

The Use and Misuse of Numbers - A Case Study of Measuring a Balanced Fund Manager's Skill

By Matthys Meintjes, Research Analyst, Advantage Asset Managers

Did you know that in South Africa, taking more risk is expected to be punished with a lower expected return? I can prove that. But, did you also know that in South Africa, taking more risk is expected to be rewarded with a higher expected return? I can also prove that.

Do these statements sound like they contradict one another?

In this article, I will discuss how they can both be true, but also both be wrong. I will also discuss a better framework in which you can evaluate future data analysis, and allow you to critically evaluate whether or not to believe the next fact 'proven' to you using some statistic or graph.

When reading articles on financial topics, or attending a presentation from one or another financial institution, we are often shown different graphs and statistics 'proving' some fact related to the topic. But, do we always understand the assumptions made when constructing that graph or statistic?

When reading about, or hearing about, some statistic or graph proving the latest fact about investments, it is always important to ask yourself three important questions:

- *Do I know what data was used, and why the specific data was chosen for the analysis?*
- *Can I explain the results from the analysis either intuitively or using theory - in other words, do the results make sense?*
- *Do I understand the assumptions under which the analysis was done, and do these assumptions hold for the period in which I want to use the results?*

To illustrate how important these questions are, and how you can go about using them to critically evaluate graphs and statistics shown to you, I will discuss a recent research study that we did to determine the value-add of the South African balanced manager, and what effect not understanding these three questions could have had on the results. I will also discuss at how very different the results from an analysis can be, by just changing the time frames. Finally, I will also look at ways of overcoming the bias you introduce into your analysis by using specific time frames.

The study

In a recent report by Allan Gray (published in their Q1 2009 newsletter) titled "Stock Selection Versus Asset Allocation in Balanced Portfolios", Maresh Cooper and Chris du Toit analysed the median South African balanced manager in terms of risk/return characteristics, and tried to attribute the median performance due to stock selection, as well as asset allocation.

Building on this study, we developed a model to address two problems that we encountered while choosing balanced fund managers to use in a multi-manager solution, namely building a benchmark against which to measure the balanced manager's performance, as well as attributing where the performance above the benchmark came from.

During the course of this study, we encountered issues that again highlighted the importance of always keeping the three questions in mind. The first of these issues was related to the data used in the study.

Understanding the data used in the analysis

As inputs into our model, we decided to use managers and performances from the Alexander Forbes Large Manager Watch survey. Because of limitations in the data, and to make the results comparable between managers, we decided to only focus on local equity, bonds and cash. This meant that we could not use the performances of each manager as stated in the survey, because we had to strip out performances from global building blocks, property as well as other asset classes.

This extraction of the performances introduced additional error into the data. While we regard this as unavoidable, by doing this it became very important that we could answer the question: *Do I know what data was used, and why the specific data was chosen for the analysis?*

By not answering this question, we would run the risk of making conclusions based on the results that are not entirely relevant, because we don't understand the data that were used to generate the results.

In addition, after doing this we also need to answer the following question: *Do I understand the assumptions under which the analysis was done, and do these assumptions hold for the period in which I want to use the results?*

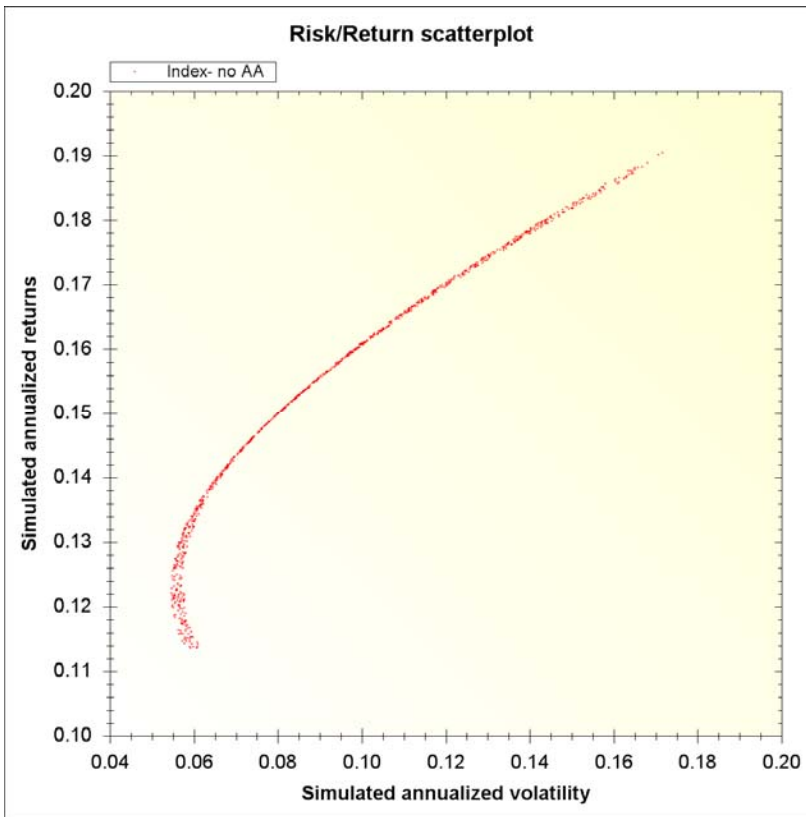
When we stripped out the performances from all the asset classes except local equities, bonds and cash, we made an implicit assumption that there were no intra-month effects that could add or detract from performance. While this means that there will always be a mismatch between our built-up performances and the actual performances of the managers, the important part is that we explicitly state the assumption that we made. This allows us to better understand under which conditions the model holds, as well as where the model may be improved in the future by relaxing the assumptions.

The next issue that we encountered is a very common issue in any analysis involving historical data:

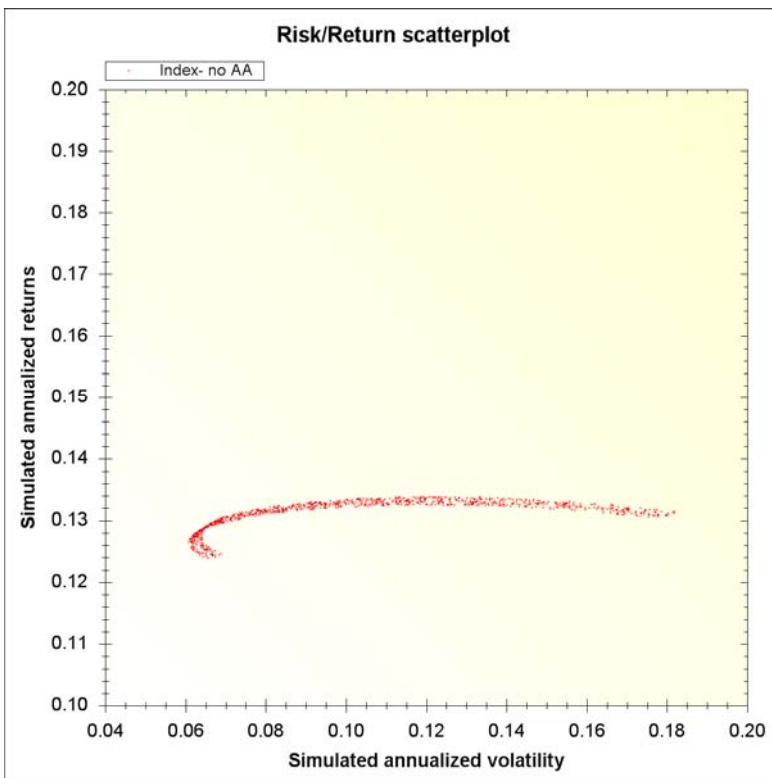
Which timeframe to use

The question of which timeframe to use is extremely important, as the timeframe you use can greatly impact the results you achieve. Remember when I said that I could prove that taking risk is being penalized in South Africa, as well as prove that taking risk is being rewarded. Let me show you how that can be, and what we did to remove this contradiction from our results.

Have a look at the following graph, which gives an indication of the results we achieved when we did the study up to June 2008, using only the index returns and allocations.



The graph looks like you would expect, with an upwards slope. In other words, this graph tells you that more risk will be rewarded with a higher expected return, exactly as modern portfolio theory states. But, now look at the graph of the index returns and allocations, but up to March 2009.



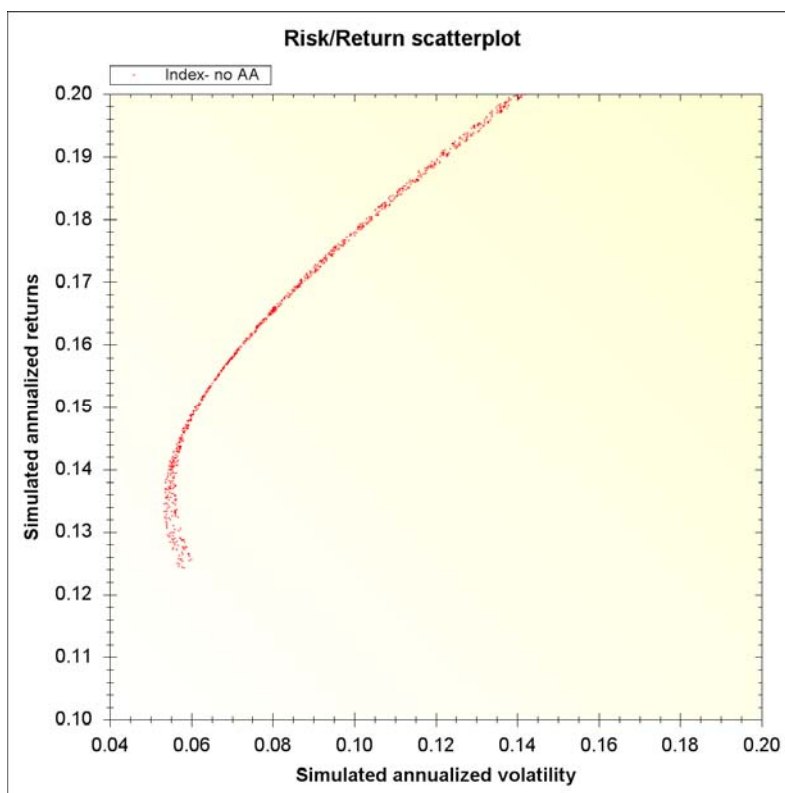
This graph, which uses the exact same data, just for a longer time frame, basically tells you to throw all of the theory out of the window. The downwards sloping trend indicates that by taking more risk, you are actually being penalized by having a smaller expected return.

While this is what happened in the past 9 years, this is not what we should be expecting to happen. Before just taking the results at face value, you need to ask yourself: *Can I explain the results from the analysis either intuitively or using theory - in other words, do the results make sense?*

When doing any data analysis using historical data, we always need to use common sense and theory to rationalise our findings. In other words, if we want to build a model to give us an indication of which managers should be expected to add value above a passive benchmark, we need to do something about the data to account for the extreme outliers embedded in the data. Remember, in statistical terms, history is but one sample out of a population of infinite size - *what happened happened, but don't expect it to happen in exactly the same way again.*

To account for this problem, and because we can't go back in time and resample history again, we used a technique called bootstrapping. I won't go into the exact details of how this technique works, but in short, bootstrapping is a statistical technique that allows one to generate a range of possible 'what-ifs'. By using this technique, we created 1000 different time series of possible returns, or in other words, we created 1000 different histories of what could have happened, under the same conditions as the previous 9 years.

The following graph gives an indication of the results after using the bootstrapping technique.



The results on this graph make sense, and we can explain it using modern portfolio theory. Thus, we are happy using these results, because we can intuitively understand the results, and we can explain it using theory.

But, again we also need to answer the question: *Do I understand the assumptions under which the analysis was done, and do these assumptions hold for the period in which I want to use the results?*

By using additional statistical techniques, like bootstrapping, we are also introducing additional assumptions into our model. It is always important to fully understand the assumptions that these techniques introduce to our model, to ensure that we understand under which conditions the model would work.

In Conclusion

So, in conclusion - what did we learn about the issues that need to be investigated when doing any data analysis, and when interpreting any results?

Firstly, do not just take fancy graphs and results at face value. Always interrogate the results, and determine whether or not they make sense. Results that do not make intuitive sense, and that cannot be explained using theory, should always be regarded with extreme suspicion.

Secondly, you should always ensure that you understand the data that was used to generate the results. Different datasets and different timeframes can generate widely different results. To properly use the results, it is very important to understand how the data were used, and why the specific data and timeframe were chosen.

Lastly, it is very important to always understand the assumptions under which the results were obtained. Understanding these assumptions will allow you to make a better judgement about whether or not the results are relevant to your specific case.

ENDS