Bubble VaR, a countercyclical value-at-risk approach

VaR and financial crises

It is a great irony that our advancement in risk management has created a new kind of risk—the risk of risk measurement—the misuse of (frequentist) statistics. The credit crisis has debunked some well-established risk models used by banks, in particular Value-at-risk (VaR), an essential tool used for the purpose of regulatory capital and banks' own economic capital. In fact, the risk of using VaR models has been forewarned by Danielsson et al. as early as 2001 in a response paper to the Basel authority. Nonetheless, VaR has become the industry standard because of its practical simplicity and a lack of agreeable alternatives.

VaR is the loss quantile of the P&L distribution of a portfolio, statistically determined using data over a sample period (of multiple years) and estimated over a specified time horizon (10 days for market risk) at a certain confidence level (for example at 97.5% confidence level i.e. a 2.5% quantile is used). The crisis has shown that, as a risk metric intended to measure extreme losses or *tail* risk, VaR is just "too little, too late". How far off are we?

Figure 1 shows the 97.5% VaR for the Dow Jones index; the P&L's which exceeded the VaRs for longs and shorts are shown as vertical bars. During the credit crisis in 2007, the exceedences are a lot larger than VaR, in fact, very much larger than that expected by the commonly assumed normal distribution. To get an idea if the exceedences are realistic, consider Table 1 which shows the top-10 largest daily losses in history. Certainly crises have occurred more often than suggested by VaR models. These led to the popular idea of Black Swans (see Taleb (2007)), events of low probability and high devastation located at the extreme tail of the distribution. Taleb found that such events in financial markets are not statistically reproducible (atypical), a phenomenon he termed *extremistan*. Without the element of reproducibility, Black Swans are not amenable to statistical quantification; the usage of frequentist statistics as applied in VaR will lead to precise but misleading (inaccurate) numbers. Because of finite sampling, a mathematical quantile can always be determined, but the expected loss beyond this quantile is an elusive number which may never converge under extremistan. This illusion of precision could undermine risk management preparedness by putting risk controllers in a comfort zone.

Secondly, VaR is always late in crisis detection—it increases sharply only *after* the initial sharp selloff of a crisis. The cause, the use of a recent rolling-window in such models means that this metric will always respond late to large market movements that mark the onset of a regime shift. Thus, VaR is an effective analytics during "peace time" when changes are gradual, but is useless for crisis warning. As VaR is used to determine regulatory capital, the burden on banks will rise sharply in a crisis situation after the fact, forcing banks to reduce positions in a falling market. Conversely, VaR is muted in a bullish market and the benign capital requirement encourages balance sheet expansion and the use of leverage. This phenomenon of low volatility during rallies and high volatility during sell-downs is called the *leverage effect* and is well known. Figure 2 shows such an inverse relationship between the S&P and VIX indices. Thus, regulatory risk model has the unintended effect of amplifying the boom-bust cycle; this *procyclicality* risk is highlighted by the *Turner Review* (2009). The Review calls for a reformed capital regime which is overtly countercyclical—reserving more capital during a boom which can later be used to cushion losses during the bust phase.

Thirdly, VaR is symmetric—it does not capture directional risks, only changes in volatility. In Figure 1, the VaR at the base of the rally (early 2005) is the *same* (about 2%) as that at the peak of the market in Oct 2007. But shouldn't the risk of a crash be highest at the peak? This fear of crash is not only intuitive but is also reflected by the markets. Since Bates (1991) first observed this, many studies show that the option implied skew is highest near the top of the market just before major crashes; this is evident in Figure 3 which shows the S&P index and its option implied skew during the 2008 crisis. But VaR could not capture this fear because *statistical* skewness of a distribution does not reflect directional risk very well, due to distortion from trading around support/ resistance levels. To understand this microstructure, consider a pegged currency attack. Speculators who sell against the peg will bring prices down to test the peg *gradually* because the currency is supported by opposing traders (and the central bank) who bet that the peg will hold. Conversely, each time the peg holds, short covering will likely see quick upward spikes. This causes occasional positive skewness in the distribution even though the real directional risk is downwards. Since trading can be seen as a battle between bulls and bears for resistances and supports, statistical skewness is highly dependent on price levels. Figure 4 illustrates the statistical skew for the S&P index—there is no obvious pattern near the market peak, the realized distribution is virtually symmetric.

This leads to the fourth weakness, VaR does not distinguish between long and short positions. In Figure 1, the VaR just before the crisis (Oct 2007) has effectively the same values for longs and shorts. But a crash can only happen downwards! (never up). So shouldn't the risk be higher for longs? Likewise, at the trough of the market, shorts should be more at risk to a (rapid) bounce, which can only happen upwards. Thus, the present day capital regime does not penalize longs for chasing an asset bubble nor recognize that the *crash* risks of opposing positions are unequal. Since the banking system is profit maximizing and capital efficient, the rules incentivize banks to chase the latest hot assets collectively; it is not macro-prudential.

Misuse of statistics

VaR models are appealing because they are mathematically tractable, meaning a risk controller can test his results and state with precision that his VaR is "true" within certain confidence. (Thinking in terms of an existence of a "true" value waiting to be uncovered is a hallmark of the frequentist school of thought.) Tractability comes with a high price—you have to assume returns are *independent and identically distributed* (i.i.d). The abstraction of i.i.d. is just a sleight of hand of modelers to forecast VaR over the next time horizon—if i.i.d. is true, then it seems acceptable to infer that tomorrow's quantile is identical to the quantile calculated from the most recent past data sample.

I.i.d. also allows a risk controller to perform back testing to "prove" that a VaR of say 99% confidence level does indeed contain on average one exceedence per 100 days of observations. In fact, back testing is nothing more than a test of the various nuances of i.i.d. In addition, when residuals are i.i.d., standard statistical estimation methods (such as regression) can produce consistent and unbiased estimates.

Ironically, the same i.i.d. assumption is the root-cause of the failures of VaR. The practice of statistics and the notion of precision (as opposed to accuracy) hinges on reproducibility of observations, which is questionable when financial markets are in crisis mode. In a crisis, contagion and positive feedback among market participants (panic selling) cause price to spiral in a self-reinforcing way. The return series become serially correlated, which leads to a breakdown of the i.i.d. assumption. This explains why high volatilites and back testing breaks are clustered together during crises, which makes back testing a tricky task—back testing is meant to validate independence of tail losses, but it is exactly at stressful periods of interest to VaR that i.i.d. is violated.

When i.i.d. is violated, the aggregate distribution becomes fat-tailed (leptokurtic) because the central limit theorem (CLT) no longer holds. The CLT states that even if variables are non-normally distributed individually, *as long as they are i.i.d.*, their joint distribution becomes normally distributed when many variables are aggregated. Even models based on extreme value theory (EVT), which attempts to model extreme tails, rest on the i.i.d. requirement. This could explain why VaR models are so often surprised by the actual loss magnitudes witnessed during crises (as seen in Table 1)—the distributional models do not work when i.i.d. is broken.

The ideal of i.i.d. is the reason risk models are based on returns as opposed to raw prices. Empirical research provides the comfort that returns exhibit this ideal characteristic during regular most-of-the-time periods (even though VaR is more concerned with irregular crisis periods). The process of differencing to arrive at returns means that useful information on cycles and trends are lost from the VaR input. Figure 5 shows the classical decomposition of a price series into three major components—trend, cycle and residual. The residual (returns) is the input for VaR. It is not surprising why VaR cannot distinguish between peaks and troughs, longs and shorts—it is missing cyclical information in its input. VaR is unable to reflect the obvious asymmetry of a crash. Without this directional element, VaR necessarily underestimates fat-tail losses.

If we ask a retail investor who has never heard of Markowitz portfolio theory, what is his risk after buying a stock, his answer will naively (and quite correctly) be downside (directional) risk, not volatility. Arguably, volatility is only a risk for an option, and only in the special case of a delta-hedged option. The pervasiveness of thinking in terms of volatility in risk management is a legacy of Markowitz (1952). If directional risk is truly potent, it seemed odd that we are leaving out trends from risk models.

By working exclusively with the i.i.d. returns, VaR models have gained the elegance of tractability and precision, unfortunately, at the expense of accuracy during times of crises. Keynes dictum is apt in risk management, “it is better to be roughly accurate than to be precisely wrong”.

It’s in the Cycle

In Figure 5, the clustering and fat-tailness of residuals are evident during crises. Academic research has attempted to explain these phenomena by modeling the process or the distribution of returns (*residual*), as in the GARCH and the EVT literature.

We posit that these phenomena are caused by the *cycle* component instead. Crashes are just breaks in the cycle—asymmetric and often sharp—which give rise to fat-tailed returns. The trend is driven by real economic growth and the i.i.d. residual is due to trading under efficient market conditions. The leverage effect is explained by observing that a cycle break is more common downward than upward. Furthermore, cycle compression during a crisis—i.e. shortening of periodicity and widening of amplitude—increases serial correlation and leads to volatility clustering. Figure 6 illustrates this; we simulate a stylized trend, cycle and i.i.d. noise. A cycle break is introduced as a vertical drop, and cycle compression is introduced in the shaded zone. We combine all three elements into a price series and take its returns. The returns in the lower panel shows manifested fat-tail events, clustering and serial correlation—similar to what is often observed in financial time series.

If crashes indeed originate from the cycle, we can design a forward looking risk metric by studying the cycle. Under this new paradigm, the larger a bubble forms on the upside (downside), the larger the risk of a crash (bounce), the long (short) positions are more risky to a crash (bounce). It follows that at the peak of a bubble, longs should be penalized most (in terms of risk capital), and at the trough, shorts should be penalized most. We propose a new metric called *bubble value-at-risk* (*buVaR*) which recognizes that "crashes" can only happen in the countertrend direction and the risks of longs and shorts are unequal.

VaR robustified

BuVaR starts with the premise that crashes are extremistan, thus, can never be measured *precisely*. The goal is to robustify VaR against the four weaknesses mentioned earlier for the purpose of risk capital. In other words, buVaR should provide a larger buffer to cover tail losses one step ahead of a potential crisis such that banks can build up a countercyclical safety buffer.

We first need to design a measure of the cyclical euphoria which we shall call the *bubble* (B). If the *bubble* is on the upside (B>0) then we inflate the return distribution on the negative side, which would penalize longs. If the *bubble* is on the downside (B<0), we inflate the positive returns which would penalize shorts. At time t, for risk factor X_n, its daily return variable $R_n = \ln(X_n/ X_{n-1})$ undergoes a transformation:

$$R_n \rightarrow \begin{cases} \Delta_t R_n & if \;\; sign(R_n) \neq sign(B_t) \\ R_n & if \;\; sign(R_n) = sign(B_t) \end{cases} \qquad (1)$$

where Δ_t (≥1) the *inflator* is a function of B_t and n is the scenario number in the historical simulation VaR approach. We shall use a 1-year sample window, so n=1,..,250.

BuVaR is not statistically estimated but is derived using heuristic boundary arguments. It is reasonable to expect the tail loss to be above VaR (a known underestimate) and below some structural upper limit. This upper limit exists because of the imposition of circuit breakers by exchanges which restrain daily loss (to say below 10%). By arbitrage argument, related derivatives in the OTC markets will also be restrained. In between the bounds, it is reasonable that the expected loss rises monotonically with the *bubble*.

The *bubble* is computed using a method called rank filtering using a rolling-window of four years of daily *prices* (not returns). This process creates a cycle measure which has many desirable characteristics, such as being in synch with (or even leading) the market cycle, and reducing false positives where a bubble evolves into a long-term trend, etc. The inflator is then given by:

$$\Delta_t = Min\left(\frac{\psi}{2\sigma_t}, \exp\left\{\left(\frac{Abs(B_t)}{B_{\max}}\right)^{\omega_2} \ln\left(\frac{\psi}{2\sigma_t}\right)\right\}\right) \tag{2}$$

where:

- Ψ : average of 5 most extreme (absolute) returns in all available history of that asset, capped by a circuit-breaker if applicable.
- $B_{\max}$: largest absolute B_n observed in all history of that asset
- σ_t : standard deviation of returns of the last 250 days
- ω_2 = 0.5

The form of (2) ensures that buVaR ranges between VaR and the upper limit, and grows with the *bubble*. For technical details please see Wong (2011a).

The inflated distribution(s) of the risk factor(s), R_n is then used to reprice a portfolio at scenario *n*, giving the P&L at scenario *n*. For *n*=1,2..,250, a distribution of P&L, *y*, can be sampled. The buVaR at confidence level *q*% is the *expected shortfall* of the distribution *y* estimated over a 1-day horizon at (1-*q*) coverage:

$$BuVaR_q = \mathrm{E}(y \mid y < \mu) \text{ where } \Pr(y < \mu) = 1 - q \tag{3}$$

Figure 7 compares the buVaR and conventional expected shortfall for the Nasdaq index; a few stylized facts are noteworthy. First, buVaR inflates in the direction of the bubble formation to discourage speculators—it peaks and troughs in synch with the market, and often months ahead of major reversals and crashes. Second, consider the period 1995-96, when the rally proves to be sustainable (and not a real bubble), buVaR has an in-built mechanism to deflate in order not to penalize a long-term trend. This ensures that banks are not discouraged from sustainable investments, only frothy ones. Third, during the sharp falls at the onset of crises, buVaR is designed *not* to deflate so that the penalty on longs is sustained. This discourages banks from averaging down during a crash and is prudential. Similar behavior is seen in Figure 8, the buVaR for crude oil futures. Comprehensive tests of buVaR are done in Wong (2011b) covering various asset classes, and performed on specific extreme events in financial history. The results are characteristically similar.

BuVaR is reliable for risk factors that are quoted in price and yield terms, where trends and cycles are clearly defined. But for variables that are fairly stationary such as option volatility and spreads, buVaR is unnecessary. Without a cycle element, the rank filtering process will measure a false positive. In addition, credit spread has a unique asymmetric (upward) mean-fleeing behavior when credit deterioration occurs. This risk is specific to an issuer and not directly related to the overall market cycle. Credit spreads are thus modeled using a different approach called *Credit buVaR*; see Wong (2011c).

A countercyclical risk metric

We summarize the design advantages of buVaR:

a. It is overtly countercyclical: it is preemptive of a crisis or crash, and often peaks months in advance. Risk capital is built up ahead of a crisis.

b. It reflects risk asymmetry, that the risks of longs and shorts are unequal. It penalizes positions that are chasing the bubble.

c. It gives a thicker capital buffer against fat-tail events: when an asset bubble is forming, buVaR becomes a multiple of VaR and provides better protection against the risk of a crash in the *countertrend* direction.

In effect, buVaR weighs one side of today's distribution with a parameter which is recalibrated daily using cyclical information. Since the cyclical parameter is modeled using a non-i.i.d. method, tractability is compromised and there is no unique "true" solution. It also means back testing cannot be used to give comfort that the model is precise and unbiased. We argue that i.i.d. is violated anyway under stressful market conditions, during which precision is misleading.

BuVaR is a more *accurate* (as distinct from precise) risk metric than VaR because it gives a best guess of the estimated loss between VaR (which is known to understate risks) and a credible upper bound. We stressed that buVaR does not produce a mathematically "true" estimate, but instead, provides a more robust (thicker and more timely) number for the purpose of risk capital. The higher capital buffer can be seen as a model reserve against our lack of knowledge of fat-tail phenomena.

Since buVaR requires the use of daily prices, its application is primarily for the trading book or for market risk, where data is abundant. It acts as a cycle dampener and is macro-prudential because positions that are "chasing the bubble" become increasingly costly in regulatory capital compared to contrarian positions.

BuVaR is also a deviation from conventional frequentist thinking. It adopts a philosophy that extreme tails are inherently immeasurable. Questions of "true risk" are empty, and a risk metric's value lies in its usefulness. The important question to ask is: used in a given application, will buVaR promote prudence and desirable behavior at financial institutions?

References

Danielsson, J., Embrechts, P., Goodhart, C., Keating, C., Muennich, F., Renault, O., Shin, H.S. (2001) 'An Academic Response to Basel II', Special Paper 130, ESRC Research Centre.

Financial Service Authority (2009) 'The Turner Review: A regulatory response to the global banking crisis', Financial Service Authority.

Taleb, N. N. (2007) 'The Black Swan: The Impact of the Highly Improbable', Random House (US) and Penguin (UK).

Bates, D. (1991) 'The Crash of '87: Was it Expected? The Evidence from Options Markets', *Journal of Finance,* Vol. 46 (3), 1009-1044.

Wong, M. (2011a) 'Market BuVaR: A countercyclical risk metric', forthcoming in *Journal of Risk Management in Financial Institutions*, Vol 4(4), 1-14.

Wong, M. (2011b) 'Bubble Value at Risk: Extremistan and Procyclicality', Immanuel Consulting.

Wong, M. (2011c) ‘Credit BuVaR: Asymmetric spread VaR with default’, *Journal of Risk Management in Financial Institutions*, Vol 5(1), 86-95.

About the author

Max Wong is a quantitative risk manager at a UK bank, a risk professional with 15 years of experience in financial services, and book-author of “*Bubble value at risk: extremistan and procyclicality*”. Max’s current research is in the area of financial regulatory reform and innovative risk management. The views expressed in this article are his own and do not represent the views of any organizations that he is affiliated to. The source code for the computations is available at: www.bubble-value-at-risk.com Email: max.wong@bubble-value-at-risk.com

Table 1: Top-10 daily losses in the history of Dow Jones index

Event Date	Daily log return	Mean number of years between occurrences (if normally distributed, assuming volatility of 25% p.a.)
19-Oct-87	-25.6%	1.86E+56
26-Oct-87	-8.4%	69,074
15-Oct-08	-8.2%	37,326
1-Dec-08	-8.0%	19,952
9-Oct-08	-7.6%	5,482
27-Oct-97	-7.5%	3,258
17-Sep-01	-7.4%	2,791
29-Sep-08	-7.23%	1,684
13-Oct-89	-7.16%	1,346
8-Jan-88	-7.10%	1,120

Figure 1: 97.5% VaR of Dow Jones index calculated using Riskmetrics method, and VaR exceedences for long/ short sides. The same stylized facts also hold for other VaR methods.

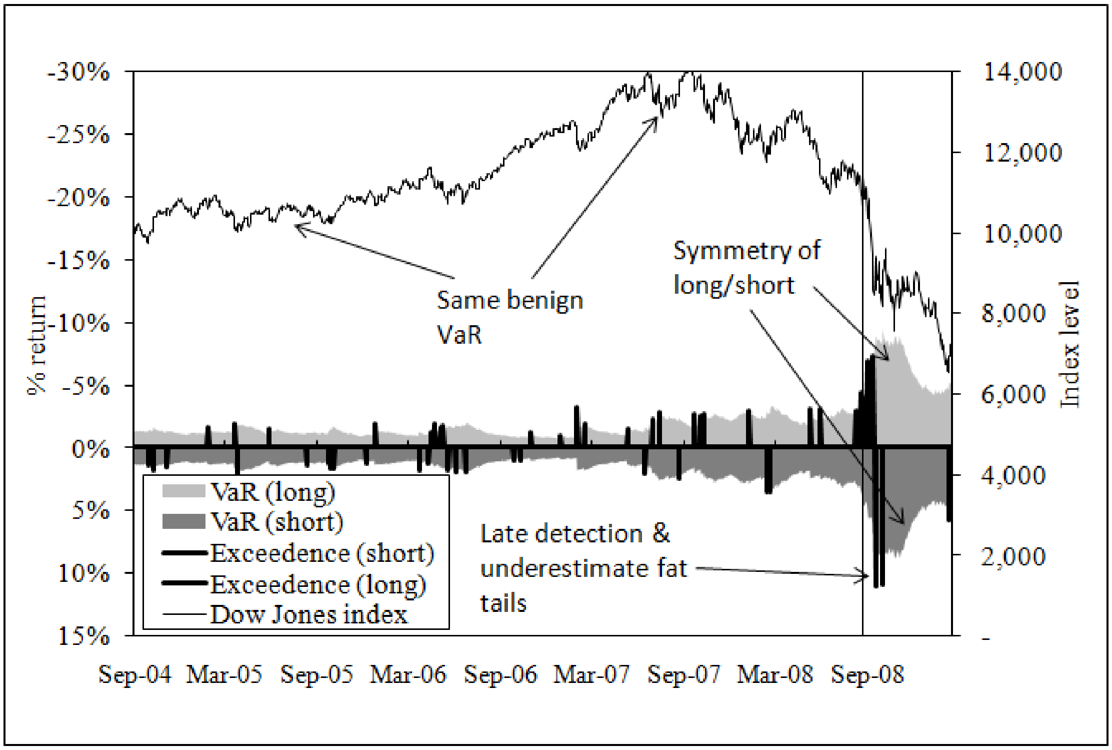

Figure 2: The leverage effect, or inverse relationship between price and volatility

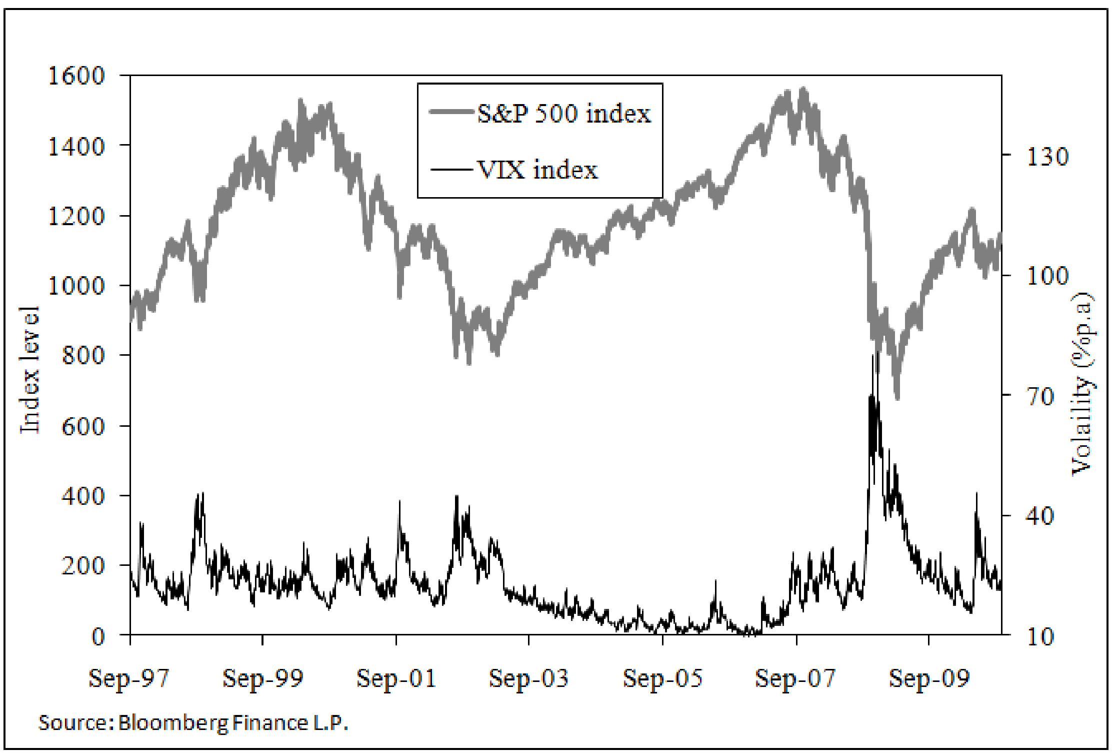

Figure 3: Equity index and its implied volatility skew (the difference between 3-month option volatilities of 90% and 110% moneyness)

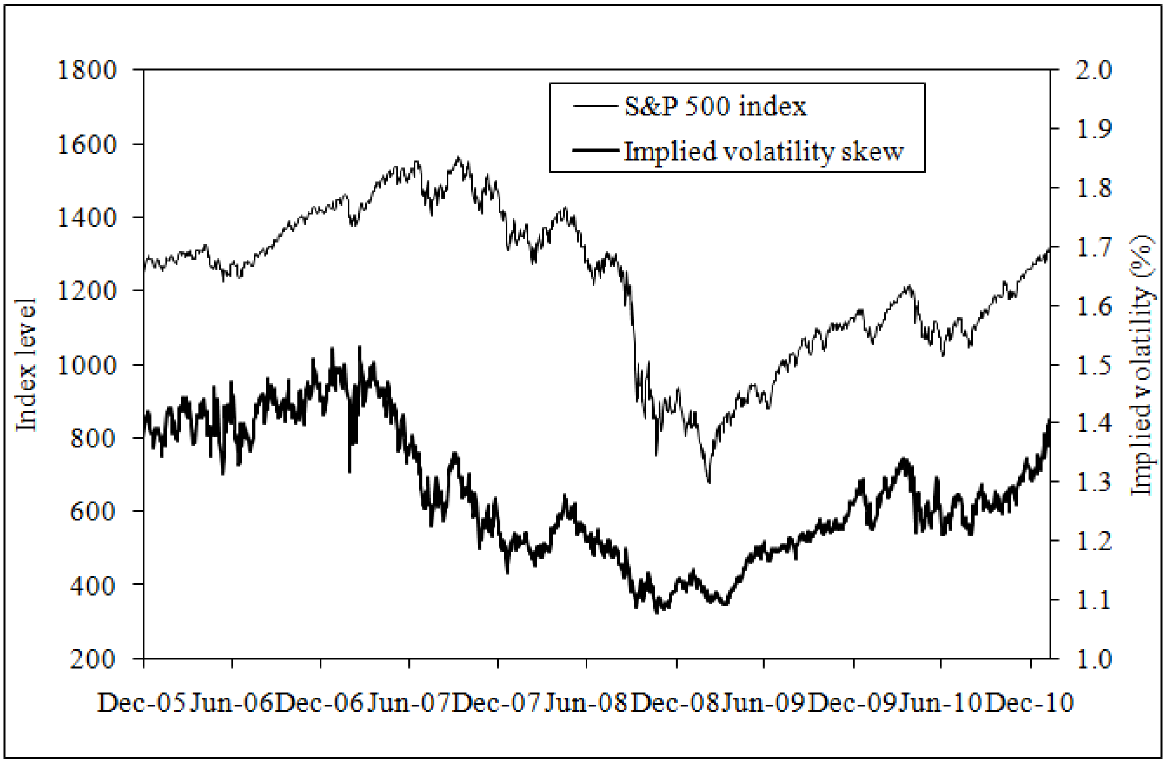

Figure 4: Equity index and its statistical skewness (calculated using 1 year rolling prices)

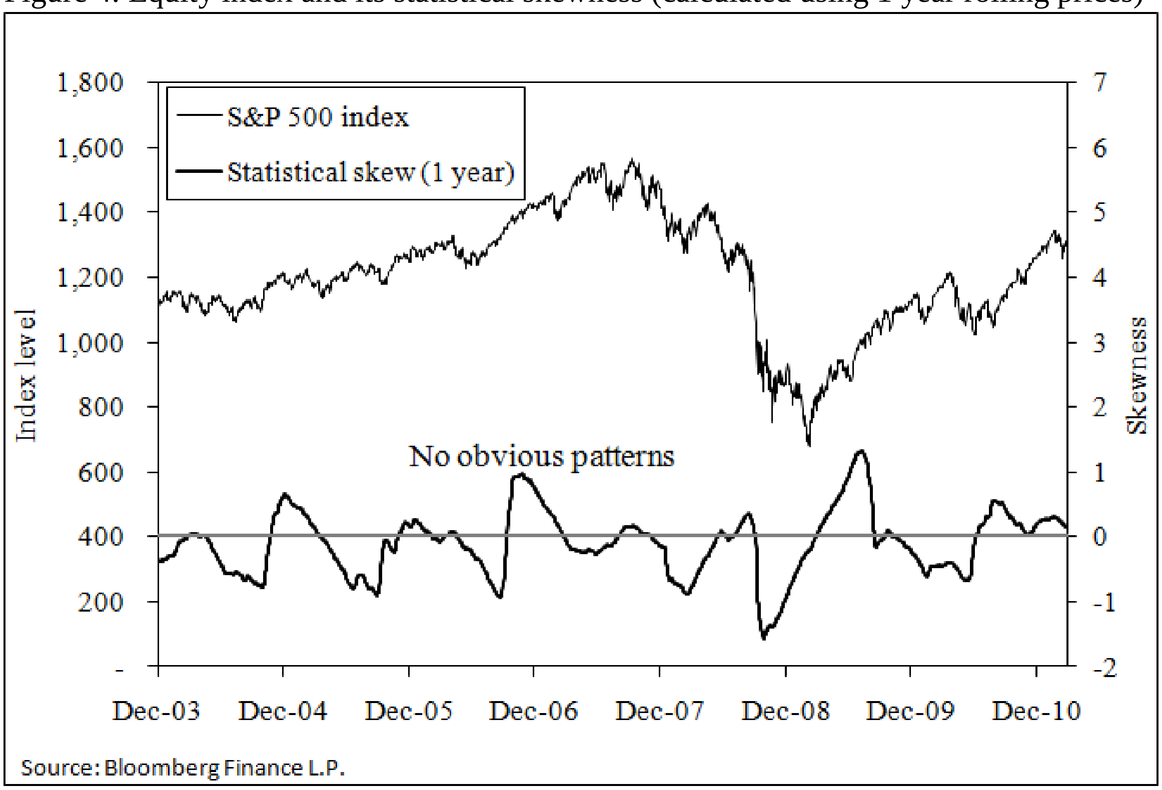

Figure 5: Classical decomposition of a price series

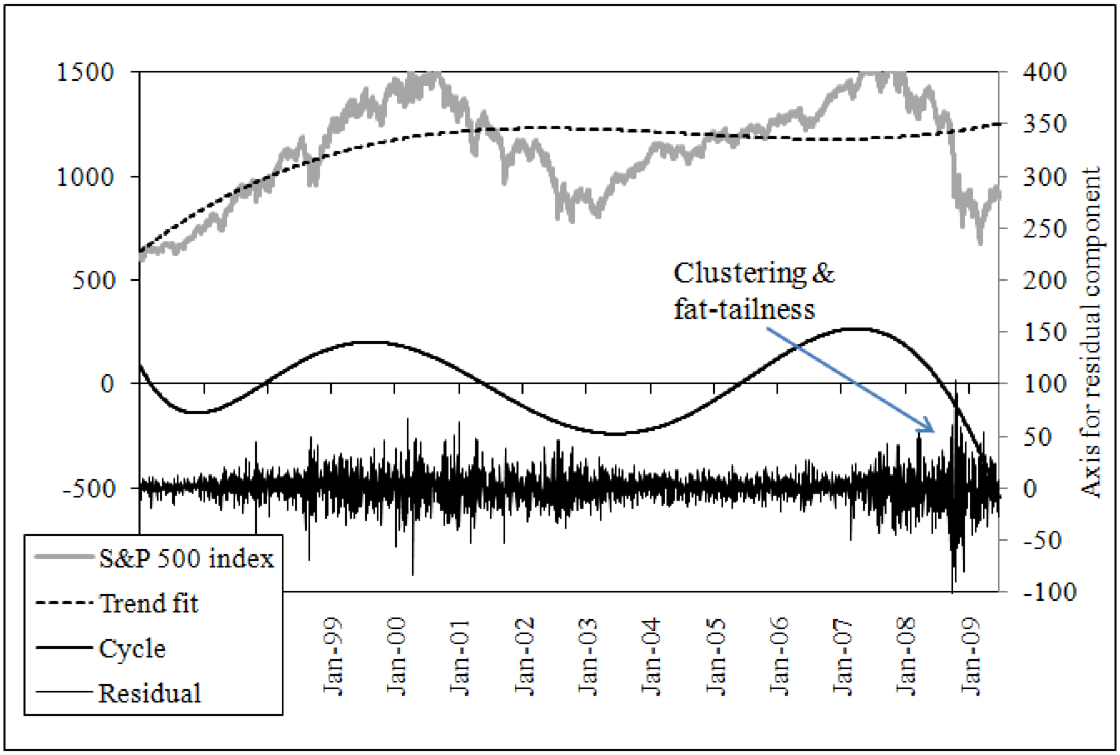

Figure 6: Illustrates how cycle breaks and compression can cause fat-tails and serial correlation in returns

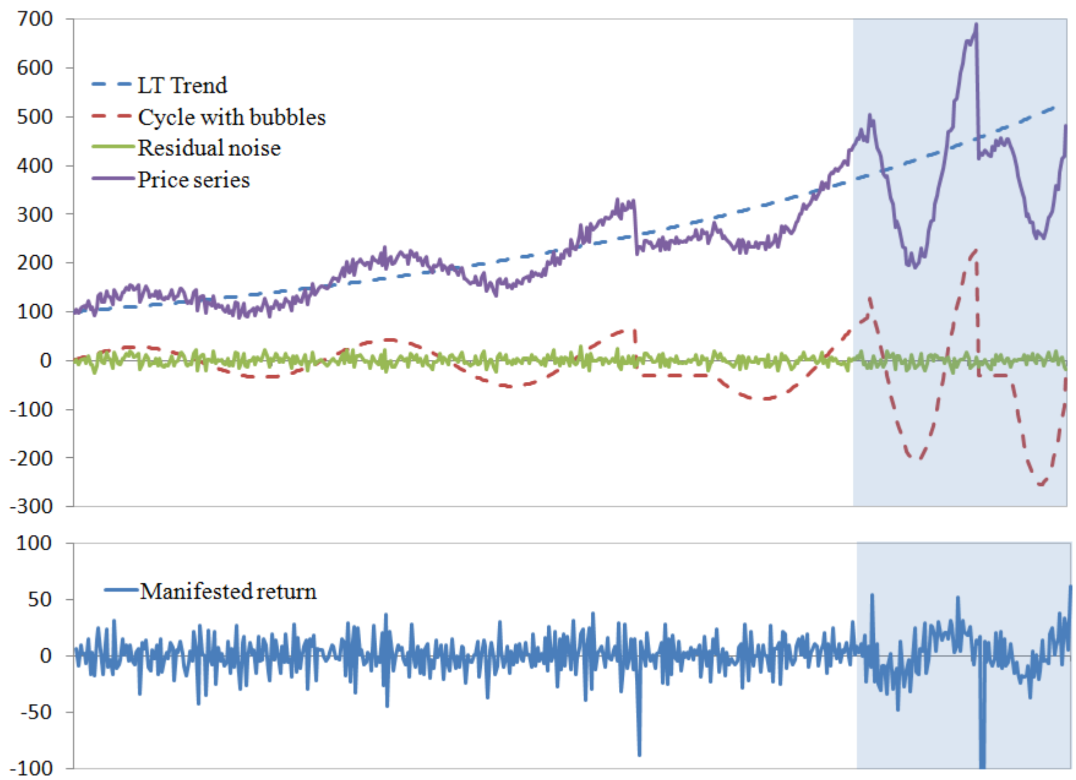

Figure 7: BuVaR vs. expected shortfall for Nasdaq index

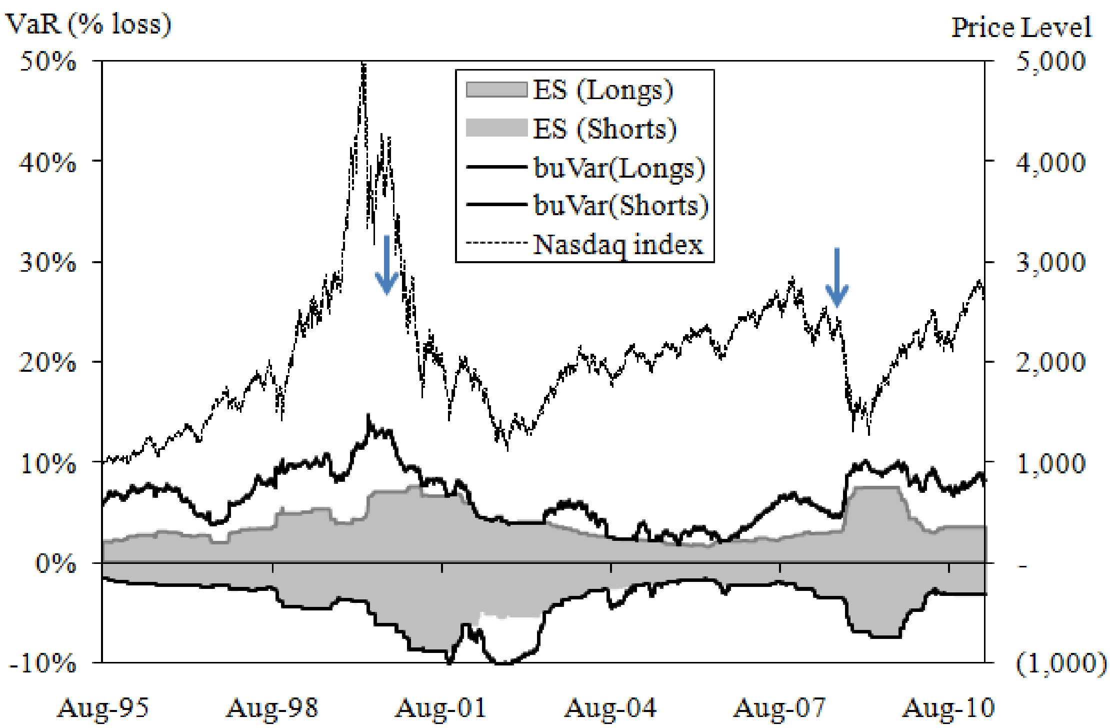

Figure 8: BuVaR vs. expected shortfall for crude oil futures (near contract)