Python-implemented Techniques for Reading the 'Tea Leaves' of Past Investment Performance & Risk Management of Funds

Table of Contents

Abbreviations

bp, Basis Point

bps, Basis Points

CDF, Cumulative distribution function

EP, Existing Portfolio

IM, Investment Management Firm

NAV, Net Asset Value

NS, New Trading Strategy

p.a., per annum

PDF, Probability density function

PRIIPs, Packaged Retail and Insurance-Based Investment Products

UCITS, Undertakings for Collective Investment in Transferable Securities

Acknowledgements

The authors wish to sincerely thank:

(i) Brian McCarthy, Will Sparks, and Henry Xu for their comments on different sections of the paper and their helpful suggestions to make the paper and the accompanying Python program easier to read;

(ii) Emily Shiel for her help in checking much of the output of the program independently using Microsoft Excel spreadsheets and the cross references in the paper; and

(iii) Andrew Caslin for his help with the proof reading of the paper.

All remaining errors are those of the authors.

The views expressed in the paper are not necessarily those of the Society of Actuaries in Ireland nor those of the employers of the authors.

Contact Information

John Caslin: john@caslin.eu

David Kavanagh: dwckavanagh@gmail.com

Executive Summary

The paper, the data sets, and the Python code accompanying the paper explore a wide range of graphical and numerical techniques for analysing the past performance of an investment fund in an attempt to validate the claims of an investment management firm, assess the quality of the fund's risk management, generate agenda points for a due diligence meeting with the investment manager, and assess the efficacy of adding an investment fund to an existing portfolio.

Spreadsheet Analyses
The production of the range of graphs and the statistics needed for such analysis from a data file containing the daily returns of an investment fund is an extremely labour intensive task particularly where: (i) several funds with different trading history lengths and different currencies of denomination are involved; and (ii) spreadsheets are used to carry out the analysis.

Python
Python together with its extensive range of libraries is closer to how analysts think, it is faster to build than spreadsheets, is more flexible, and it reduces the scope for error. Perhaps most importantly of all, it has a relatively shallow learning curve. Python can be learnt for free on-line through massive open on line courses and the web provides a phenomenal range of answers to virtually every problem one is likely to encounter.

No Single Statistic for Selecting Funds
There is no single statistic that might confirm a decision to invest in a fund. The implications of a wide range of statistics needs to considered as part of an analysis of fund when considering it for investment.

The range of statistics covered in the paper and the program broadly focus on the analysis of: (i) return; (ii) risk; and (iii) those that may be useful in portfolio construction where decisions such as whether to add a new fund to an existing portfolio must be considered.

The return statistics examined in the paper and the programs include rolling, n-year rates of return, the empirical probability density function of daily returns, empirical plot of ordered pairs, the number of days accounting for a percentage of total return, and the Omega ratio.

The risk statistics examined in the paper and the programs include the skewness of daily returns, the kurtosis of daily returns, extreme standardised daily returns, maximum peak-to-trough fall in value, and rolling n-day volatility.

The correlation statistics examined in the paper and the programs include not just correlation of returns but correlation of returns squared which is a proxy measure for volatility correlation.

Closing Remark
Finally, we are not suggesting that the range of statistics in the paper is exhaustive when considering a fund for investment. We are simply illustrating some of the statistics and how

to produce them quickly and efficiently in a robust manner using Python and its range of libraries.

Introduction

Investment due diligence on a trading strategy has two major facets: quantitative due diligence and qualitative due diligence.

This paper focuses almost exclusively on the quantitative aspects of investment due diligence on a trading strategy. It does not cover in any significant detail the qualitative aspects of investment due diligence and it does not cover the implications of such an analysis for the process of selecting an investment management firm ("IM"). For an insight into qualitative due diligence, the reader is referred to an earlier paper which was presented to the Society of Actuaries in Ireland on 13 February 2018 and which was entitled: "What questions might pension funds, insurance companies & investment advisors reasonably ask in an investment due diligence process?

Graphical and numerical analysis of past returns are essential to developing the due diligence questions in relation to past performance and risk management and ultimately to deciding to allocate assets for management to an IM or not.

The due diligence team ought to request daily price data as it gives more observations for assessing performance and risk management.

Enquiries ought to be made as to the valuation time upon which the daily prices are based as this is important for correlation analysis with other funds and indices.

The due diligence team ought to check that the valuation of the fund is carried out independently of the IM of the fund using independent sources of information on holdings in the fund and of their prices.

Software

While spreadsheets could be used to carry out the analysis in this paper, they require some reconfiguring for each new data set if only because the length of the data set varies.

The authors have developed a software program written in Python which can analyse any data set provided it is in a relatively straight forward format and propose to make the software readily available via the website of the Society of Actuaries in Ireland.

Description of Data File used by Python

To use the Python programme to carry out the analysis of past performance and risk management, each fund's daily prices must be on a separate spreadsheet and the sheet's name should reflect the name of the fund.

For each fund, column A in the spreadsheet should contain the dates and column B should contain the prices. Row 1 should contain the column titles, namely, date and price; and the dates and prices should start on row 2.

Overview of the Programme's Output

The software allows due diligence teams to obtain a battery of performance, risk, risk-return trade-off, and correlation statistics in relation to a fund in a matter of seconds.

The paper discusses the possible implication of the battery of statistics. A number of data sets are used to illustrate possible interpretations of the data.

Data Sets

In this paper, we examine three different types of funds. A brief description of the investment objectives, instruments used, and investment time horizon of each fund is given below. Each fund is assigned a label which is used to refer to that fund throughout the paper.

IGC Bond Fund

The objective of IGC Bond Fund is to maximise the total return in a manner that is consistent with the preservation of capital by investing in a diversified portfolio of investment grade fixed income securities across the globe.
For efficient portfolio management purposes, the fund may use a range of fixed income derivatives including bond futures, interest rate swaps, total return swaps, money market derivatives, mortgage derivatives, credit default swaps, and options.
The portfolio management is discretionary rather than systematic.
The fund does not target any specific absolute return nor does it target any specific risk range.

European Equity Fund

The objective of the European Equity Fund is to outperform the FTSE World Ex UK Total Return index in EUR over rolling three year periods by holding a highly-concentrated portfolio of companies domiciled in or having significant operations in Continental Europe.
For efficient portfolio management purposes, the fund may use a range of derivatives including options, futures, forward contracts, and swaps on financial indices and currencies.
The portfolio management is discretionary rather than systematic.
The fund does not target any specific absolute return nor does it target any specific risk range.

Multi-Asset Fund

The objective of the fund is to provide capital growth by investing:

(i) at least 50% in government bonds, corporate bonds of investment grade credit ratings and high yield bonds;
(ii) at most 20% in equities; and
(iii) any balance in cash and money market instruments.

Derivative instruments may be used for both investment purposes and efficient portfolio management purposes. The fund may use an extensive range of derivatives.
The fund may use derivatives to leverage the exposure of the fund up to twice its value.
The portfolio management is discretionary rather than systematic.
The fund does not target any specific absolute return nor does it target any specific risk range.

Data Tests

A number of tests ought to be run on the data to ensure that it is in fact daily data. Three statistics that may be helpful in this regard are as follows:

1. The maximum number of days between pricing dates;
2. The average number of days between pricing dates; and
3. The standard deviation of the number of days between pricing dates.

For a data set of at least 1,000 days, if the maximum value of the three statistics quoted above is greater than 5 days, 1.45 days, and 0.9 days respectively a closer examination of the data may be required to ensure that there are no major gaps in the daily returns history.

Some data sets quote a NAV on each weekday regardless of whether it is a valuation day. Such data sets can usually be identified by figure for the maximum number of days between pricing dates being equal to 3. Such data sets may also have lower values of the other two statistics than data sets that only quote a NAV for fund valuation days.

A snippet of the Python code to run the tests on a data set is shown below.

Python Code Snippet

```
max_price_gap_days = ((df_prices.Date - df_prices.Date.shift(1)) / np.timedelta64(1, 'D')).max()
min_price_gap_days = ((df_prices.Date - df_prices.Date.shift(1)) / np.timedelta64(1, 'D')).min()
mean_price_gap_days = ((df_prices.Date - df_prices.Date.shift(1)) / np.timedelta64(1, 'D')).mean()
std_price_gap_days = ((df_prices.Date - df_prices.Date.shift(1)) / np.timedelta64(1, 'D')).std()
```

Table 1 below summarises the tests results for the three funds considered in this paper.

Table 1

Fund	Number of Daily Returns	Maximum Number of Days between Pricing Dates	Average Number of Days between Pricing Dates	Standard Deviation of Number of Days between Pricing Dates
IGC Bond	3,780	3	1.40	0.800
European Equity	2,743	5	1.45	0.889
Multi-Asset	2,751	5	1.44	0.888

The data set for the Multi-Asset fund lists NAVs for actual fund valuation dates. By way of contrast, the data sets for the other two funds list NAVs for every weekday whether that weekday is a valuation day for the fund or not. This feature of the IGC Bond fund and the European Equity fund is identified by the fact that the maximum number of days between pricing dates is just 3.

'Cleaning' the Data Set to Remove NAVs for Days which were not Valuation Dates

If the data set contains NAVs for every weekday whether that weekday is a valuation day for the fund or not the maximum number of days between pricing dates will show up as 3. While it is possible to identify some weekdays which are probably 'redundant' NAVs, Christmas Day and New Year's Day are examples of such days, it is difficult to remove such 'redundant' NAVs without knowing the exact days on which the fund was not priced.

The problem is particularly acute for funds with low mean daily returns and low standard deviations of daily returns such as highly-rated government bond funds. Using a data file of public holidays in the jurisdiction of the administrator of the fund and identifying unchanged NAVs across consecutive days, a reasonable attempt can be made to eliminate a high proportion of the redundant NAVs.

The 'cleaning' process removes redundant trading days which have zero return. The impact of a 'cleaning' process like that outlined above on some of the major statistics we examine in this paper is shown in Table 2.

Table 2

Statistic	Impact of Cleaning Process
Number of days in data set	REDUCED (zero return days removed)
Maximum peak-to-trough fall in value	UNCHANGED (uses only two NAVs in data set which are never redundant)
Omega ratio	UNCHANGED (only uses positive and negative returns)
Mean daily return	INCREASED (sum of daily returns unchanged but number of days reduced)
Annualised return	Generally unchanged unless the first or last day in the data set is removed in the cleaning process.
Average value of rolling 20-day volatility	INCREASED (fewer 20-day periods with zero returns)

We examined the impact of such a cleaning process for the European Equity fund. While the number NAVs was reduced from 2,831 to 2,743 there were no major differences between the cleaned and the original data set.

Understanding of the Investment Strategy

It's essential to understand the investment strategy of the fund and how it has changed over time in studying the past performance.

To examine or monitor the performance of an IM, one needs to have some broad concept of the likely distribution of the investment strategy's daily returns. Is the distribution of daily returns close to the normal distribution, or is it characterised by a small number of large positive daily returns and large number of near zero or small negative returns or vice-versa?

Frequency of Valuation

If a trading strategy has operated for a relatively short period of time there may only be say, six observations of daily performance available. One would be concerned about the validity of any conclusions drawn from a sample of size six. However, in six months there are approximately 120 trading days, and that is a much larger sample from which to draw conclusions regarding investment performance and risk management.

Number and Distribution of Trades

Even if one has say 1,000 observations of daily data, this may not be enough from which to draw any conclusions. The 1,000 observations may be based on a limited number of large trades. Where an IM's returns are based on a small number of large trades it is difficult to distinguish lucky IMs from skilful IMs.

By contrast, there is a much better chance of distinguishing luck from skill in cases where the 1,000 observations are based on many trades of roughly equal size. Put simply, it is important to understand how many trades the daily returns data represent and the distribution of the size of trades.

Calendar Days and Trading Days

For the purpose of defining daily returns, we distinguish between trading days and calendar days. Weekdays which are not public or other types of holiday are normally trading days whereas Saturdays and Sundays are generally not trading days in the major markets of the world. Two successive trading days may: (i) co-inside with two successive calendar days; or (ii) be separated by one or more calendar days.

Daily Return - Definition

The return for a trading day may defined as, r_t, the ratio of the NAV of the fund on trading day (t) to the NAV on trading day (t-1) less 1. However, in this paper we define the return for a trading day as the log to the base e (ln) of the ratio of the NAV on trading day (t) to the NAV on trading day (t-1), i.e. $\ln(NAV_t/NAV_{(t-1)}) = \ln(1+r_t)$. In this context, it may be useful to recall that if r_t is small, $\ln(1+ r_t)$ is approximately equal to r_t.

Python Code Snippet

```
df_returns['Return'] = np.log(df_returns.Price) - np.log(df_returns.Price.shift(1))
```

In terms of operations involving $\ln(1+r_t)$ in the Python programme accompanying this paper, it is also useful to note that $\ln(1+r_t) = \ln(NAV_t/NAV_{(t-1)}) = \ln(NAV_t) – \ln(NAV_{(t-1)})$.

For numerical tests on $\ln(1+r_t)$ such as checking if $\ln(1+r_t) =1$, checking if:

$$\ln(NAV_t) – \ln(NAV_{(t-1)}) = 0$$

is likely to be numerically safer when used in a computer programme.

All daily returns used and all graphs and statistics produced by the Python program are the log to the base e of one plus the actual daily return, $\ln(1+ r_t)$. Thus, when the program takes the sum of terms of the form $\ln(1+ r_t)$, it is effectively chain linking the product of terms of the form $(1+ r_t)$ when the result is transformed using the exponential function.

What is a Daily Return?

While a daily return is defined as the $\ln(NAV_t/NAV_{(t-1)})$, we have not considered questions like:

1. Is the change in NAV between Friday's valuation and the following Monday's valuation a daily return?
2. If New Year's Day, a day when most major markets around the world are closed, fell on say a Wednesday, is the change in NAV between Tuesday's valuation and Thursday's valuation a daily return?
3. Suppose there are two public holidays in a row such as Christmas Day and St. Stephen's Day and these two public holidays fell on Monday and Tuesday. If the fund were valued at the close of business on the preceding Friday and next valued on the succeeding Wednesday, is the change in NAV between the Friday and the Wednesday a daily return?

Daily returns ought to be reasonably uniform in nature. Where there are several calendar days between trading days due to a run of bank holidays and a weekend, the return generated between close of business on trading day (t-1) to trading day t may not be representative of a typical trading day simply because of the range of news events which might impact the NAV of a fund in the intervening number of calendar days.

In this paper and the accompanying Python program, we have regarded the returns in the three cases listed above as daily returns. However, some analysts may develop rules to eliminate daily returns unless they meet certain criteria. A sample of such criteria is as follows:

- Remove the 'daily return' from the data set if the number of calendar days between[1] trading day t and trading day (t-1):
    - Is greater than 4; or
    - If it is equal to 3 and trading day t is NOT a Monday; or
    - If it is equal to 4 and trading day t is NOT a Monday or a Tuesday.

Using these sample criteria, the answer to questions 1, 2, and 3 above are YES, YES, and NO respectively.

Length of Track Record

The due diligence team will wish to assess the length of the track record of the fund. This statistic ought to be calculated in either years or months. The length of a track record is needed to properly compare the maximum peak-to-trough fall in value of the strategy with other strategies as the maximum peak-to-trough fall in value is a time dependant variable. Track record length is also an indicator of the size of the sample and therefore the reliability of any conclusions drawn. The number of daily returns is one less than the number of NAV prices as x consecutive, trading-day, NAV prices only permit the computation of (x-1) daily returns.

A snippet of the Python code to compute the mean number of trading days per year is shown below. One would expect the answer to be between 252 days (equity funds) and 259 days (currency funds).

Python Code Snippet

```
start_date = df_prices.Date.min()
end_date = df_prices.Date.max()
term_days = (end_date - start_date) / np.timedelta64(1, 'D')
term_years = term_days / 365.25
num_returns = len(df_prices) - 1
mean_days_per_year = num_returns / term_years
```

Table 3 summarises the start dates, end dates, and number of daily returns for the three funds.

[1] For the purpose of this test, the number of days between: (i) close of business on say a Friday trading day where the following Monday is a bank holiday; and (ii) close of business on the next succeeding trading day, Tuesday, is regarded as a difference of three calendar days.

Table 3

Fund	Start Date of NAV Prices	End Date of NAV Prices	Length of Track Record (Number of Daily Returns)
IGC Bond	15 September 2003	12 March 2018	3,780
European Equity	1 June 2007	9 April 2018	2,743
Multi-Asset	20 April 2007	8 March 2018	2,751

Performance Metrics

Mean and Median Daily Return

The mean of the daily returns should be calculated and the median daily return identified. As a descriptive statistic, the mean daily return is of limited value if a distribution has more than three quarters of its observations above or below the `average'. So, it is important to measure the degree of skewness[2] of the distribution of daily returns before relying on the mean as a descriptive statistic for a distribution. In a highly skewed distribution, the median is a much more descriptive statistic than the mean. There is a pervasive rule of thumb which states: *if the mean exceeds the median, the distribution is said to have positive skewness, and if the mean is less than the median, it is said to have negative skewness*. It is important to be aware that this rule of thumb does not work in all cases. In the case of the three funds we examine in this paper, the IGC Bond fund is an example where the rule of thumb does not work as the mean exceeds the median yet the skewness of daily returns is negative. For quantitative due diligence purposes, the skewness of daily returns ought to be calculated using the statistical formula rather than rely on the rule of thumb.

A snippet of the Python code to compute the above statistics is shown below.

Python Code Snippet

```
mean_return = df_returns.Return.mean()
mean_return_annual = mean_return * mean_days_per_year
median_return = df_returns.Return.median()
median_return_annual = median_return * mean_days_per_year
std_return = df_returns.Return.std()
std_return_annual = std_return * mean_days_per_year ** 0.5
skew_return = df_returns.Return.skew()
kurt_return = df_returns.Return.kurtosis()
```

Table 4 shows the mean and median daily returns and the skewness of the three funds.

Table 4

Fund	Mean Daily Return (bps)	Median Daily Return (bps)	Skewness of Daily Returns
IGC Bond	1.957	0	-0.609
European Equity	2.106	8.072	-0.522
Multi-Asset	2.769	4.618	-0.754

[2] While some distributions of daily return data may be symmetric about a central value, others may exhibit a pattern where the left and the right tails of the distribution are not perfectly symmetric. Skewness is a measure of the extent of asymmetry in a distribution. A distribution of daily returns that is skewed to the right has a long tail that extends to the right and is termed positively skewed.

Charts 1, 2, and 3 below plot the empirical probability density functions for the three funds.

Chart 1

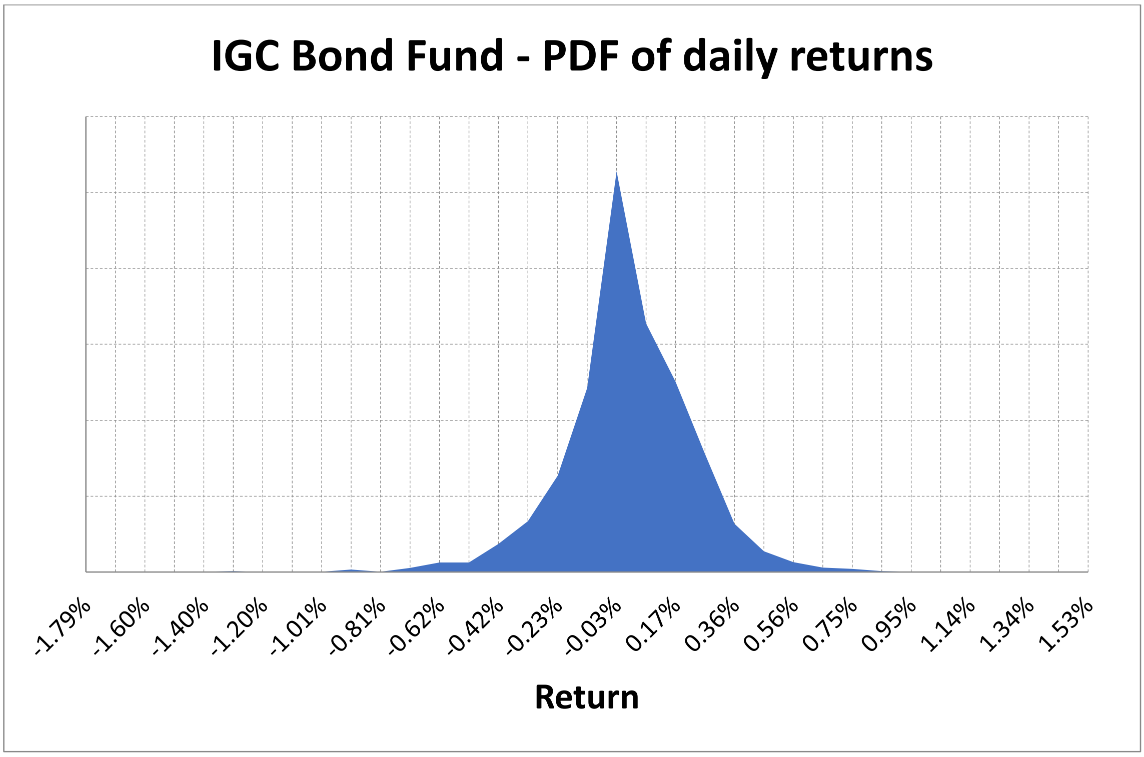


Chart 2

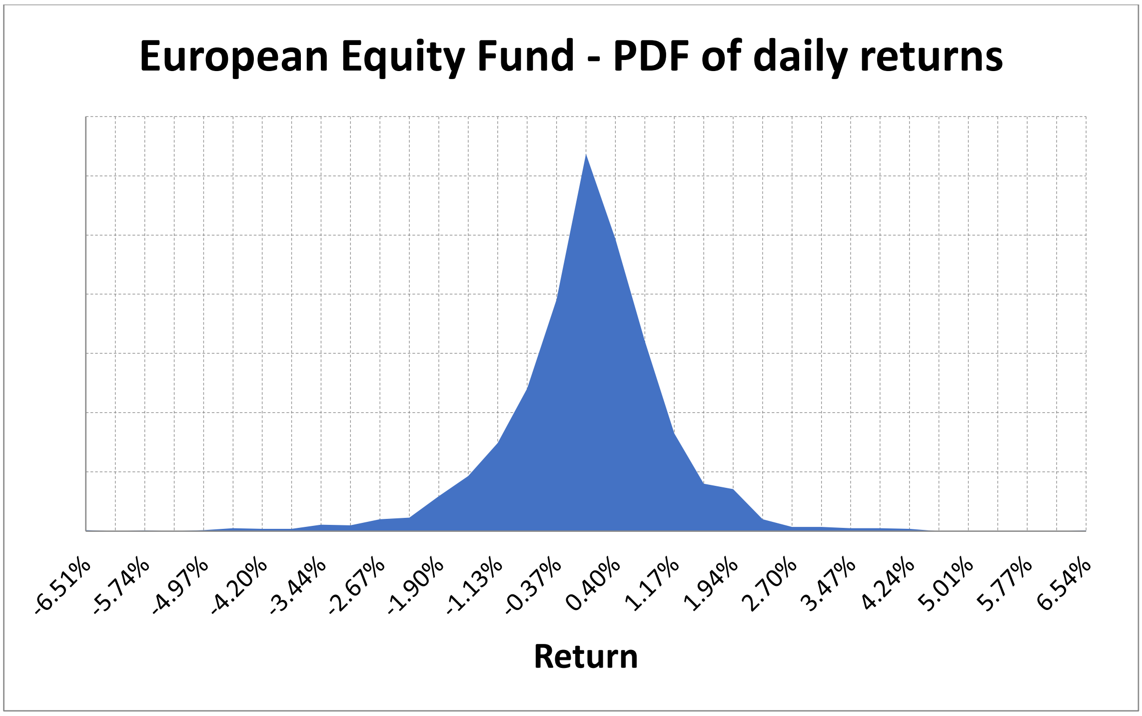
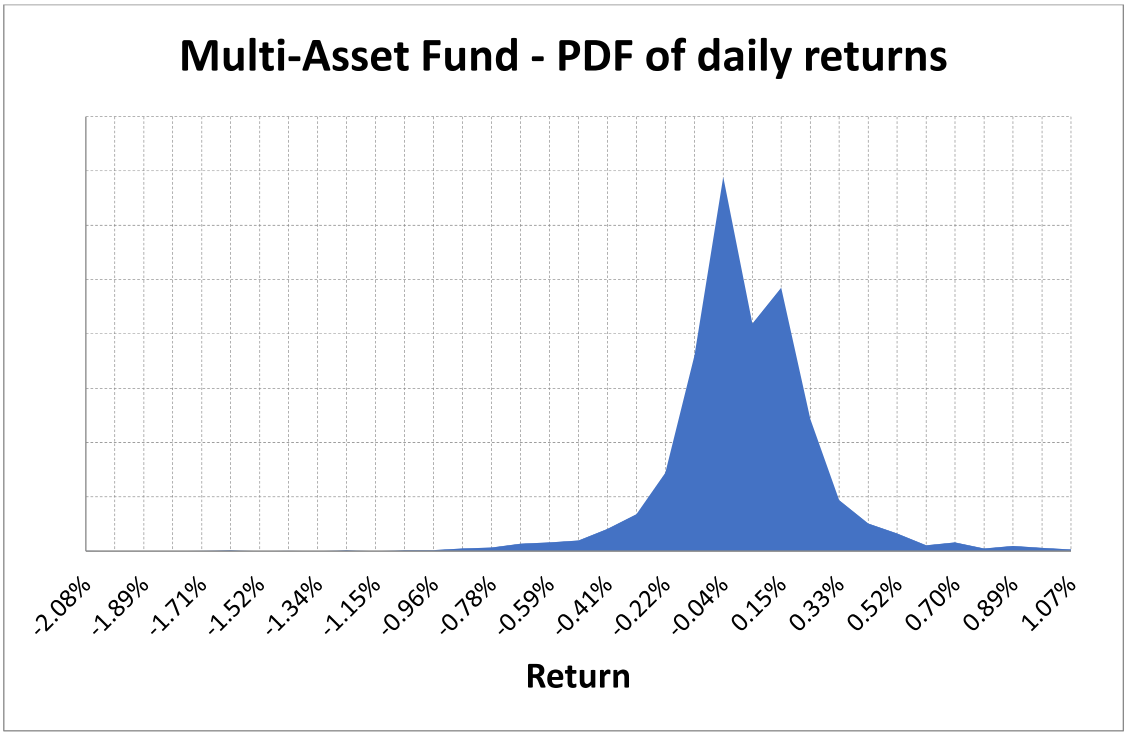


Chart 3



Plot of Ordered Pairs

Plot the ordered pairs (highest daily return, lowest daily return), (second highest daily return, second lowest daily return), etc. and examine the slope of the trend line to further assess skewness.

Chart 4

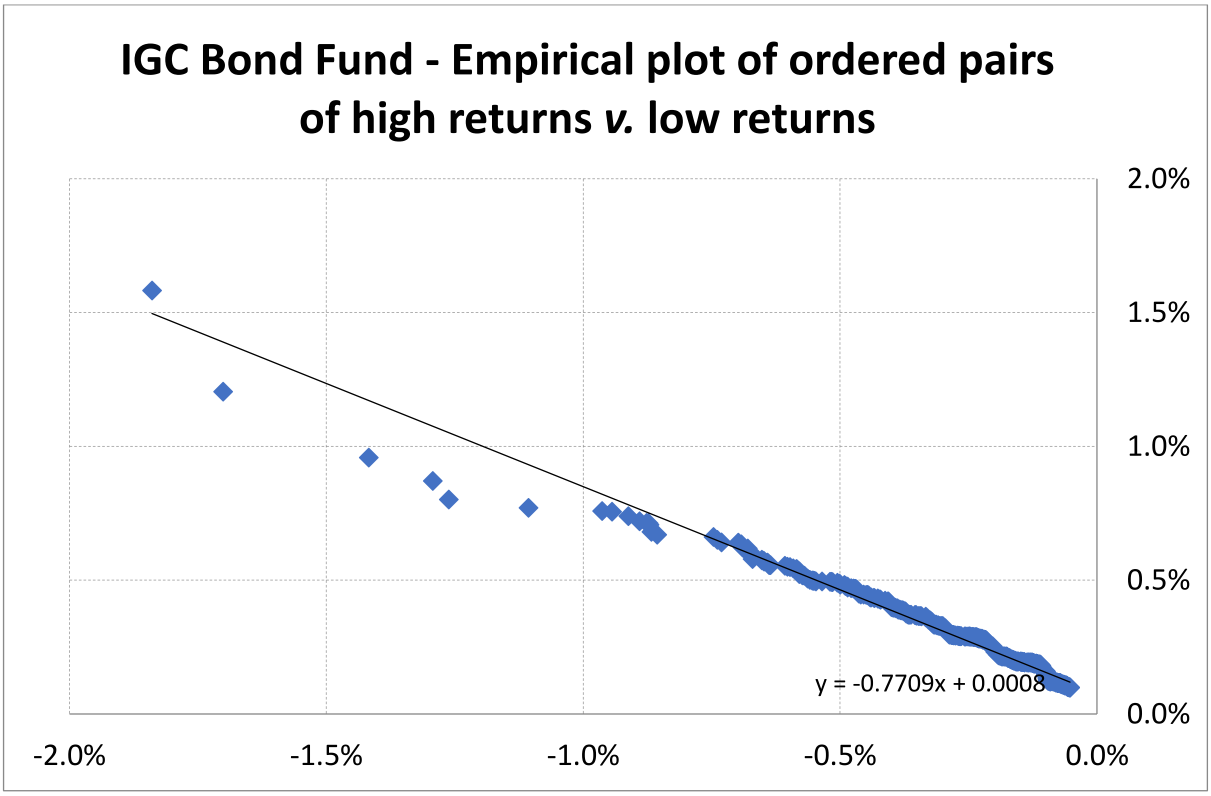


Chart 4 shows that there are a number of ordered pairs, (lowest, highest), (second lowest, second highest), etc., of daily returns for which the absolute value of the large negative daily returns is greater than those of the large positive daily returns.

Chart 5

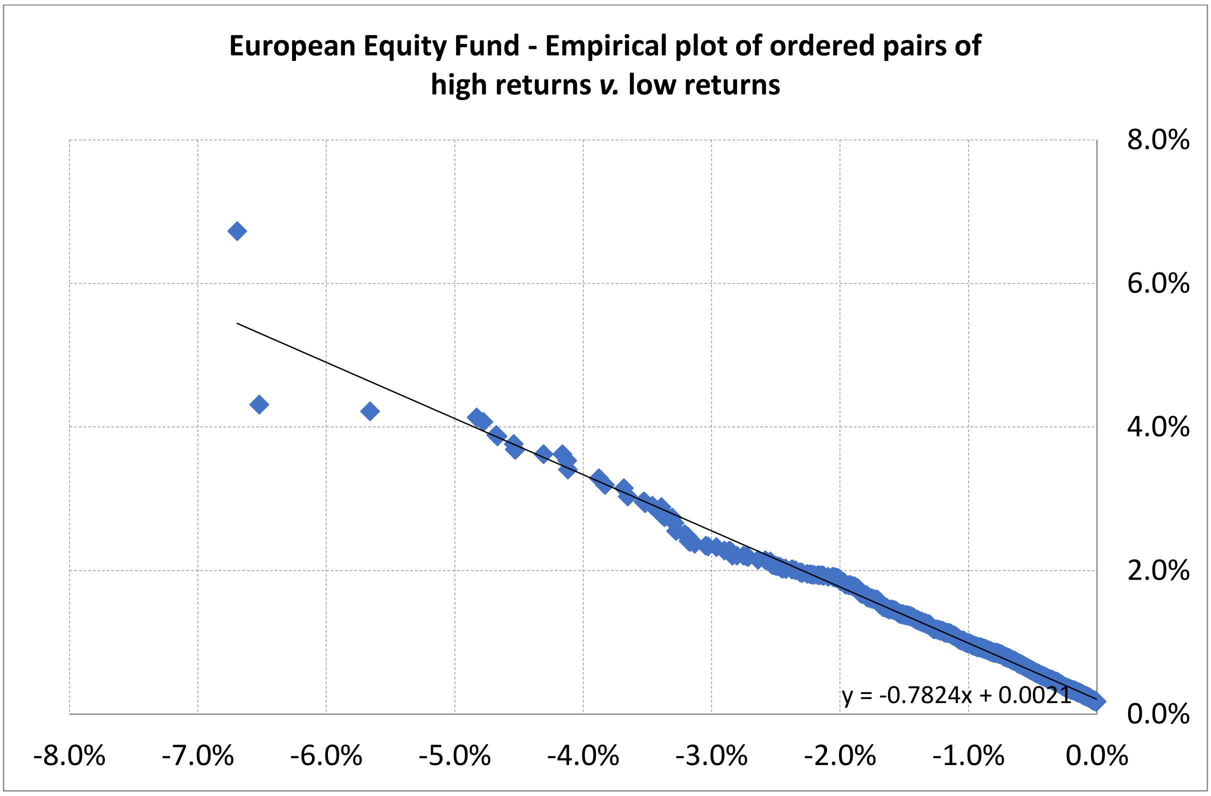


Chart 6

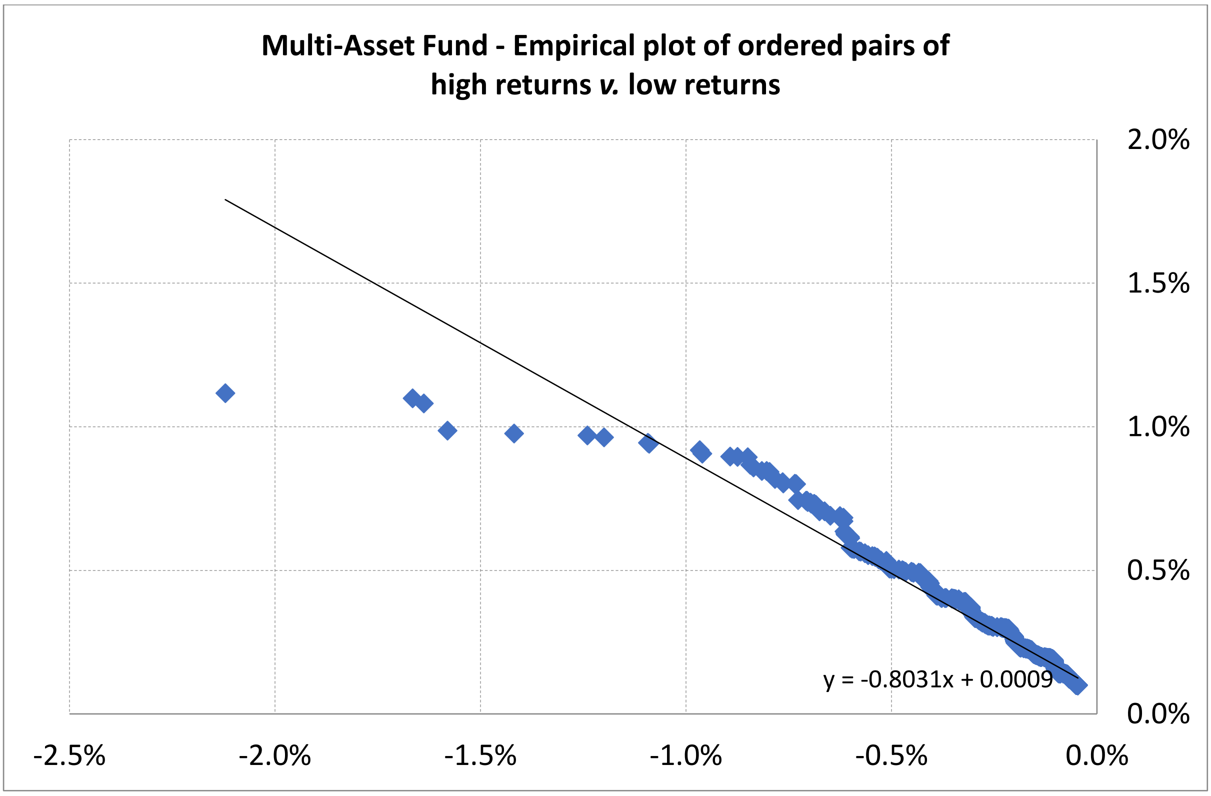


Chart 6 shows that there is a strong trend for the ordered pairs, (lowest, highest), (second lowest, second highest), etc., of daily returns to illustrate higher absolute values of large negative daily returns than large positive daily returns.

Charts 4, 5, and 6 illustrate that when the absolute value of the daily return is large, it tends to be more often negative than positive and this is particularly the case for the multi-asset fund. Further, we have fitted a regression line to the points in the three charts of plots of ordered returns. The slope of all three lines is negative and the absolute value of each slope is less than one indicating that most positive returns are of a lesser absolute value than that of most negative returns.

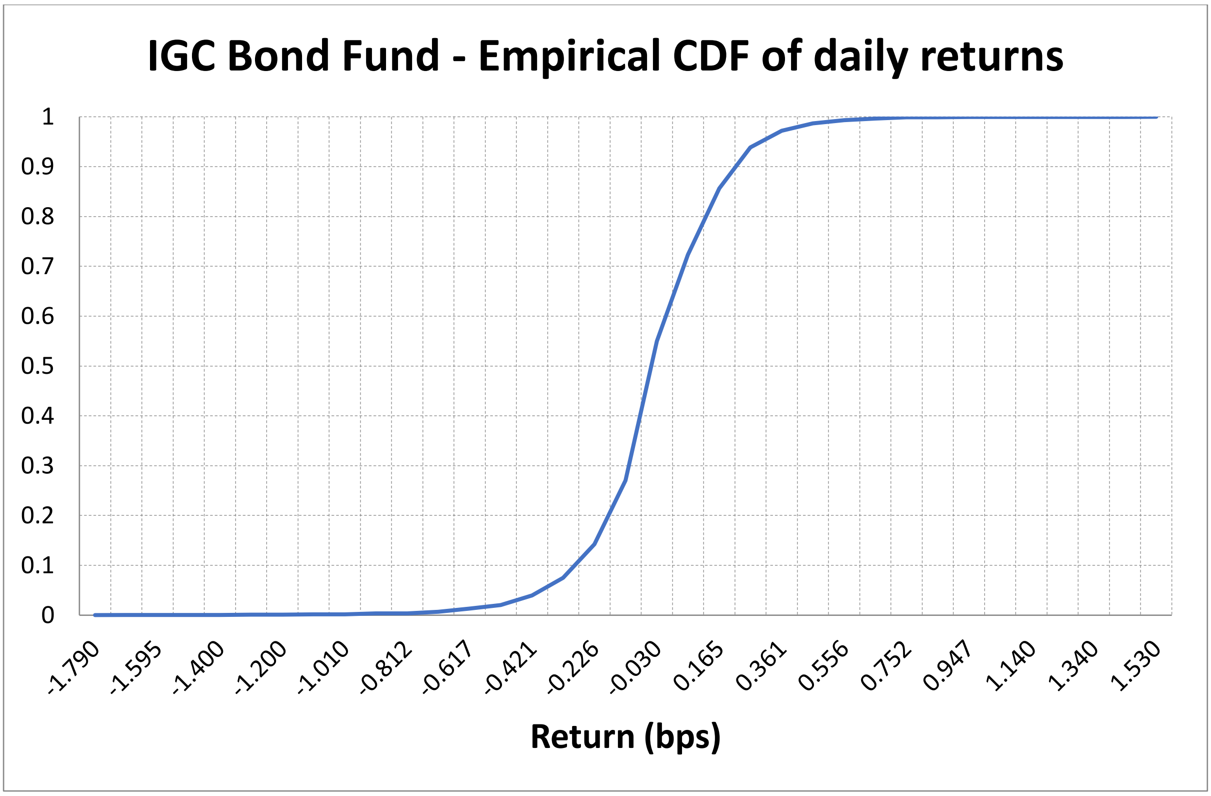
Empirical Cumulative Distribution Function

The empirical cumulative distribution function of a data set can be sketched as follows:

(1) Rank the observations from the lowest to the highest.
(2) For each ordered value of the data, calculate the number of ordered data values less than or equal to that value. Call this number i_j.
(3) For each ordered value of the data, calculate the proportion of ordered values less than or equal to that value.
Call this proportion i_j/n, where n is the total number of observations.
The empirical cumulative distribution function, $F_n(x) = i_j/n$.
Plot the values i_j/n against their corresponding ordered value $X_{(j)}$, where $X_{(j)}$ is the jth ordered value of the sample. The general co-ordinate of a point on the plot is $(X_{(j)}, j/n)$.

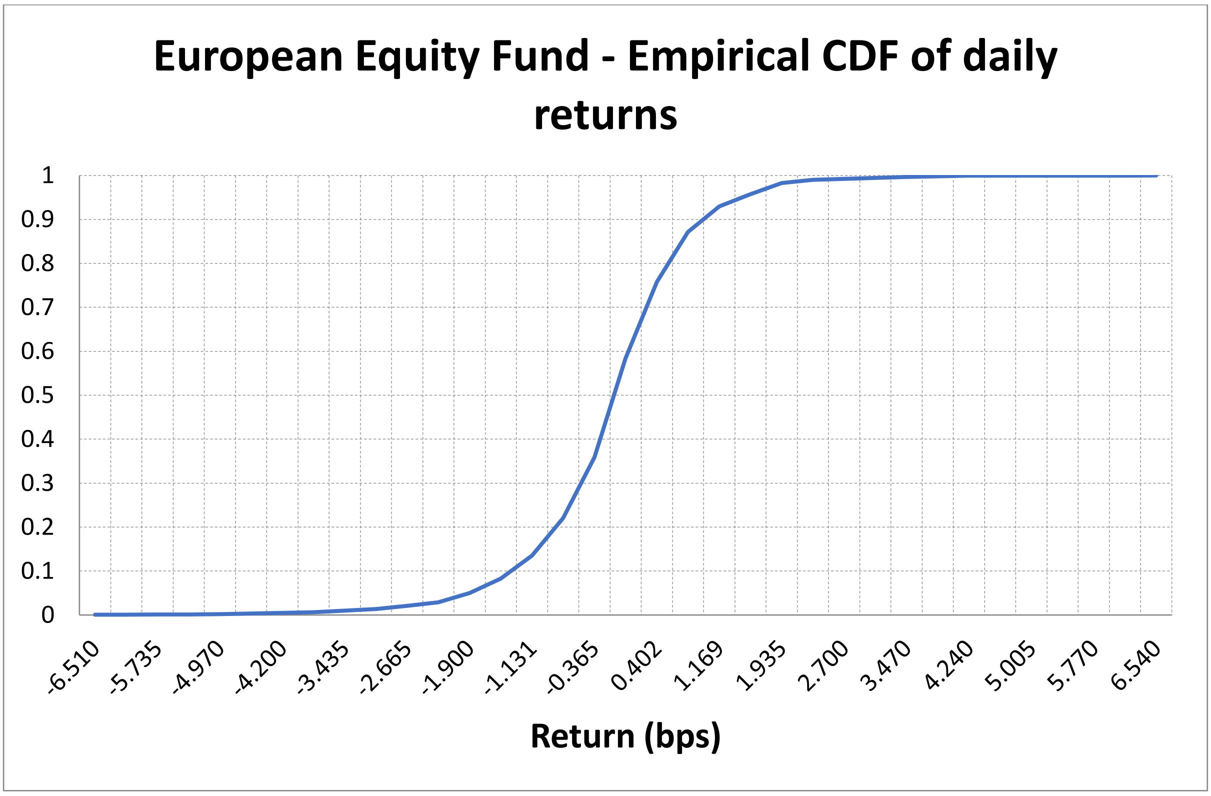
Charts 7, 8, and 9 illustrate the empirical cumulative distribution functions (CDFs) for the set of funds we examine in the paper.

Chart 7



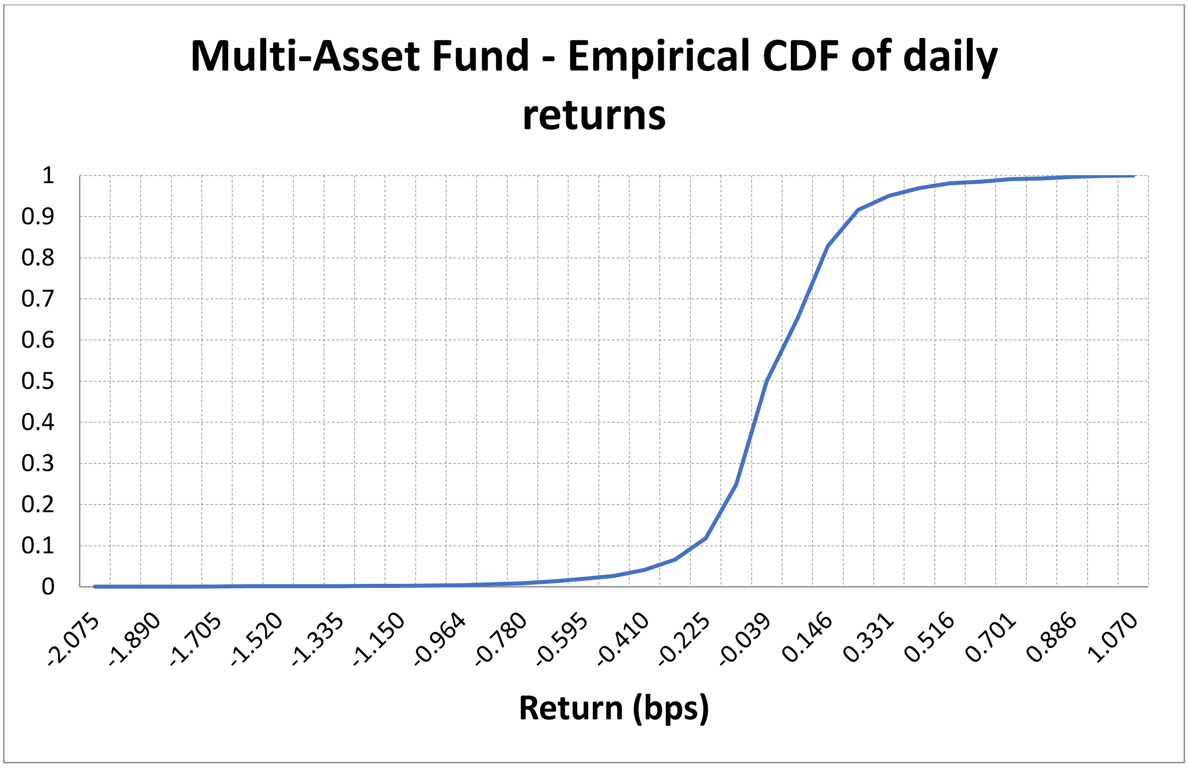
There are no major outliers on the upside or the downside however the graph has a long flat portion which is almost parallel to the horizontal axis on the left-hand side. The skewness of the distribution shown in Table 4 is -0.609. The empirical plot of high returns versus low returns shown in Chart 4 shows a number of ordered pairs, (lowest, highest), (second lowest, second highest), etc., of daily returns where the absolute values of the high negative daily returns are larger than those of the high positive daily returns.

Chart 8



There are perhaps two major outliers one on the upside and one on the downside. The skewness of the distribution shown in Table 4 is -0.522. The empirical plot of high returns versus low returns shown in Chart 5 shows few ordered pairs, (lowest, highest), (second lowest, second highest), etc., of daily returns where the absolute values of the high negative daily returns are larger than those of the high positive daily returns.

Chart 9



The empirical CDF of the Multi-Asset Fund illustrates a very long trail on the left of zero relative to that on the right of zero.

This is consistent with the empirical plot of high returns versus low returns shown in Chart 6 where there are a number of ordered pairs, (lowest, highest), (second lowest, second highest), etc., of daily returns where the absolute values of the high negative daily returns are significantly larger than those of the high positive daily returns.

The skewness of daily returns of the Multi-Asset fund is -0.754 as shown in Table 4. The Multi-Asset fund also has the highest absolute value of skewness of daily returns among the three funds discussed in this paper. The due diligence team might reasonably seek some explanation from the IM as to likely reasons for this negative skewness.

Kurtosis

In statistics, Kurtosis is the fourth central moment of a distribution. The estimate of the index of kurtosis derived from a sample of observations is given by the formula:

$$Kurtosis\,(r_1,\ldots,r_n) = \frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}\left(\frac{r_i-\bar{r}}{\sigma}\right)^4 - \frac{3(n-1)^2}{(n-1)(n-3)}$$

where n is the number of observations, and r(bar) is replaced by the sample mean and σ by the sample standard deviation. For n > 1;000, the sampling distribution of this index is approximately normal, with mean zero and standard error √(24/n).

In the formula above[3], the second term tends to 3 for very large n. The kurtosis of the normal distribution calculated using the first term of the formula tends to 3 for large n. The second term is subtracted from the first term to standardise the kurtosis of the normal distribution to zero.

In Table 5 below, all the kurtosis figures quoted for the three funds are 'excess' kurtosis relative to that of the normal distribution or put another way the second term in the equation below has been subtracted in calculating the kurtosis figures.

Table 5

Fund	Kurtosis
IGC Bond	5.9
European Equity	4.0
Multi-Asset	8.6

The IGC Bond fund exhibits higher kurtosis than the European Equity fund while the Multi-Asset fund has the highest kurtosis of the three funds.

If we take the daily returns for each of the three funds and standardised them by stating them in terms of the number of standard deviations from the relevant average daily return, we find that the Multi-Asset fund has one standardised daily return which is more than 9 in absolute value terms. This is a nine standard deviations move to the downside. The absolute value of the highest downside standardised daily return for IGC Bond fund and the European Equity fund is 8.6 and 6.4 respectively.

Kurtosis is very sensitive to large standardised daily returns. The difference between raising a figure like 9 to the power of four (6,561) and 8.6 to the power of four (5,470) in the first part of the formula above contributes an extra 1,090 to the summation when calculating the kurtosis of the Multi-Asset fund and the IGC fund respectively.

Kurtosis is a measure of the 'thickness' of the tails of a probability distribution relative to that of a normal distribution with the same standard deviation. Thick tails contain "surprises" for investors where surprises are defined as a daily return which is large relative to the usual periodic returns we expect from the trading strategy. Investors may like a large positive return in a period but these may be indicative of the potential for large negative returns which tend to be less welcomed by investors.

Kurtosis may be considered as a measure of the extent of variation in the risk of the fund. Kurtosis may be indicative of how well the IM is controlling the overall risk of the fund. The extent of variation in the risk of the fund is examined in the context of the annualised, rolling 20-day volatility of the fund.

The negative skew and positive excess kurtosis of the three funds warn us that we cannot rely on the standard deviation as a measure of risk. If we are trying to avoid 'negative' surprises,

[3] Source: http://www.styleadvisor.com/resources/statfacts/kurtosis. Microsoft Excel uses the same formula in calculating kurtosis.

large daily losses, then we should not rely on standard deviation as our measure of risk in constructing a portfolio consisting of these funds.

Number of Days Accounting for Various Percentages of Total Return for a Period

Table 6 shows the percentage of daily returns by number accounting for 90%, 80%, and 75% of the total return of the funds over the length of their track records.

Table 6

Fund	Percentage of Days Accounting for Certain Percentages of Total Return		
	75%	**80%**	**90%**
IGC Bond	2.6%	2.9%	3.3%
European Equity	0.4%	0.4%	0.5%
Multi-Asset	3.1%	3.4%	4.1%

Table 6 perhaps illustrates the difficulty in market timing trading strategies particularly in relation to European equities. The Multi-Asset fund invests at least 50% in bonds so it is not surprising that its results are closer to those of the IGC Bond fund than to that of the European Equity Fund.

The snippet of Python code required to produce the above table is shown below.

Python Code Snippet

```
total_percs = (90, 80, 75)
```

```
total_return = df_returns.Return.sum()
df_returns.sort_values(by = 'Return', ascending = False, inplace = True)

total_perc_num_days = []
total_perc_num_days_prop = []

for total_perc in total_percs:
    num_days = len(df_returns.loc[df_returns.Return.cumsum() < total_return * total_perc / 100]) + 1
    num_days_prop = num_days / num_returns
    total_perc_num_days.append(num_days)
    total_perc_num_days_prop.append(num_days_prop)
```

Ranked Daily Returns

Charts 10, 11, and 12 show the daily returns of the three funds ranked from lowest to highest. The charts provide a picture of the size of losses on the worst trading days and the size of gains on the best trading days as well as illustrating the balance between negative daily returns, zero daily returns, and positive daily returns.

Chart 10

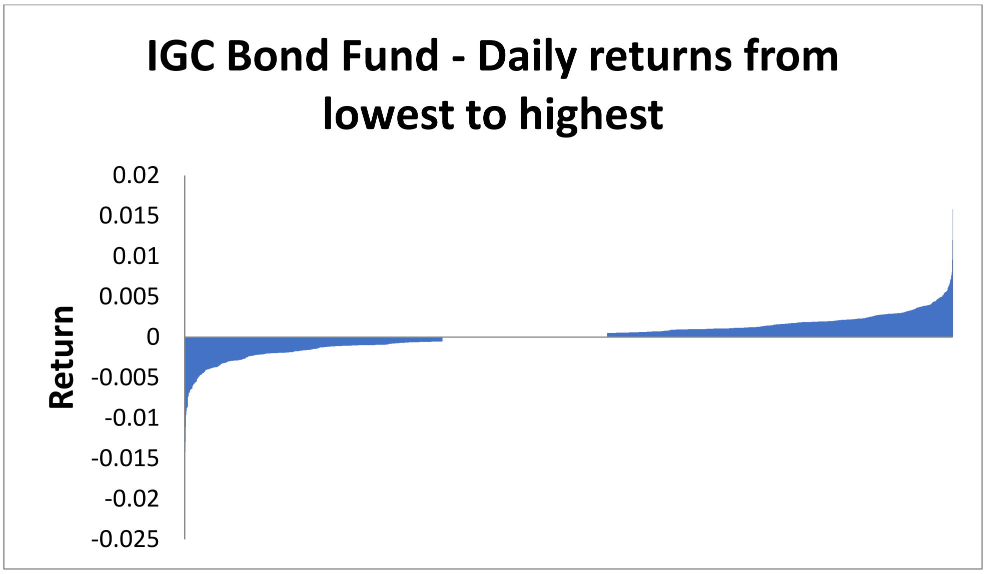


Chart 10 shows a high proportion (21.5%) of daily returns which are zero.

Chart 11

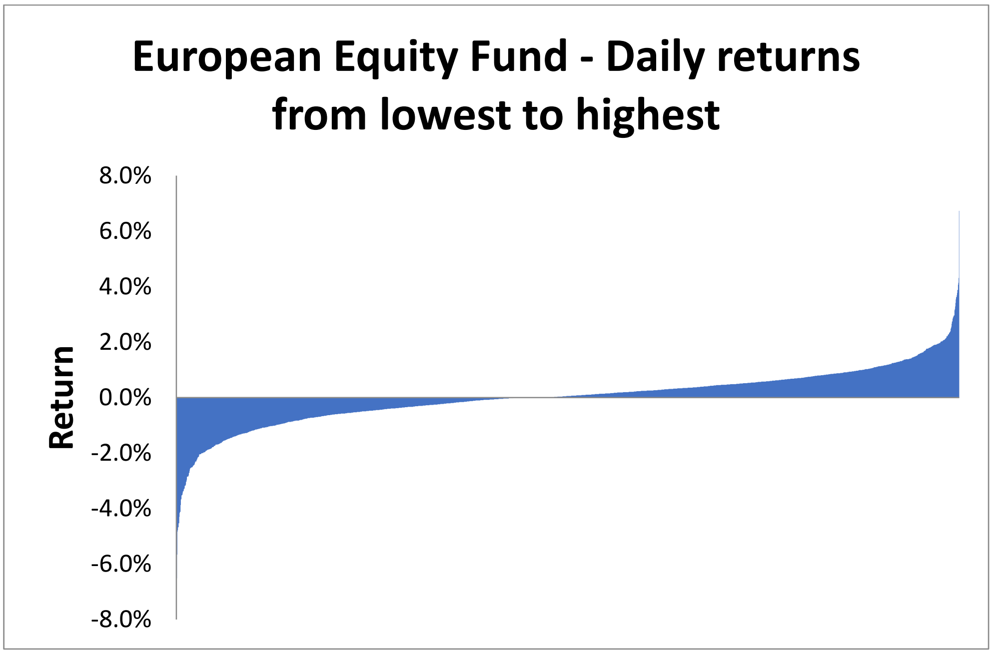


Chart 11 shows a relatively low proportion (3.6%) of daily returns which are zero.

Chart 12

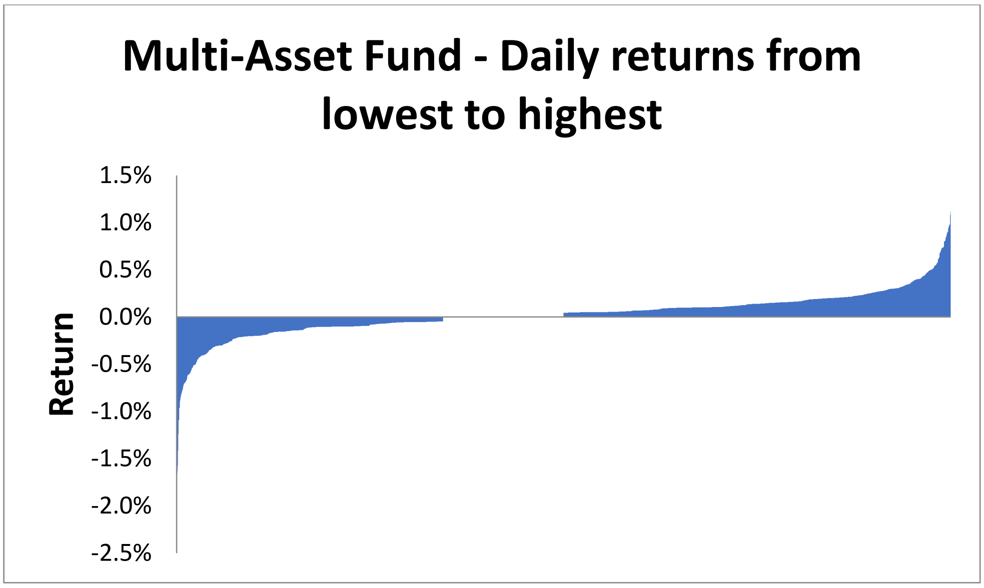


Chart 12 shows a high proportion (15.6%) of daily returns which are zero.

The IGC Bond fund and the Multi-Asset fund which invests at least 50% in bonds both have relatively high proportions of zero-return days compared to the European Equity.

Absolute Performance

There is no "all-weather" asset class that always makes money and there is no "all-weather" trading strategy that always makes money. What we are looking for is a trading strategy where: (i) the combination of the percentage of positive daily returns and the average positive daily return; exceeds (ii) the absolute value of the combination of the percentage of negative daily returns and the average negative daily return over its trading history.

Table 7

Fund	Percentage of Positive Daily Returns	Average Positive Daily Return (bps)	Percentage of Negative Daily Returns	Average Negative Daily Return
IGC Bond	45.0%	19.07	33.5%	-19.74
European Equity	54.2%	71.73	45.2%	-81.41
Multi-Asset	50.0%	18.62	34.4%	-19.01

Table 7 shows the European Equity fund as having a slightly higher percentage of positive days and a much higher percentage of negative days compared with the other two funds. The average return of the European Equity fund on positive and negative days is at least 3.5 times those of the IGC Bond fund and the Multi-Asset fund which invests at least 50% in bonds. In the case of all three funds, the absolute value of the average negative daily return is greater than the average positive daily return. The difference between the absolute value of the average negative daily return and the average positive daily return is greatest for the European Equity fund.

For a trading strategy to be repeatable, we must consider why the losing parties are likely to continue to engage in such losing trades in the long-run. One possible explanation that might be considered is that the losing parties are engaged in hedging; they are focused not on making a profit but rather on making the outcome of the combination of their underlying exposures and the hedging trade more certain. Thus, the trading strategy is providing a form of insurance for the losing parties and may be sustainable albeit with some losses from time to time. Other possible explanations for repeatable net positive investment returns may be: (i) the IM has legally-obtained, superior knowledge when compared with other market players; or (ii) the IM is being compensated for taking liquidity risk by, for example, earning bid-ask spreads as a kind of market maker.

Omega Ratio for a Threshold Return of Zero

The Omega ratio for a threshold return of zero per day is defined as the ratio of the cumulative probability of an outcome above zero to the cumulative probability of an outcome below zero. The ratio can also be calculated for threshold levels other than zero; for example, it could be calculated for a threshold level of say, 6bp per day.

The Omega ratio partitions the set of daily returns into gains and losses above and below the investor's threshold level. The Omega ratio is the ratio of returns above the threshold level to returns below the threshold level. The higher the Omega ratio, the greater the probability that the threshold return will be at least met and possibly exceeded.

Table 8 shows the Omega ratio for a threshold return of zero for the three funds. It has been estimated as the ratio of the area above the curve to the right of zero divided by the area under the curve to the left of zero. Days on which the return was zero do not enter the computation.

Table 8

Fund	Omega Ratio
IGC Bond	1.295
European Equity	1.057
Multi-Asset	1.423

The Python code snippet to compute the Omega ratio is set out below.

Python Code Snippet

```
omega_ratio = -df_returns.loc[df_returns.Return > 0, 'Return'].sum() / df_returns.loc[df_returns.Return < 0, 'Return'].sum()
```

Sharpe Ratio

The Sharpe ratio is defined as follows:

$$(R_p - r_f)/\sigma_p$$

where R_p is the annualised return on the fund, r_f the risk-free rate of interest, and σ_p is the annualised standard deviation of the returns of the fund. Throughout this paper, in computing Sharpe ratios, we have assumed that the risk-free rate of interest is zero.

Sharpe Ratio versus Omega Ratio

The Sharpe ratio is computed using the first two moments of the distribution of returns and focuses on absolute returns rather than returns in excess of a threshold value ("TV"). By contrast, the Omega ratio:

[Cumulative probability of an outcome above the TV] / [Cumulative probability of an outcome below the TV]

captures all the moments of the distribution and can focus on absolute returns or returns above a TV.

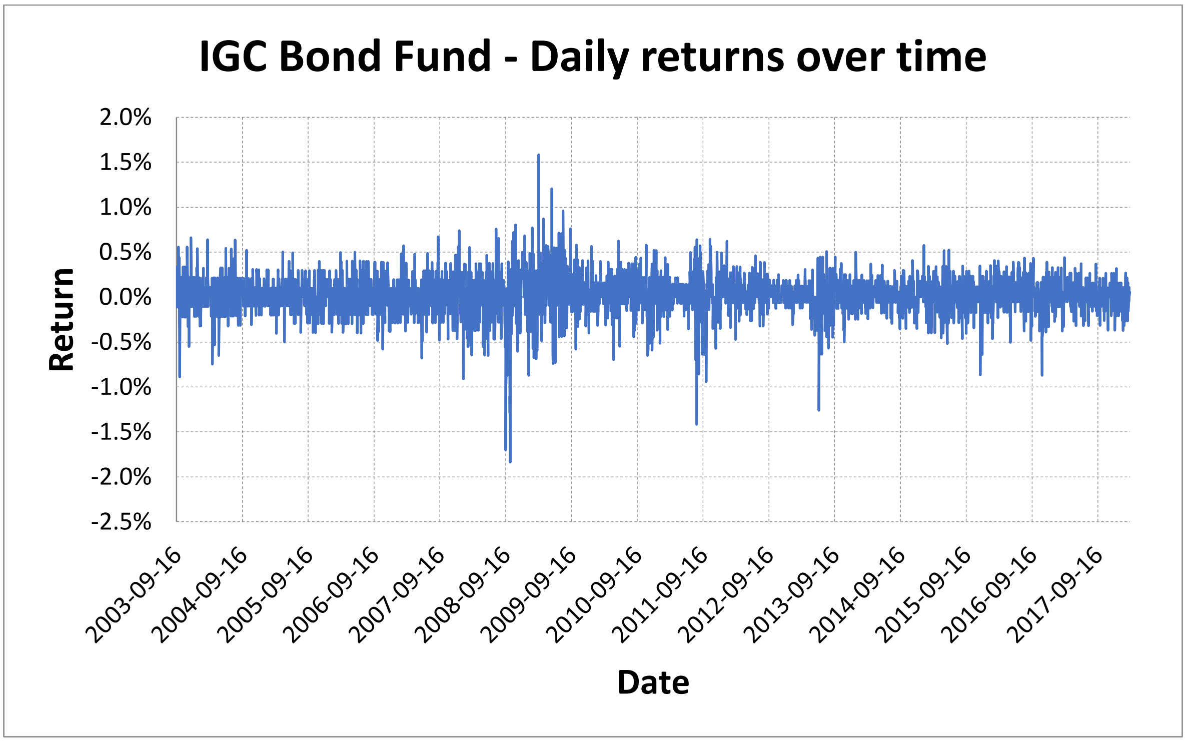
Given the obvious lack of normality in the PDFs of the three funds shown in Charts 1, 2, and 3, the Omega ratio is more appropriate than the Sharpe ratio in assessing the risk-adjusted return of the funds.

Plot of Daily Returns against Time

In a plot of daily returns against time, the spikes of positive and negative returns can easily be identified and we can illustrate how the trading strategy behaved during known periods of extreme market events in the past.

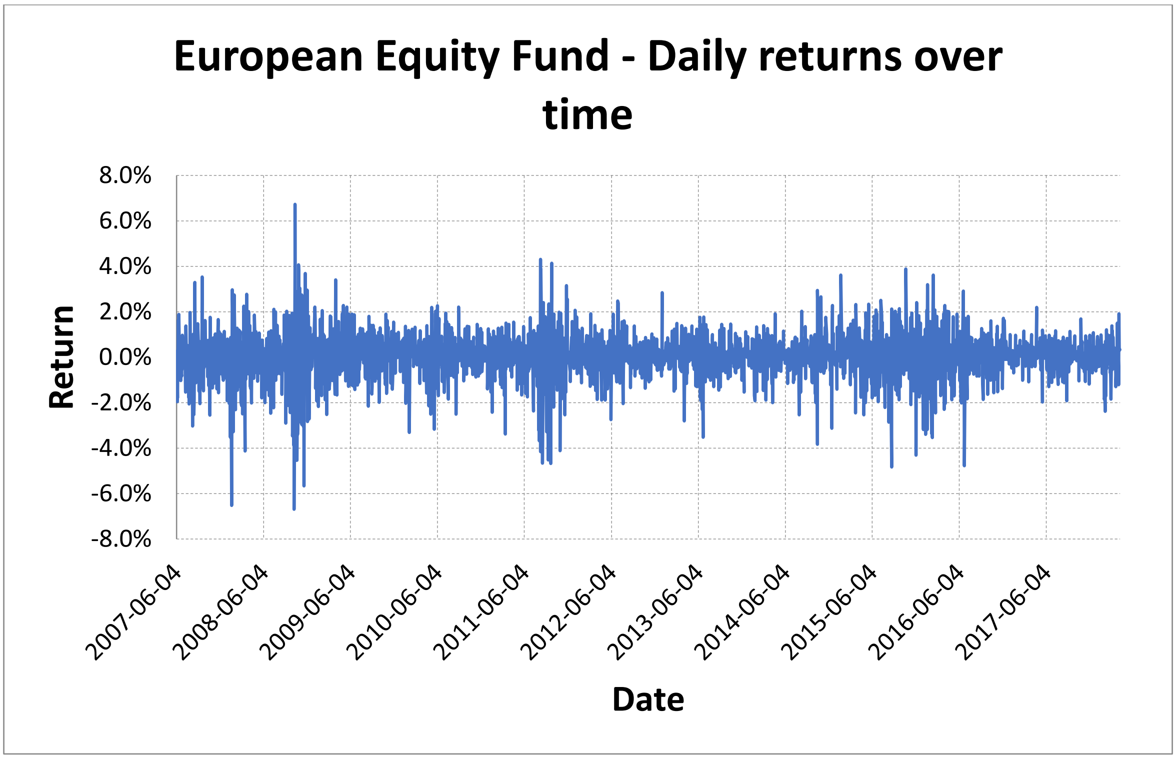
Charts 13, 14, and 15 below illustrate the concept for the set of funds we examine in the paper. An explanation should be sought from the IM for both large negative spikes and large positive spikes in daily returns.

Chart 13



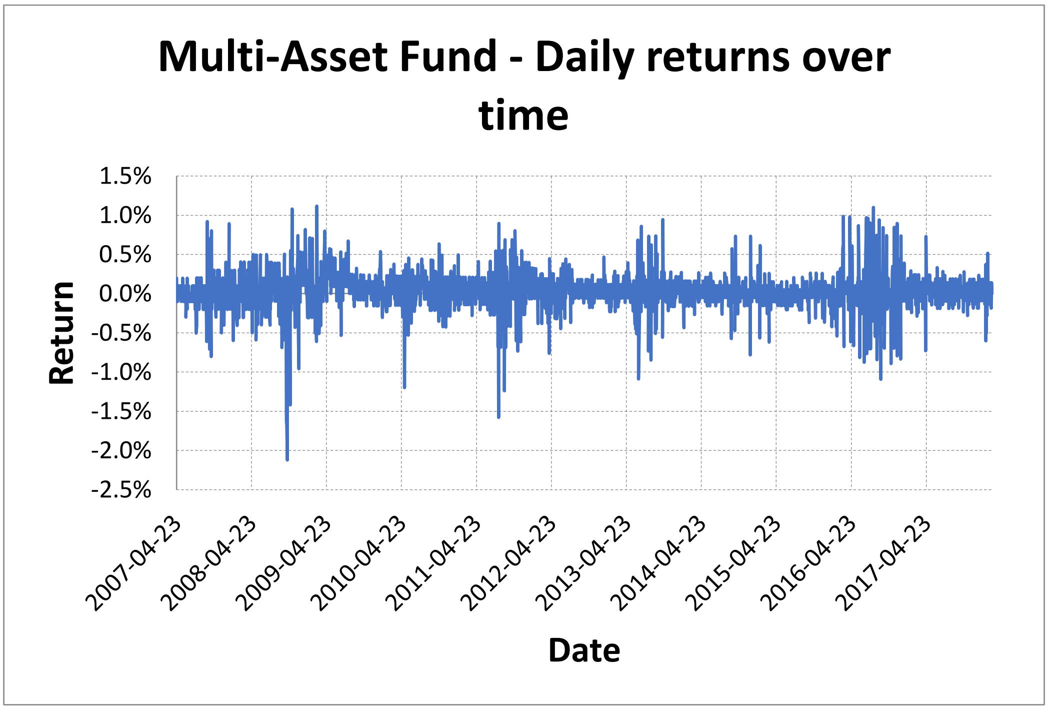
Daily losses have exceeded 1% of the NAV on four occasions and exceeded 1.5% on two occasions.

Chart 14



Daily losses have exceeded 4% of the NAV on more than ten occasions illustrating that the European Equity fund is significantly more volatile than the IGC Bond fund.

Chart 15



Daily losses have exceeded 1% of the NAV on seven occasions and 2% on one occasion. Given the fund's ability to invest up to 20% in equities and to leverage, the size and frequency of daily losses in excess of 1% is greater than that of the IGC Bond fund.

Rolling Rates of Return

Table 9 shows the actual annualised returns of the three funds over the history of their data sets.

Table 9

Fund	Actual Annualised Return
IGC Bond	5.1%
European Equity	5.3%
Multi-Asset	7.0%

Statistics like the annualised return since-start-date use only two of the data points from the data set and don’t provide a full picture of the nature of the investment returns.

Rolling rates of return provide a better insight into the nature of a fund’s investment returns and the return experience of different time-cohorts of investors.

Some funds target specific return objectives. For example, a fund may target a return of 5% per annum above the return on cash over rolling three year periods. Examining the rolling 3-year rates of return for such a fund will quickly reveal how successful the IM has been at achieving the fund’s objective.

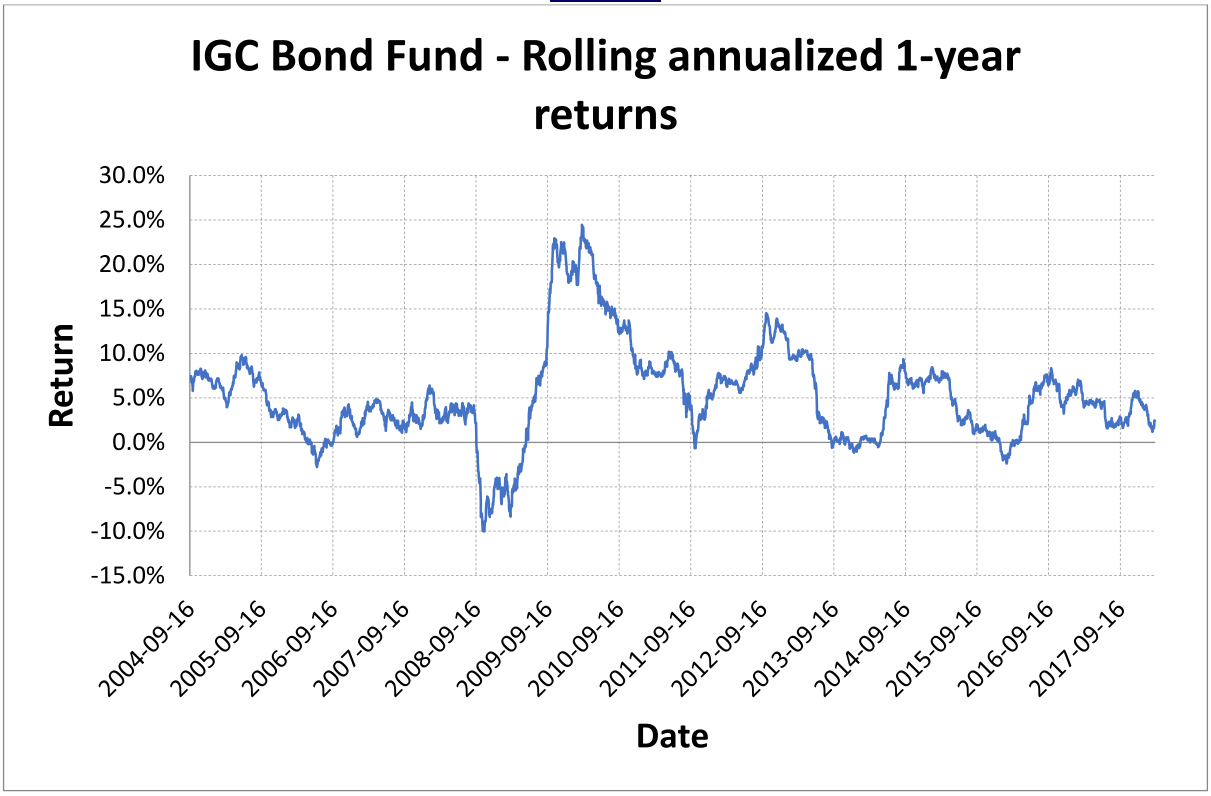
Rolling rates of return are not independent observations of returns over the time periods in question as consecutive rolling periods built from daily data differ only by two daily returns; all the other daily returns are common to both of the consecutive rolling rates of return. For example, in consecutive, rolling, 1-year, annualised rates of return, 250 of the 252 trading days are common to both of the consecutive periods.

Charts 16 to 24 show the annualised rolling daily rates of return over one, three, and seven-year time periods and illustrate the variation in return over these time periods.

Python Code Snippet

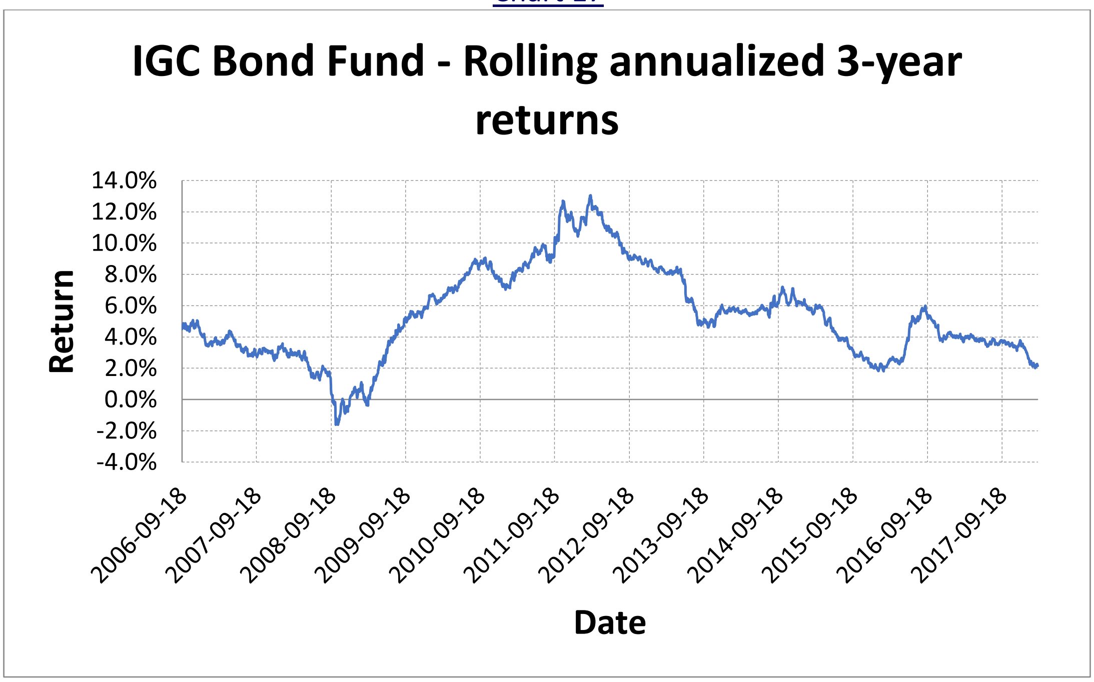
```
roll_years = (1, 3, 5, 7)
show_ann_roll_years = True
    for roll_year in roll_years:
        if roll_year <= term_years:
            first_roll_date = df_prices.Date.min()
            last_date = df_prices.Date.max()

            if show_ann_roll_years:
                roll_year_col = 'Rolling annualized ' + str(roll_year) + '-year return'
                df_returns[roll_year_col] = np.nan
            else:
                num_pos_periods = 0
                num_non_pos_periods = 0

# Find the date exactly roll_year years before the last pricing date. If the last pricing date is
# February 29th and roll_year is not divisible by 4, then there will be no date exactly roll_year
# years before the last pricing date, so use February 28th instead.
            try:
                last_roll_date = datetime.datetime(last_date.year - roll_year,
                                                   last_date.month,
                                                   last_date.day)
            except:
                last_roll_date = datetime.datetime(last_date.year - roll_year,
                                                   last_date.month,
                                                   last_date.day - 1)

# Loop through every date (whether or not it was a pricing date) and compare the price applicable
# to that date with the price applicable exactly roll_year years later to determine whether the
# return over that particular period was positive.
            for start_roll_date in pd.date_range(start = first_roll_date, end = last_roll_date, freq = '1D'):
                start_price = df_all_prices.Price.loc[start_roll_date]

# Find the date exactly roll_year years after the start date. If the start date is February
# 29th and roll_year is not divisible by 4, then there will be no date exactly roll_year years
# after the start date, so use February 28th instead.
                try:
                    end_roll_date = datetime.datetime(start_roll_date.year + roll_year,
                                                      start_roll_date.month,
                                                      start_roll_date.day)
                except:
                    end_roll_date = datetime.datetime(start_roll_date.year + roll_year,
                                                      start_roll_date.month,
                                                      start_roll_date.day - 1)

                end_price = df_all_prices.Price.loc[end_roll_date]

                if show_ann_roll_years:
                    df_returns.loc[
                        df_returns.Date == end_roll_date, roll_year_col] = (np.log(end_price) - np.log(start_price)) / roll_year
```

Chart 16



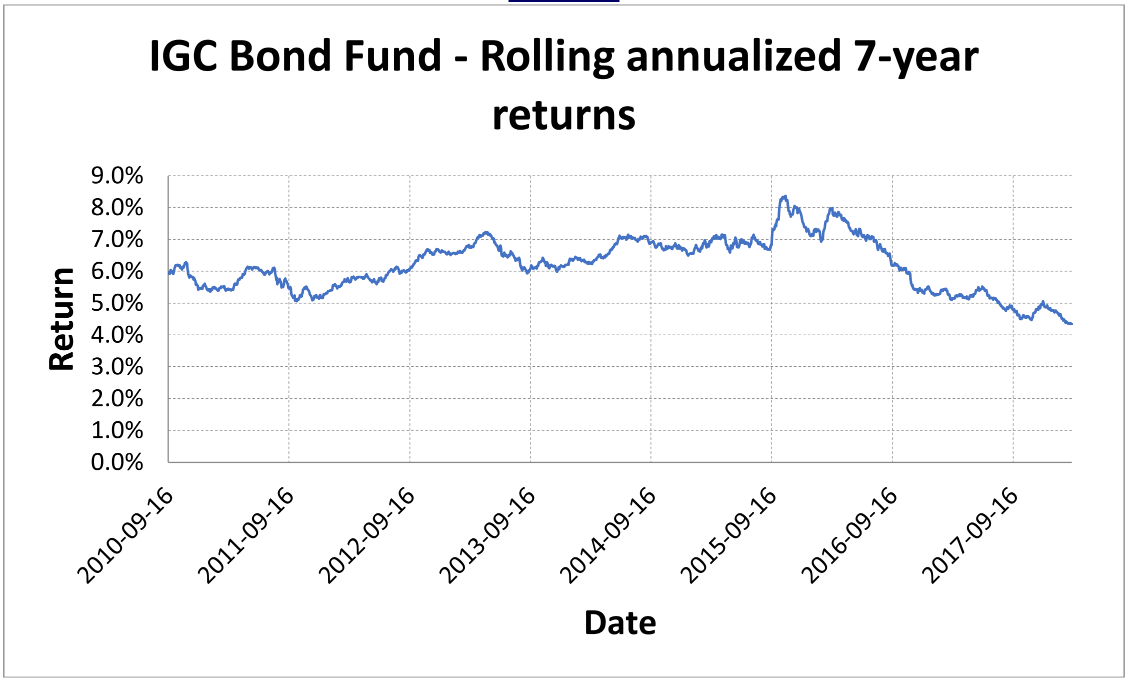
Just over 88% of rolling, 1-year, annualised rates of return for the IGC Bond fund are positive. The rolling, 1-year, annualised rates of return vary between -10% p.a. and +24% p.a.

Chart 17



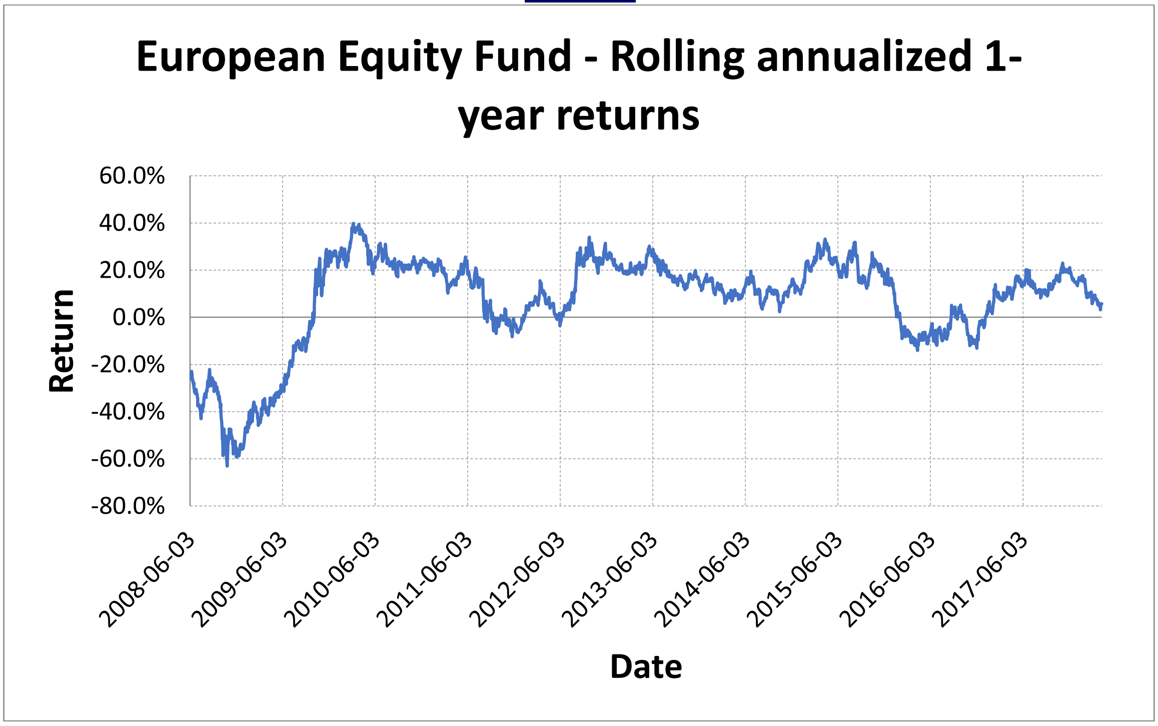
Almost 98% of the rolling, 3-year, annualised returns for the IGC Bond fund are positive; the exception was the period ending between mid-September 2008 and end-March 2009. The rolling, 3-year, annualised returns exhibit significant variability ranging from about -2% p.a. to +13% p.a.

Chart 18



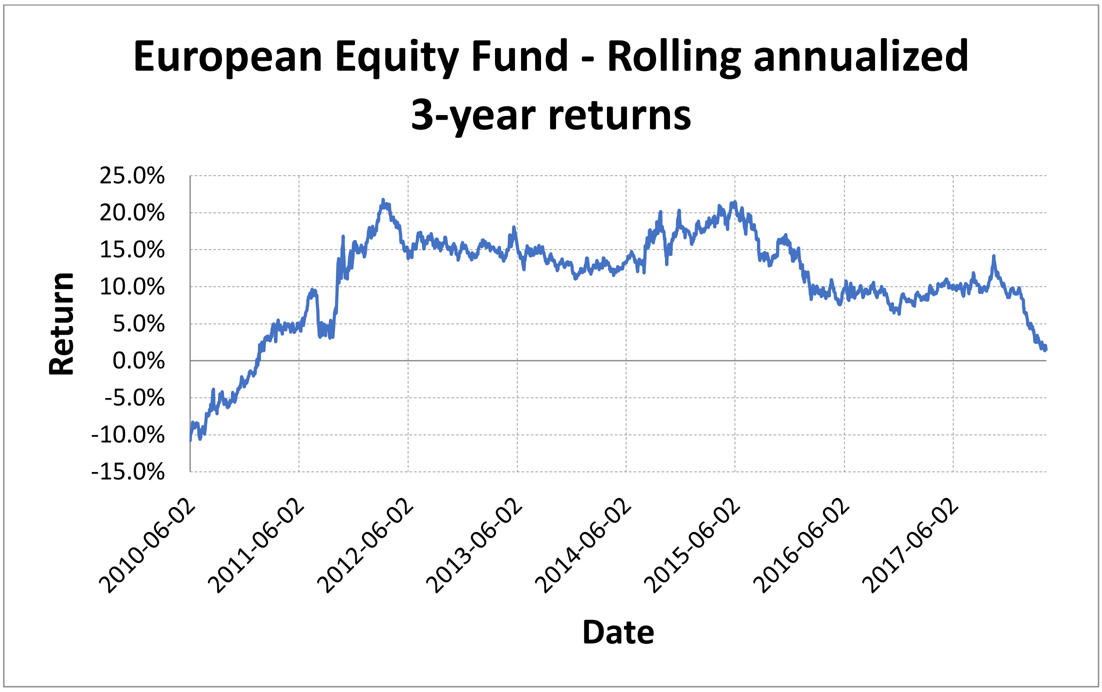
All the rolling, 7-year, annualised returns for the IGC Bond fund are positive. Up until the end of October 2015, rolling, 7-year, annualised returns for the IGC Bond fund were relatively stable and broadly remained in the 5% to 7% per annum region. Since then, the rolling, 7-year, annualised returns have been on a downward trend.

Chart 19



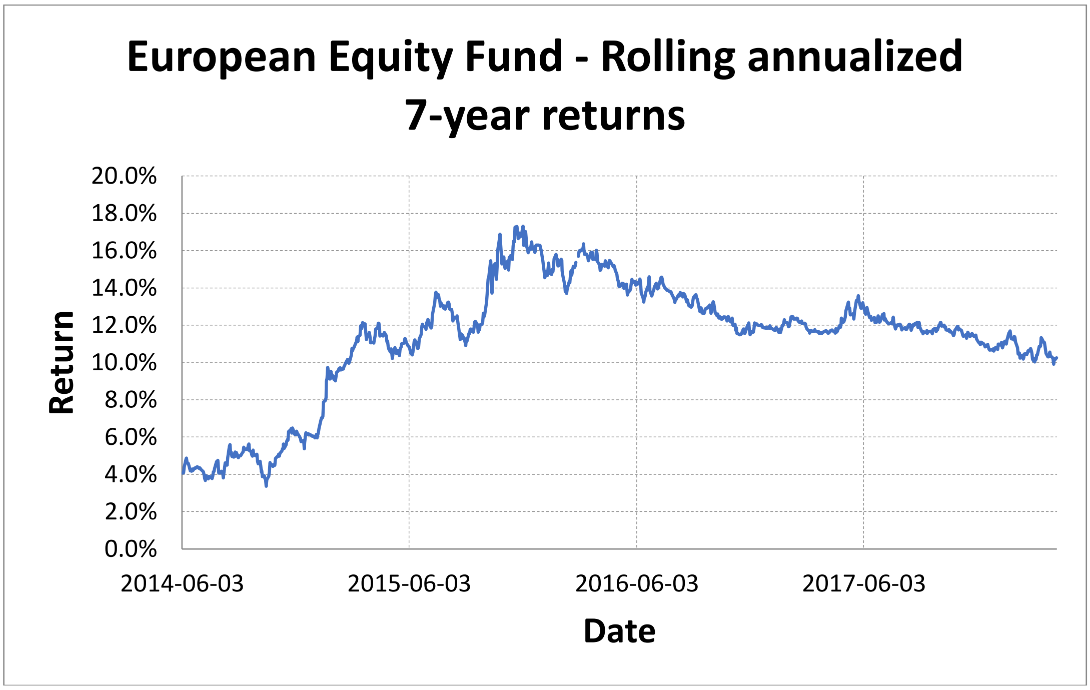
Just under 75% of rolling, 1-year, annualised rates of return are positive for the European Equity Fund. There is enormous volatility in the rolling, 1-year, annualised rates of return as they vary from -63% for the 1-year period ending 27 October 2008 to +39% for the 1-year period ending 9 March 2010.

Chart 20



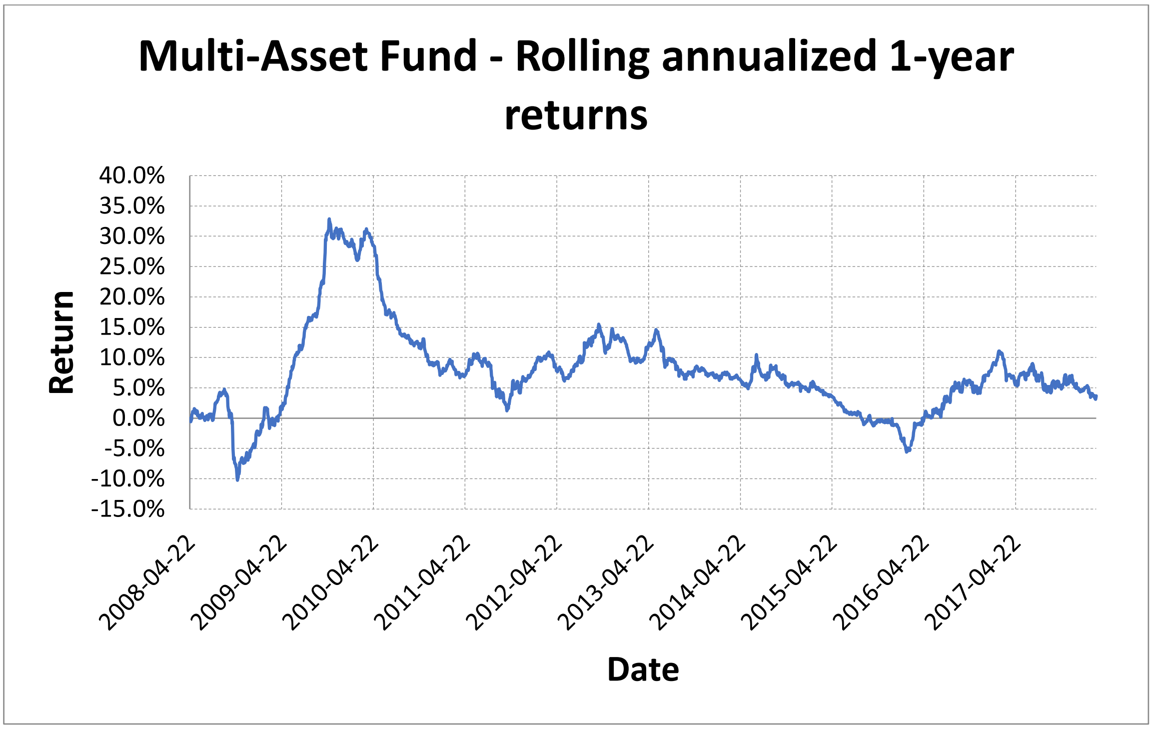
Just over 92% of rolling, 3-year, annualised returns for the European Equity Fund are positive. Up until 17 January 2011, rrolling, 3-year, annualised returns were broadly negative but have remained positive since that date. Since the end of May 2015, the rolling, 1-year, annualised rates of return for the European Equity fund have been declining.

Chart 21



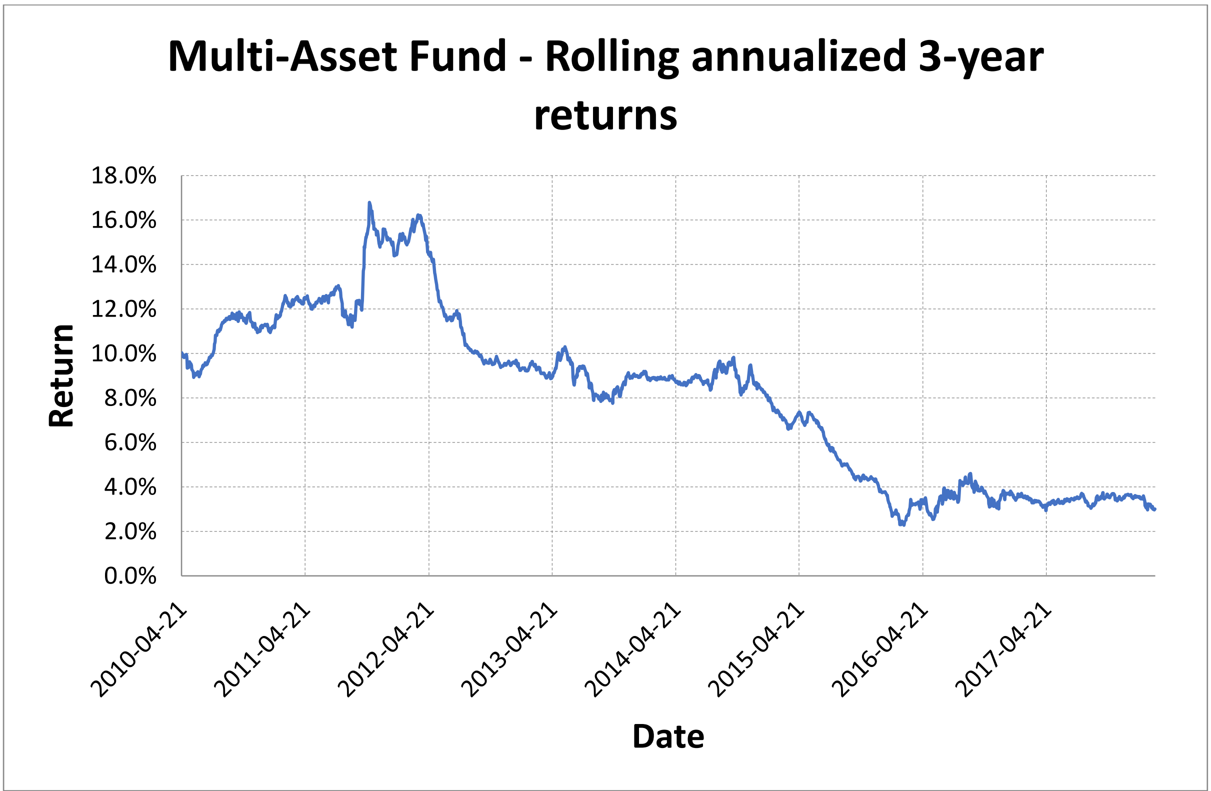
All the rolling, 7-year, annualised returns for the European Equity fund are positive. However, since early December 2015, they have been exhibiting a downward trend falling from highs of over 17% p.a. to 10% p.a.

Chart 22



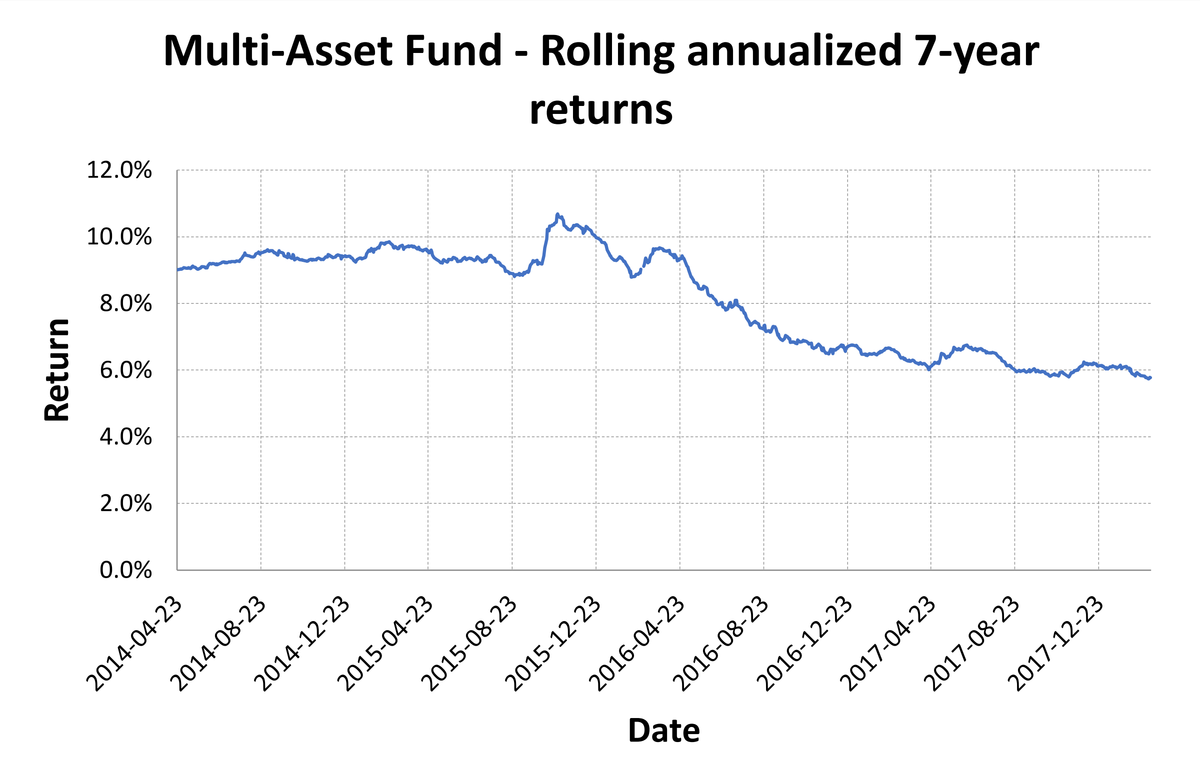
Almost 88% of rolling, 1-year, annualised rates of return for the Multi-Asset fund are positive. The rolling, 1-year, annualised rates of return of the Multi-Asset fund range from -10.2% p.a. to just over +32.5%.

Chart 23



All the rolling, 3-year, annualised returns for the Multi-Asset fund are positive although they have been exhibiting a downward trend since early November 2011.

Chart 24



All the rolling, 7-year, annualised returns for the Multi-Asset fund are positive although they have been exhibiting a downward trend since the end of October 2015. The rolling, 7-year, annualised returns range from 8.8% p.a. to 10.7% p.a. which is quite a narrow range.

Fees as a Percentage of Actual Return

It is important to examine fees as a percentage of actual return and fees as a percentage of expected return. In this section, we define fees to be the sum of the annual management charge and the other ongoing costs. These fees are usually expressed as a percentage of assets under management. Portfolio transaction charges are excluded from the definition of fees as UCITS will not be subject to the PRIIPs legislation until the beginning of 2020.

Fees as a percentage of actual return is simply the total fees paid as a percentage of the returns generated over a given period, usually the period since inception of the trading strategy.

Table 10 illustrates the concept for the set of funds we examine in the paper. Returns generated may have to be chain-linked to allow for new money invested and withdrawals of existing money.

Table 10

Fund (1)	Fees (2)	Actual Annualised Return (3)	Fees as a Percentage of Actual Return (4) = (2)/(3)
IGC Bond	0.49%	5.1%	9.6%
European Equity	0.83%	5.3%	15.7%
Multi-Asset	0.92%	7.0%	13.1%

The fees as a percentage of actual annualised return are highest for the European Equity fund and lowest for the IGC Bond fund.

Fees as a Percentage of Expected Return

Few single-asset trading strategies have a long-term Sharpe ratio more than 0.4[4]. Using a Sharpe ratio of 0.4 and the realised volatility[5] of the trading strategy, we can calculate the expected return of the trading strategy. For the Multi-Asset fund, we assume a Sharpe ratio of 0.5 in view of the diversification of risk inherent in the Multi-Asset portfolio.

For example, if the realised volatility of the trading strategy were say 2%, then the expected return would be 0.8% (2%*0.4). If the fees expressed in basis points were say 4 basis points, then fees as a percentage of expected return would be 5% (4/80).

This measure of fees provides a basis for the discussion of fees with an IM if fees are seen to be high relative to expected returns.

Table 11

Fund (1)	Fees (2)	Realised Volatility (3)	Estimated Long-term Sharpe Ratio (4)	Expected Long-term Return (5)=(3)*(4)	Fees as a Percentage of Expected Return (6)=(2)/(5)
IGC Bond	0.49%	3.5%	0.4	1.4%	35%
European Equity	0.83%	16.9%	0.4	6.8%	12%
Multi-Asset	0.92%	3.8%	0.5	1.9%	48%

Fees as a percentage of expected annualised return are highest for the Multi-Asset fund and lowest for the European Equity fund.

Relative to Table 10, Table 11 shows a much wider dispersion between expected returns and the actual returns realised over the period of the data sets and this has led to a different view of the relative ranking of the funds in terms of fees. The European Equity fund has moved

[444] See for example *The Sharpe Ratio Frontier* by David H. Bailey and Marcos M. Lopez de Prado. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1821643. Accessed 18 September 2018.

[5] Assuming that the realised volatility is a reasonable guide to the long-term volatility of the trading strategy.

from the most expensive on the fees as a percentage of actual return measure to the cheapest on the fees as a percentage of expected returns measure.

Risk Metrics

Looking at the realised historic volatility of a fund without an understanding of the extent of variation in volatility may lead to surprises for investors in relation to the size of future peak-to-trough falls in value.

The size of peak-to-trough falls in value of a trading strategy over any given time horizon is driven primarily[6] by the base-line level of risk at which the trading strategy operates and the extent of variation in that base-line level of risk.

The due diligence team will wish to conduct an examination of the historic, annualised, rolling 20-day daily volatility of a fund which will provide an insight into the extent of variation in the volatility of the fund.

Apart from the average level of volatility, one of the key drivers of the size of the maximum peak-to-trough fall in the value of a fund over any time period is the extent of variation in volatility around its average level.

Rolling n-day volatility can also be used to assess the extent to which a fund which is specified to operate within a certain risk band actually adheres to the stated volatility band specified for an n-day period.

Volatility

It is important to assess the annualised standard deviation of daily returns of the trading strategy to gain some idea of the base-line level of risk of the strategy. We shall refer to the standard deviation of daily returns as the volatility of daily returns.

Before measuring the standard deviation of returns, we need to understand the background to the observations of returns that we have collected. If the portfolio valuations on which the returns are calculated are in any way subjectively determined by the manager or an associate, then little value can be placed on the standard deviation as a measure of variability. This could arise where a trading strategy uses illiquid securities and where the daily mark-to-market value is determined by the IM. The values of such illiquid securities could easily be smoothed to reduce the variance of published returns where the manager determines the value of securities.

The due diligence team would need to find out if an IM follows any practice like reducing the exposure of the portfolio to risk for the remainder of a year after a `good' month or `good' months. If this is the case, then the standard deviation is not based on as many observations as might at first appear.

To gain some idea of the variation in volatility, the annualised, rolling, 20-day volatility of the trading strategy can be monitored and plotted on a graph such as that shown in Charts 25, 26, and 27 for the set of funds we examine in the paper.

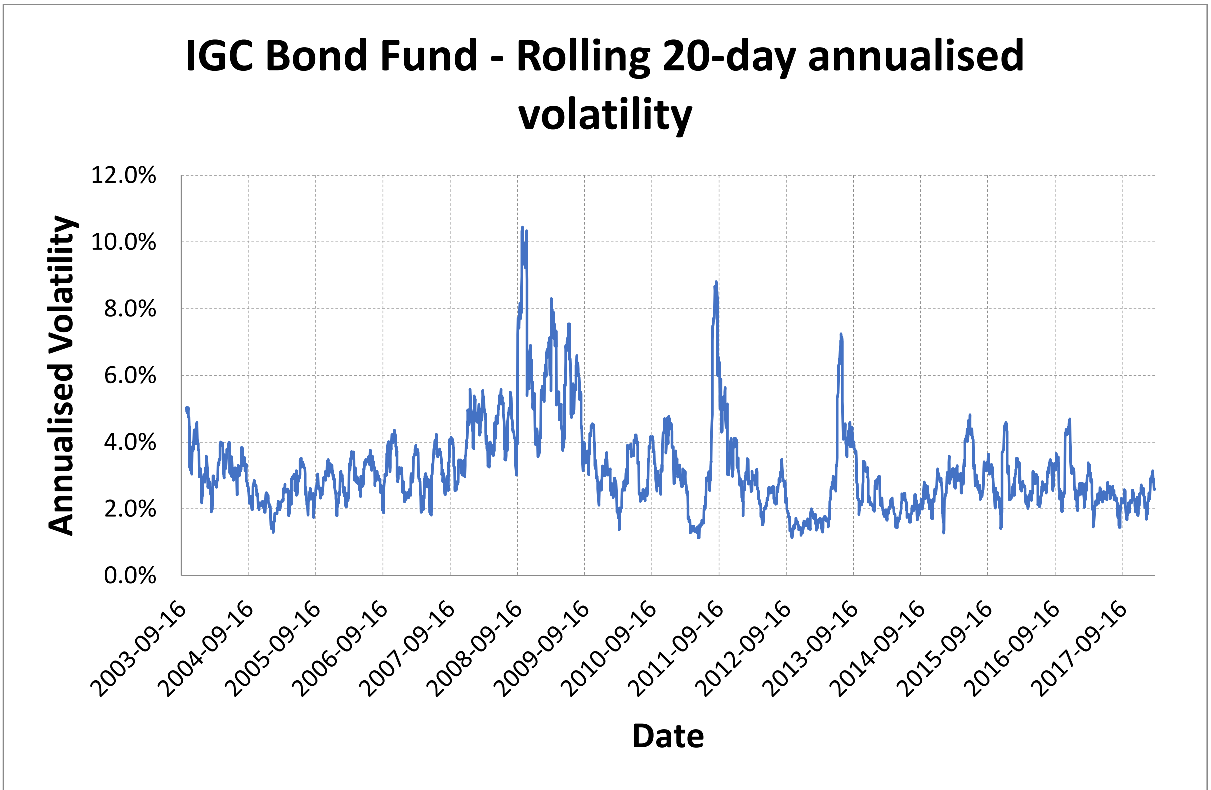
[6] The return of the trading strategy also plays a part in determining the size of peak-to-trough falls in value. However, increasing the expected return does not significantly dampen the size of peak-to-trough falls in value.

A chart such as Chart 25 can indicate trends in the risk of the portfolio and identify if the IM is taking enough risk, too much risk or keeping the risk level within a relatively tight range. Again, Charts 25, 26, and 27 illustrate this for the set of funds we examine in the paper. A snippet of the Python code to produce such charts is given below.

Python Code Snippet

```
roll_days = (5, 20)
for roll_day in roll_days:
    roll_sum = df_returns.Return.rolling(roll_day).sum()
    roll_day_col = 'Rolling ' + str(roll_day) + '-day return'
    df_returns[roll_day_col] = roll_sum
    roll_day_col = 'Rolling ' + str(roll_day) + '-day annual return'
    df_returns[roll_day_col] = roll_sum * mean_days_per_year / roll_day
    roll_day_col = 'Rolling ' + str(roll_day) + '-day annual standard deviation'
    df_returns[roll_day_col] = df_returns.Return.rolling(roll_day).std() * mean_days_per_year ** 0.5
```

Chart 25

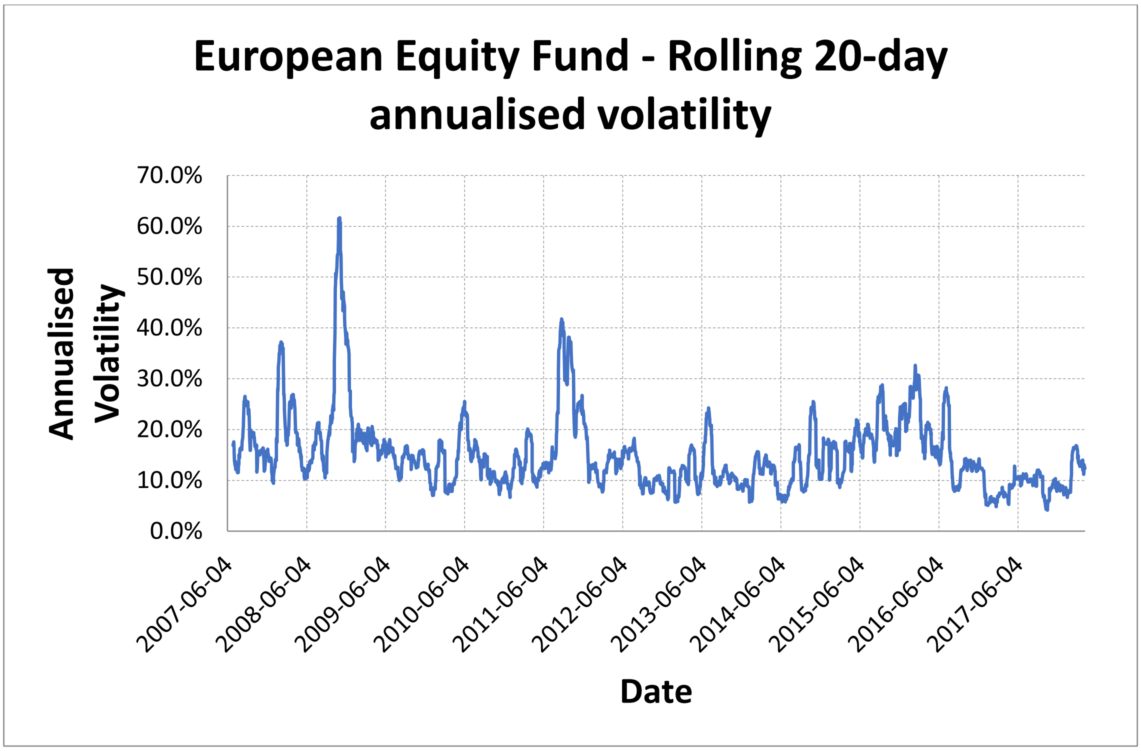


The annualised volatility of the IGC Bond fund over its entire data set is 3.5% as per Table 11.

The main risks in a bond portfolio are interest-rate risk, credit risk, currency risk, and liquidity risk.

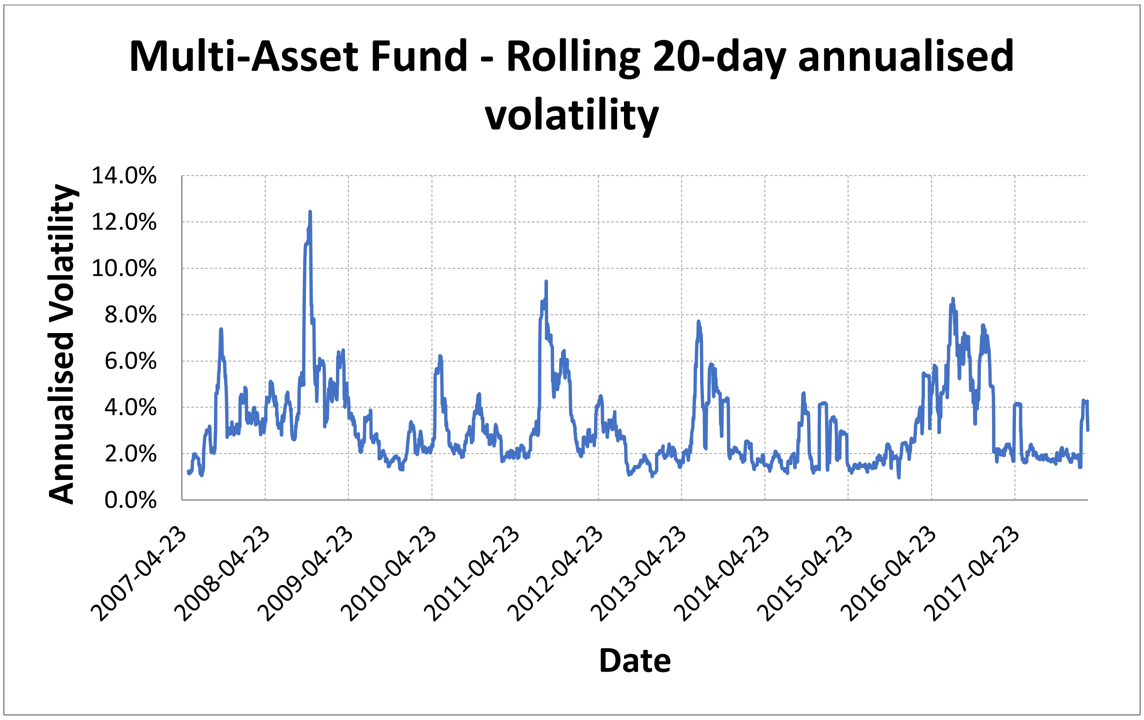
Chart 25 shows that the rolling 20-day volatility varies significantly around this long-term average level of 3.50%. Indeed, during the period that the IGC Bond fund suffered its largest peak-to-trough fall in value, September 2008, rolling 20-day volatility rose significantly above its long-term average level. Chart 25 suggests that there is no significant attempt on the part of the IM to control the volatility of the fund within a tight range.

Chart 26



The annualised volatility of the European Equity fund over its entire data set is 16.9% as per Table 11. Chart 26 shows that the rolling 20-day volatility varies significantly around this long-term average level of 16.9% rising at one point to over 60%. Chart 26 suggests that there is no significant attempt on the part of the IM to control the volatility of the fund; the fund volatility seems to vary in line with market volatility.

Chart 27



The annualised standard deviation of the Multi-Asset fund is 3.8%. During the period that the Multi-Asset fund suffered its largest peak-to-trough fall in value, June 2007 to March 2009, the rolling 20-day volatility spiked above 7% and 15%.

Rolling, Annualised, 5-day Volatility

Charts 28, 29, and 30 show the rolling, 5-day, annualised volatility for the three funds and provide a more short-term measure of variation in portfolio risk. The variation in annualised volatility in these charts is more extreme than in charts 25, 26, and 27 simply due to the smaller sample size used in 5-day volatility measures than in 20-day volatility measures.

Large spikes in 5-day volatility may be indicative of a pricing or valuation error and ought to be investigated in light of market conditions by the due diligence team.

Chart 28

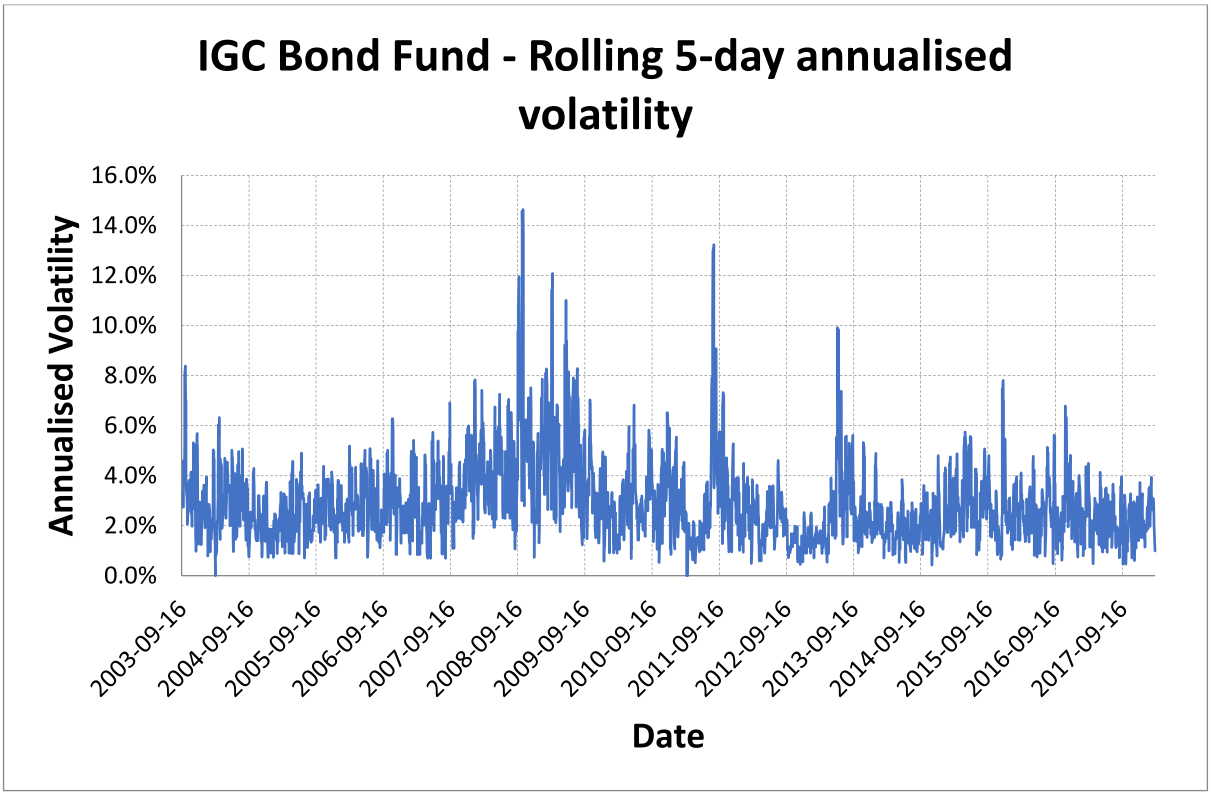


Chart 29

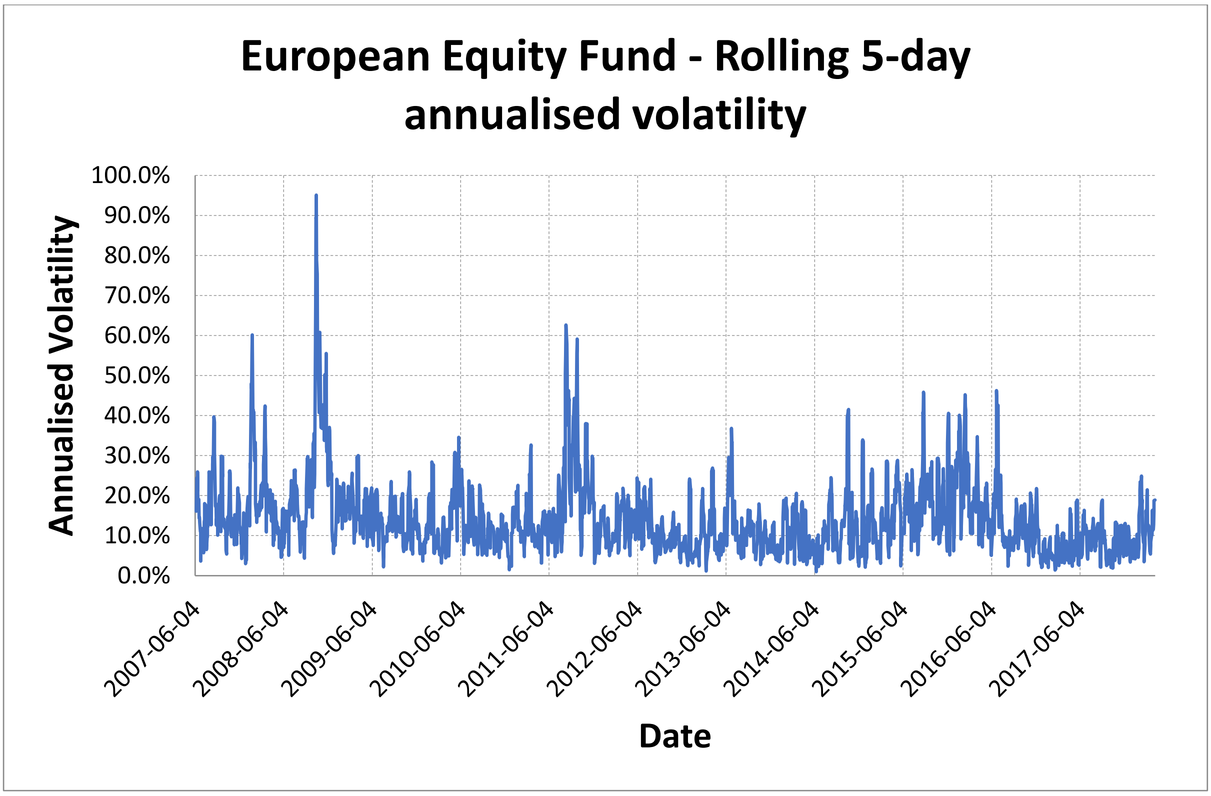
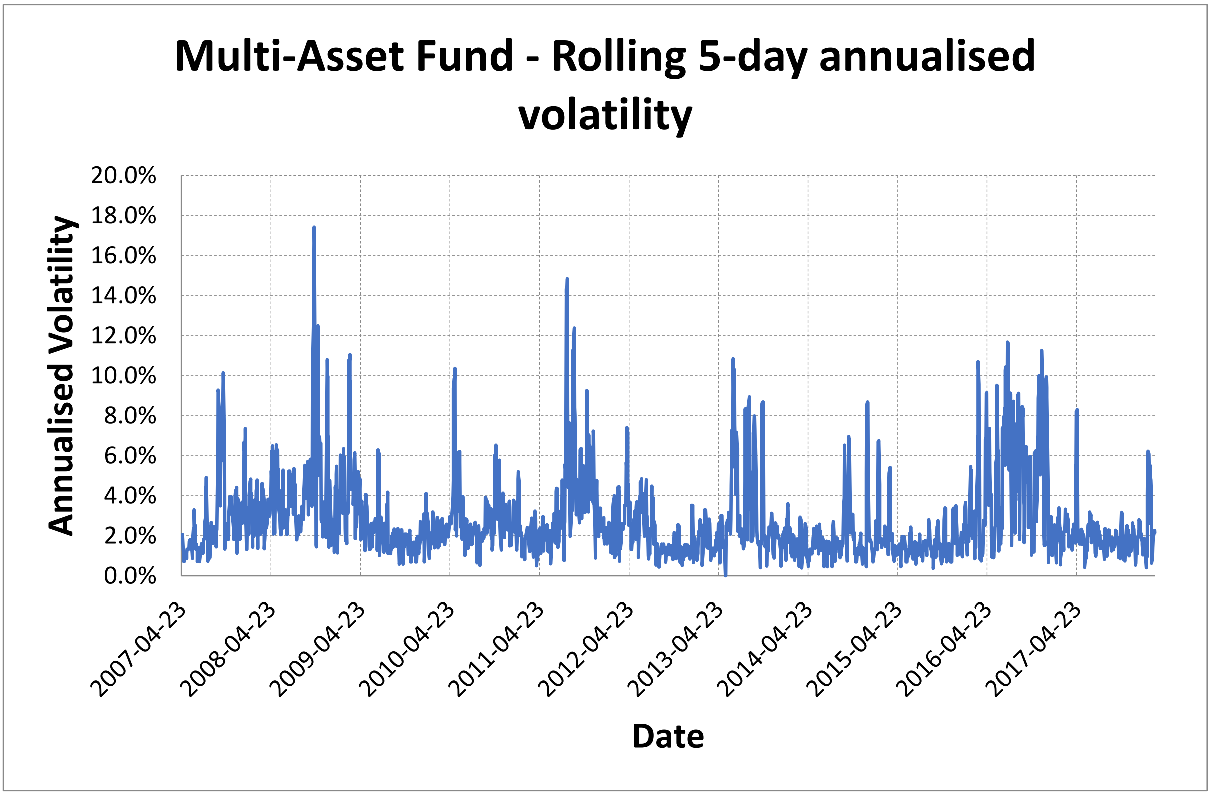


Chart 30



Risk and Performance Fees

When reviewing funds that charge performance fees, it is important to realise that performance fees may change the risk taking behaviour of the IM. Variations in risk may have implications for the investor especially where the investor has included the fund in a portfolio which is designed to achieve a certain risk target.

For these reasons, it is important to monitor the annualised, rolling 5-day and 20-day volatility of the fund and examine the way in which volatility varies with the fund's 'distance' from the threshold at which performance fees accrue and the time period remaining until the performance fee becomes payable by the fund to the IM.

Any behaviour where it is suspected that the variations in the risk of the fund cannot be justified for investment reasons needs to be discussed with the IM.

Ranking of Daily Returns – Risk Issues

Ranking the daily returns in descending order quickly shows percentage of the total return over the observation period that was achieved in just one, two, three, or four days. Where a high proportion of the total return over the observation period is achieved over just a few days, this may indicate a significant variation in the level of risk taken by the portfolio from day to day, and would warrant a discussion with the IM on volatility management strategy.

Maximum Peak-to-trough Fall in Value

The maximum peak-to-trough fall in value is principally[7] a function of time, volatility, and variation in volatility. When comparing the maximum peak-to-trough fall in value statistic across a number of different funds the length of the time period over which the statistic is assessed should be the same as the likely size of the maximum peak-to-trough fall in value varies with time.

The maximum peak-to-trough fall in value statistic for the set of funds we are considering in this paper is shown in Table 12. All three funds are evaluated over the same time period, namely, 4 June 2007 to 8 March 2018, a period of approximately 10.77 years. A snippet of the Python code to compute the maximum peak-to-trough fall in value is shown below.

Python Code Snippet

```
    old_peak = 0
    max_peak_to_trough = 0

    for dummy_date in np.arange(len(df_prices) - 1):
        peak = df_prices.Price[dummy_date]

        if peak > old_peak:
            old_peak = peak
            trough = df_prices.Price[dummy_date + 1:].min()
            peak_to_trough = np.log(peak) - np.log(trough)
            if max_peak_to_trough < peak_to_trough: max_peak_to_trough = peak_to_trough

    if max_peak_to_trough < 0: max_peak_to_trough = 0

# Calculate the non-logarithmic equivalent of the maximum peak-to-trough fall.
    max_peak_to_trough_non_log = 1 - np.exp(-max_peak_to_trough)
```

[7] The mean return also affects the size of peak-to-trough falls in value but its impact is significantly less than any of the other three factors identified.

Table 12

Fund	Annualised Standard Deviation of Daily Returns	Excess Kurtosis	Maximum Peak-to-trough Fall in Value[8]
IGC Bond	3.5	5.9	12%
European Equity	16.9	4.0	52%
Multi-Asset	3.8	8.6	12%

The maximum peak-to-trough fall in value statistics for the European Equity fund and the Multi-Asset fund are roughly the same multiple of annualised standard deviation of daily returns (the "Multiple"), 3.1 and 3.2 times respectively.

The Multiple for the IGC Bond, 3.4 times, is higher than for the other two funds. Compared with the European Equity fund, the excess kurtosis statistic for the IGC Bond would certainly lead one to expect a higher element of 'surprise' in the size of maximum peak-to-trough fall in value relative to the annualised standard deviation of daily returns of the fund.

Surprisingly, the Multiple for the Multi-Asset fund, 3.2 times, is lower than one might have expected from the excess kurtosis statistic of the fund.

[8] The maximum peak-to-trough fall in value shown is the non-logarithmic value. The output from the programs accompanying the paper contains both the logarithmic and non-logarithmic values for the maximum peak-to-trough fall in value statistic.

Risk & Return Measures

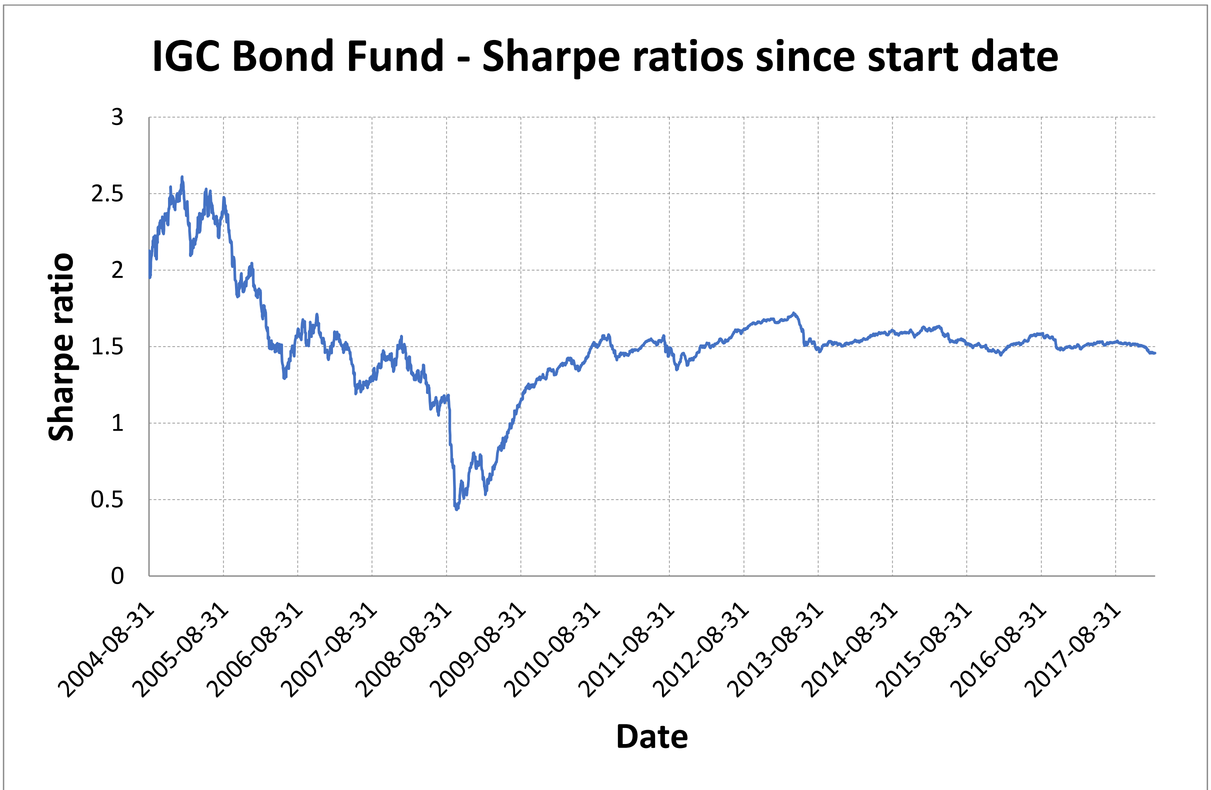
Sharpe Ratio

Charts 31, 32, and 33 graph the trend in the Sharpe ratio[9] of the set of funds we examine in the paper. The first Sharpe ratio is calculated for a period of one year from the start date of the data set; subsequent Sharpe ratios cover the period from the start date of the data set to the date shown. Such graphs are useful in assessing risk-adjusted returns. The Python code snippet for producing the graphs is given below.

Python Code Snippet

```
df_sharpe = df_returns[['Date', 'Return']]
df_sharpe['Days since start'] = (df_sharpe.Date - start_date) / np.timedelta64(1, 'D')
df_sharpe['Annualized return since start'] = df_sharpe.Return.cumsum() * 365.25 / df_sharpe['Days since start']
df_sharpe['Annualized standard deviation since start'] = df_sharpe.Return.rolling(
    len(df_sharpe), min_periods = 2).std() * mean_days_per_year ** 0.5
df_sharpe['Sharpe ratio since start'] = df_sharpe['Annualized return since start'] / df_sharpe[
    'Annualized standard deviation since start']
df_sharpe['Annualized standard deviation since start'].fillna('', inplace = True)
df_sharpe['Sharpe ratio since start'].fillna('', inplace = True)
df_sharpe.drop(['Return', 'Days since start'], axis = 1, inplace = True)
```

Chart 31



The period from end August 2005 to end August 2009 exhibits a declining since-start-date, Sharpe ratio. It is interesting to investigate if the decline was due to an increase in since-start-date risk or a fall in the since-start-date return. A review of the period shows that since-start-date risk was reasonably stable whereas since-start-date return declined. The since-start-date Sharpe ratio has held steady around 1.5 since the end of August 2010, a period which has been characterised by a bull run in investment-grade credit bonds.

[9] For the purpose of calculating Sharpe ratios in this paper, we have assumed that the risk-free rate of interest is zero.

Chart 32

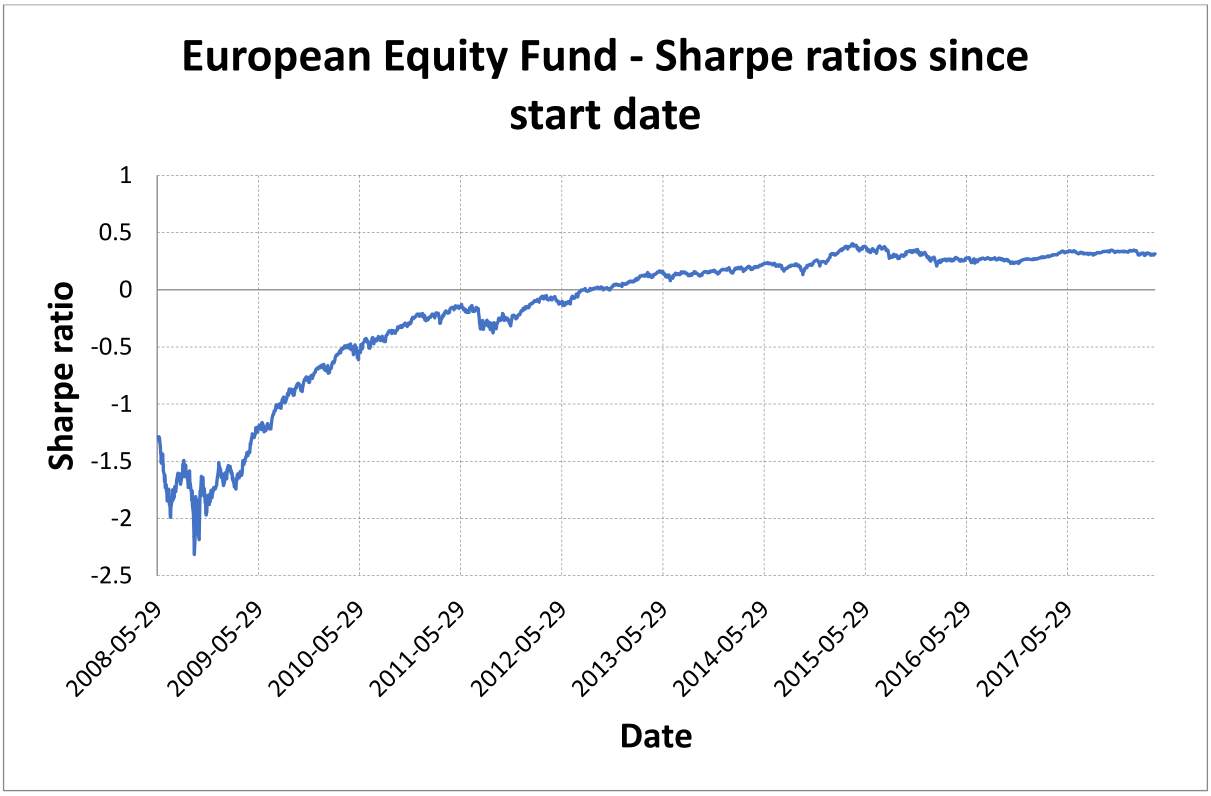
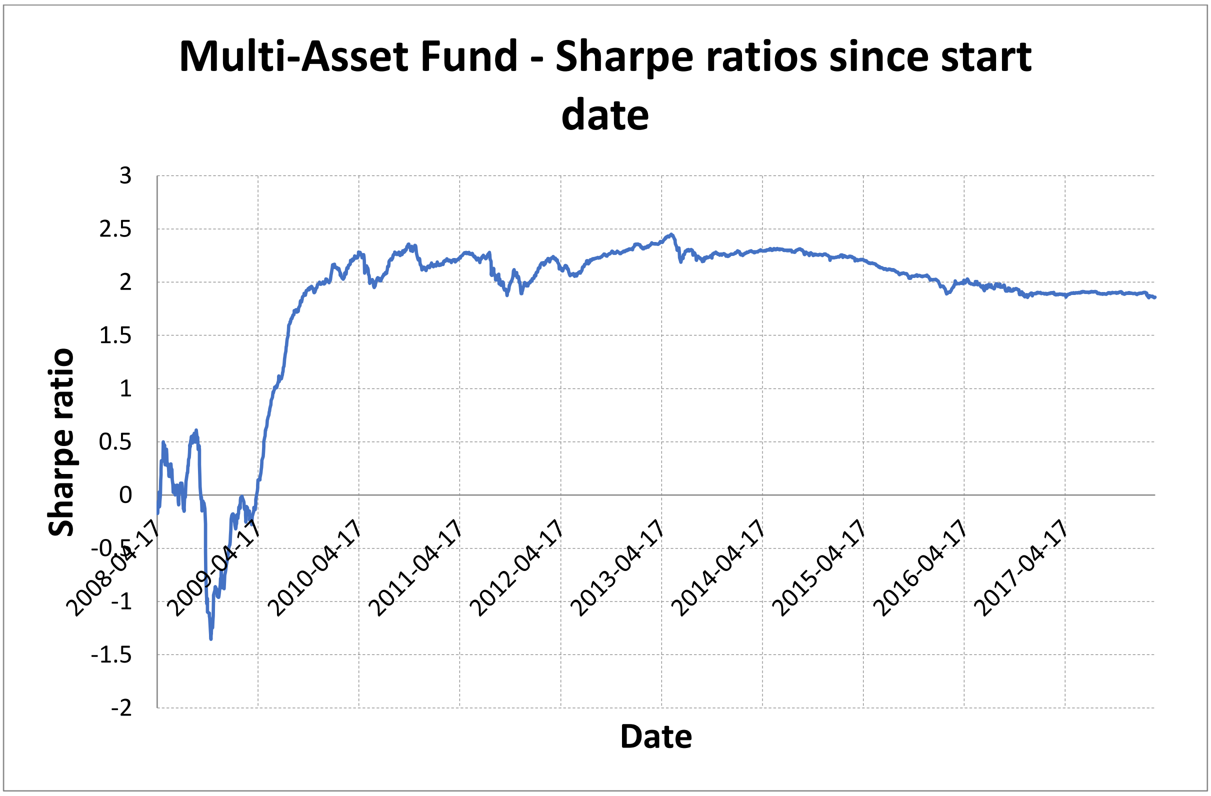


Chart 33



Such graphs allow a comparison between investment managers and prompt discussions about the reasons for long-term trends in Sharpe ratio.

Correlation Analysis

In this section of the paper we consider how:

1. The IGC Bond fund and the Multi-Asset fund are correlated; and
2. The European Equity fund and the Multi-Asset fund are correlated.

This type of analysis might be carried out if we were considering a portfolio consisting of either the IGC Bond fund or the European Equity fund and the Multi-Asset fund. Given that the annualised return of the IGC Bond fund and the European Equity fund is lower than that of the Multi-Asset fund, any combination of the Multi-Asset fund and one of the other two funds would have resulted in a lower annualised return. However, we are more interested in finding out if the addition of either the IGC Bond fund or the European Equity fund to the Multi-Asset fund would have produced a combination of funds which had a better risk-adjusted return profile than the Multi-Asset fund alone.

Valuation Time and Valuation Sources

Not surprisingly, if we calculate the correlation of daily returns between the EURO STOXX 50 index with a daily valuation time of 0900 hours GMT with that of the EURO STOXX 50 index with a daily valuation time of 1300 hours GMT, we will find that the index is not 100% correlated with itself. Similarly, it is important when calculating the correlation of daily returns for two different funds that enquiries are made as to: (i) the valuation times of the two funds; and (ii) whether those valuation times have varied over the life of the two funds in question.

Enquiries might also be made as to the source of valuation data for the two funds. Different sources can give rise to different valuations for the same fund at the same valuation time. This is particularly the case when some of the instruments to be valued are OTC derivatives or are otherwise traded in OTC markets.

3x3 Tables

Table 13 illustrates the possible combinations of daily returns of two funds, the IGC Bond fund and the Multi-Asset fund. The two funds have a common history of daily returns covering 2,751 days.

Table 13

		Multi-Asset Fund Classification of Daily Return		
		Positive	Negative	Zero
IGC Bond Fund Classification of Daily Return	Positive	29.8%	11.3%	6.6%
	Negative	11.2%	17.8%	5.7%
	Zero	9.1%	5.3%	3.2%

On almost 30% of days, the return of one of the two funds in Table 13 is zero. Perhaps this is not surprising; the Multi-Asset fund allocates at least 50% of its NAV to bonds while the IGC Bond fund allocates 100% of its NAV to bonds and the IGC Bond fund has a high proportion of zero return days. The shaded cells cover 48.7% of daily returns where a combination of the two funds does not lose money or makes a positive return. Both funds have negative days at the same time on 17.8% of days.

In 22.5% of days one fund's negative return is at least partially compensated by a positive return from the other fund.

The Python code snippet below illustrates one means of calculating the figures in the table.

Python Code Snippet

```
pos_pos_days = len(df_returns.loc[(df_returns.Return1 > 0) & (df_returns.Return2 > 0)])
pos_pos_prop = pos_pos_days / num_returns
pos_zero_days = len(df_returns.loc[(df_returns.Return1 > 0) & (df_returns.Return2 == 0)])
pos_zero_prop = pos_zero_days / num_returns
pos_neg_days = len(df_returns.loc[(df_returns.Return1 > 0) & (df_returns.Return2 < 0)])
pos_neg_prop = pos_neg_days / num_returns
zero_pos_days = len(df_returns.loc[(df_returns.Return1 == 0) & (df_returns.Return2 > 0)])
zero_pos_prop = zero_pos_days / num_returns
zero_zero_days = len(df_returns.loc[(df_returns.Return1 == 0) & (df_returns.Return2 == 0)])
zero_zero_prop = zero_zero_days / num_returns
zero_neg_days = len(df_returns.loc[(df_returns.Return1 == 0) & (df_returns.Return2 < 0)])
zero_neg_prop = zero_neg_days / num_returns
neg_pos_days = len(df_returns.loc[(df_returns.Return1 < 0) & (df_returns.Return2 > 0)])
neg_pos_prop = neg_pos_days / num_returns
neg_zero_days = len(df_returns.loc[(df_returns.Return1 < 0) & (df_returns.Return2 == 0)])
neg_zero_prop = neg_zero_days / num_returns
neg_neg_days = len(df_returns.loc[(df_returns.Return1 < 0) & (df_returns.Return2 < 0)])
neg_neg_prop = neg_neg_days / num_returns
```

Chart 34 below shows a scatter plot of the results from Table 13.

Chart 34

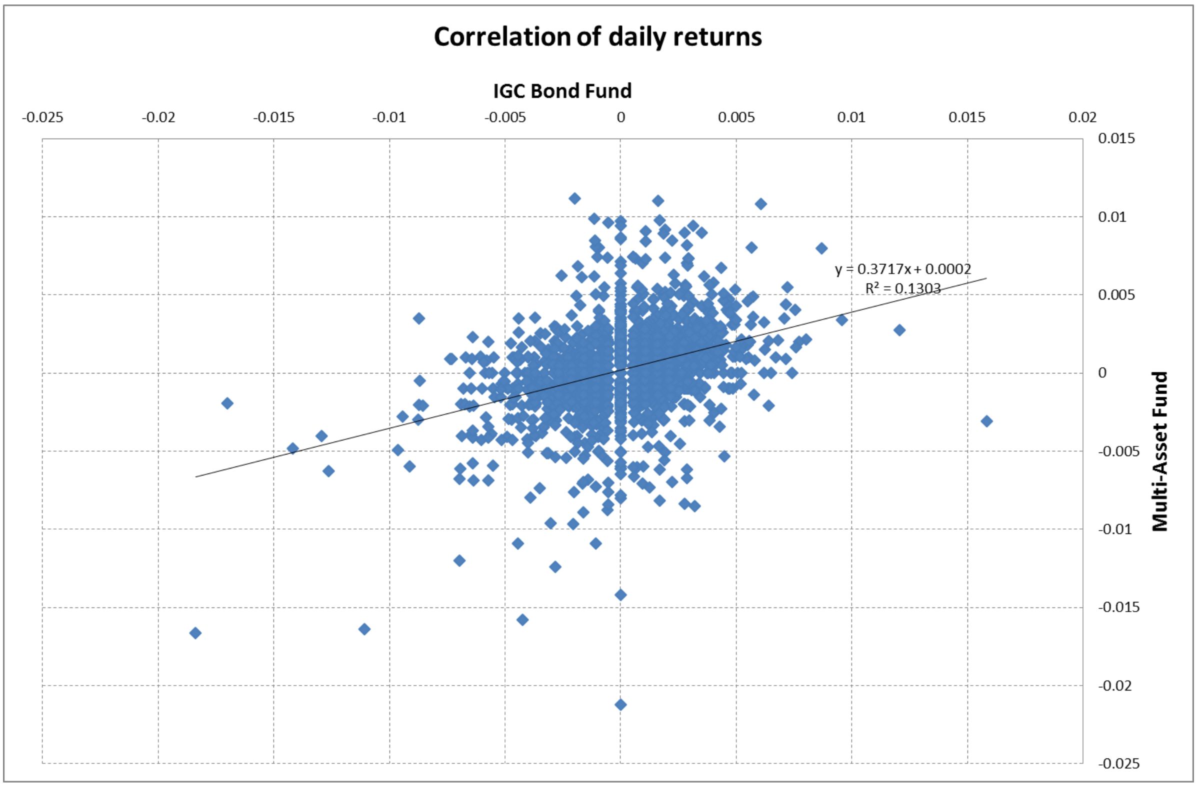


Table 14 illustrates the possible combinations of daily returns of two funds, the European Equity fund and the Multi-Asset fund. The two funds have a common history of daily returns covering 2,723 days.

Table 14

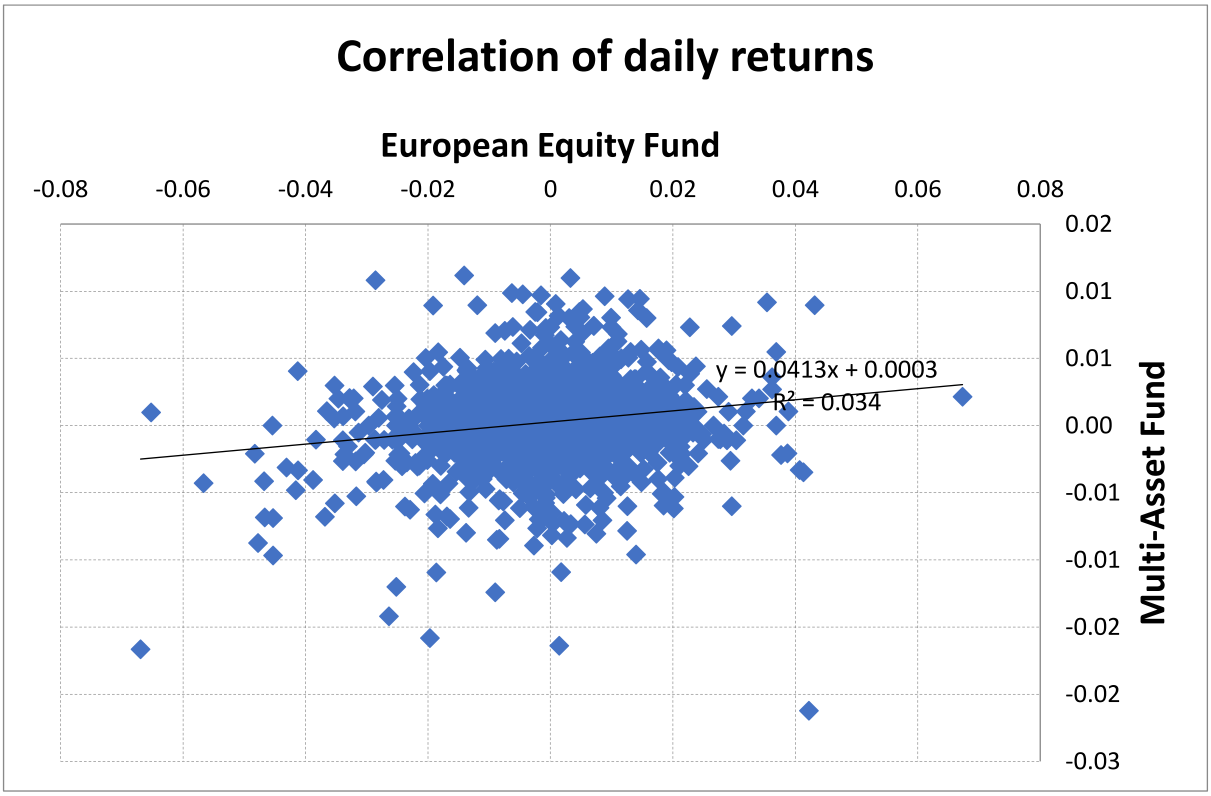
		Multi-Asset Fund Classification of Daily Return		
		Positive	Negative	Zero
European Equity Fund Classification of Daily Return	Positive	30.2%	15.0%	9.1%
	Negative	19.6%	19.5%	6.1%
	Zero	0.3%	0.1%	0.1%

On 15.7% of days, the return of one of the two funds in Table 14 is zero. The shaded cells cover 39.7% of daily returns where the fund does not lose money or makes a positive return. Both funds have negative days at the same time on 19.5% of days. In 34.6% of days one fund's negative return is at least partially compensated by a positive return from the other fund.

By contrast with the figures in Table 13, there is a significantly higher percentage (34.6% versus 22.5%) of days where the negative return of one fund is at least partially compensated for by the positive return of the other fund.

Chart 35 below shows a scatter plot of the results from Table 14.

Chart 35



Correlation of Returns

Table 15 shows the correlation matrix for the returns of the three funds, the IGC Bond fund, the Multi-Asset fund, and the European Equity fund.

The Python code snippet for calculating the correlations is provided below.

Python Code Snippet

```
correl = df_returns['Return1'].corr(df_returns['Return2'])
df_returns['Squared return1'] = df_returns['Return1'] ** 2
df_returns['Squared return2'] = df_returns['Return2'] ** 2
squared_correl = df_returns['Squared return1'].corr(df_returns['Squared return2'])
```

Table 15
Correlations of Returns

	IGC Bond	European Equity	Multi-Asset
IGC Bond	1	0.159	0.361
European Equity		1	0.184
Multi-Asset			1

The correlation between the IGC Bond fund and the Multi-Asset fund is 0.361. This may be partly due to the high percentage of days on which one or other of these funds has a return of zero (30%) and partly due to the relatively low percentage of days (22.5%) on which a loss on one fund is at least partially compensated for by a gain on the other fund.

The correlation between the European Equity fund and the Multi-Asset fund is 0.184 which is relatively low. On just 15.7% of days, the return of one of these two funds is zero so this is not as likely as in the case of the other combination of funds to be a contributory factor to the low volatility. The percentage of days (34.6%) on which a loss on one fund is at least partially compensated for by a gain on the other fund is higher for this combination of funds than the last combination and this may account for the relatively low correlation figure.

The Python code snippet below illustrates a means of producing a chart of the variation in correlation over time; rolling 60-day correlations in this case.

Python Code Snippet

```
for roll_day in roll_days:
    col_name = 'Rolling ' + str(roll_day) + '-day correlation of returns'
    df_returns[col_name] = df_returns['Return1'].rolling(window = roll_day).corr(df_returns['Return2'])
    df_returns[col_name].fillna('', inplace = True)
    col_name = 'Rolling ' + str(roll_day) + '-day correlation of squared returns'
    df_returns[col_name] = df_returns['Squared return1'].rolling(window = roll_day).corr(df_returns['Squared return2'])
    df_returns[col_name].fillna('', inplace = True)
```

Chart 36 below shows the variation over time in the rolling 60-day correlation of returns between the IGC Bond fund and the Multi-Asset fund. The rolling, 60-day correlation is positive almost all the time but it exhibits significant oscillation between 0.6 and zero while spiking at above 0.7 on one occasion.

Chart 36

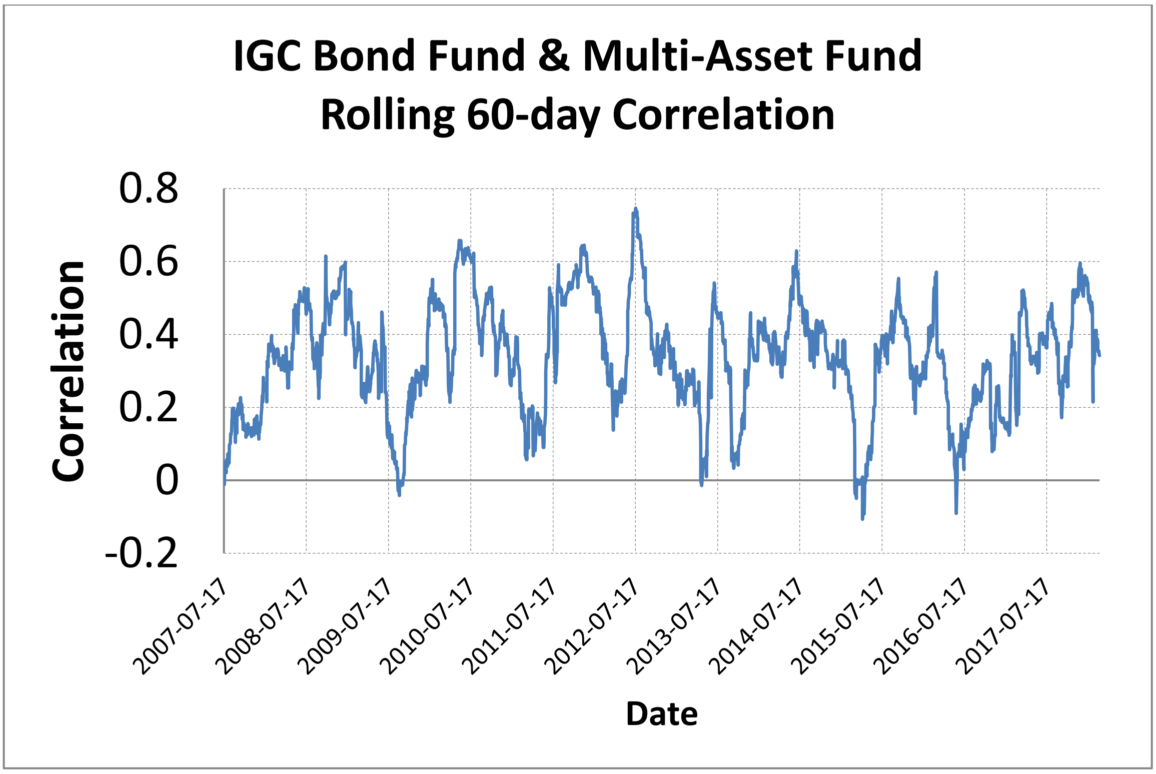
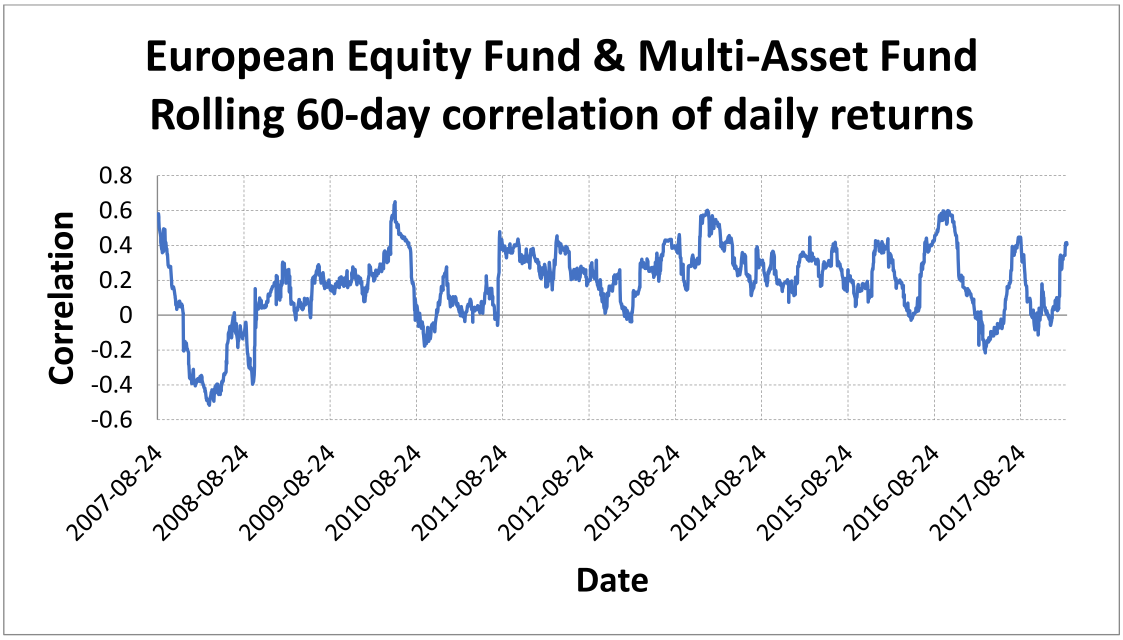


Chart 37 below shows the variation over time in the rolling 60-day correlation of returns between the European Equity fund and the Multi-Asset fund. The rolling, 60-day correlation is positive most of the time but it did turn negative during part of the Global Financial Crisis. When not negative, it exhibits significant oscillation between 0.6 and zero but with a lower frequency than the IGC Bond fund and Multi-Asset combination.

Chart 37



Correlation of Returns Squared

It is useful to measure the extent of volatility correlation between two funds. A proxy measure for volatility correlation is the correlation of the squares of the daily returns. The correlation of the squares of daily returns indicates the extent to which two funds have large moves in absolute value terms at the same time.

In cases where the correlation of daily returns is high and the correlation of daily returns squared is also high, the risk of large (+ve, +ve) and large (-ve, -ve) returns by the two funds is high. By contrast, where the correlation of daily returns is low and the correlation of daily returns squared is high, the combinations of days with large absolute returns are more likely to be of the large (+ve, -ve) and large (-ve, +ve) type returns for the funds. Table 16 illustrates the correlations of returns squared for the three funds covered in this paper.

Table 16
Correlations of Returns Squared

	IGC Bond	European Equity	Multi-Asset
IGC Bond	1	0.258	0.245
European Equity		1	0.259
Multi-Asset			1

Chart 38 shows a plot of the returns squared for the IGC Bond fund and the Multi-Asset fund. All entries in the chart are positive because the returns have been squared.

Chart 38

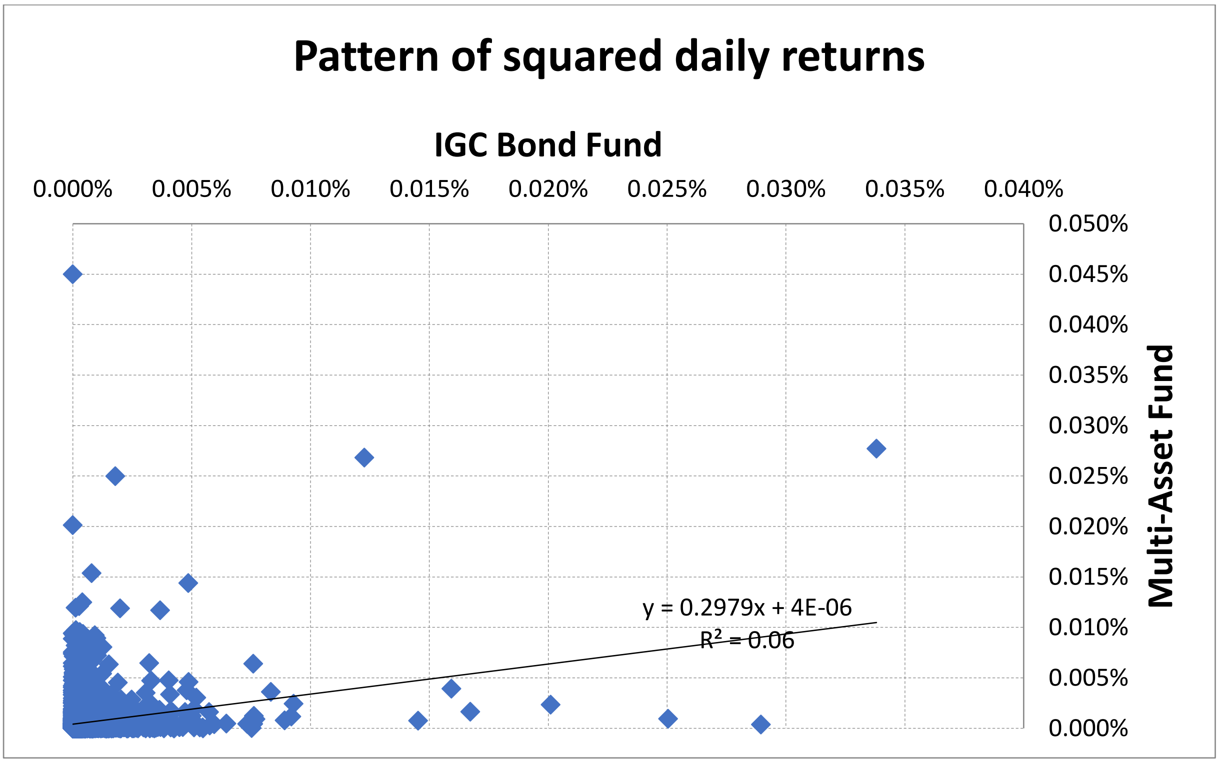
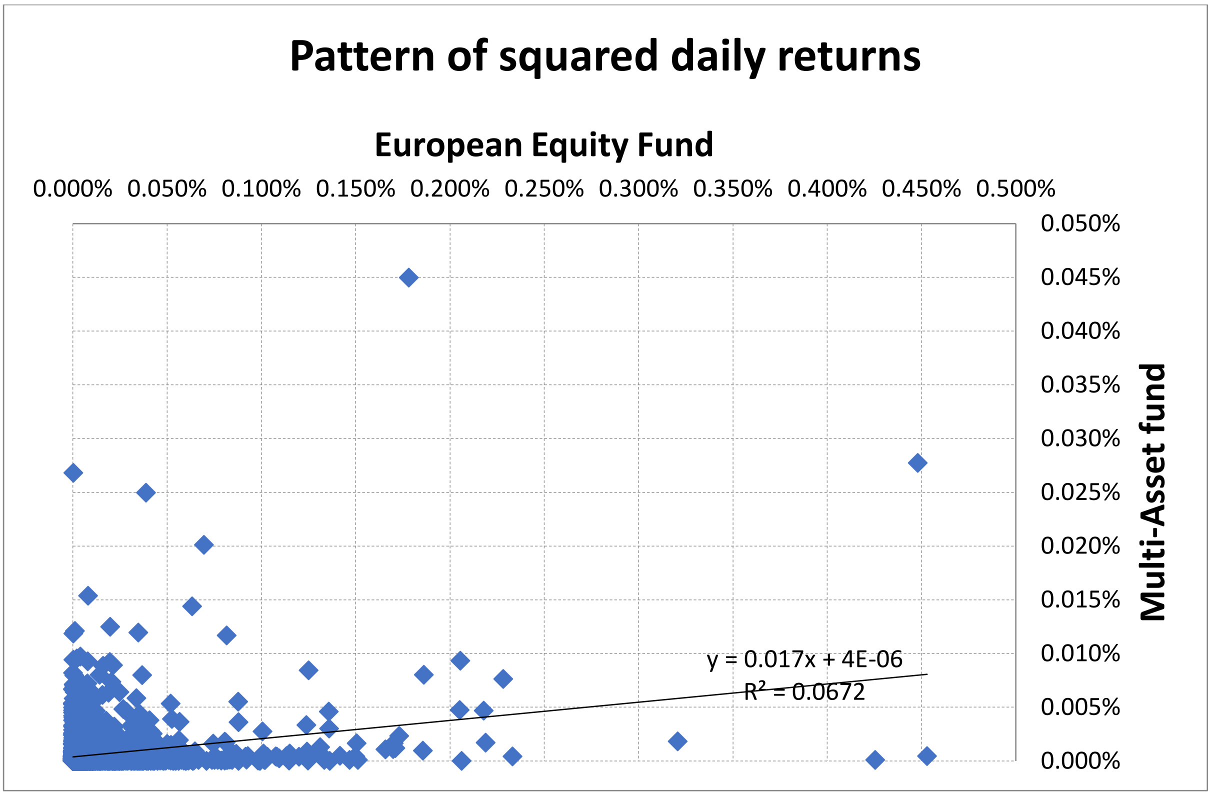


Chart 39 shows a plot of the returns squared for the European Equity fund and the Multi-Asset fund. Again, all entries in the chart are positive because the returns have been squared.

Chart 39



There is not a significant number of entries in the upper right-hand quadrant of either of charts 38 and 39 indicating that the volatility correlation between the two different combinations of funds is not very high.

Rolling 60-day Correlations of Returns Squared

Charts 40 and 41 below show the variation over time in the rolling 60-day correlation of returns squared between the IGC Bond fund and the Multi-Asset fund and between the European Equity Fund and the Multi-Asset fund.

Chart 40

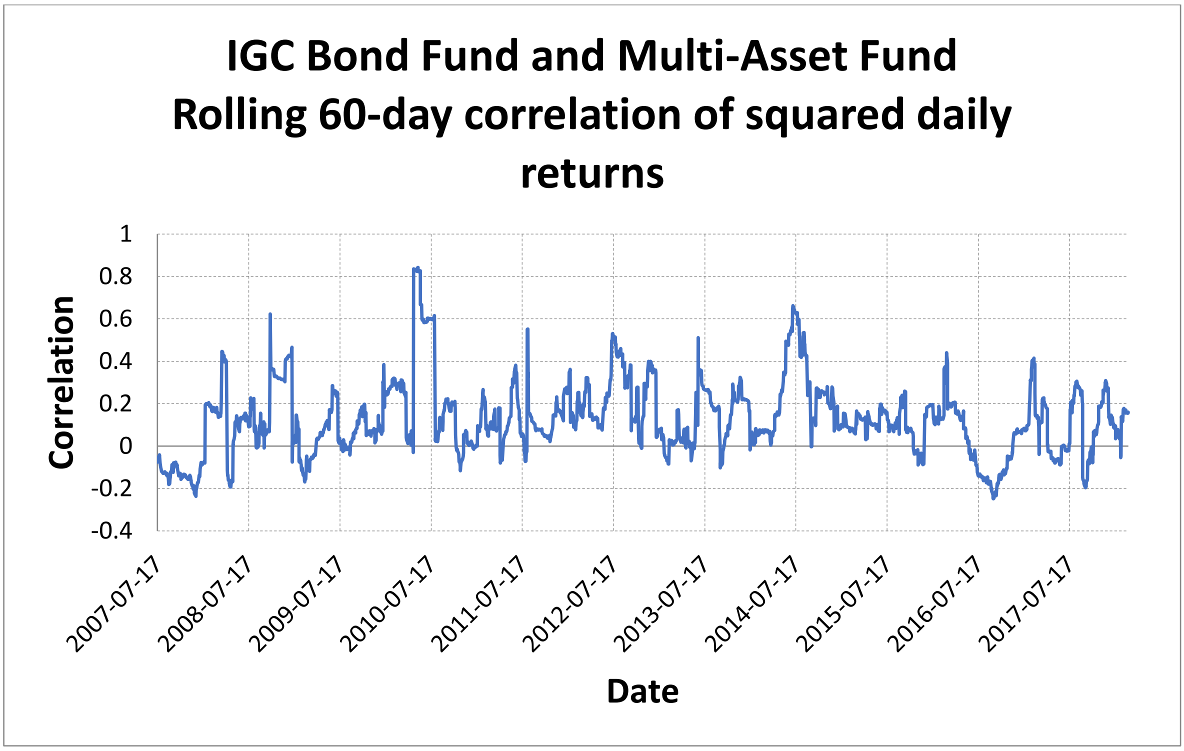
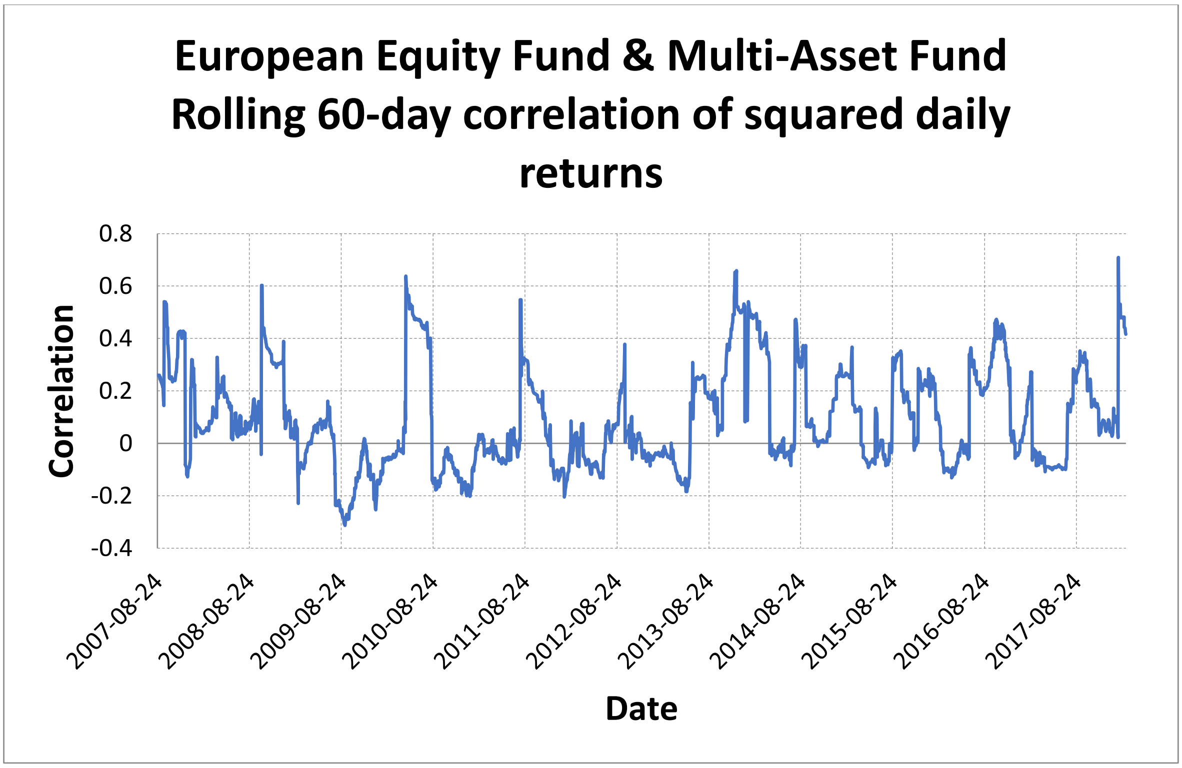


Chart 41



Classical Mean Variance Test

Classical mean-variance analysis has a test for deciding whether adding a new trading strategy ("NS") to an existing portfolio ("EP") will produce a combination of funds with a higher Sharpe ratio than the existing portfolio. Under the test, a new trading strategy should be added to the existing portfolio and the combination of the NS and the EP will have a higher Sharpe ratio that the EP if the following condition is met:

$$[\text{Sharpe Ratio}]_{NS} > [\text{Sharpe Ratio}]_{EP} * \rho_{(NS,EP)}$$

where $\rho_{(NS,EP)}$ is the correlation between the new trading strategy and the existing portfolio.

Were the returns of the new trading strategy and that of the existing portfolio normally distributed, this test would indicate that one could add the IGC Bond fund to the Multi-Asset fund to arrive at a combination of the IGC Bond fund and the Multi-Asset fund which has a higher Sharpe ratio than that of the Multi-Asset fund alone. Table 17 has the details.

Table 17

New Strategy	Correlation to Multi-Asset Fund	Sharpe Ratio of New Strategy	$[\text{Sharpe Ratio}]_{EP}*\rho(NS,EP)$ ($[\text{Sharpe Ratio}]_{EP}=1.858$)	Result of Test $[\text{Sharpe Ratio}]_{NS} > [\text{Sharpe Ratio}]_{EP}*\rho(NS,EP)$
IGC Bond	0.361	1.459	1.858*0.361 = 0.671	ADD
European Equity	0.184	0.315	1.858*0.184 = 0.342	DO NOT ADD

According to this test which again is based on the assumption that the new trading strategy and that of the existing portfolio are normally distributed, adding the European Equity fund to the Multi-Asset fund will not lead to a combination of the Multi-Asset fund and the European Equity fund with a higher Sharpe ratio than the Multi-Asset fund.

If the long-term Sharpe ratio of the Multi-Asset fund is assumed to be 0.5 and that of the two single strategy funds is assumed to be 0.4, then the test would suggest that both funds should be added.

It is important to bear in mind that one would not decide to add a fund to an existing portfolio solely based on this test. One needs to consider the implications of all the statistics and graphs in this paper in making such a decision. As we saw earlier, the negative skew and positive excess kurtosis of the three funds warn us that we cannot rely on standard deviation as a measure of risk in mean-variance analysis. If we are trying to avoid 'negative' surprises, large daily losses, then we should not rely on standard deviation as our measure of risk in portfolio construction.

Python Programs

There are two Python programs accompanying this paper; one program produces an analysis of an individual fund whereas the other covers a number of different types of correlation for a pair of funds.

The Python scripts contains many comments to make them accessible to individuals who are new to Python code. The Python script for analysing an individual fund reads in data from a spreadsheet containing the dates and NAV prices of the fund to calculate the fund's daily returns and a range of summary statistics relating to these returns. The Python script for analysing the different types of correlation between a pair of funds also reads in data from a spreadsheet with the prices and the dates for the two funds in the same spreadsheet. Both programs write the results to a spreadsheet.

Python Version

The script is written in Python 3.6 and uses the following libraries:

- pandas (version 0.22.0),
- SciPy (version 1.0.0),
- Xlrd (1.0.0),
- XlsxWriter (0.9.6),
- numpy (1.14.2),
- statsmodel (0.8.0), and
- datetime.

Change of Variables

The following comment:

> *# You can change the ten variables below without having to adjust the rest of the script.*

appears in the code for the analysis of a single fund. In the code for the analysis of the different types of correlation between two funds, the following comment appears:

> *# You can change the five variables below without having to adjust the rest of the script.*

These comments identify the variables in the programs that may be changed without any need to adjust other parts of the programs.

Location of Input File

When setting the location and name of the spreadsheet containing the date and price data, it is useful to recall that Python uses the / symbol rather than \ to separate folders and sub-folders.

Input File Layout

The column titles in the spreadsheet containing the input data might be headed "Date" for dates and "NAV" for prices. The program does recognize other column titles.

Output Files

The default output files are EXCEL spreadsheets with “.xlsx” extensions and they are written to the same location as the input file.