How relevant are Black Swan and Outlier events when planning your investments?

Mohammed A. El-Kuwaiz
mohammed.el-kuwaiz@derayah.com

Summary:

Over the past few years, the ‘Black Swan’ theory has become extremely popular to explain many phenomena in the world of investing and Capital Markets. The theory stipulates that returns are not normally distributed into a neat bell-curve, and that returns have significantly more outliers than people had previously thought. The implication of this is that improbable events are not as improbably as people often calculate based on their traditional models and that traditional risk management, by using models that assume normal distributions, grossly under-estimate the risks they are bearing.

We have given this some thought and research. And while we discovered that the theory holds true in the short-term, that it does not apply for investors who seek to invest over the long run. The rationale for this is that while returns are not normally distributed for short observation periods (i.e. daily returns), containing many more outliers that is predicted in a normal distribution, as you move to longer observation periods (i.e. annually, ten-years) returns become more and more normally distributed.
Introduction: Black Swan Defined

The Black Swan Theory was penned as a concept and popularized by author and probability scholar Nassim Nicholas Taleb in his book of the same name and his previous book “Fooled By Randomness” to explain the existence and occurrence of high-impact, hard-to-predict, and rare events that are beyond the realm of normal expectations. Unlike the philosophical "black swan problem", the "Black Swan Theory" (capitalized) refers only to unexpected events of large magnitude and consequence and their dominant role in history. Such events are considered extreme outliers.

The Black Swan Theory and Black Swan Events subsequently became quite popular in the world of investing and capital markets and was often used as a criticism of traditional investment methods of looking at risk and returns, mainly through the lens of Modern Portfolio Theory (MPT). For an explanation of the criticism, please refer to the following interview with Nassim Nicholas Taleb:

**The Quarterly:** You question many of the underpinnings of modern financial theory. If you were the dean of a business school, how would you overhaul the curriculum?

**Nassim Nicholas Taleb:** I would tell people to learn more accounting, more computer science, more business history, more financial history. And I would ban portfolio theory immediately. It’s what caused the problems. Frankly, anything in finance that has equations is suspicious. I would also ban the use of statistics because unless you know statistics very, very well, it’s a dangerous, double-edged sword. And I would ban linear regression. All these things don’t work.

**The Quarterly:** What are your concerns with statistics and portfolio theory?

**Nassim Nicholas Taleb:** The field of statistics is based on something called the law of large numbers: as you increase your sample size, no single observation is going to hurt you. Sometimes that works. But the rules are based on classes of distribution that don’t always hold in our world. All statistics come from games. But our world doesn’t resemble games. We don’t have dice that can deliver. Instead of dice with one through six, the real world can have one through five—and then a trillion. The real world can do that. In the 1920s, the German mark went from three marks to a dollar to three trillion to a dollar in no time. That’s why portfolio theory simply doesn’t work. It uses metrics like variance to describe risk, while most real risk comes from a single observation, so variance is a volatility that doesn’t really describe the risk. It’s very foolish to use variance.

---

Methodology:

In conducting our research, we took the daily closing price for the S&P 500 index over 60 years (from 1/1/1950 to 31/12/2009). From this data, we calculated 5 sets of return series:

- Daily returns (14,940 observations)
- Rolling\(^2\) weekly returns (14,936 observations)
- Rolling monthly returns (14,920 observations)
- Rolling annual returns (14,701 observations)
- Rolling ten-year returns (12,497 observations)

We then calculated the mean and standard deviation for each return series and the expected observations based on the data that should lie beyond 3, 4, and 6 standard deviations from the mean, respectively (based on normal distribution). We finally compared these expected observations with the actual observations to see to what extent the data fit the prediction over various time periods.

What we found was that although there were many more outliers than would be predicted by normal distribution if you look at the daily return series, the data falls closer and closer to the bounds of normal distribution (with fewer outliers) as you look at series of longer observation periods.

This is depicted in the table below:

<table>
<thead>
<tr>
<th>Outliers /Accuracy</th>
<th>3 sigma</th>
<th>4 sigma</th>
<th>6 sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expected</td>
<td>Observed</td>
<td>Difference</td>
</tr>
<tr>
<td>Daily</td>
<td>40</td>
<td>192</td>
<td>152</td>
</tr>
<tr>
<td>Weekly</td>
<td>40</td>
<td>155</td>
<td>115</td>
</tr>
<tr>
<td>Monthly</td>
<td>40</td>
<td>156</td>
<td>116</td>
</tr>
<tr>
<td>Annual</td>
<td>40</td>
<td>28</td>
<td>-12</td>
</tr>
<tr>
<td>Ten Yearly</td>
<td>34</td>
<td>0</td>
<td>-34</td>
</tr>
</tbody>
</table>

\(^2\) We are aware of some of the issues associated with taking data for overlapping periods, but due to small number of data points in the case of non overlapping periods, we resort to this approach. Our findings also suggested that it would hold true for non overlapping period as well.
As you can see from the table, there are very few outliers in the annual and ten-year data series. This can be further demonstrated in the following graph:

![Deviation from normal distribution across different time ranges](image)

This is consistent with intuition in that, returns over shorter periods are subject to the emotions and instincts of market participants, who are all too often swayed by greed and fear, making them less rational and therefore less normally distributed. As such, capital markets at these short observation periods have been compared to a complex adaptive system that is often prone to volatile swings. But the longer term you go, the less applicable raw emotions are and the more the market is guided by more rational fundamentals that fit more with the normal distribution model.
Conclusions:

From the preceding analysis we can conclude the following:

- Investors can safely use normal distribution, bell-curve-shaped models when analyzing returns and risks over the long run (i.e. 1 year or more), since the actual data complies reasonably well with the model.

- However, investors should NOT use normal distribution, bell-curve shaped models when analyzing returns and risks over the short term (i.e. for periods and movements of less than one year, and especially for daily and weekly periods), as this could lead to systematically under-estimating the risks.

- As a result of the 2 previous points, investors should generally stay away from leverage for the purpose of investment. This is because investors who are investing their own equity can easily tolerate a high degree of volatility in the short-term in the pursuit of long term gains. However, for an investor who has leverage, especially high levels of leverage, all it takes is one of these black-swan events over a short term period to expose him to margin calls and potential liquidation of his position with a substantial loss of capital. This issue is even more true in cases where leverage comes in the form of margin or other securitized loans where the investment itself is collateral for the loan. In other instances where the loan is guaranteed by some other asset, or better yet an unrelated cash-flow, the risk is somewhat mitigated although not completely eliminated.
Empowering the Investor