Alternative Investment Analyst Review

WHAT A CAIA MEMBER SHOULD KNOW
Understanding Drawdowns
Galen Burkhardt, Ryan Duncan, and Lianyan Liu

CAIA SUBMISSION
Liquid Real Assets
Anton Loukine, CAIA

RESEARCH REVIEW
Crude Oil Price Forecasting Techniques: a Comprehensive Review of Literature
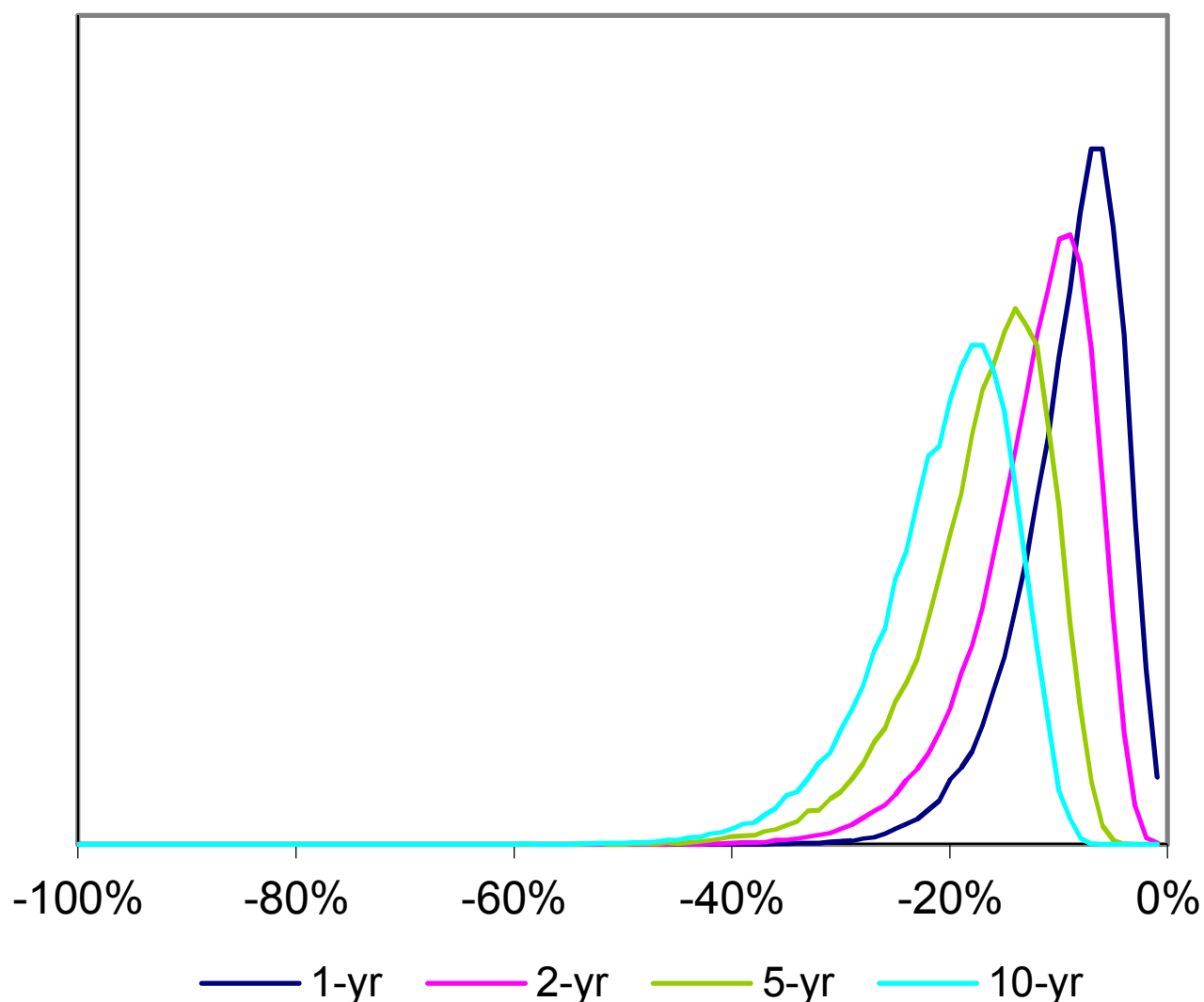*Niaz Bashiri Behmiri and Jose Ramos Pires Manso*

TRADING STRATEGIES
Alternative Inflation Hedging Portfolio Strategies: Going Forward Under Immoderate Macroeconomics
Nicolas Fulli-Lemaire

CAIA MEMBER PERSPECTIVE
The Next Wave of Futurization
Kathryn M. Kaminski, CAIA

Understanding Drawdowns

Galen Burkhardt
Senior Advisor, Newedge USA, LLC.

Ryan Duncan
Global Co-Head, Newedge Alternative Investment Solutions' Advisory Group

Lianyan Liu
Quantitative Analyst, Newedge Alternative Investment Solutions' Advisory Group

1. Introduction

Within the universe of hedge funds and commodity trading advisers (CTAs), one of the most widely quoted measures of risk is peak-to-trough drawdown. Our experience suggests that investors do not have a widely accepted way of forming expectations about just how much managers who are in business over long periods of time might be expected to lose. Rather, we find that 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. Drawdown as a measure of risk has failed to attract the same kind of research and attention that is devoted to other common measures, such as return volatility, VaR, or Sharpe ratios.

Our purpose in this article is to illustrate that it is possible to get a reasonable fix on what drawdown distributions should look like. This is not a trivial problem. Any manager for whom the standard deviation of returns is large enough to produce a loss in any given investment period will experience drawdowns. Most managers are in drawdown most of the time. And managers who have been in business a long time may well have experienced more and bigger drawdowns than those with short track records.

If it is possible to predict how drawdowns should behave, then we can address two important kinds of questions.

Looking back over a manager's track record, does his drawdown history make sense? That is, does the frequency and size of his drawdowns look reasonable and does his maximum drawdown accord with what we would expect?

Looking forward, what kinds of drawdowns should we expect over any given investment horizon? How many drawdowns should he experience? How big? How likely is it that his largest drawdown going forward will be greater than his maximum drawdown so far? And, if it is bigger, how much bigger?

What we show in this article is that the three most important determinants of drawdowns are length of track record, mean return, and volatility of returns. The acid test, we think, 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.

2. Drawdown Defined

Drawdown measures the change in the value of a portfolio from any newly established peak (or high water mark) to a subsequent trough (or low water mark).
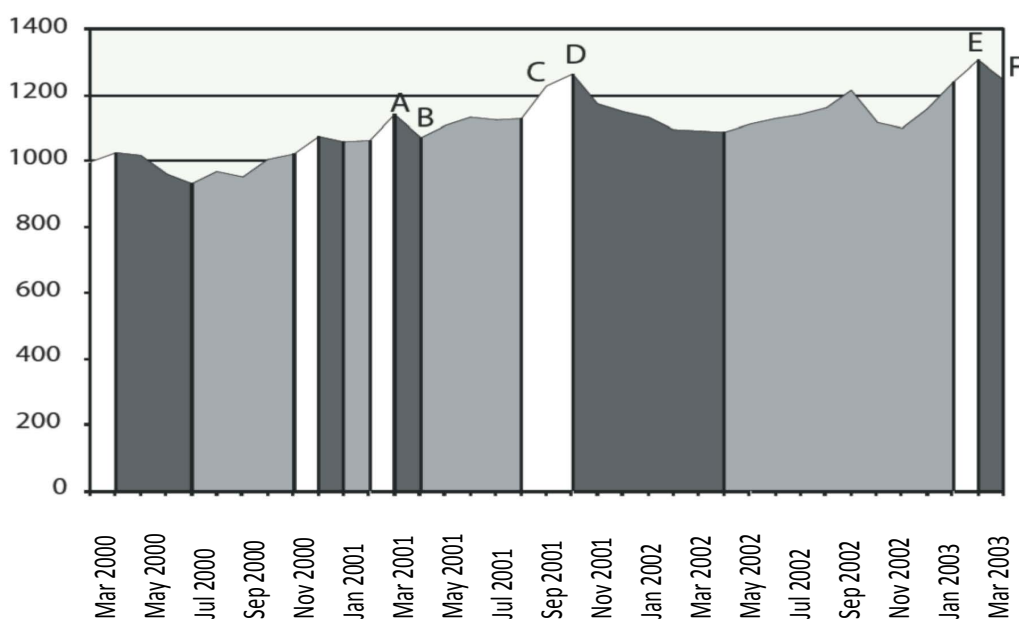

Exhibit 1 Net Asset Value History with Sample High And Low Water Marks
Source: Barclay Hedge, Newedge Prime Brokerage Research

One of the things that make drawdowns interesting is that they depend so much on the sequence of a manager's returns. The usual summary statistics such as mean and volatility of returns reflect nothing of the sequence in which the returns occur. The sequence is critical, however, for drawdowns. Two managers with identical means and volatilities of returns can experience very different drawdowns.

In practice, a drawdown is defined as the percent change in a manager's net asset value from a high water mark to the next low water mark. A net asset value qualifies as a high water mark if it is higher than any previous net asset value and if it is followed by a loss. Thus, points A and D in Exhibit 1 are high water marks, but point C is not, even though net asset value at that point is higher than it has ever been. Point C is not a high water mark because it is followed by a gain.

A net asset value qualifies as a low water mark if it is the lowest net asset value between two high water marks. Point B qualifies as a low water mark. Or, if one is at the end of a data series, a low water mark is simply the lowest net asset value following the last high water mark. For example, point F, which follows the newly established high water mark at E, would be a low water mark for the purpose of calculating drawdowns even though the manager's net asset value has not yet reached a new high water mark.

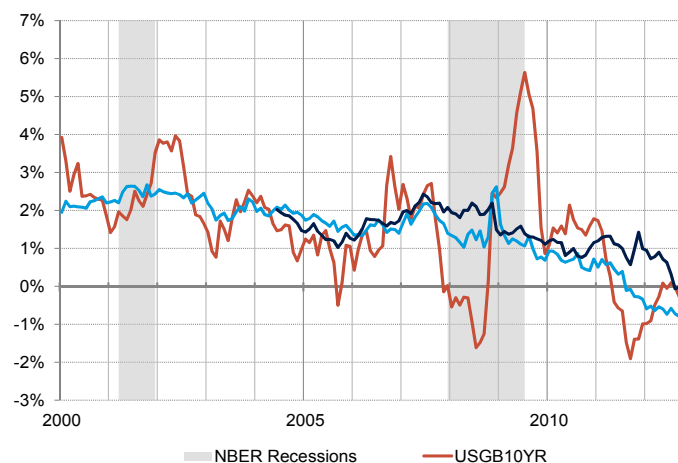A manager's maximum drawdown is simply the largest of these drawdowns.

Exhibit 2 The Distribution of All Drawdowns

3. What Should Drawdowns Look Like?

Because realized drawdowns are the result of sequences of returns and depend entirely on the paths that a manager's net asset value can follow, the only practical way to discover what drawdowns should look like is to simulate as many net asset value paths as one needs to produce reasonable looking distributions. In what follows, we have used Monte Carlo simulations in which we have controlled for the length of track
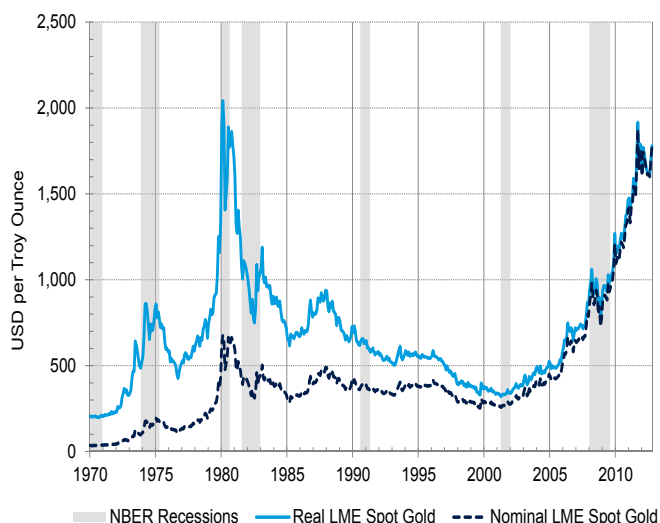

Exhibit 3 The Distribution of Maximum Drawdowns

record, the distribution of returns, deleveraging when in a drawdown, and survival. 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 Exhibit 2. In this exhibit, and in all of our drawdown exhibits, we show drawdowns as negative percent changes, and so we see in Exhibit 2 a high frequency of small drawdowns and a small frequency of large drawdowns.

Also, even 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 Exhibit 3.

3.1. What Forces Shape Drawdown Distributions?

In our simulations, we were able to control the return generating process for most of the parameters that would seem to make sense. In particular, we controlled for:

- Length of track record
- Mean return
- Volatility of returns
- Skewness
- Kurtosis
- Deleveraging 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.

3.2. The Distribution of All Drawdowns

From Exhibit 2, 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.

Mean return and the volatility of returns, however, matter a great deal. Exhibit 4 shows a manager's average or expected drawdown versus length of track record for different values of the four key moments of the
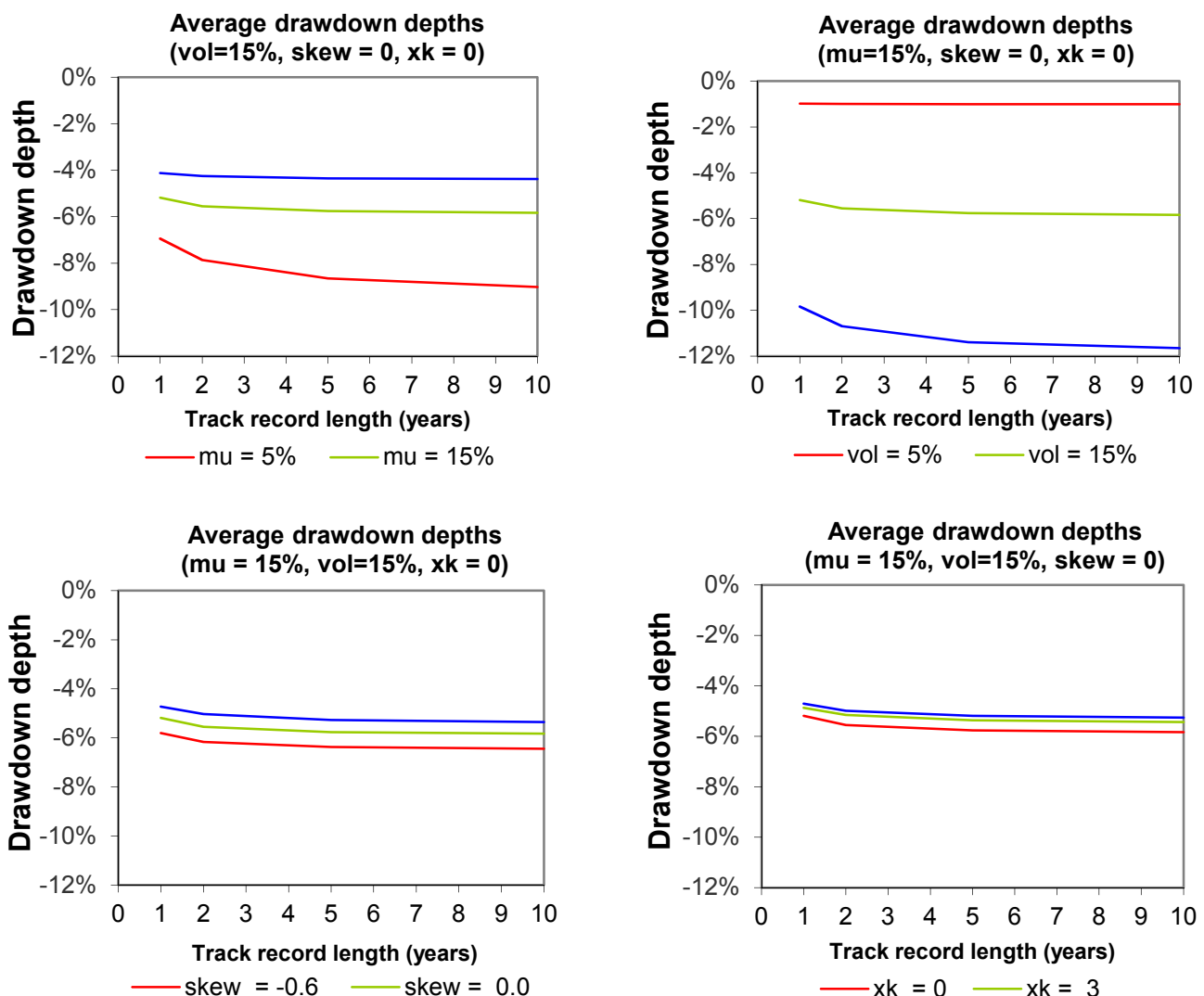

Exhibit 4 Track Record, Mean Return, and Volatility of Returns Have the Greatest Effect on Expected Drawdowns

return distribution. In the top panel, for example, we have varied the manager's mean return while holding volatility, skewness, and excess kurtosis (xk) constant. As one would expect, higher mean returns lead to smaller expected drawdowns.

From the second panel in Exhibit 4, it is also apparent that the volatility of returns 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 sequences of returns. As a result, even though the distribution from which any given return is drawn may be higher skewed or exhibit fat tails, the result of cumulating returns produces a random variable that tends (as suggested by the central limit theorem) to be more normally distributed.

3.3. The Distribution of Maximum Drawdowns
A manager's worst drawdown is taken from a distribution that is highly sensitive to length of track, mean return, and volatility of returns. Skewness and kurtosis, on the other hand, tend not to matter much.

The likelihood of any given drawdown is largely independent of how long a manager is in business. The likelihood of experiencing a drawdown that is bigger than anything experienced so far increases with every passing day. As a result, as shown in Exhibit 3, increases in the length of track record shift the entire maximum drawdown distribution to the left.

Exhibit 5 shows how the distribution is affected by mean returns and the standard deviation of returns. The values we have chosen here correspond roughly to the range of values we observe in our data base of CTA returns. As the upper panel shows, high returns tend to produce smaller maximum drawdowns, while the lower panel shows how increases in the volatility of returns increases the likelihood of large maximum drawdowns.

Exhibit 6 compares the effects of changing each of the four return characteristics on a manager's expected maximum drawdown. In each panel, it is apparent that length of track record matters more than it did with the expected value of all drawdowns. Each of these curves is considerably steeper here than those in Exhibit 4. It is also apparent that mean return and volatility of returns matter a lot, while skewness and kurtosis matter hardly at all.

To put the importance of these things in perspective, we calculated the partial effect of each moment on expected maximum drawdown and multiplied the partials by the standard deviation of each moment as measured in
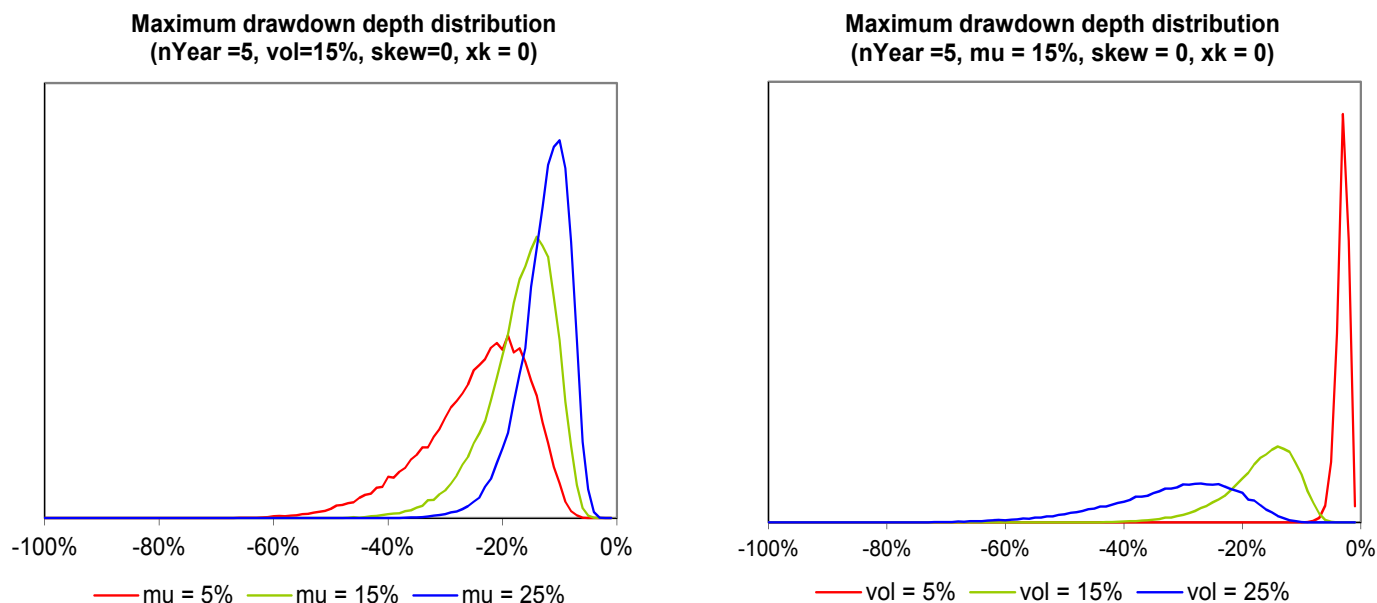

**Maximum drawdown depth distribution
(nYear =5, vol=15%, skew=0, xk = 0)**

mu = 5%   mu = 15%   mu = 25%

**Maximum drawdown depth distribution
(nYear =5, mu = 15%, skew = 0, xk = 0)**

vol = 5%   vol = 15%   vol = 25%

Exhibit 5 The Effects of Returns and Volatility on Maximum Drawdowns

our database. The results are shown in Exhibit 7, which shows that volatility of returns is by far and away the most important variable across managers. Variation in mean returns is a strong second. In contrast, skewness and kurtosis rank very low on the list of things that matter for explaining why different managers have different drawdowns.

4. The Core Drawdown Function

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. However, given the sizes of their respective 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 σ is the standard deviation of returns and μ is the mean return. A manager's drawdowns, when divided by the volatility of returns, can be written as a function of the manager's modified Sharpe ratio (i.e., the ratio of mean return to the standard deviation of returns).
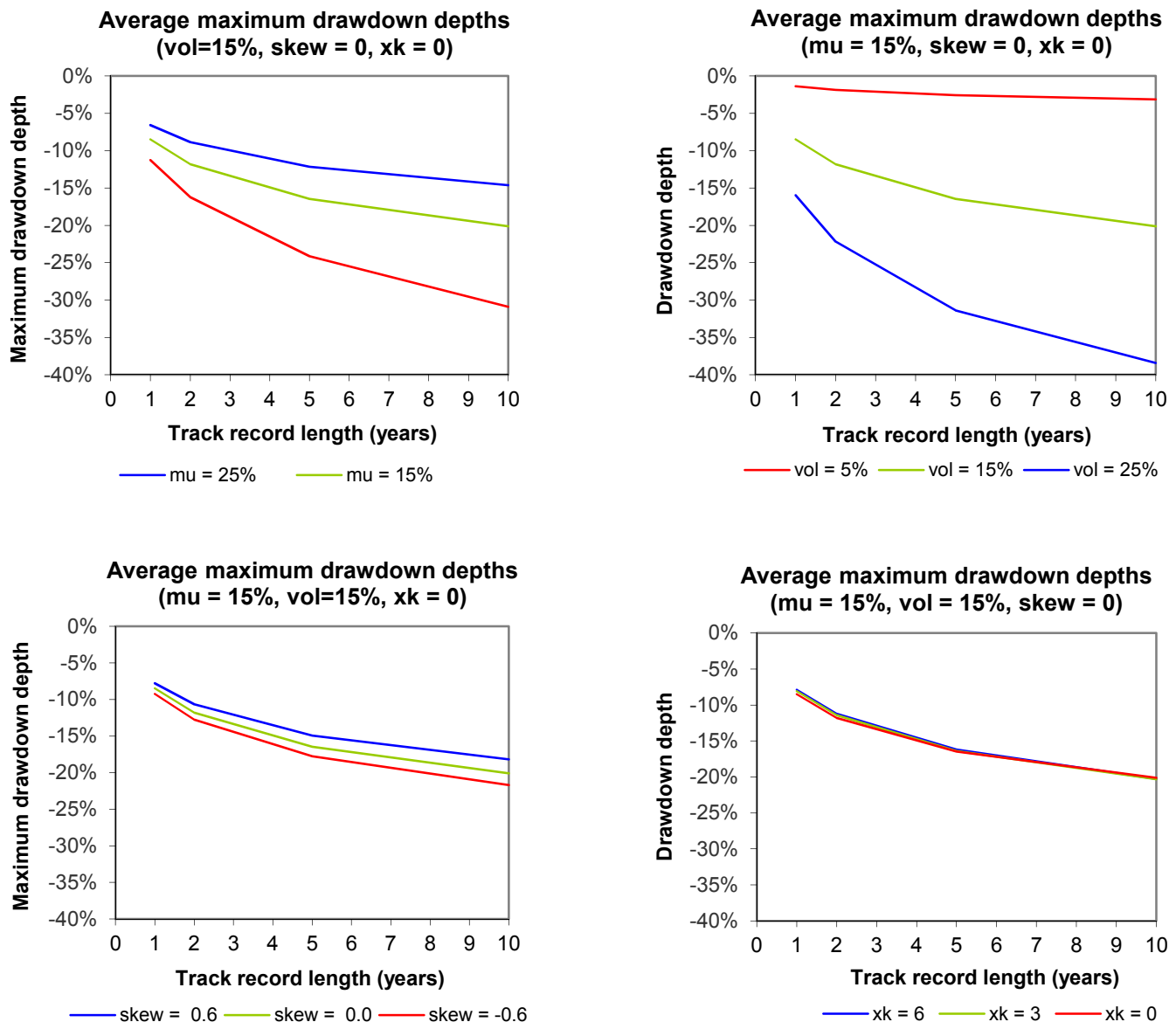

Exhibit 6 The Three Most Important Variables for Maximum Drawdown Are Length of Track Record, Mean Return, and Volatility

Exhibit 7 How important are the four moments for expected maximum drawdowns?

Moment	Potential Effect		
	Partial Effect* (%)	Standard Deviation of Moment	Combined Effect (%)
Mean	68.0	0.139	9.5
Standard Deviation	-191.0	0.133	-25.4
Skewness	2.7	1.035	2.8
Kurtosis	0.1	5.914	0.4

*local partial of expected maximum drawdown with respect to each moment
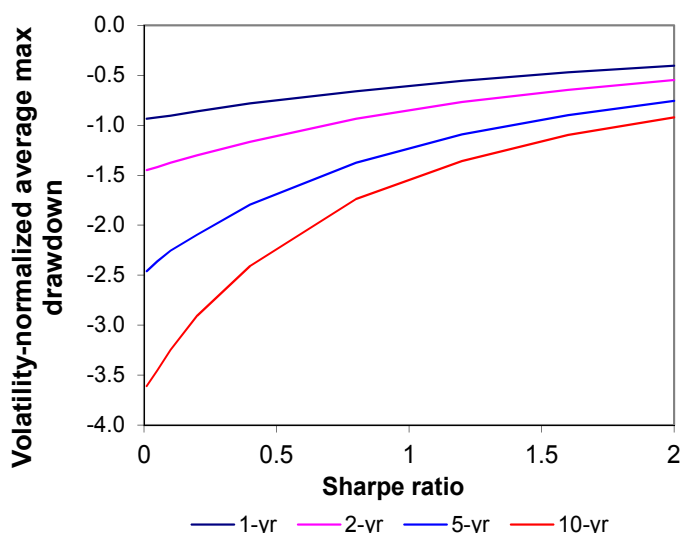Source: Newedge Prime Brokerage Research

Exhibit 8 The Shape of the Relationship between Expected Maximum Drawdowns and Returns When Both Are Normalized For Volatility

The shape of this function is illustrated in Exhibit 8 for track records ranging from one to 10 years. The curvature is consistent with our intuition 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 we are concerned only about the sizes of drawdowns, as opposed to drawdowns per unit of volatility, this relationship can be rewritten as:

$$DD = \sigma f(\mu/\sigma)$$

which suggests the following:

- a doubling of both mean return and volatility (which would leave the modified Sharpe ratio unchanged) will exactly double expected maximum drawdowns
- a doubling of volatility alone will more than double expected maximum drawdowns
- one would have to more than double mean return to compensate for a doubling of volatility

These points help to illustrate the differences between drawdown, volatility of returns, and a modified Sharpe ratio as measures of risk. All three are related, but provide different perspectives. Two managers with the same volatility of returns will have different expected drawdowns if their mean returns are different. Two managers with identical modified Sharpe ratios will have different expected drawdowns if their return volatilities are different.

5. Empirical Drawdown 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 using CTA returns from the Barclay's database: Using return histories for all managers with a 1-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 2-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 5-year and 10-year track records. By design, this approach produces different drawdown histories than those actually reported by the CTAs in our database. It has the advantage, however, of considering all managers under the same market conditions.

The results of these efforts are shown in Exhibit 9. The distributions of all drawdowns are shown in the upper panel, while the distributions of maximum drawdowns are shown in the lower panel.

6. Reconciling Theoretical and Empirical Distributions
The distribution of all drawdowns shown in Exhibit 9 look about the way we would expect. The distributions of maximum drawdowns, on the other hand, posed a
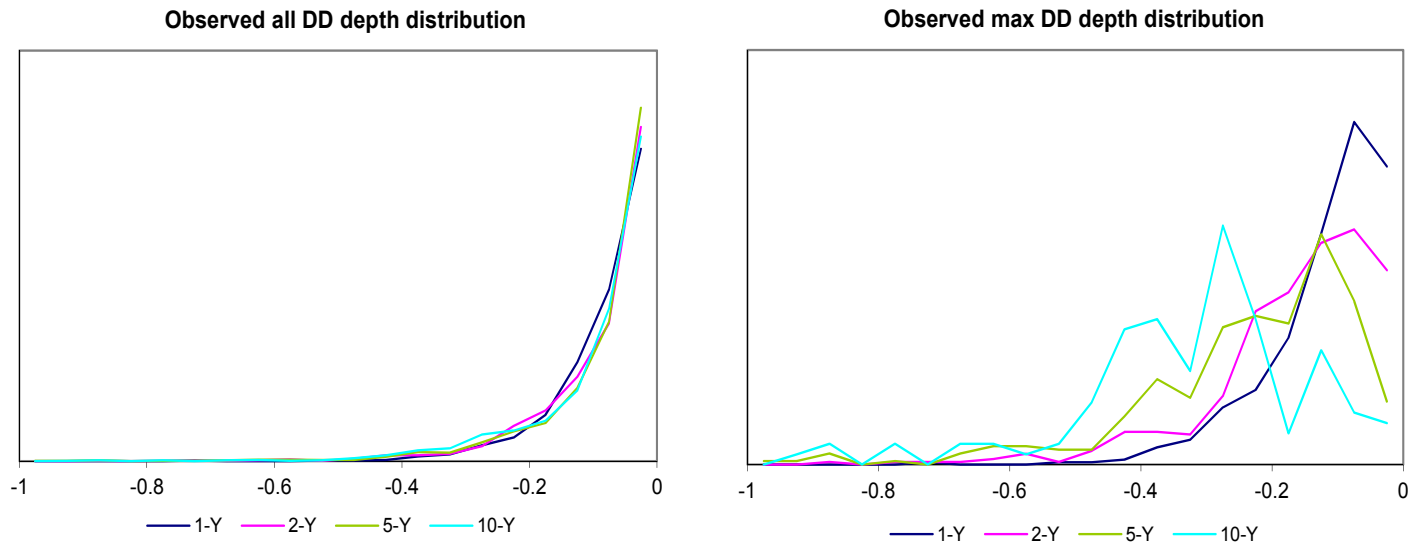
Exhibit 9 Observed Drawdown Distributions

real challenge. First, they are irregularly shaped. Second, Exhibit 10 shows that that the observed distribution, derived from the actual maximum drawdowns for all CTAs with 10-year track records, does not fit well with the theoretical distribution (labeled "discrete").

The problem is that the observed drawdown distribution peaks at a much lower level of drawdowns than does the theoretical. One plausible explanation for the difference in the shapes of the two distributions is that managers may deleverage when they are in drawdown (scale back the risk they take) and thereby avoid the larger drawdowns they would experience if they were to keep the volatility of returns constant.

Exhibit 10 shows that it is possible to pull the theoretical drawdown distribution to the right by allowing for deleveraging when simulating returns. In this case, we scaled the manager's mean and volatility of returns as:

$$\mu' = f \times \mu$$
$$\sigma' = f \times \sigma$$

where

$f = 1 - [abs(drawdown)]1/2$

so that if a manager's current drawdown were 5%, the scaling factor would be .29 (= 1 - .51/2). The new distribution of maximum drawdowns that results is labeled "Adapted" in Exhibit 10 and peaks very close to where it should.

One drawback to this approach is that the evidence on deleveraging is largely anecdotal. We know managers who attest to the fact that they scale back risk when in drawdown. We know other managers, however, who say that they do not. And we have not been able yet to find any evidence in the volatilities of managers' returns that suggests that they deleverage when in drawdown.

Another drawback to this approach is, while it produces a probability distribution that generally looks like what we observe, it greatly under-predicts the several large drawdowns that we observe in the data.

A better solution seems to lie in the fact that managers exhibit very different volatilities of returns. This is borne out in Exhibit 11, which shows a scatter plot of mean returns and their corresponding volatilities for those managers for whom we have 10 years of performance data. For the purposes of this exercise, we grouped the managers into three broad volatility groups: (1) low (0–12.5%), (2) medium (12.5–25%), and (3) high (25–50%).

Using the group returns and group volatilities, we simulated the three maximum drawdown distributions shown in the upper panel of Exhibit 12. 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
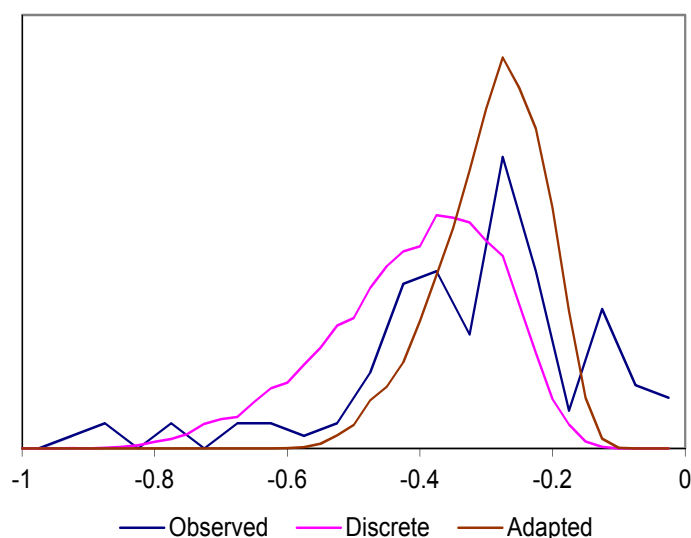
10-Y max DD depth distribution

Exhibit 10 First Cut At Explaining Observed Maximum Drawdowns

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 Exhibit 12, the composite exhibits some of the irregular shape that we see in the observed distribution of drawdowns.

7. Putting a Manager's Drawdown Experience in Perspective

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 Exhibit 13 for a manager with a 10-year track record, a mean return in excess of 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 more than 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.

8. What About Future Drawdowns?

What kinds of drawdowns might an investor expect going forward? This work suggests that 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 corresponding conditional maximum drawdown.

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? If we are willing to assume a mean return and volatility (e.g., 12% and 20%), we find in Exhibit 14 that 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.

Observed volatility-return relation
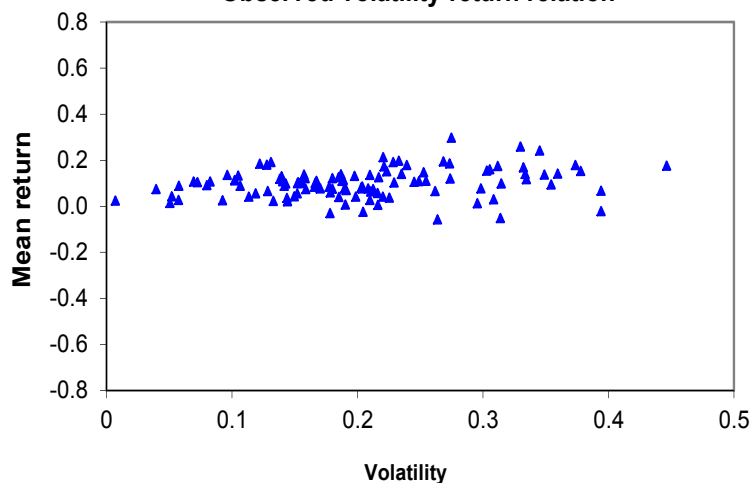
Managers grouped by volatility			
Volatility range	Number of managers	Group return	Group volatility
0.0% ~ 12.5%	19	8.3%	8.6%
12.5% ~ 25.0%	64	9.4%	18.8%
25.0% ~ 50.0%	9	12.3%	32.4%

Exhibit 11 The Range Of Returns and Volatilities for Managers with a 10-Year Track Record

14

10-Y max DD depth distribution
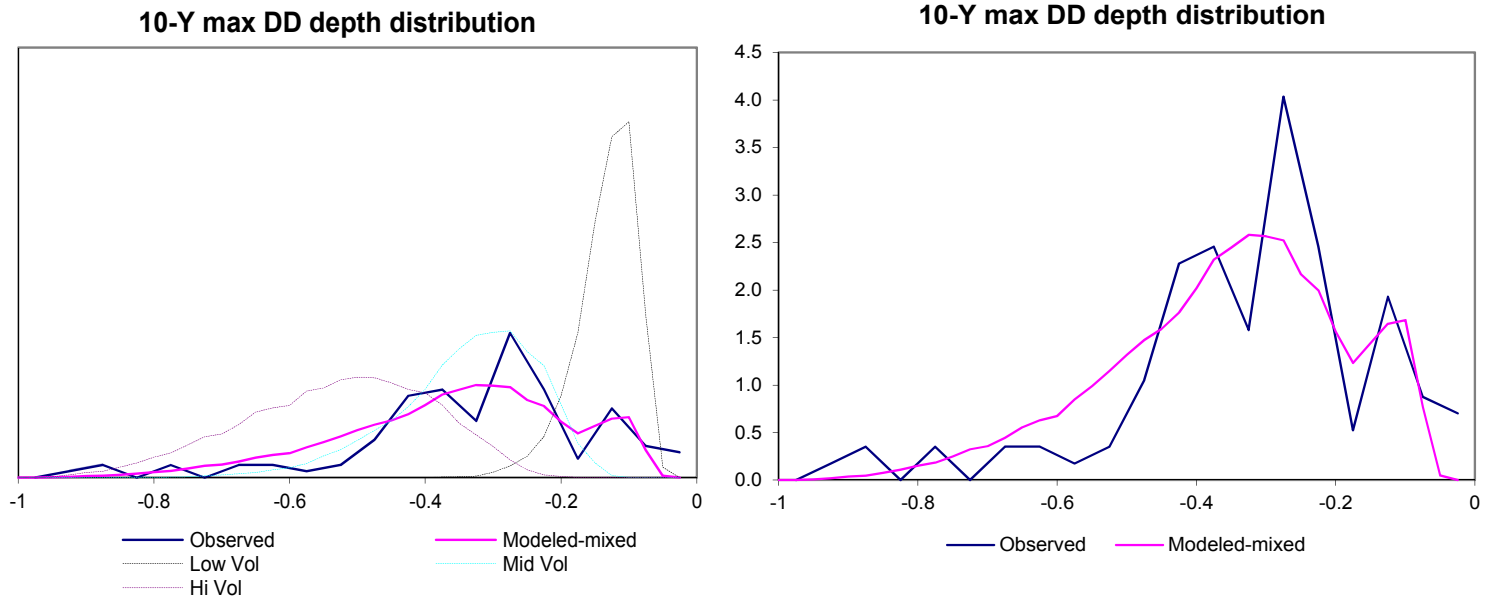
10-Y max DD depth distribution

Exhibit 12 A Composite Maximum Drawdown Distribution

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 that 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.

9. Further Questions

Our conversations with clients and colleagues about this work have generated several questions that deserve a closer look. How important, for example, are serially correlated returns? How reliable are our estimates of return volatilities? Would the analysis be better if one had daily rather than monthly return data? Would this analysis work as well for hedge funds as it does for commodity trading advisers? What happens if a manager's return volatility changes in response to drawdowns?
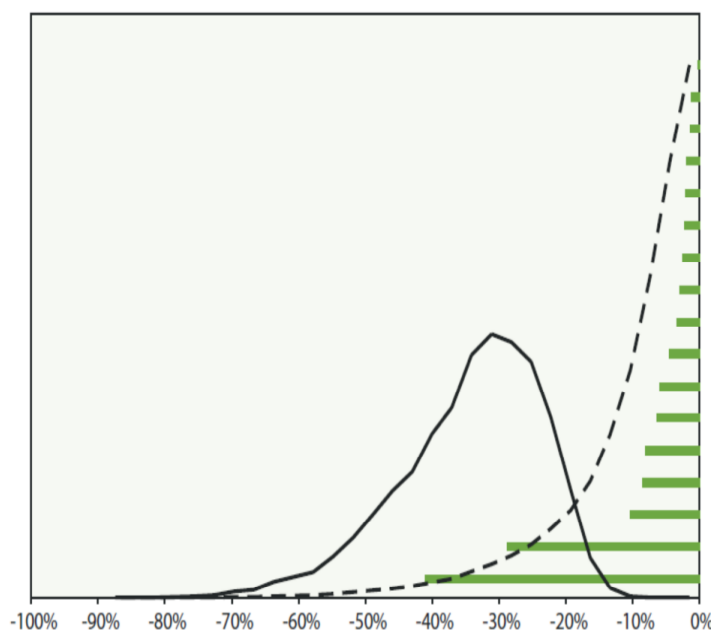
Our preliminary thoughts suggest the following. First, serially correlated returns could have a measurable effect on drawdown distributions, but we have found no evidence of serial correlation in CTA returns. Second, volatility estimates based on monthly return data can be subject to very large statistical 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. Thirdwhile we know that some managers deleverage when in drawdown, the evidence on CTAs as a class is ambiguous.

Exhibit 13 Assessing a Manager's Drawdown Experience

Exhibit 14 How much worse can it get?

(Mean return = 12%, Volatility = 20%)

Investment Horizon (Years)	Probability of a Drawdown if > 41%	Expected drawdown if > 41%
1	0.1%	-44.1%
2	1.7%	-45.8%
3	4.3%	-47.1%
4	6.8%	-47.3%
5	9.5%	-47.8%
6	12.9%	-48.2%
7	15.3%	-48.4%
8	17.8%	-48.8%
9	20.5%	-49.1%
10	23.4%	-49.0%

Source: Barclay Hedge, Newedge Prime Brokerage Research

Technical Note

To simulate net asset value 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 generalized lambda distribution. The values of skewness and excess kurtosis used in this note were roughly consistent with the range of values we observed for CTAs in our database. From the return series, we construct net asset value series. And from these, we derive the simulated drawdowns that are used to produce the theoretical drawdown distributions. A typical run usually requires 10,000 iterations to produce a smooth distribution.

Authors' Bios

Galen Burghardt is a Senior Advisor for Newedge USA, LLC. He is the lead author of The Treasury Bond Basis and The Eurodollar Futures and Options Handbook, which are standard texts for users of financial futures. His latest book, written with Brian Walls, is Managed Futures for Institutional Investors: Analysis and Portfolio Construction (Bloomberg/Wiley, 2011). His work experience includes: adjunct professor of finance in the University of Chicago's Graduate School of Business (now the Booth School); director of research for the futures divisions of Calyon Financial, Dean Witter, and Discount Corporation of New York; head of financial research for the Chicago Mercantile Exchange; and economist in the Capital Markets Section of the Federal Reserve Board, where he gained access to the working world of finance and futures. His PhD in economics is from the University of Washington in Seattle.

Ryan Duncan is the Global Co-Head of the Newedge Alternative Investment Solutions' Advisory Group. His responsibilities entail the development of investor-focused research on a wide variety of hedge fund strategies including managed futures, global macro, commodity, currency, and volatility trading. He has been the author and co-author on several papers focused on the performance characteristics of hedge fund strategies and portfolio construction issues such as Superstars versus teamwork and There are known unknowns. Additionally, he manages the Newedge Index committee and has worked to develop a number of hedge fund indices as well as design customized performance and risk reporting for clients. Such experience differentiates Ryan as a sought after advisor to institutional investors.

Prior to joining Newedge, he developed software for the Weiland Financial Group. Ryan holds a BBA in both Economics and Management Information Systems from The University of Iowa and is a CAIA designee.

Lianyan Liu is a Quantitative Analyst for the Newedge Alternative Investment Solutions' Advisory Group. A seasoned member of the team, he has been a key contributor to the research the firm has done on alternative investments. Lianyan's recent interests include portfolio construction, trading strategies, and fund manager evaluation. Before joining Newedge in 2000, he had done research in the areas of nuclear physics, nuclear oil well logging, and radiation oncology physics. He holds a PhD in Nuclear Engineering from North Carolina State University, and an MBA with High Honors from The University of Chicago Booth School of Business. He is a member of Chartered Alternative Investment Analysts.