ii) provides a good approximation if the portfolio is tiny compared to the market ($N^* >> N$). We provide simulation evidence in the External Appendix that for the particular example showed in Figure 6 (i.e., for $T = 360$, $N = 50$, $N^* = 1000$), these approximating assumptions are fairly accurate.

C.3 Simulation results

The results in Panels C and D of Figure 5 are based on simulations. For a given set of $(u, d, N, T)$ values, we implement the following procedure:

i. For iteration $j = 1$, generate $N \times T$ iid realizations of $x_{ti}$ (for $i = 1, ..., N$ and $t = 1, ..., T$), where $x_{ti}$ can take the values $u$ or $d$ with equal probability. Using the simulated $x_{ti}$, calculate the following objects:

- $X_{Ti} = \prod_{t=1}^{T} x_{ti}$ for $i = 1, ..., N$,
- $X_{T_p} = \prod_{t=1}^{T} \left( \frac{1}{N} \sum_{i=1}^{N} x_{ti} \right)$,
- $X_T (k)$ as the $k$-th largest element of $\{X_{T1}, ..., X_{TN}\}$

ii. Let $I_{jk} = I \left( X_{T_p} > X_T (k) \right)$, where $I ()$ is the indicator function. That is, $I_{jk}$ takes the value one if the portfolio compound return is larger than the $k$-th largest individual compound return in iteration $j$.

iii. Repeat (i) to (ii) a large number of times, $j = 1, ..., J$ (we use $J = 200,000$ iterations). The probability that the portfolio beats its $k$-th best constituent can be estimated as

$$P \left( X_{T_p} > X_T (k) \right) = \frac{1}{J} \sum_{j=1}^{J} I_{jk}.$$
References


Figure 1: Skewness of compound returns

The graphs show the skewness of compound returns, $Skew(X_T)$, as a function of the compounding horizon, $T$, when single-period returns are iid. The values are calculated using equation (7). The expected value of the single-period gross return is $\mu = 1.01$ in all cases, the volatility of the single-period return, $\sigma$, is varied across the panels (see above each panel), while different single-period skewness values are represented by the different lines (see legends).

A. $\sigma = 0.05$

- Log-Normal
- $Skew=0$
- $Skew=2$
- $Skew=4$

B. $\sigma = 0.08$

- Log-Normal
- $Skew=0$
- $Skew=2$
- $Skew=4$

C. $\sigma = 0.11$

- Log-Normal
- $Skew=0$
- $Skew=2$
- $Skew=4$

D. $\sigma = 0.17$

- Log-Normal
- $Skew=0$
- $Skew=2$
- $Skew=4$
Figure 2: Properties of the single-stock strategy

The top two graphs show quantiles of compound returns from the single-stock strategy as a function of the compounding horizon $T$. The bottom two graphs show the probability of the single-stock strategy beating the risk-free asset (Panel C) or the equal-weighted market portfolio (Panel D) as a function of the horizon $T$. In all the graphs, the lines show quantiles or probabilities calculated via the log-normal approximation (i.e., via equation (18) in Panels A and B, equation (20) in Panel C, and equation (23) in Panel D). For the log-normal approximation, the single-period mean and volatility of $\mu = 1.0102$ and $\sigma = 0.186$ are used. The round markers show quantiles or probabilities estimated directly from the single-stock bootstrap exercise in Section 2 (the corresponding values are also reported in Table 1).
Figure 3: Strategies in the mean-volatility space

The round markers in each graph present the mean (y-axis) and volatility (x-axis) of the single-period (monthly) gross return of different strategies. The mean and volatility values are the same as the ones reported in columns “μ” and “σ” of Table 4 (the panels in this figure and in Table 4 correspond directly to each other). The diamond marker in all graphs correspond to the monthly return on the market portfolio; the equal-weighted market portfolio in Panels A and B (with \( \mu_m = 1.0105 \), \( \sigma_m = 0.058 \), and \( \psi_m = 0.0088 \)) and the value-weighted market portfolio in Panel C (with \( \mu_m = 1.0090 \), \( \sigma_m = 0.045 \), and \( \psi_m = 0.0080 \)). The curves represent mean-volatility combinations for which the value of \( \psi \), calculated via equation (14), is constant. Specifically, they correspond to \( \psi = 0 \) (solid line), \( \psi = r_f = 0.0026 \) (dashed line), and \( \psi = \psi_m \) (dotted line).

A. Unconditional single-stock strategy

B. Portfolio strategies (equal-weighted)

N=2,5,10,25,50,100 (right to left)

C. Portfolio strategies (value-weighted)

N=2,5,10,25,50,100 (right to left)
Figure 4: Return sequences in the binomial model for two stocks and two periods
The figure shows all possible return-sequence combinations in the binomial model with two stocks and two periods. The return of stock $i$ in period $t$ can take two values, $u$ (“up”) or $d$ (“down”) with equal probability. Within each cell of the matrix, the top sequence corresponds to stock 1, while the bottom sequence corresponds to stock 2. The two highlighted cells correspond to the only cases when the equal-weighted portfolio of the two stocks has a higher compound return after the two periods than any of the two stocks (when $u = 1.18$ and $d = 0.84$).
Figure 5: Probability of the portfolio beating its constituents

The graphs show the probability that the total compound return on the equal-weighted portfolio of $N$ stocks, $X_{T_p}$, is higher than the compound return on its $k$-th best performing constituent, $X_T(k)$, in the binomial model (i.e., $X_T(1)$ is the total return of the best performing constituent, $X_T(2)$ is the total return on the second best constituent, etc.). The $x$-axis on each graph corresponds to the investment horizon, $T$, and the panels present different portfolio sizes, $N$ (as indicated above each panel). The single-period (monthly) gross return on stock $i$ in period $t$ can take two values, $u = 1.18$ or $d = 0.84$ with equal probability, and the single-period returns are iid both across stocks and time. Panels A and B show analytical results, and the results in Panels C and D are obtained via simulations. The analytical formulas and the simulation procedure are given in Appendix C.
Figure 6: Probability of the portfolio beating the $k$-th best stock on the market

The figure shows the probability that the total compound return on the equal-weighted portfolio of 50 stocks, $X_{tp}$, is higher than the compound return on the $k$-th best performing stock, $X^*_T (k)$, out of 1,000 identical stocks in the binomial model. The x-axis corresponds to $k$. The investment horizon is 30-years ($T = 360$ periods). The single-period (monthly) gross return on stock $i$ in period $t$ can take two values, $u = 1.18$ or $d = 0.84$ with equal probability, and the single-period returns are iid both across stocks and time. The analytical formula for calculating $P (X_{tp} > X^*_T (k))$ is given in Appendix C.2.
Table 1: Distribution of long-horizon returns from a single-stock strategy

The table shows descriptive statistics of the total gross return, over different investment horizons, from the single-stock strategy that invests in a single new random stock in each period from the universe of CRSP stocks. For each horizon, the total compound return of the strategy is simulated in a bootstrap-like manner, using 200,000 repetitions, and the statistics in the table are calculated over these simulated returns. The sample period is from January 1987 to December 2016. The columns “Mean”, “Std”, and “Skew” report the mean, standard deviation, and skewness of the compound returns, respectively. The column “Impl Skew” shows the implied skewness of compound returns calculated using the monthly moments and an \textit{iid} assumption (equation (7) in Section 3.1). The columns “p10”, “Median”, and “p90” correspond to the 10th, 50th, and 90th percentiles of the compound returns. The columns “\%>Rf”, “\%>VW”, and “\%>EW” show the percent of simulated single-stock strategies that have higher total return than the risk-free asset, the value-weighted market portfolio, and the equal-weighted market portfolio, respectively, over the same period.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Mean</th>
<th>Std</th>
<th>Skew</th>
<th>Impl Skew</th>
<th>p10</th>
<th>Median</th>
<th>p90</th>
<th>%&gt;Rf</th>
<th>%&gt;VW</th>
<th>%&gt;EW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 month</td>
<td>1.0102</td>
<td>0.186</td>
<td>3.63</td>
<td>3.63</td>
<td>0.83</td>
<td>1.00</td>
<td>1.18</td>
<td>48.8</td>
<td>46.6</td>
<td>46.4</td>
</tr>
<tr>
<td>1 year</td>
<td>1.13</td>
<td>0.82</td>
<td>5.21</td>
<td>4.66</td>
<td>0.48</td>
<td>0.97</td>
<td>1.91</td>
<td>44.7</td>
<td>38.8</td>
<td>38.5</td>
</tr>
<tr>
<td>5 years</td>
<td>1.83</td>
<td>4.96</td>
<td>44.4</td>
<td>68.9</td>
<td>0.11</td>
<td>0.76</td>
<td>4.05</td>
<td>37.4</td>
<td>28.1</td>
<td>25.6</td>
</tr>
<tr>
<td>10 years</td>
<td>3.02</td>
<td>22.3</td>
<td>115.5</td>
<td>3269</td>
<td>0.04</td>
<td>0.49</td>
<td>5.50</td>
<td>29.2</td>
<td>20.1</td>
<td>17.1</td>
</tr>
<tr>
<td>20 years</td>
<td>9.51</td>
<td>326.9</td>
<td>232.1</td>
<td>$1.0 \times 10^7$</td>
<td>0.01</td>
<td>0.22</td>
<td>7.35</td>
<td>21.5</td>
<td>11.5</td>
<td>8.7</td>
</tr>
<tr>
<td>30 years</td>
<td>20.9</td>
<td>1044.9</td>
<td>339.2</td>
<td>$3.2 \times 10^{10}$</td>
<td>0.00</td>
<td>0.12</td>
<td>7.88</td>
<td>17.6</td>
<td>6.4</td>
<td>5.5</td>
</tr>
</tbody>
</table>
The table shows the skewness of compound returns, $Skew(X_T)$, for various compounding horizons $T$ (in different rows), when single-period returns are iid. The values are calculated using equation (7). The expected value of the single-period gross return is $\mu = 1.01$ in all cases, the volatility of the single-period return, $\sigma$, is varied across the columns, while the skewness of the single-period return, $\gamma$, is varied across the panels of the table.

<table>
<thead>
<tr>
<th>$T$</th>
<th>$\sigma = 0.02$</th>
<th>$\sigma = 0.05$</th>
<th>$\sigma = 0.08$</th>
<th>$\sigma = 0.11$</th>
<th>$\sigma = 0.14$</th>
<th>$\sigma = 0.17$</th>
<th>$\sigma = 0.20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. $\gamma = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>12</td>
<td>0.19</td>
<td>0.48</td>
<td>0.78</td>
<td>1.10</td>
<td>1.46</td>
<td>1.86</td>
<td>2.31</td>
</tr>
<tr>
<td>60</td>
<td>0.46</td>
<td>1.23</td>
<td>2.26</td>
<td>3.89</td>
<td>6.89</td>
<td>13.2</td>
<td>28.0</td>
</tr>
<tr>
<td>120</td>
<td>0.66</td>
<td>1.93</td>
<td>4.24</td>
<td>10.2</td>
<td>30.8</td>
<td>123</td>
<td>625</td>
</tr>
<tr>
<td>240</td>
<td>0.97</td>
<td>3.36</td>
<td>11.7</td>
<td>68.5</td>
<td>74.5</td>
<td>$1.4 \times 10^4$</td>
<td>$3.8 \times 10^5$</td>
</tr>
<tr>
<td>360</td>
<td>1.22</td>
<td>5.19</td>
<td>32.1</td>
<td>513</td>
<td>$2.0 \times 10^4$</td>
<td>$1.6 \times 10^6$</td>
<td>$2.3 \times 10^8$</td>
</tr>
<tr>
<td>B. The log-normal case: $\gamma = \frac{\sigma^2}{\mu^2} \left( \frac{\sigma^2}{\mu^2} + 3 \right)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.06</td>
<td>0.15</td>
<td>0.24</td>
<td>0.33</td>
<td>0.42</td>
<td>0.51</td>
<td>0.60</td>
</tr>
<tr>
<td>12</td>
<td>0.21</td>
<td>0.52</td>
<td>0.86</td>
<td>1.23</td>
<td>1.65</td>
<td>2.14</td>
<td>2.75</td>
</tr>
<tr>
<td>60</td>
<td>0.47</td>
<td>1.26</td>
<td>2.33</td>
<td>4.09</td>
<td>7.50</td>
<td>15.3</td>
<td>36.3</td>
</tr>
<tr>
<td>120</td>
<td>0.67</td>
<td>1.95</td>
<td>4.35</td>
<td>10.8</td>
<td>35.1</td>
<td>161</td>
<td>1031</td>
</tr>
<tr>
<td>240</td>
<td>0.97</td>
<td>3.40</td>
<td>12.1</td>
<td>75.7</td>
<td>960</td>
<td>$2.3 \times 10^4$</td>
<td>$1.0 \times 10^6$</td>
</tr>
<tr>
<td>360</td>
<td>1.23</td>
<td>5.25</td>
<td>33.54</td>
<td>595</td>
<td>$2.9 \times 10^4$</td>
<td>$3.6 \times 10^6$</td>
<td>$1.0 \times 10^9$</td>
</tr>
<tr>
<td>C. $\gamma = 2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>12</td>
<td>0.77</td>
<td>1.09</td>
<td>1.46</td>
<td>1.88</td>
<td>2.39</td>
<td>3.03</td>
<td>3.84</td>
</tr>
<tr>
<td>60</td>
<td>0.73</td>
<td>1.59</td>
<td>2.86</td>
<td>5.15</td>
<td>10.2</td>
<td>23.4</td>
<td>65.4</td>
</tr>
<tr>
<td>120</td>
<td>0.87</td>
<td>2.28</td>
<td>5.23</td>
<td>14.4</td>
<td>57.1</td>
<td>351</td>
<td>3272</td>
</tr>
<tr>
<td>240</td>
<td>1.13</td>
<td>3.85</td>
<td>15.3</td>
<td>125</td>
<td>2489</td>
<td>$1.1 \times 10^5$</td>
<td>$1.0 \times 10^7$</td>
</tr>
<tr>
<td>360</td>
<td>1.37</td>
<td>5.95</td>
<td>46.1</td>
<td>1260</td>
<td>$1.2 \times 10^5$</td>
<td>$3.7 \times 10^7$</td>
<td>$3.4 \times 10^{10}$</td>
</tr>
<tr>
<td>D. $\gamma = 4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
</tr>
<tr>
<td>12</td>
<td>1.36</td>
<td>1.71</td>
<td>2.14</td>
<td>2.68</td>
<td>3.38</td>
<td>4.32</td>
<td>5.61</td>
</tr>
<tr>
<td>60</td>
<td>1.00</td>
<td>1.95</td>
<td>3.50</td>
<td>6.61</td>
<td>14.5</td>
<td>40.4</td>
<td>149</td>
</tr>
<tr>
<td>120</td>
<td>1.07</td>
<td>2.64</td>
<td>6.33</td>
<td>20.1</td>
<td>105</td>
<td>996</td>
<td>$1.7 \times 10^4$</td>
</tr>
<tr>
<td>240</td>
<td>1.29</td>
<td>4.36</td>
<td>19.8</td>
<td>230</td>
<td>8263</td>
<td>$8.9 \times 10^5$</td>
<td>$2.7 \times 10^8$</td>
</tr>
<tr>
<td>360</td>
<td>1.51</td>
<td>6.77</td>
<td>65.9</td>
<td>3083</td>
<td>$7.3 \times 10^5$</td>
<td>$8.4 \times 10^8$</td>
<td>$4.5 \times 10^{12}$</td>
</tr>
</tbody>
</table>
Table 3: 30-year compound returns in the binomial model

The table shows 30-year ($T = 360$) compound returns from the binomial model. The single-period return can take two values, $u$ (“up”) or $d$ (“down”) with equal probability, $\pi = 0.5$, and the returns are iid across time. $M$ is a random variable representing the number of $u$ realizations throughout the 360 periods. The first column shows some selected values that $M$ can take. The second column shows the general formula for the 30-year compound return, $X_T$, when $M = m$ (see equation (10)). The third column shows the probability $P(M \leq m) = \sum_{j=0}^{m} \binom{T}{j} \pi^j (1 - \pi)^{T-j}$. The final two columns show specific values of $X_T$, for two different parameterizations (described in the column headers).

<table>
<thead>
<tr>
<th>$M = m$</th>
<th>$X_T$</th>
<th>$P(M \leq m)$</th>
<th>specific $X_T$ values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\mu = 1.01$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\sigma = 0.17$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$u = 1.18$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$d = 0.84$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$ud = 0.991$</td>
</tr>
<tr>
<td>150</td>
<td>$(ud)^{150} u^{60}$</td>
<td>0.0009</td>
<td>$7.6 \times 10^{-6}$</td>
</tr>
<tr>
<td>160</td>
<td>$(ud)^{160} d^{40}$</td>
<td>0.020</td>
<td>$2.3 \times 10^{-4}$</td>
</tr>
<tr>
<td>170</td>
<td>$(ud)^{170} d^{20}$</td>
<td>0.158</td>
<td>0.007</td>
</tr>
<tr>
<td>175</td>
<td>$(ud)^{175} d^{10}$</td>
<td>0.318</td>
<td>0.037</td>
</tr>
<tr>
<td>180</td>
<td>$(ud)^{180}$</td>
<td>0.521</td>
<td>0.204</td>
</tr>
<tr>
<td>185</td>
<td>$(ud)^{175} u^{10}$</td>
<td>0.720</td>
<td>1.114</td>
</tr>
<tr>
<td>190</td>
<td>$(ud)^{170} u^{20}$</td>
<td>0.866</td>
<td>6.096</td>
</tr>
<tr>
<td>200</td>
<td>$(ud)^{160} u^{40}$</td>
<td>0.985</td>
<td>182.4</td>
</tr>
<tr>
<td>210</td>
<td>$(ud)^{150} u^{60}$</td>
<td>0.999</td>
<td>5459</td>
</tr>
<tr>
<td>360</td>
<td>$v_{360}$</td>
<td>1</td>
<td>$7.5 \times 10^{25}$</td>
</tr>
</tbody>
</table>

51
Table 4: Properties of 30-year compound returns from various strategies

The first four columns of the table show descriptives of the single-period returns of different bootstrapped strategies: mean ($\mu$), standard-deviation ($\sigma$), and $\psi$ and $\eta$ calculated via equation (14). The last four columns show descriptives of the 30-year compound returns from the same strategies. The columns labeled “actual” under “% $> \text{Rf}$”, “% $> \text{VW}$”, and “% $> \text{EW}$” show the percent of simulated strategies that have higher total return than the risk-free asset and the value- or equal-weighted market portfolio, respectively, over the 30-year period. The columns labeled “implied” show the corresponding probabilities implied by the log-normal approximation and the single-period parameters in the first four columns. The bootstrap procedure is described in Appendix A and the number of simulations is set to 200,000. The sample is the CRSP stocks and the sample period is from January 1987 to December 2016.

<table>
<thead>
<tr>
<th>single-period returns</th>
<th>30-year returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$\sigma$</td>
</tr>
<tr>
<td>A. Unconditional single-stock strategy</td>
<td>1.0102</td>
</tr>
<tr>
<td>B. Portfolio strategies (equal-weighted)</td>
<td></td>
</tr>
<tr>
<td>N=2</td>
<td>1.0103</td>
</tr>
<tr>
<td>N=5</td>
<td>1.0103</td>
</tr>
<tr>
<td>N=10</td>
<td>1.0105</td>
</tr>
<tr>
<td>N=25</td>
<td>1.0105</td>
</tr>
<tr>
<td>N=50</td>
<td>1.0105</td>
</tr>
<tr>
<td>N=100</td>
<td>1.0104</td>
</tr>
<tr>
<td>C. Portfolio strategies (value-weighted)</td>
<td></td>
</tr>
<tr>
<td>N=2</td>
<td>1.0093</td>
</tr>
<tr>
<td>N=5</td>
<td>1.0096</td>
</tr>
<tr>
<td>N=10</td>
<td>1.0096</td>
</tr>
<tr>
<td>N=25</td>
<td>1.0095</td>
</tr>
<tr>
<td>N=50</td>
<td>1.0093</td>
</tr>
<tr>
<td>N=100</td>
<td>1.0092</td>
</tr>
</tbody>
</table>
Table 5: Properties of 30-year compound returns from various strategies (early samples)
The table shows the same descriptive statistics as Table 4 (see the caption of Table 4), but for different sample periods. Panel A corresponds to the 30-year period from January 1957 to December 1986, while Panel B corresponds to the 30-year period from January 1927 to December 1956.

<table>
<thead>
<tr>
<th></th>
<th>single-period returns</th>
<th>30-year returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu$</td>
<td>$\sigma$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### A. Jan 1957 - Dec 1986

#### Unconditional single-stock strategy

|                      | 1.0124 | 0.136 | 0.0034 | 0.134 | 42.9  | 42.1   | 24.3   | 24.4   | 13.4   | 13.1   |

#### Portfolio strategies (equal-weighted)

<table>
<thead>
<tr>
<th>N</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\psi$</th>
<th>$\eta$</th>
<th>% $&gt;\text{Rf}$</th>
<th>% $&gt;\text{VW}$</th>
<th>% $&gt;\text{EW}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.0122</td>
<td>0.105</td>
<td>0.0068</td>
<td>0.103</td>
<td>69.9</td>
<td>64.2</td>
<td>41.4</td>
</tr>
<tr>
<td>5</td>
<td>1.0124</td>
<td>0.080</td>
<td>0.0093</td>
<td>0.079</td>
<td>94.0</td>
<td>86.0</td>
<td>65.7</td>
</tr>
<tr>
<td>10</td>
<td>1.0124</td>
<td>0.069</td>
<td>0.0101</td>
<td>0.068</td>
<td>99.5</td>
<td>92.9</td>
<td>82.2</td>
</tr>
<tr>
<td>25</td>
<td>1.0123</td>
<td>0.062</td>
<td>0.0104</td>
<td>0.061</td>
<td>100.0</td>
<td>95.9</td>
<td>96.4</td>
</tr>
<tr>
<td>50</td>
<td>1.0123</td>
<td>0.059</td>
<td>0.0105</td>
<td>0.058</td>
<td>100.0</td>
<td>96.8</td>
<td>99.7</td>
</tr>
<tr>
<td>100</td>
<td>1.0124</td>
<td>0.058</td>
<td>0.0107</td>
<td>0.057</td>
<td>100.0</td>
<td>97.6</td>
<td>100.0</td>
</tr>
</tbody>
</table>

#### Portfolio strategies (value-weighted)

<table>
<thead>
<tr>
<th>N</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\psi$</th>
<th>$\eta$</th>
<th>% $&gt;\text{Rf}$</th>
<th>% $&gt;\text{VW}$</th>
<th>% $&gt;\text{EW}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.0114</td>
<td>0.099</td>
<td>0.0066</td>
<td>0.097</td>
<td>64.2</td>
<td>63.9</td>
<td>34.8</td>
</tr>
<tr>
<td>5</td>
<td>1.0104</td>
<td>0.075</td>
<td>0.0076</td>
<td>0.074</td>
<td>83.3</td>
<td>76.8</td>
<td>44.5</td>
</tr>
<tr>
<td>10</td>
<td>1.0101</td>
<td>0.064</td>
<td>0.0080</td>
<td>0.064</td>
<td>91.4</td>
<td>83.3</td>
<td>48.2</td>
</tr>
<tr>
<td>25</td>
<td>1.0096</td>
<td>0.055</td>
<td>0.0080</td>
<td>0.055</td>
<td>97.4</td>
<td>87.1</td>
<td>50.0</td>
</tr>
<tr>
<td>50</td>
<td>1.0095</td>
<td>0.051</td>
<td>0.0082</td>
<td>0.050</td>
<td>99.4</td>
<td>90.2</td>
<td>50.8</td>
</tr>
<tr>
<td>100</td>
<td>1.0092</td>
<td>0.048</td>
<td>0.0081</td>
<td>0.047</td>
<td>99.9</td>
<td>90.7</td>
<td>51.4</td>
</tr>
</tbody>
</table>

### B. Jan 1927 - Dec 1956

#### Unconditional single-stock strategy

|                      | 1.0145 | 0.163 | 0.0017 | 0.160 | 65.1  | 53.7   | 25.4   | 20.7   | 13.7   | 11.5   |

#### Portfolio strategies (equal-weighted)

<table>
<thead>
<tr>
<th>N</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\psi$</th>
<th>$\eta$</th>
<th>% $&gt;\text{Rf}$</th>
<th>% $&gt;\text{VW}$</th>
<th>% $&gt;\text{EW}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.0146</td>
<td>0.136</td>
<td>0.0056</td>
<td>0.133</td>
<td>91.1</td>
<td>75.1</td>
<td>42.9</td>
</tr>
<tr>
<td>5</td>
<td>1.0147</td>
<td>0.115</td>
<td>0.0056</td>
<td>0.133</td>
<td>99.8</td>
<td>88.9</td>
<td>67.4</td>
</tr>
<tr>
<td>10</td>
<td>1.0150</td>
<td>0.107</td>
<td>0.0093</td>
<td>0.106</td>
<td>100.0</td>
<td>93.4</td>
<td>83.7</td>
</tr>
<tr>
<td>25</td>
<td>1.0148</td>
<td>0.103</td>
<td>0.0096</td>
<td>0.101</td>
<td>100.0</td>
<td>95.0</td>
<td>97.2</td>
</tr>
<tr>
<td>50</td>
<td>1.0148</td>
<td>0.101</td>
<td>0.0098</td>
<td>0.099</td>
<td>100.0</td>
<td>95.6</td>
<td>99.8</td>
</tr>
<tr>
<td>100</td>
<td>1.0149</td>
<td>0.100</td>
<td>0.0100</td>
<td>0.098</td>
<td>100.0</td>
<td>96.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

#### Portfolio strategies (value-weighted)

<table>
<thead>
<tr>
<th>N</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\psi$</th>
<th>$\eta$</th>
<th>% $&gt;\text{Rf}$</th>
<th>% $&gt;\text{VW}$</th>
<th>% $&gt;\text{EW}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.0123</td>
<td>0.122</td>
<td>0.0050</td>
<td>0.120</td>
<td>86.9</td>
<td>74.1</td>
<td>32.1</td>
</tr>
<tr>
<td>5</td>
<td>1.0113</td>
<td>0.097</td>
<td>0.0066</td>
<td>0.096</td>
<td>97.7</td>
<td>87.2</td>
<td>36.9</td>
</tr>
<tr>
<td>10</td>
<td>1.0101</td>
<td>0.086</td>
<td>0.0064</td>
<td>0.085</td>
<td>99.7</td>
<td>88.9</td>
<td>38.6</td>
</tr>
<tr>
<td>25</td>
<td>1.0101</td>
<td>0.079</td>
<td>0.0070</td>
<td>0.078</td>
<td>100.0</td>
<td>93.3</td>
<td>40.4</td>
</tr>
<tr>
<td>50</td>
<td>1.0100</td>
<td>0.074</td>
<td>0.0073</td>
<td>0.074</td>
<td>100.0</td>
<td>95.0</td>
<td>42.4</td>
</tr>
<tr>
<td>100</td>
<td>1.0099</td>
<td>0.073</td>
<td>0.0073</td>
<td>0.072</td>
<td>100.0</td>
<td>95.4</td>
<td>44.5</td>
</tr>
</tbody>
</table>
External Appendix to “Compound Returns”

Adam Farago and Erik Hjalmarsson*
1 Skewness under serial correlation

In this section, we analyze the effect of serial dependence on the skewness of compound returns. Since compounding involves multiplication rather than summation, the exact effects of serial dependence on the compound returns is extremely difficult to derive. Therefore, we rely on a heuristic approximation based on the log-normal case. Our results can be summarized by the following proposition:

**Proposition 1** Suppose \( x \) and \( x_t, t = 1, \ldots, T \) are log-normally distributed random variables with mean \( \mu \) and variance \( \sigma^2 \), and let \( X_T = \prod_{t=1}^{T} x_t \) be the corresponding compound returns. Further, let \( VR \) denote the ratio between the long-run and short-run variance of \( x_t \), such that
\[
VR \equiv \frac{LR.Var(x)}{Var(x)} = \frac{\sum_{j=\infty}^{\infty} Cov(x_t, x_{t+j})}{Var(x_t)}. \tag{1}
\]

The skewness of \( X_T \) in the case when the \( x_t \)-s are serially correlated can be approximated as
\[
Skew(X_T) = \left( \left( 1 + \frac{VR \times \sigma^2}{\mu^2} \right)^T + 2 \right) \left( \left( 1 + \frac{VR \times \sigma^2}{\mu^2} \right)^T - 1 \right)^\frac{1}{2}. \tag{2}
\]

**Proof.** Let \( x_t \) be log-normally distributed with parameters \( \psi \) and \( \eta \). That is, the log-returns \( y_t \equiv \log (x_t) \) are normally distributed with mean \( \psi \) and volatility \( \eta \). Assume further that \( y_t \) follows a linear (infinite moving average) process, such that
\[
y_t = \psi + u_t, \tag{3}
\]
and
\[
u_t = C(L) \epsilon_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j}. \tag{4}
\]
The innovations \( \epsilon_t \) are assumed to be \( iid \) standard normal, i.e., \( \epsilon_t \sim N(0,1) \).

The compound return over \( T \) periods is given by \( X_T = \prod_{t=1}^{T} x_t \) and the log-compound returns satisfy,
\[
Y_T = \log(X_T) = \sum_{t=1}^{T} y_t = \psi T + \sum_{t=1}^{T} u_t. \tag{5}
\]
Using the BN decomposition (Beveridge and Nelson, 1981), we can write

\[ u_t = C(L) \epsilon_t = C(1) \epsilon_t + \tilde{\epsilon}_{t-1} - \tilde{\epsilon}_t, \tag{6} \]

where

\[ \tilde{\epsilon}_t = \tilde{C}(L) \epsilon_t = \sum_{j=0}^{\infty} \tilde{c}_j \epsilon_{t-j} \quad \text{and} \quad \tilde{c}_j = \sum_{s=j+1}^{\infty} c_s. \tag{7} \]

\[ C(1) = \sum_{j=0}^{\infty} c_j \] denotes the so-called long-run moving average coefficient. The process \( Y_T \) can therefore be written as,

\[ Y_T = \psi_T + C(1) \sum_{t=1}^{T} \epsilon_t + \sum_{t=1}^{T} (\tilde{\epsilon}_{t-1} - \tilde{\epsilon}_t) = \psi_T + C(1) \sum_{t=1}^{T} \epsilon_t - \tilde{\epsilon}_T, \tag{8} \]

using the fact that \( \sum_{t=1}^{T} (\tilde{\epsilon}_{t-1} - \tilde{\epsilon}_t) = \tilde{\epsilon}_0 - \tilde{\epsilon}_T \) and imposing \( \tilde{\epsilon}_0 = 0 \).

The BN decomposition decomposes the process into a drift component \((\psi_T)\), a martingale component \(\left( C(1) \sum_{t=1}^{T} \epsilon_t \right)\), and a transitory component \((\tilde{\epsilon}_T)\). For large \( T \), the permanent (martingale) component has a variation that is of an order of magnitude greater than the transitory component, and will therefore dominate the stochastic properties of \( Y_T \). We can therefore write the “long-run” part of \( Y_T \) as

\[ Y_T^{LR} \equiv \psi_T + C(1) \sum_{t=1}^{T} \epsilon_t \approx Y_T. \tag{9} \]

Since \( \epsilon_t \overset{iid}{\sim} N(0,1) \), it follows that \( \sum_{t=1}^{T} \epsilon_t \sim N(0,T) \) and \( Y_T^{LR} \sim N(\psi_T, C(1)^2 T) \). Thus, from the definition of the log-normal distribution, \( e^{Y_T^{LR}} \sim LN(\psi_T, C(1)^2 T) \). That is, since \( \log(X_T) = Y_T \approx Y_T^{LR}, X_T \approx LN(\psi_T, C(1)^2 T) \). The parameters \( \psi_T \) and \( C(1)^2 T \) pin down the distribution, and therefore also the skewness, of the compound returns as discussed previously.

In order to assess the effects of serial dependence on compound returns vis-à-vis the \( iid \) setting, consider the case where \( u_t \) is \( iid \). In this case, \( \log(X_T) = Y_T^{LR} \sim N(\psi_T, \eta^2 T) \), where the equality between \( \log(X_T) \) and \( Y_T^{LR} \) is now exact. Compared to the serially correlated case, the mean parameter of the distribution of \( Y_T^{LR} \) is the same, but the
variance is different. The effect of the serial dependence is therefore summarized by the differences between the variance of $Y_{LR}^{T}$ in the serially correlated case, and the variance of $Y_{LR}^{T}$ in the iid case.

Note that the (short-run) variance of $y_t$ is given by,

$$\eta^2 = \text{Var} (y_t) = \text{Var} (u_t) = \sum_{j=0}^{\infty} c_j^2,$$

and the so-called long-run variance of $y_t$ is given by

$$LR.\text{Var} (y_t) = \sum_{j=-\infty}^{\infty} \text{Cov} (y_t, y_{t+j}) = C (1)^2 = \left( \sum_{j=0}^{\infty} c_j \right)^2.$$  \hspace{1cm} (11)

The variance of $Y_{LR}^{T}$ in the iid case is thus equal to $T \times \text{Var} (y_t)$, whereas the variance of $Y_{LR}^{T}$ in the serially correlated case is equal to $T \times LR.\text{Var} (y_t)$. Define the variance ratio, of the long-run variance of $y_t$ over its short-run variance,

$$VR = \frac{LR.\text{Var} (y_t)}{\text{Var} (y_t)} = \frac{\left( \sum_{j=0}^{\infty} c_j \right)^2}{\sum_{j=0}^{\infty} c_j^2}.$$

A given serial correlation structure $\{c_j\}_{j=0}^{\infty}$ reduces or increases the long-run variance of $y_t$, relative to the iid case, by a factor given in the expression above. The impact of serial correlation on skewness in compound returns is therefore evaluated by comparing the skewness implied for compound returns when using the short-run variance and the skewness implied when using the long-run variance.

The variances in equation (12) correspond to log-returns, $y_t$, whereas the inputs into the skewness formula in Corollary 1 of the main text correspond to variances of simple returns, $x_t$. Since $x_t \sim LN (\psi, \eta^2)$, the relationship between the parameter $\eta^2$ and $\text{Var} (x) = \sigma^2$ is given by $\sigma^2 = \mu^2 \left( e^{\eta^2} - 1 \right)$, where $\mu = E [x]$. Defining $\sigma_{VR}^2 \equiv \mu^2 \left( e^{VR\eta^2} - 1 \right)$ and using the first-order Taylor approximation $e^{VR\eta^2} - 1 \approx VR\eta^2$ (which is a good approximation, since the typical values of $\eta^2$ in our context are close to zero), it can be shown
that
\[ \sigma^2_{VR} = \frac{e^{VR\eta^2} - 1}{e^{\eta^2} - 1} \approx \frac{VR\eta^2}{\eta^2} = VR. \]

That is, \( \sigma^2_{VR} \approx VR\sigma^2 \). As the log-variance shifts by a factor \( VR \), so does the variance of simple returns, up to a first order approximation. In order to assess the effect of serial correlation on skewness, one can therefore also equally well calculate the variance ratio of the simple returns, rather than for the log-returns.

Finally, applying Corollary 1 of the main text to the log-normal case, it can be shown that if \( x_t \sim LN(\psi, \eta^2) \), then
\[ Skew(X_T) = \left( \left( 1 + \frac{\sigma^2}{\mu^2} \right)^T + 2 \right) \left( \left( 1 + \frac{\sigma^2}{\mu^2} \right)^T - 1 \right)^{\frac{1}{2}}. \quad (13) \]

For a specific variance ratio, \( VR \) (calculated from either simple returns or log-returns), skewness of the compound returns can be calculated as in equation (2).

There is a large literature suggesting that returns are mean-reverting over longer horizons, which implies that \( VR < 1 \) (see, for instance, Fama and French (1988), Poterba and Summers (1988), Cecchetti, Lam, and Mark (1990), Cutler, Poterba, and Summers, (1991), Siegel (2008), and Spierdijk, Bikker, and van den Hoek (2012)).\(^1\)

Proposition 1 essentially implies that if single-period returns have mean \( \mu \), volatility \( \sigma \), and are non-iid, the skewness of the resulting compound returns behaves as if the single-period returns were iid with mean \( \mu \) and volatility \( \sqrt{VR} \times \sigma \). If, for example, the non-iid single period returns have \( \sigma = 0.17 \) and \( VR = 0.8 \), the resulting \( Skew(X_t) \) can be well approximated by the iid formula with volatility parameter equal to \( \sqrt{0.8 \times 0.17} \approx 0.152. \)

---

\(^1\)The presence of mean reversion in stock returns is not universally accepted, however, and other studies argue against it; for instance, Richardson and Stock (1989), Kim, Nelson, and Startz (1991), and Richardson (1993).

\(^2\)In practical applications, the long-run variance of \( x_t \) can be estimated through various estimators. The long-run variance is equal to the sum of all autocovariances and can be estimated by essentially calculating a sample analogue of \( LR.Var(x_t) = \sum_{j=-\infty}^{\infty} Cov(x_t, x_{t+j}) \), as in the Newey and West (1987) estimator. The long-run variance is also equal to \( 2\pi f_{x_t}(0) \), where \( f_{x_t}(\cdot) \) is the spectrum, or spectral density, of \( x_t \), and one can form an estimator as the average periodogram (sample spectrum) across frequencies close to zero. Recent work by Müller and Watson, 2017, 2018, and Lazarus, Lewis, Stock, and Watson, 2016, suggest that long-run (or low-frequency) components of the data are most efficiently extracted by considering a small set of frequencies or basis components around the zero frequency.
Figure A1 illustrates the effects of serial dependence on the skewness of compound returns. The effects of $VR = 0.9$ and $VR = 0.8$ are compared to the benchmark iid case within each panel, and the volatility of the single-period returns is varied across the panels. The conclusions are similar to those obtained when looking at the effect of single-period skewness. When $\sigma$ is low, the effect of serial dependence on long-horizon skewness is small. When $\sigma$ is high, the effect of serial dependence can be sizable, but only in the range of extreme skewness levels, where interpretation of the different skewness values is not straightforward any more. To that extent, the effect of serial dependence is of second order importance compared to the effect of single-period return volatility.

2 Properties of skewness and quantile estimates

2.1 Skewness estimation

In this section, we explore the distribution of the skewness estimator

$$g \equiv \frac{\frac{1}{n} \sum_{i=1}^{n} (z_i - \bar{z})^3}{\left(\frac{1}{n} \sum_{i=1}^{n} (z_i - \bar{z})^2\right)^{3/2}}. \quad (14)$$

As discussed in the main text, Wilkins (1944) shows that there is an upper limit to the absolute value of $g$, which depends solely on the sample size $n$:

$$|g| \leq \frac{n - 2}{\sqrt{n - 1}}. \quad (15)$$

We focus on the estimation of skewness for long-horizon compound returns from individual stocks, because in this case, as we will see, the estimator $g$ becomes problematic. We start with a Monte Carlo simulation to show the finite sample distribution of the estimator. Later we also derive its asymptotic distribution.

For the simulation exercise, we assume that the monthly return, $x$, is log-normal and set the mean and volatility to $\mu = 1.01$ and $\sigma = 0.17$, respectively.\textsuperscript{3} The horizon is set to

\textsuperscript{3}We could have used any distributional assumption for the single-period returns. We chose the log-normal distribution so that the results are comparable with our discussion on quantiles in External Appendix 2.2. The conclusions are qualitatively the same if we use the normal distribution or a more
30 years, i.e., $T = 360$. For a given sample size $n$, we carry out the following simulation:

1. Simulate iid realizations of $x_t$ for $t = 1, ..., T$ and calculate the 30-year compound return using these monthly return realizations.

2. Repeat step (1) $n$ times to get a sample of compound returns (with sample size $n$), and estimate the skewness of the compound returns using the estimator $g$.

3. Repeat steps (1) and (2) 10,000 times to get a sample of the skewness estimates.

Panels A and B of Figure A2 show the distribution of the skewness estimates for two sample sizes, $n = 20,000$ and $n = 200,000$. The vertical line on each graph represents the upper limit of $g$ from equation (15). With the distributional assumptions on the single-period return used for the simulation, the skewness of the 30-year compound returns is $3.6 \times 10^6$ according to Proposition 1 of the main text. Therefore, the upper limit of the estimator $g$, which is 141.4 for $n = 20,000$ (Panel A of Figure A2) and 447.2 for $n = 200,000$ (Panel B of Figure A2), is clearly binding and the estimator $g$ is severely downward biased. This exercise highlights why simulation-based estimates of the skewness of long-horizon compound returns from individual stocks can be highly misleading.

In order for the upper limit on $g$ not to be binding and to possibly estimate a skewness of $3.6 \times 10^6$, a sample of $n \geq 1.13 \times 10^{13}$ would be needed. Since it is not feasible to provide simulation evidence at such a large sample size, we turn to asymptotic results. Let the $k$-th central moment of the variable $z$ be denoted by $\mu_k$, and its sample analogue by $m_k$, i.e.,

$$
\mu_k \equiv E \left[ (z - E[z])^k \right] \quad \text{and} \quad m_k \equiv \frac{1}{n} \sum_{i=1}^{n} (z_i - \bar{z})^k .
$$

Then the skewness of $z$, and its estimator from equation (14), $g$, are defined as

$$
Skew(z) = \frac{\mu_3}{\mu_2^{3/2}} \quad \text{and} \quad g = \frac{m_3}{m_2^{3/2}} .
$$

skewed distribution for the single-period returns.
Provided the third moment of $z$ exists, $m_2$ and $m_3$ trivially converge to $\mu_2$ and $\mu_3$, respectively, by a law of large numbers. Further, from Serfling (1980, page 72), as $n \to \infty$,

$$
\sqrt{n} \begin{bmatrix} m_2 - \mu_2 \\ m_3 - \mu_3 \end{bmatrix} \xrightarrow{d} N \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mu_4 - \mu_2^2 & \mu_5 - 4\mu_2\mu_3 \\ \mu_5 - 4\mu_2\mu_3 & \mu_6 - \mu_3^2 - 6\mu_2\mu_4 + 9\mu_2^3 \end{bmatrix}. \tag{18}
$$

By the delta method, and provided the sixth moment of $z$ exists, $g$ satisfies

$$
\sqrt{n} (g - \text{Skew}(z)) \xrightarrow{d} N \left( 0, \frac{1}{\mu_2} \left( \mu_6 - 6\mu_2\mu_4 + 9\mu_2^3 - \mu_3^2 \right) - 3\frac{\mu_3}{\mu_2} \left( \mu_5 - 4\mu_2\mu_3 \right) + \frac{9}{4\mu_2^2} \left( \mu_4 - \mu_2^2 \right) \right). \tag{19}
$$

That is, the skewness estimator $g$ is consistent and asymptotically normally distributed, with an asymptotic variance that is a function of the central moments up to order 6.\footnote{The asymptotic normality result in Serfling (1980) is derived for the $iid$ case, although the result should extend to more general cases as long as sufficient conditions for a central limit theorem apply.}

As is implied by Proposition 1 of the main text, the skewness (and other higher order moments) in compound returns can be extremely large at long horizons. The higher-order moments present in the variance formula for $g$ in equation (19) is therefore a warning sign that the variance of the skewness estimator (based on the asymptotic approximation) might be very large for long-horizon compound returns.

Panel A of Figure A3 shows two-standard error bounds of the skewness estimator, as a function of the horizon on which returns are compounded, for sample sizes $n = 20,000$, $n = 200,000$, and $n = 10^{16}$. The standard errors are calculated via the asymptotic approximation in equation (19) using the assumption of $iid$ log-normally distributed one-period returns with $\mu = 1.01$ and $\sigma = 0.17$.\footnote{The asymptotic approximation can be applied to any distribution for the single-period returns where the first 6 central moments are known. The log-normal distribution is used for illustration so that the results are comparable with our discussion on quantiles in Appendix 2.2. The conclusions are qualitatively the same if we use the normal distribution or a more skewed distribution for the single-period returns. When $x \sim LN(\psi, \eta^2)$, the non-central moments are given by $E[ x^k ] = e^{k\psi + \frac{1}{2}k^2\eta^2}$, and the central moments can be calculated from the non-central moments. The central moments for the product process $X_T$ can then be obtained via the formula in Proposition 1. Finally, using the central moments of $X_T$ in the variance formula from (19) provides the standard error estimates used in Panel A of Figure A3.} The two-standard error bounds on the skewness estimator widen very quickly with the horizon. As shown in the previous simulation exercise, the actual distribution of $g$ is very far from the asymptotic approximation for the
sample sizes \( n = 20,000 \) and \( n = 200,000 \) (see Panels A and B of Figure A2), so let us focus on the case when the sample size is much larger. For \( n = 10^{16} \), the upper limit from equation (15) is not binding and the estimator \( g \) should be mostly unbiased if its actual distribution is close to the asymptotic approximation in equation (19). However, as seen in Panel A of Figure A3, the skewness estimator is still essentially uninformative for horizons of 10 years or more due to the large standard errors of the estimator.

Overall, the results in this section show that in the case of individual stocks, direct estimation of the skewness, in returns compounded over 10 or more years, is next to meaningless. For practically relevant sample sizes, the estimator is severely downward biased, while for considerably larger sample sizes (if they were practically feasible), the enormous standard errors make the estimates highly unreliable. These results also imply that one cannot put much trust into simulation results concerning skewness of long-horizon returns.

### 2.2 Quantile estimation

The general formula for the sample quantile is given by

\[
\hat{q}_\alpha \equiv \inf \left\{ w : \frac{1}{n} \sum_{i=1}^{n} I \{ z_i \leq w \} \geq \alpha \right\}.
\]

(20)

The asymptotic distribution of \( \hat{q}_\alpha \), as \( n \to \infty \), is given by

\[
\sqrt{n} (\hat{q}_\alpha - q_\alpha) \xrightarrow{d} N \left( 0, \frac{\alpha (1 - \alpha)}{f_z^2 (q_\alpha)} \right),
\]

(21)

where \( f_z (\cdot) \) is the density of the random variable \( z \) (Serfling, 1980). Using the iid log-normal assumption on the single-period returns (with \( \mu = 1.01 \) and \( \sigma = 0.17 \)), Panels B to D in Figure A3 show quantiles of the distribution of compound returns as a function of the horizon, together with the two-standard error bounds of \( \hat{q}_\alpha \) for sample sizes \( n = 20,000 \) and \( n = 200,000 \). The standard errors are calculated using the asymptotic approximation in (21), and Panels B, C, and D correspond to \( \alpha \) levels of 0.9, 0.99, and 0.999, respectively. The results in Figure A3 reveal that the two-standard error bounds for the quantile
estimates are considerably narrower than in the case of the skewness estimator. In general, the standard error of the quantile estimator increases with the horizon, and is larger for quantiles far out into the tail. However, even for the 99.9-th percentile of the 30-year compound returns (Panel D), \( \hat{q}_\alpha \) is fairly precisely estimated with a sample size of \( n = 200,000 \), in large contrast to the skewness estimator in Panel A.

The results in Figure A3 show that based on the asymptotic approximation in equation (21), the quantiles of long-horizon compound returns can be fairly precisely estimated. However, as seen in the case of the skewness estimator, the asymptotic approximation can be far away from the actual distribution of the estimator. To show that this is not the case for the quantile estimates, we carry out the same simulation exercise as in the beginning of Section 2.1 to get the finite sample distribution of \( \hat{q}_\alpha \). The results are reported in Panels C to F of Figure A2 for the \( \alpha = 0.99 \) and \( \alpha = 0.999 \) quantiles (we would expect more discrepancies for quantiles far out in the tail). The histograms represent the distribution from the Monte Carlo simulation, while the solid lines show the corresponding asymptotic approximation from equation (21). The simulated finite sample distributions are close to their asymptotic counterparts, especially for the larger sample size (\( n = 200,000 \)).

3 Additional results for the binomial model

3.1 Derivation details for Appendix B.1 of the main text

The joint distribution of the random variables \( (L_{uu}, L_{ud}, L_{du}, L_{dd}) \) is a multinomial distribution and has the pmf

\[
P(L_{uu} = l_{uu}, L_{ud} = l_{ud}, L_{du} = l_{du}, L_{dd} = T - l_{uu} - l_{ud} - l_{du}) = \frac{T!}{l_{uu}!l_{ud}!l_{du}!(T - l_{uu} - l_{ud} - l_{du})!} \pi_{uu}^{l_{uu}} (\pi - \pi_{uu})^{l_{ud} + l_{du}} (1 - 2\pi + \pi_{uu})^{T - l_{uu} - l_{ud} - l_{du}}
\] (22)
Using (22),

\[ P(L_{ud} + l_{du} = l) = \sum_{l_{ud}=0}^{l} P(L_{ud} = l_{ud}, L_{du} = l - l_{ud}) \]

\[ = \frac{T!}{l! (T - l)!} (1 - 2\pi + 2\pi_{uu})^{T-l} \sum_{l_{ud}=0}^{l} \frac{l!}{l_{ud}!(l - l_{ud})!} (\pi - \pi_{uu})^{l_{ud}} (\pi_{uu})^{l - l_{ud}} \]

\[ = b(l; T, 2(\pi - \pi_{uu})) \]  

where the binomial expansion was used again to go from line 3 to 4. Using (23) again,

\[ P(L_{ud} = l_{ud} | L_{ud} + L_{du} = l) \]

\[ = \frac{P(L_{ud} = l_{ud}, L_{ud} + L_{du} = l)}{P(L_{ud} + L_{du} = l)} = \frac{P(L_{ud} = l_{ud}, L_{du} = l - l_{ud})}{P(L_{ud} + L_{du} = l)} \]

\[ = \frac{T!}{l_{ud}!(l - l_{ud})!(T - l)!} (\pi - \pi_{uu})^{l} (1 - 2\pi + 2\pi_{uu})^{T-l} \frac{l!}{l_{ud}!(l - l_{ud})!} (2\pi - 2\pi_{uu})^{l} (1 - 2\pi + 2\pi_{uu})^{T-l} \]

\[ = \frac{l!}{l_{ud}!(l - l_{ud})!} (\pi - \pi_{uu})^{l} \frac{l!}{l_{ud}!(l - l_{ud})!} 0.5^{l - l_{ud}} 0.5^{l_{ud}} \]

\[ = b(l_{ud}; l, 0.5) \]  

(25)
Using (23), it is also straightforward to show that

\[
P(L_u = 0, L_d = 0) = \frac{T!}{0!0!T!} (\pi - \pi_{uu})^0 (1 - 2\pi + 2\pi_{uu})^T = (1 - 2\pi + 2\pi_{uu})^T.
\] (26)

3.2 A portfolio beating its \textit{k-th} best constituent for \(N = 3\)

Let us assume there are \(N = 3\) stocks. The joint distribution of the single-period returns is described by the following table

<table>
<thead>
<tr>
<th>Probability</th>
<th>(x_{t1})</th>
<th>(x_{t2})</th>
<th>(x_{t3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\pi_{uuu}) = (\frac{1}{2}) (\pi_{uu})</td>
<td>(u)</td>
<td>(u)</td>
<td>(u)</td>
</tr>
<tr>
<td>(\pi_{uud}) = (\frac{1}{2}) (\pi_{uu})</td>
<td>(u)</td>
<td>(u)</td>
<td>(d)</td>
</tr>
<tr>
<td>(\pi_{udu}) = (\frac{1}{2}) (\pi_{uu})</td>
<td>(u)</td>
<td>(d)</td>
<td>(u)</td>
</tr>
<tr>
<td>(\pi_{udd}) = (\pi - \frac{3}{2}) (\pi_{uu})</td>
<td>(u)</td>
<td>(d)</td>
<td>(d)</td>
</tr>
<tr>
<td>(\pi_{duu}) = (\frac{1}{2}) (\pi_{uu})</td>
<td>(d)</td>
<td>(u)</td>
<td>(u)</td>
</tr>
<tr>
<td>(\pi_{dud}) = (\pi - \frac{3}{2}) (\pi_{uu})</td>
<td>(d)</td>
<td>(u)</td>
<td>(d)</td>
</tr>
<tr>
<td>(\pi_{ddu}) = (\pi - \frac{3}{2}) (\pi_{uu})</td>
<td>(d)</td>
<td>(d)</td>
<td>(u)</td>
</tr>
<tr>
<td>(\pi_{ddd}) = (1 - 3\pi + \frac{5}{2}) (\pi_{uu})</td>
<td>(d)</td>
<td>(d)</td>
<td>(d)</td>
</tr>
</tbody>
</table>

The returns on all three stocks are identically distributed and have the same distribution as before (\(u\) with probability \(\pi\) and \(d\) with probability \(1 - \pi\)). The pair-wise return correlation is the same as in the \(N = 2\) case, i.e.,

\[
Corr(x_{ti}, x_{tj}) = \frac{(\pi_{uu} - \pi^2)}{\pi (1 - \pi)}.
\] (27)

Consider \(T\) periods and assume that the joint distribution of \((x_{t1}, x_{t2}, x_{t3})\) is \textit{iid} across time for \(t = 1, ..., T\). Let us introduce the following random variables: \(L_{uuu}\) is the number of periods where \(x_{t1} = x_{t2} = x_{t3} = u\), \(L_{uud}\) is the number of periods where \(x_{t1} = x_{t2} = u\) and \(x_{t3} = d\), etc. Note that \(L_{uuu} + L_{uud} + L_{uda} + L_{udd} + L_{duu} + L_{ddu} + L_{ddd} = T\).
The compound return on the three stocks and on the equal-weighted portfolio are

\[
X_{T1} = u^{L_{uuu} + L_{uud} + L_{udu} + L_{ddu} + L_{ddu} + L_{ddd}}
\]

\[
X_{T2} = u^{L_{uuu} + L_{uud} + L_{udu} + L_{ddu} + L_{ddu} + L_{ddd}}
\]

\[
X_{T3} = u^{L_{uuu} + L_{uud} + L_{udu} + L_{ddu} + L_{ddu} + L_{ddd}}
\]

\[
X_{Tp} = u^{L_{uuu} \left( \frac{2u + d}{3} \right) L_{uud} + L_{udu} + L_{ddu} \left( \frac{u + 2d}{3} \right) L_{udd} + L_{udd} + L_{ddd}}
\]

(28)

Denote the random vector \( L \equiv [L_{uud} L_{udu} L_{udd} L_{duu} L_{dud} L_{ddd}]^\top \) and the vector \( l \equiv [l_{uud} l_{udu} l_{udd} l_{duu} l_{dud} l_{ddd}]^\top \) (note that the “uuu” and “ddd” cases are omitted). It is rather straightforward to show that for a given value of the vector \( l \),

\[
X_{Tp} > X_{T1} \iff a_1^\top l > 0
\]

\[
X_{Tp} > X_{T2} \iff a_2^\top l > 0
\]

\[
X_{Tp} > X_{T3} \iff a_3^\top l > 0
\]

(29)

with

\[
a_1^\top \equiv \begin{bmatrix}
\log \left( \frac{2u + d}{3u} \right) & \log \left( \frac{2u + d}{3u} \right) & \log \left( \frac{u + 2d}{3u} \right) & \log \left( \frac{2u + d}{3u} \right) & \log \left( \frac{u + 2d}{3u} \right) & \log \left( \frac{u + 2d}{3u} \right)
\end{bmatrix}
\]

\[
a_2^\top \equiv \begin{bmatrix}
\log \left( \frac{2u + d}{3u} \right) & \log \left( \frac{2u + d}{3u} \right) & \log \left( \frac{u + 2d}{3u} \right) & \log \left( \frac{2u + d}{3u} \right) & \log \left( \frac{u + 2d}{3u} \right) & \log \left( \frac{u + 2d}{3u} \right)
\end{bmatrix}
\]

\[
a_3^\top \equiv \begin{bmatrix}
\log \left( \frac{2u + d}{3d} \right) & \log \left( \frac{2u + d}{3d} \right) & \log \left( \frac{u + 2d}{3d} \right) & \log \left( \frac{2u + d}{3d} \right) & \log \left( \frac{u + 2d}{3d} \right) & \log \left( \frac{u + 2d}{3d} \right)
\end{bmatrix}
\]

(30)
The joint distribution of \([L_{uuu} L_{uud} L_{udu} L_{udd} L_{duu} L_{dud} L_{ddu} L_{ddd}]\) is a multinomial distribution, and therefore for any \(l\) such that \(\mathbf{1}^\top l \leq T\),

\[
P (L = l) = \sum_{l_{uuu}=0}^{T - \mathbf{1}^\top l} P (l_{uuu}, l_{uud}, l_{udu}, l_{udd}, l_{duu}, l_{dud}, l_{ddu}, T - \mathbf{1}^\top l - l_{uuu})
= \sum_{l_{uuu}=0}^{T - \mathbf{1}^\top l} \frac{T!}{l_{uuu}! l_{uud}! l_{udu}! l_{udd}! l_{duu}! l_{dud}! l_{ddu}! (T - \mathbf{1}^\top l - l_{uuu})!} \times \prod_{\nu=1}^{T-1} \prod_{\alpha=0}^{l_{\nu}} \pi_{\nu} \prod_{\gamma=0}^{l_{\gamma}} \pi_{\gamma} T - \mathbf{1}^\top l - l_{uuu},
= \frac{T!}{l_{uud}! l_{udu}! l_{udd}! l_{duu}! l_{dud}! l_{ddu}! (T - \mathbf{1}^\top l)!} \prod_{\nu=1}^{T-1} \prod_{\alpha=0}^{l_{\nu}} \pi_{\nu} \prod_{\gamma=0}^{l_{\gamma}} \pi_{\gamma} T - \mathbf{1}^\top l - l_{uuu},
= \frac{T!}{l_{uud}! l_{udu}! l_{udd}! l_{duu}! l_{dud}! l_{ddu}! (T - \mathbf{1}^\top l)!} \prod_{\nu=1}^{T-1} \prod_{\alpha=0}^{l_{\nu}} \pi_{\nu} \prod_{\gamma=0}^{l_{\gamma}} \pi_{\gamma} T - \mathbf{1}^\top l - l_{uuu},
\]

where \(\mathbf{1}\) denotes an appropriately sized vector of ones.

The probability that the portfolio beats the \(k\)-th best constituent is

\[
P (X_{T_p} > X_T (k)) = \sum_{\mathbf{1} \geq 0, \mathbf{1}^\top \mathbf{1} \leq T} P (L = l) I (X_{T_p} > X_T (k) \mid L = l), \quad \text{(32)}
\]

where \(I ()\) is the indicator function and the summation goes over all possible values of the vector \(l\) such that \(l \geq 0\) (every element of the vector is non-negative) and \(\mathbf{1}^\top l \leq T\) (the sum of the elements is not greater than \(T\)). The formula for \(P (L = l)\) is given in equation (31), and using (29),

\[
I (X_{T_p} > X_T (k) \mid L = l) = I ([I (a_1^\top l > 0) + I (a_2^\top l > 0) + I (a_3^\top l > 0)] = 4 - k) , \quad \text{(33)}
\]
3.3 Additional simulation results

3.3.1 Additional results to Figure 5 of the main text

The results in Figure 5 of the main text are based on the binomial model, where the single-period return, $x_{ti}$, can take only two values $u = 1.18$ and $d = 0.84$ with equal probability. We now provide simulation evidence that the conclusions regarding the behavior of long-run compound returns do not hinge on this assumption about the single period return.

We redo the exact same simulation exercise as the one described in Appendix C.3 of the main text, with the only exception that the single period return is normally distributed instead, with its mean and variance being the same as in the main text. In particular,

$$x_{ti} \sim N(1.01, 0.17^2). \quad (34)$$

Figure A4 shows the probability that the portfolio compound return is higher than compound return on constituent stocks as a function of the investment horizon, and compares the results from the binomial and normal models. In particular, the graphs on the left are exactly the ones from Figure 5 of the main text (based on analytical formulas for $N = 2$ and $N = 3$, and on simulations for $N = 50$). The graphs on the right in Figure A4 show the corresponding results when the assumption in equation (34) is used instead (the results are based on simulations). As it can be seen, the corresponding results are practically identical for horizons longer than 5 years ($T > 60$).

3.3.2 Additional results to Figure 6 of the main text

The results in Figure 6 of the main text are based on the analytical formula derived in Appendix C.2 of the main text. When deriving the analytical formula, we make the approximating assumptions that (i) $X_{T_p}$ is log-normal, and (ii) $X_{T_p}$ and $X^*_k$ are independent. Figure A5 recreates Figure 6 from the main text by showing the probability that an equal-weighted portfolio of 50 stocks beats the $k$-th best from 1,000 stocks over a 30-year investment horizon. The solid line corresponds to our analytical formula (i.e., it is exactly the same line as the one in Figure 6 of the main text). The dashed line
shows simulation results that do not use the approximating assumptions and where the single-period return on all 1,000 stocks can take only two values $u = 1.18$ and $d = 0.84$ with equal probability. The dotted line shows simulation results that do not use the approximating assumptions and where the normal distribution in equation (34) is used instead for all single-period returns. The three lines are very close to each other, showing that (i) the approximating assumptions used for the analytical formula are fairly accurate, and (ii) the results do not hinge on the particular choice of the single-period return distribution as long as its moments are fixed.
References


Figure A1: Skewness of compound returns - the effect of serial correlation

The graphs show the skewness of compound returns as a function of the compounding horizon, $T$, when single-period returns might be serially correlated. The values are calculated using equation (2). The single-period gross return is assumed to be log-normal with mean $\mu = 1.01$ and standard deviation $\sigma$ that varies across the panels (see above each panel). The lines correspond to cases where single-period returns are independent across time ($iid$), or are serially correlated ($VR = 0.9$ or 0.8). $VR$ is the ratio between the long-run and short-run variance of the single-period return, defined via equation (12).

A. $\sigma = 0.05$

B. $\sigma = 0.08$

C. $\sigma = 0.11$

D. $\sigma = 0.17$
The histograms in these graphs represent the distribution of skewness estimates (Panels A and B) and quantile estimates (Panels C to F) of 30-year ($T = 360$) compound returns from individual stocks. The single-period (monthly) gross returns are assumed to be iid log-normal with mean $\mu = 1.01$ and volatility $\sigma = 0.17$. The histograms are the results of Monte Carlo simulations described in External Appendix 2.1. The distribution of the skewness and quantile estimates correspond to two sample sizes, $n = 20,000$ (panels on the left) and $n = 200,000$ (panels on the right). The vertical lines in the two top graphs correspond to the upper limit on the skewness estimator in equation (15). The curves in Panels C to F are based on the asymptotic distribution of the quantile estimates in equation (21).
Figure A3: Asymptotic two-standard error bounds for skewness and quantile estimators

The graphs show two-standard error bounds of the skewness estimator (in Panel A) and quantile estimator (in Panels B to D for various $\alpha$-quantiles), as a function of the horizon over which the underlying returns are compounded, for different sample sizes (see legends). In all the graphs, the solid line shows the skewness or quantile of the compound returns. The other lines show the two-standard error bounds, for which the standard errors are calculated via the asymptotic approximation (equation (19) for skewness and equation (21) for the quantiles). It is assumed that single period returns are iid log-normally distributed with mean $\mu = 1.01$ and volatility $\sigma = 0.17$. 

A. Skewness

B. Quantile, $\alpha = 0.9$

C. Quantile, $\alpha = 0.99$

D. Quantile, $\alpha = 0.999$
Figure A4: Probability of the portfolio beating its constituents

The graphs on the left are exactly the same as the ones in Figure 5 of the main text (see the caption there for details), where the single-period (monthly) gross return on stock $i$ in period $t$, $x_{it}$, can take two values, $u = 1.18$ or $d = 0.84$ with equal probability. The graphs on the right show corresponding results, based on simulations, when $x_{it} \sim N (1.01, 0.17^2)$ instead.
Figure A5: Probability of the portfolio beating the $k$-th best stock on the market

The figure shows the probability that the total compound return on the equal-weighted portfolio of 50 stocks, $X_{T_p}$, is higher than the compound return on the $k$-th best performing stock, $X^*_T(k)$, out of 1,000 identical stocks. The line labeled “Binomial-analytical” is based on our approximate analytical formula, and is exactly the same as the one in Figure 6 of the main text (see the caption there for details). The line labeled “Binomial-simulated” shows simulation results that do not use the approximating assumptions and where the single-period return on all stocks can take only two values $u = 1.18$ and $d = 0.84$ with equal probability. The line labeled “Normal-simulated” shows simulation results that do not use the approximating assumptions and where the distribution in equation (34) is used instead for all single-period returns.