Deciphering drawdown

When selecting a hedge fund manager, one risk measure investors often consider is drawdown. How should drawdown distributions look? Carr Futures' Galen Burghardt, Ryan Duncan and Lianyan Liu share some insights from their research to show investors how to begin to answer this tricky question

Despite peak-to-trough drawdown being a widely quoted measure of risk for hedge funds and commodity trading advisers (CTAs), investors don't appear to have a widely accepted way of forming expectations about just how much managers might lose.

Drawdown measures the change in the value of a portfolio from a defined peak to a subsequent trough. Investors tend to monitor a manager's worst or maximum drawdown, with only informal or anecdotal information about the manager's average annual or previous year's returns. We will show that it is possible to get a reasonable fix on what drawdown distributions should look like. This, in turn, allows a manager's track record to be scrutinised to see if their drawdown history is reasonable. Moreover, knowledge of drawdown distributions allows investors to estimate the magnitude and frequency of future drawdowns.

Our analysis shows that the three most important determinants of drawdowns are length of track record, mean return and volatility of returns. The acid test, we believe, is that our simulated drawdown distributions do a very good job of explaining the kinds of drawdown patterns that CTAs have exhibited over the past 10 years.

In practice, a drawdown is defined as the percentage change in a manager's net asset value (NAV) from a high water mark to the next low water mark. An NAV qualifies as a high water mark if it is higher than any previous NAV, and is followed by a loss. It qualifies as a low water mark if it is the lowest NAV between two high water marks. Or, if one is at the end of a data series, a low water mark is simply the lowest NAV following the last high water mark.

What should drawdowns look like?

Realised drawdowns are the result of sequences of returns and depend entirely on the paths that a manager's NAV can follow. So, the only practical way to discover what drawdowns should look like is to simulate as many NAV paths as needed to produce reasonable looking distributions. We shall use Monte Carlo simulations in which we have controlled for length of track record, the distribution of returns, and de-leveraging when in drawdown.

The resulting drawdown distributions have two basic shapes. For a given return distribution and length of track record, the frequency and size of a manager's entire collection of drawdowns will look like the distribution shown in figure 1. Drawdowns in all exhibits are presented as negative percentage changes. So, for example, figure 1 displays a high frequency of small drawdowns and a small frequency of large drawdowns.

Though any given manager can only have one worst drawdown, it still makes sense to think of the distribution from which that worst or maximum drawdown was drawn. Or, if we think of several managers, all of whom have the same or very similar track records and return characteristics, we can think about what the distribution of their various worst drawdowns should look like. An example of what the distribution of maximum drawdowns should look like is provided in figure 2.

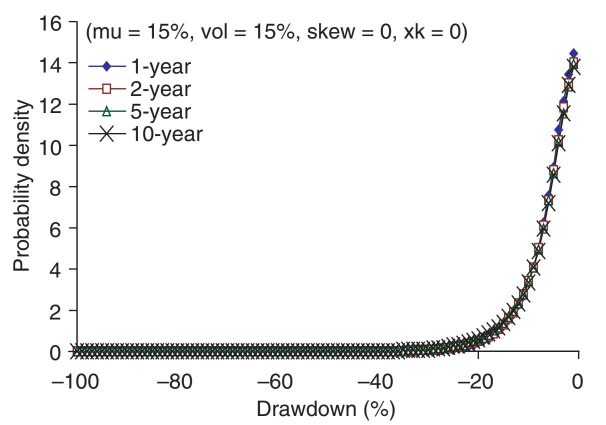
In our simulations, we were able to control the return-generating process for most of the factors that would seem to make sense. In particular, we controlled for: length of track record; mean return; volatility of returns; skewness; kurtosis; and de-leveraging when in drawdown. Of these, the only three that have any empirical importance seem to be length of track record, mean return and the volatility of returns. The rest tend not to matter much, in some cases because the effect of a change in the variable is small and in others because the range of the variable is small.

From figure 1, we can see that length of track record matters very little to the distribution of all drawdowns. In other words, the likelihood of experiencing a drawdown of any given size is largely independent of how long a manager is in business.

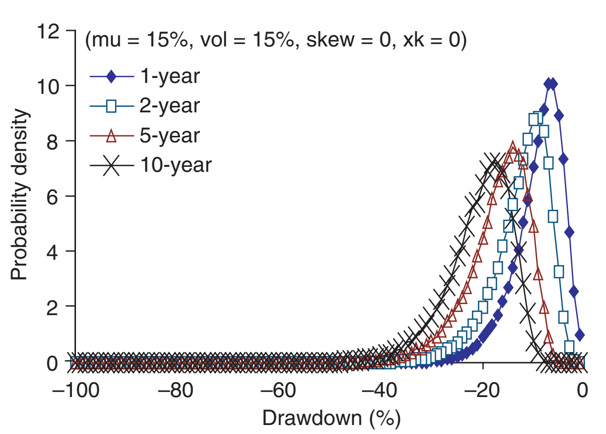
Though not presented here for reasons of brevity, we found that volatility of returns and the mean return matter a lot. As one would expect, higher mean returns lead to smaller expected drawdowns. Volatility of returns also has a large influence over a manager's drawdowns. Higher volatility leads to larger expected drawdowns.

Skewness and kurtosis, on the other hand, matter very little, at least given the range of values for skewness and kurtosis that we have observed in CTA returns over the past 10 years. The most plausible reason for this seems to be that drawdowns are the result of adding together

1. The distribution of all drawdowns



2. The distribution of maximum drawdowns



sequences of returns. As a result, even though the distribution from which any given return is drawn may be highly skewed or exhibit fat tails, the result of adding returns together produces a random variable that tends – *à la* the central limit theorem – to be more normally distributed.

Distribution of maximum drawdowns

As mentioned, the likelihood of any given drawdown is independent of how long a manager is in business. But the likelihood of experiencing a drawdown that is larger than anything experienced so far increases with every passing day. As a result, as shown in figure 2, increases in the length of track record shift the entire maximum drawdown distribution to the left.

We also found (but do not show here for reasons of brevity) that maximum drawdown distributions are highly sensitive to the kinds of differences in the mean and standard deviation of returns that we observe across managers. In contrast, observed differences in the skewness and kurtosis of returns hardly matter at all.

To put these things in perspective, we calculated the partial effect of each variable (mean, volatility, skew and kurtosis) and multiplied the partials by the standard deviation of each variable as measured in our database. We ended up with something like the following: mean return = 19%; volatility = 51%; skewnesss = 6%; kurtosis = 1%. Thus, differences in volatility are more than twice as important as differences in mean returns, and mean returns are more than three times as important as skewness, with kurtosis way down the list.

If higher returns produce smaller drawdowns while higher volatilities produce larger drawdowns, then one can trade off one for the other to produce the same expected drawdowns. But given the sizes of their respect effects, it can take a lot of extra return to make up for a little extra volatility.

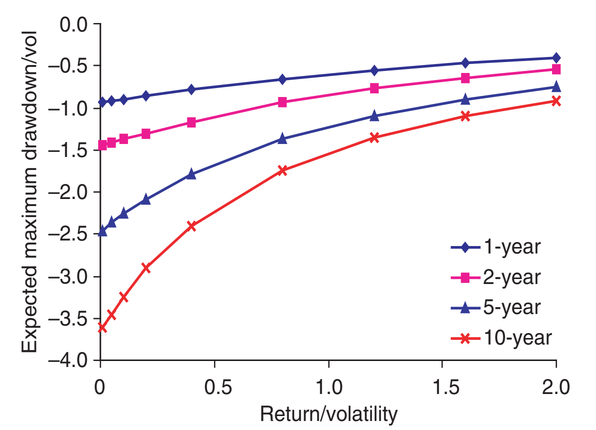
Even though there is no clean, analytical function that relates drawdowns to a manager's returns, the relationship between a manager's drawdowns and his returns and risks can be described by $DD/\sigma = f(\mu/\sigma)$, where DD denotes drawdown, σ is the standard deviation of returns and μ is the mean return. So, when divided by the volatility of returns, a manager's drawdowns can be cast as a function of the manager's modified Sharpe ratio, that is, the ratio of mean return to the standard deviation of returns.

The shape of this function is illustrated in figure 3 for track records ranging from one to 10 years. The curvature bears out our sense that volatility matters more than mean return. A doubling of a manager's mean return while holding return volatility constant will reduce expected drawdown per unit of volatility, but by less than half. In turn, a doubling of volatility while holding mean return constant will more than double expected maximum drawdown per unit of volatility.

If instead, we are concerned only about the sizes of drawdowns, this relationship can be rewritten as $DD = \sigma f(\mu/\sigma)$. This suggests the following conclusions. A doubling of both mean return and volatility – which would leave the modified Sharpe ratio unchanged – will exactly double expected maximum drawdowns. Also, a doubling of volatility alone will more than double expected maximum drawdowns. Finally, the mean return would have to be more than doubled to compensate for a doubling of volatility.

An important corollary is then that two managers with the same volatility of returns will have different expected drawdowns, if their mean returns are different.

3. Relationship between maximum drawdowns and returns when normalised for volatility

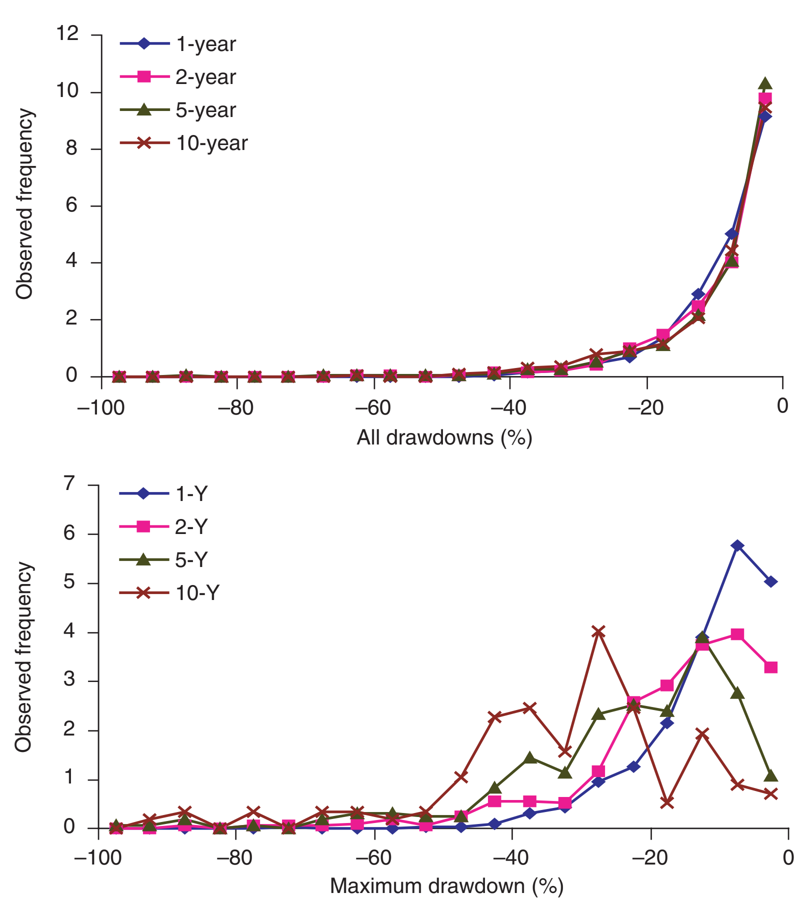


Similarly, two managers with identical modified Sharpe ratios will have different expected drawdowns if their return volatilities are different.

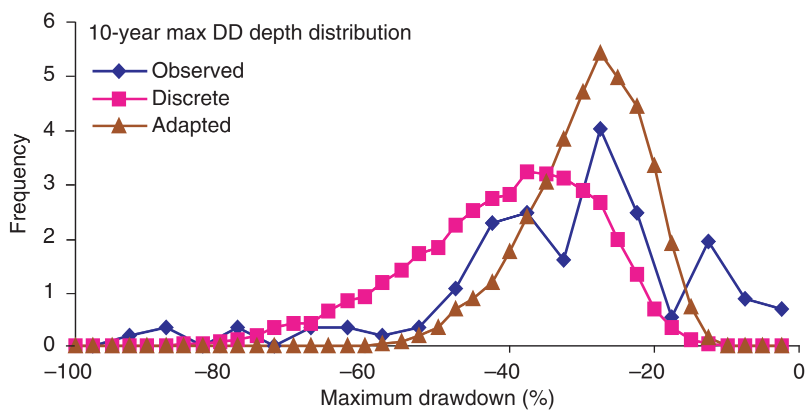
Empirical distributions

To see whether this approach could be used to explain the drawdown patterns we observe, we constructed drawdown histories for one, two, five and 10 years

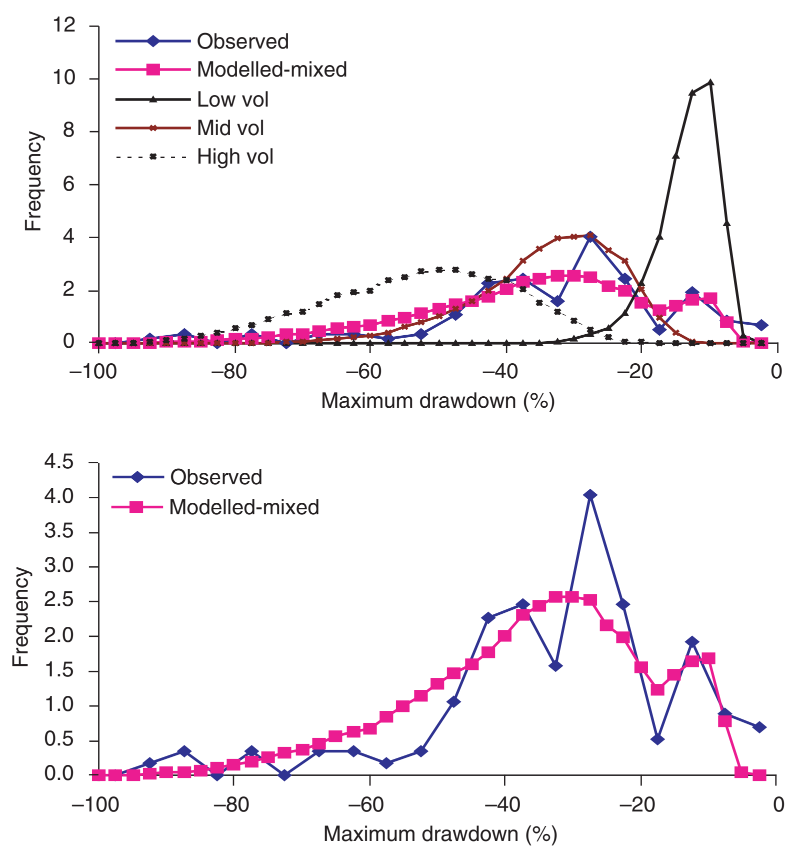
4. Observed drawdown distributions



5. First cut at explaining observed maximum drawdowns



6. A composite maximum drawdown distribution



in the following way using CTA returns from the Barclays database. Using return histories for all managers with a one-year track record as of November 2002, we determined what their drawdowns would have been had they all started from scratch at the end of November 2001. Then, for all managers who had a two-year track record as of November 2002, we determined what their drawdowns would have been had they all started fresh at the end of November 2000. And so forth for five-year and 10-year track records. By design, this approach produces different drawdown histories than those actually reported by the CTAs in the database. The advantage, however, is that it puts all managers up against the same market conditions. The results of these efforts are shown in figure 4. The distributions of all drawdowns are shown in the upper panel, while the distributions of maximum drawdowns are shown in the lower panel.

Reconciling theoretical and empirical distributions

The distributions of all drawdowns shown in figure 4 are largely as expected. But the distributions of maximum drawdowns seem puzzling at first glance. They are irregularly shaped. Also, as shown in figure 5, the observed distribution does not fit well with the theoretical distribution – which is labelled discrete – derived from the actual distribution of returns for all CTAs with a 10-year track record.

The problem is that the observed drawdown distribution peaks at a much lower level of drawdowns than the theoretical does. One plausible explanation for the different shapes is that managers may de-leverage when they are in drawdown – thereby avoiding future larger drawdowns that could be experienced if they were to keep the volatility of returns constant.

Figure 5 shows it is possible to pull the theoretical drawdown distribution to the right by allowing for de-leveraging when simulating returns. In this case, we scaled the manager's mean and volatility of returns as $\mu' = f \times \mu$ and $\sigma' = f \times \sigma$, where $f = 1 - [abs(drawdown)]^{1/2}$. For example, if a manager's current drawdown were 50%, the scaling factor would be 0.29 [= $1 - 0.5^{1/2}$]. The resultant new distribution of maximum drawdowns is labelled 'adapted' in figure 5 and peaks just about where it should.

One drawback to this approach is that the evidence on de-leveraging is largely anecdotal. We know managers who attest to the fact that they scale back risk when in drawdown. But we also know other managers who say that they do not. We have not yet been able to find any evidence in the volatilities of managers' returns that suggests they de-leverage when in drawdown. Also, while this approach produces a mass of probabilities that looks like what we observe, it greatly under-predicts the several large drawdowns we observe in the data.

A better approach is to divide managers into one of three groups: low volatility (0% to 12.5%), medium volatility (12.5% to 25%) and high volatility (25% to 50%). Using the group returns and group volatilities, we simulated the three maximum drawdown distribution shown in the upper panel of figure 6. Then, using the numbers of managers in each of the three groups, we produced a composite distribution that is a weighted average of the three separate distributions.

The resulting composite distribution

has some attractive features. First, because of the presence of the low volatility group, the composite distribution peaks about where it should. Second, because of the presence of the high volatility group, the composite distribution allows for a sufficiently high probability of large drawdowns. And third, as shown in the lower panel of figure 6, the composite even exhibits some of the irregular shape we see in the observed distribution of drawdowns.

At this point, we think it is reasonable to draw two theoretical drawdown distributions for any given manager, both based on length of track record and the mean and volatility of returns. We have done this in figure 7 for a manager with a 10-year track record, a mean return of just over 12% and a standard deviation of returns of 20%. Over this, we have superimposed the manager's actual drawdowns, which are represented by the horizontal lines stemming from the vertical axis on the right. This particular manager has experienced 17 drawdowns over the 10 years, most of them less than 10%. The maximum drawdown was just over 40%.

Overall, this manager's actual drawdown experience is roughly in line with what we would expect. The maximum drawdown is in the upper end of the theoretical distribution, but appears to be only about one standard deviation above the mean.

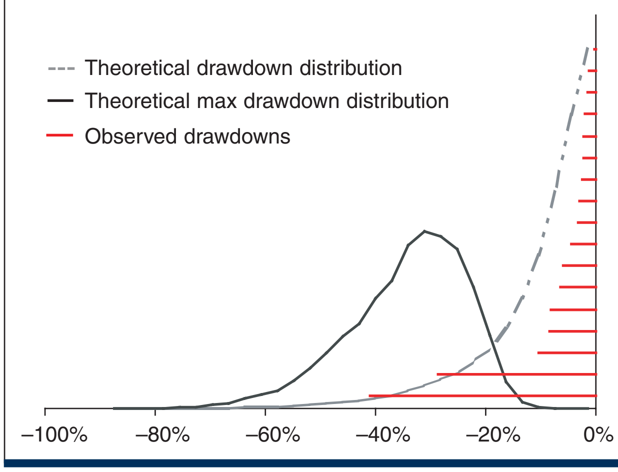
What about future drawdowns?

What kinds of drawdowns might an investor expect in the future? This work suggests we can form reasonable expectations about the size and frequency of drawdowns for any given investment horizon. We can also say something useful about the possibility that a manager will experience a larger drawdown than the maximum drawdown to date. In particular, for a given investment horizon and assumptions about the mean and volatility of returns, we can calculate the likelihood that a manager will experience a worse drawdown and a conditional maximum drawdown to go with it.

For example, how likely is it that a manager whose worst drawdown to date is 41% will have a worse drawdown over any given investment horizon? Assuming a mean return and volatility (12% and 20%, respectively), we find the probability of experiencing a drawdown greater than 41% is only 0.1% over the next year but would be 23.4% if the investment horizon is extended to 10 years. We also find that the expected value of this worse drawdown would be 44.1% if it occurs in the next year but would be 49.0% if experienced over a 10-year horizon.

In practice, we can use any target drawdown, not just the worst or maximum drawdown to date. And we can, if need be, modify the assumptions about the manager's returns to produce more realistic theoretical distributions. This would be especially useful if we think a manager's trading strategy is likely to produce a mix of volatilities over time. The drawdown distributions for high and low return volatilities have very different shapes and could produce very different probabilities of large losses than one would get using an assumption of constant volatility.

7. Assessing a manager's drawdown experience



Further research

Our conversations with clients and colleagues about this work have raised several questions and issues related to serial correlation of returns, reliability of volatility estimates, frequency of return data, applicability to hedge funds and variation of return volatility changes in response to drawdowns.

Our preliminary work on these suggests the following. First, serially correlated returns could have a measurable effect on drawdown distributions, but we have found no evidence of serial correlation in CTA returns. Volatility estimates based on monthly return data can be subject to very large statistical errors and would be much improved, at least in the case of CTAs, if we had daily return data. To the extent one can get reliable return and volatility information about hedge funds, the analysis should work well. It is much harder, though, to get the same quality information about hedge funds as one can get for CTAs. And, while we know that some managers trade differently when in drawdown, the evidence on CTAs as a class is ambiguous. □

Technical note

To simulate net asset value (NAV) series where skewness and kurtosis are zero, we draw sample returns from a lognormal return distribution. To capture skewness and kurtosis, we sample returns from a generalised lambda distribution. The values of skewness and excess kurtosis used were roughly consistent with the range of values we observed for commodity trading advisers in our database. The NAV series is constructed from the return series. The simulated drawdowns are then derived and used to produce the theoretical drawdown distributions. A typical run usually requires 10,000 iterations to produce a smooth distribution. □

Related reading

Emmanuel Acar and Shane James, *Maximum loss and maximum drawdown in financial markets*, http://www.btinternet.com/~emmanuel.acar/Mldwdn.pdf, March 1997, for a discussion of the form of the drawdown and max loss functions

Zvi Bodie, Alex Kane and Alan Marcus, chapter 24 (Portfolio performance evaluation) of *Investments*, fourth edition, 1999, Irwin/McGraw-Hill, for a discussion of how changes in portfolio mix affect return/risk measures

Alan Marcus, *The Magellan Fund and market efficiency: here's how to assess the performance of money managers*, Journal of Portfolio Management, autumn 1990, for a discussion of what the distribution of the winner's returns should look like

Galen Burghardt, Ryan Duncan and Lianyan Liu are members of the research team at Carr Futures, Crédit Agricole Indosuez's Chicago-based brokerage unit. An extended version of this paper is available via email from gburghardt@carrfutures.com