NON GAUSSIAN RETURNS: WHICH IMPACT ON DEFAULT OPTIONS RETIREMENT PLANS?

Stéphane Hamayon (Harvest)  
Florence Legros (University Paris-Dauphine)  
Yannick Pradat (Harvest)

ABSTRACT

We study the statistical characteristics of the financial assets in the long run, on the 1895-2011 period. It shows that, on long holding periods, equities provide a lower risk than other assets. Moreover, equities returns are non-Gaussian and show a mean reversion process.

Consequently, simulating a pension fund asset allocation in the life cycle framework drives to use a Cornish-Fisher VaR that leads to a different asset allocation whose characteristic is a notably larger share of equities than in the Gaussian case. Back testing with computation of replacement rates show a notable dominance of the “non-Gaussian” strategy.

Keywords: financial returns, long run, retirement pension plans, life cycle allocation, value at risk, Cornish Fisher value at risk, replacement rates.

INTRODUCTION

The European Commission (EC) has introduced proposals that aim to align European pensions’ regulation with Solvency II under which insurers must hold capital to cover risks stressed over a 12-month time horizon at a confidence level of 99% using a value-at-risk measure.

These proposals had a huge echo and opponents are numerous who argue that they do not promote long term investments.

On an other hand, demographic trends and forecasts point out that long term investments must be promoted in order to adequate savings to the need for additional pensions.

These points lead to two questions:

- First, they question the long run returns by themselves. Which asset is the most adequate to long run savings?
- Second, they question the adequacy of a 12 month time horizon for a value at risk criterion. Because one of the determining factors of VaR is the length of holding period, the question of the adequacy of such a short period will be central.

Chapter one will deal with the long run statistical characteristics of financial returns in France and in United states: for this purpose, we studied the “risk-return” properties of T-bills, bonds and stocks over the long term.

1 Corresponding author : Florence Legros, Florence.legros@dauphine.fr.

The authors especially thank Frank E. Curtis from the Department of Industrial and Systems Engineering at Lehigh University, who has personally helped us to implement the "adaptive gradient sampling algorithm for nonsmooth optimization" he describes in Curtis, F. E. and Que, X. (2012).
This study clearly shows that the returns’ distributions are clearly far from the Gaussian law as soon as are considered long holding periods. This leads to take into account Skewness and Kurtosis in the VaR criterion which means that a non parametric VaR will be more adequate that a parametric one.

The second chapter analyzes the consequences of non-Gaussian returns on default options retirement plans. It appears that optimization methods incorporating higher-order moments determine a greater proportion of risky assets compared to the mean-variance approach.

This chapter also back tests these results and computes replacement rates which would result from this alternative asset allocation.

Comparing these replacement rates distributions confirms unequivocally that the asset allocation computed with the semi parametric strategy provides higher replacement rates than the Gaussian one.

**FINANCIAL RETURNS IN THE LONG RUN IN FRANCE AND IN UNITED-STATES**

VaR quantification relies in the simplest cases on the analysis of mean and standard deviation. This step is a useful preliminary to the asset allocation strategy analysis.

Considering long term savings this chapter deals with long term statistical properties of main assets; we thus check the attributes which are often assigned to the asset classes. Historical properties are considered; this can limit the prospective significance of the exercise but enlightens some simple but fundamental questions as: are the assets returns constant over time? Are shocks permanents or transitory? How does risk evolve as the holding period increases?

**Assets performances**

This paper uses and analyses total return indexes - i.e. including dividends or interests - for equities, bonds and monetary assets. These indexes are considered on the period 1895-2011 and are provided by Global financial data. When we examine nominal returns, the numbers appear similar in the two countries. However we note that periods of War are characterized by high inflation rates in France (Table 1a). The presence of “money illusion” distorts the comparison between asset returns. However the distinction between the real and the nominal rate of return is crucial in making investment choices when investors are interested in the future purchasing power of their saving. Following Knight (1921), we assumed that world wars are governed by an unknown probability model, the difference between nominal and real return drive us to exclude war periods from the French sample. Thus, based on a standard deviation criterion, we solved the problem of hyperinflation by eliminating the war periods [1915-1920] and [1940-1948] from the French dataset. This decision appears to be very drastic if are considered the historian economists studies (Arbulu, 1998, Le Bris, 2010).

---

2We use « total return » indexes computed by Global Financial Data (www.globalfinancialdata.com). For French assets, the indexes are: monetary assets, bonds (10 years maturity) and equities (SGF-Insee, SBF250 then CAC All-Tradable). Still for French data, we use the consumer price index provided by Insee (National institute for statistics and economic studies) as deflator. For the United-States, the indexes are: monetary assets (US Treasury Bills Total Return Index), bonds (US Total Return Long-term Government Bond Index) and equities (S&P 500 Composite Total Return Index).
Friggit, 2010), but it allows a long term parallelism between French and American asset returns (table 1b).

Table 1a – Financial assets: average annual nominal rate of return (dividends included).

<table>
<thead>
<tr>
<th>Period</th>
<th>United-States</th>
<th>France</th>
</tr>
</thead>
<tbody>
<tr>
<td>1895-1914</td>
<td>4.3%</td>
<td>2.5%</td>
</tr>
<tr>
<td>1914-1950</td>
<td>2.2%</td>
<td>3.5%</td>
</tr>
<tr>
<td>1950-2011</td>
<td>5.9%</td>
<td>6.0%</td>
</tr>
<tr>
<td>1995-2011</td>
<td>3.1%</td>
<td>6.9%</td>
</tr>
<tr>
<td>1895-2011</td>
<td>4.6%</td>
<td>7.3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Period</th>
<th>United-States</th>
<th>France</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monetary assets</td>
<td>2.4%</td>
<td>-0.3%</td>
</tr>
<tr>
<td>Bonds</td>
<td>1.2%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Equities</td>
<td>5.8%</td>
<td>4.6%</td>
</tr>
</tbody>
</table>

The retrospective analysis underlines a certain bias that would result from a nominal analysis: it would clearly overestimate the so-called “riskless asset” returns compared to so-called “risky asset” returns. More, if the analysis is done in real terms (rather than nominal terms) the assets hierarchy is changed: as soon as the investor’s target is to maintain his portfolio’s in real terms, the “riskless asset” disappears and the zero loss probability is null for all the considered assets (charts 1a and 1b).

On long term horizons, the only assets that have over performed inflation are equities. Considering bonds and monetary assets, the loss probability does not decrease with the holding period. These results are consistent with those obtained by Laulanie (2003).

Graph 1a – loss profiles according to the holding period length – United States
Risk and holding period

When considering real returns, it is useful to examine the relationship between expected returns, risk and time horizon; more precisely, our purpose is to analyse the distribution of returns around their expected value through time.

- Empirical evidences

As shown before, in the long run (nearly 30 years in France, between 20 and 30 years in United-States) and in real terms, equities appear to be noticeably less risky than bonds and monetary assets. Nevertheless, they are markedly more risky in the short run.

If we assume that the distribution of log returns is reasonably normal in the long term, Tables 2a and 2b show the speed of concentration of returns around the median value is specific to each considered asset: this is notable for equities whose gain and loss perspectives are high only in the short run; in the long run, the volatility decreases and speculative returns vanish.

Table 2a- Assets real returns changes - France

<table>
<thead>
<tr>
<th>FRANCE: Changes in equities real returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holding periods (years)</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>1st centile</td>
</tr>
<tr>
<td>median</td>
</tr>
<tr>
<td>99th centile</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FRANCE: Changes in bonds real returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holding periods (years)</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>1st centile</td>
</tr>
<tr>
<td>median</td>
</tr>
<tr>
<td>99th centile</td>
</tr>
</tbody>
</table>
Apart from a slight excess return for the US Stocks market, we observe the same dynamic of risk reduction over time in the French case and in the American case. In order to measure the interest such a consideration may have for an investor, it is necessary to get back to the parametric assumption.

If the returns standard error does not decrease quicker than a Gaussian white random process, the previous conclusions have to be changed: this random process has no memory and shocks are permanent. These properties correspond to the efficient markets hypothesis by Fama (1965). According to Fama, if markets are efficient, they are not predictable. In a dynamic which has no memory, the loss probability is independent of the holding period and a risk adverse investor flees equities.

• “Dilution” of risk varies depending on the assets

If we use volatility as a proxy for risk, we can compare the historical and theoretical values over time. An overview of figures 2a to 2f, linking volatility and horizon for the considered assets, allows for several observations.

Once again, the behavior of shares differs from that of fixed income assets. For this unique asset class, volatility decreases quicker than for the Gaussian with same risk/return ratio. For equities again, the speed of risk reduction over time and the historical expected returns are similar in France and in United-States. On the other side, for bonds as for monetary assets, volatility is higher in France than in US, due to high inflation periods during the considered period.
In the case of fixed income assets, the excess of volatility relative to a Gaussian can be attributed to a non stationarity presumption. When a process is stationary at the second order, the series oscillates around its mean with constant variance. The origin of the non-stationarity of fixed income assets can come from a dependence of their expectancy and / or their variance over time.

Equities are considered as stationary by the tests (table 3) which means that the expected return is constant over time. This property distinguishes the shares of other asset classes, but
does not provide to investors, by itself, any certainty about the amount of capital at the end of
its investment period. We must recall that unit root tests are performed on the financial
indexes’ growth; the diagnosis is that stock prices are difference stationary (DS) and a DS
series does not systematically comes back to her deterministic trend after a shock because the
shock permanently affects the stochastic component of the series. That means that the equities yield is driven by a mean reversion process; nevertheless the shares’ price can be unpredictable because – once integrated – the DS series follows a random walk with drift. However for investors, holding risky assets in the long term seems to have no interest unless the prices, and not the returns are trend stationary (TS).

Table 3 – Stationarity tests

<table>
<thead>
<tr>
<th>Stationarity tests: assets real returns (1895 – 2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Augmented Dickey Fuller</strong></td>
</tr>
<tr>
<td>Log-Returns USA</td>
</tr>
<tr>
<td>Equities</td>
</tr>
<tr>
<td>Bonds</td>
</tr>
<tr>
<td>Monetary assets</td>
</tr>
<tr>
<td>Log-Returns France</td>
</tr>
<tr>
<td>Equities</td>
</tr>
<tr>
<td>Bonds</td>
</tr>
<tr>
<td>Monetary assets</td>
</tr>
</tbody>
</table>

In order to clarify the nature of the relationship between risk reduction and investment length, it remains to consider the distribution of returns over time; the inferences drawn from volatility indicators can be misleading if the variables are not correctly described by their first two moments (expectation, variance).

In this purpose, we compared the theoretical and historical distributions of returns for various holding horizons. The exercise is performed only for shares, because there is some empirical evidence that in the long run risks to stocks tend to diminish more rapidly than in the Gaussian case.

The distribution of historical returns is then compared to that of a Gaussian with same risk/return ratio. The historical returns are first computed on a 12 month rolling window over the period 1895-2011. Cumulative distributions are slightly more risky than the same Gaussian mean-variance; table 4 shows high kurtosis i.e. is an evidence for a leptokurtic distribution and negative skewness i.e. asymmetry on the loss side.

\[ \varepsilon_t \rightarrow N(0, \sigma^2) \] With \( L \) the lag operator, \( d \) the order of integration and \( b \) the drift. The process can be rewritten as \( X_t = b + X_{t-1} + \varepsilon_t = b + (b + X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \ldots = X_0 + bt + \sum \varepsilon_i \). The process has a stochastic trend, which includes a deterministic component that can account for a time series tendency to increase on average over time. Like the pure random walk, it will be characterized by a long-run forecast error variance that is increasing without bound as the forecast horizon gets sufficiently long.
In a second time, the historical returns are computed with an increasing moving window size. For a 360 month holding period (30 years – table 4), the insights derived from the study of volatility are reinforced. It is clear that the cumulative function can be considered as Gaussian: the leptokurtic shape has disappeared and the skewness now shows asymmetry on the gain side.

Table 4 – Statistics on equities’ returns distribution function on the 1895-2011 period

<table>
<thead>
<tr>
<th>Holding period 12 months</th>
<th>Holding period 360 months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>USA S&amp;P 500</td>
</tr>
<tr>
<td>Mean</td>
<td>6.20%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>20.70%</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>5.1</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.6</td>
</tr>
</tbody>
</table>

Nevertheless, the main teaching comes from the comparison with the theoretical distribution with parameters $\mu_T$, $\sigma_T/\sqrt{T}$ (Graphs 3a, b, c, d). Considering the American and the French shares, the comparison of the historical repartition function ($F$) with the theoretical one ($G$), it clearly comes out that $F$ is more concentrated than $G$. $G$ has the same distribution than $F+\epsilon$ with $E(\epsilon/F)=0$. Taking into account that the distributions have the same expectancy and that $V(F)<V(G)$, it clearly comes that a risk adverse investor will prefer $F$ to $G$.

Graphs 3 a and b – Repartition function – French equities
To sum up, we can say that over short time horizons, distribution shows a leptokurtic shape. Lengthening the holding period in the stock market lowers risk: it decreases faster than in the Gaussian case. It appears that stock returns exhibit negative autocorrelations over long horizons. This diagnosis, if proved, challenges the assumption of efficient markets and calls for alternative modeling of the stocks’ prices dynamics in which the prices’ extreme movements that can be observed in financial markets are transitory: a large change in one direction is followed by opposite movements that correct the initial deviation and bring asset prices towards their fundamental values. Because this mechanism is more likely when the holding period is long, an increased horizon leads to a risk reduction.

A quick glance at markets’ efficiency

In order to test the weak form efficiency hypothesis of the French and American stocks’ markets in a more rigorous way, we have completed our analysis by “variance ratio” tests on the logarithm of nominal equities’ prices.

Since the paper by Lo and MacKinlay (1988), the variance ratio test and related ones such as the non-parametric test by Wright (2000), for instance, became the favorite tools in order to test the efficiency hypothesis.

The intuition behind this test is that if the asset price follows a random walk, its variance increases linearly with the interval of observations. It follows that the variance of returns $r_t$ over a $q$ long period should be approximately equal to $q$ times the variance of a unit period. If we denote $Var (r_t (q))$ the variance of the sum of the first differences of the logarithm of the $q$ long period, then the variance ratio test $VR$ is:

$$VR(q) = \frac{Var(r_t(q))}{qVar(r_t)}$$

Lo and MacKinlay put in evidence that for $q=2$, $VR(2)-1=\rho(1)$, where $\rho(1)$ is the first order autocorrelation coefficient.
If the ratio is close to one (1), it suggests a random walk while a significantly different value is an indicator for a returns autocorrelation. A weak ratio – lower than one (1) – pleads for a mean reversion process.

Taking into account the high sensitivity of the test, the variance ratios have been estimates on a rolling time window. This method provides a good visibility: graphs 4a and b exhibit that the tests which are done on a yearly frequency provide variance ratios that are mainly lower than one for the French CAC All-tradable and for the S&P500, i.e. mean reversion processes for both the series.

These results can appear as fragile given the significance thresholds but they are in line with Poterba and Summers (1988), Bec and Gollier (2007) and Hamayon and Legros (2008), who all conclude that the excess changes in observed prices on the financial markets are transitory. These arguments justify the time diversification strategies on savings in the long term.

Graphs 4 a and b – Variance ratio tests– French and American indexes

Production and stock prices: a good cyclical synchronization

The price dynamic being driven by a mean reversion process, this should imply that there is a good connection between real and financial economies. Economic theory suggests that there should be a strong link between economic activity and security prices, given that the stock price is the discounted present value of the firm’s payout.

In order to empirically check this hypothesis, we compare the industrial production cycles with stocks prices. We use a Hodrick-Prescott\(^4\) filter (1981) to decompose the series between

\[ y_t = g_t + c_t \]

where \( g_t \) is the trend value in \( t \), \( c_t \) the cyclical component. In order to determine \( g_t \), the Hodrick Prescott filter finds a function which leads to:

- find the values of \( g_t \) that provide a good approximation of \( y_t \);
- find the values of \( g_t \) whose acceleration (\( \Delta g_{t+1} - \Delta g_t \)) is smallest as possible.

The optimization program is the following: \( \text{min} \sum (y_t - g_t)^2 + \lambda \sum (\Delta g_{t+1} - \Delta g_t)^2 \). In this program, \( \lambda \) is an parameter which reflects the weight granted to the trend flexibility compared to \( \lambda \) the cycles magnitude.

\(^4\) In order to avoid discontinuities in the time series, war periods are considered in this test.
\(^5\) Let’s write \( y_t = g_t + c_t \) where \( g_t \) is the trend value in \( t \), \( c_t \) the cyclical component. In order to determine \( g_t \), the Hodrick Prescott filter finds a function which leads to:
- find the values of \( g_t \) that provide a good approximation of \( y_t \);
- find the values of \( g_t \) whose acceleration (\( \Delta g_{t+1} - \Delta g_t \)) is smallest as possible.
cycles and a trend. Then we used the Bry-Boschan (1971) algorithm in order to date the recessions.

To summarize, the algorithm detects local maxima (peaks) $y_i$ when $\{y_{i+k} < y_i > y_{i-k}\}$ and minima (troughs) $y_i$ when $\{y_{i+k} > y_i < y_{i-k}\}$. In practice, the algorithm must meet certain additional requirements to eliminate false turning points: rules that avoid cycles of length or insufficient amplitude and procedures to ensure that peaks and troughs alternate (Male, 2010).

The exercise is made for France (chart 5a) and United-States (chart 5b) between January 1970 and December 2011. It detects a good cyclical correlation between the economic growth measured by the industrial production and the shares price change led by 6 months. This time lag provides the best correlation between the series: 0.53 and 0.65 for France and United-States, respectively.

If we except the period of October 1987, when the crash was mainly due to a violent increase in the American interest rate intended for stop the dollar drop, the shares prices are leading indicators of the economic activity. These results are a further argument against the random walk hypothesis of stock markets, as the price of risky assets seems to depend on the expectations on future economic growth.

Graph 5 a- Production and shares prices - France
To sum up, as soon as one is concerned by the financial real returns, American and France show a clear convergence; more this convergence holds when considering the hierarchy between assets when the holding period changes. After 22 and 27 years respectively in United-States and in France, the probability that stocks outperform the other asset classes is 100% (graphs 6a and 6b).

Graphs 6 a and b – Equities over performance probability – France and United States
Further, the analysis of the historical returns distribution shows that the standard deviation of stocks returns declines faster than for a Gaussian; the good cyclical correlation between economic growth and stocks’ prices are strong indications that accredits the existence of a mean reversion process for the stocks prices. The stocks’ returns show a clear connection with the potential economic growth and – by the way – a decreasing relationship between holding horizon and risk.

In the second part, we show how it impacts a long run allocation, considering the case of a pension fund.

**PARAMETRIC AND NON-PARAMETRIC LIFE CYCLE ASSET ALLOCATION**

Studying basic asset classes teaches a lot about the relationship between risk and holding period. It nevertheless says nothing about portfolio choices because it does not integrate the risk diversification and it impact on asset allocation.

In this second chapter we shall consider optimal asset mix under parametric and semi-parametric framework. In order to determine the hierarchy of the investment strategies, we take the example of a default option defined contribution pension scheme.

If we except the sophisticated financial tools as the portfolio that only guarantee a performance when reaching a given horizon (and that penalize the early exits such that structured funds) the default option retirement plans usually met 4 strategies which are life cycle allocation, profiled funds, constant proportion portfolio insurance (CPPI) and stop loss management.

rem : Il manque une phrase de liaison indiquant que l’on ne traite que les life cycle allocation ? Ou alors on supprime carrément la référence aux autres stratégies !

At this stage, we must specify that, for the forthcoming empirical developments, no hypothesis about fees that are linked to the financial investments has been done; they, of course, decrease the returns but do not affect their hierarchy. The only costs that will be taken into account are the costs that result from the arbitrage between assets and which decrease the yield provided by a reallocation. The hypothesis is that these costs represent 0.5% of the arbitrage, whatever its amount is.

**The choice of a life cycle allocation and the linked methodology**

During the past decades, defined benefits retirement pension plans left room to defined contributions pension plans in the firms’ supply of retirement pension plans. This phenomenon led workers to make choices in asset allocations. In response, assuming that workers usually have a lack in financial education, many pension funds have launched default option pension plans that are target date funds or life cycle pension plans, as described in Mitchell and alii (2006) or Poterba and alii (2007).

Life cycle management appears as particularly consistent in the case of pension funds because it relies on an intertemporal diversification strategy. The introduction of a time diversification in investment strategies involves answering the following question: how do households have to change the weight and structure of their asset portfolio over time?

Life cycle pension plans simplify the asset allocation by considering that, when people retire, their exposure to risk has to be completely reduced. In this vision, the equities share path decreases from a high percentage when people are young to 0% when people retire. As shown
in Antolin and alii (2010), this is a more or less sophisticated version of the traditional “vulgate” that the percentage of equities in a portfolio should be 100 less the fund participant age in years. As shown by Lavigne and Legros (2006) there is any strong argument in favor of this rule of thumb. Over all, adopting such a deterministic allocation relies on the hypothesis that savers have a very specific utility function. For example, Gollier and Zeckhausser (1997), justify this decrease in the share of risky assets with age by the fact that the time to spend before retirement i.e. the ability to reallocate a portfolio decreases with age. The flexibility factor (the portfolio possible reallocation period) outweighs the background risks (professional or family linked risks, for example) if the risk tolerance is a decreasing and convex function of the consumer resources. In this case, called “clause Duration enhances risk”, the youngs will be more risk tolerant than the olds.

This necessity to turn the savers’ utility function led us to adopt a VaR (Value at Risk) constraint into the asset allocation strategy of life cycle investment fund. VaR is the loss in a portfolio’s value that results from an adverse movement in assets prices over a specific time horizon an with a given degree of confidence. On the one hand, VaR allows to overcome the utility function problem, on the other hand, the VaR criterium includes a holding period consideration which sticks optimally with a life cycle allocation (see Lewis and Okunev, 2009).

Another characteristic is that VaR can take into account the non Gaussian distribution of some assets and the changing distribution with the holding period.

In order to get convinced of the interest of taking into account this non Gaussian distribution, we compare the structure of portfolios determined from a standard criterion of "mean-variance" and an allocation which is based on the all moments i.e. of higher orders than 2. For this purpose, the risk definition initially given by Markowitz is enlarged in order to include the asymmetrical distribution of the returns (i.e. the Skewness) and the extreme events (i.e. the Kurtosis of returns). In the short run, the optimal allocations which are deducted from the Markowitz framework differ from the optimal allocations which can be obtained when using risk measures that integrate the upper moments (charts 7a and b); this convinces that the adjustments that are linked to the extreme risks have to be taken into account.

Graphs 7a and 7b – Efficiency frontiers with or without moments of order higher than 2 – France
In the usual approach, in order to determine the evolution of the structure of optimal allocations, one relies on the theory of expected utility and assumes that the investor’s preferences remain constant over time. The classical maximization problem in a mean variance space, is written as:

\[
\text{Max } \mu - \lambda_1 \sigma^2
\]

With:

\(\mu\) = the portfolio returns’ expectancy  
\(\sigma^2\) = the portfolio returns’ variance  
\(\lambda_1\) = a risk aversion parameter

The agent’s preferences are, in this case, summarized by the risk aversion.

As soon as are introduced the Skewness (S) and the Kurtosis (K) in the agent’s utility function in order to take into account the non-Gaussian type returns’ distribution, the program has to be written:

\[
\text{Max } \mu - \lambda_1 \sigma^2 + \lambda_2 S - \lambda_3 K
\]

Where:

\(\lambda_1, \lambda_2, \lambda_3\), are respectively the variance aversion (risk aversion), the asymmetry preference (Skewness preference) and the Kurtosis aversion (extreme events aversion).

The portfolio optimization with high moments is a much more complicated problem than the classical optimization one. In this last case, the investor aims at maximize both the expected yield and the Skewness (the odd-order moments) while minimizing the variance and the Kurtosis (the even-order moments). In this type of maximization program generally solved by a polynomial goal programming (PGP), the portfolio selection depends on the goals and preferences of the investor. Prigent and Mhiri (2010) show after Lai, Yu and Wang (2006), that the arbitrage between these goals is more demanding (and sometimes simply intractable) than in a mean variance dimension.

To circumvent this difficulty, i.e. incorporating higher-order moments in the life cycle strategy, we used a criterion of Value at Risk (VaR) to define the risk of loss accepted by the agents.
Selecting a measure which is unequivocal facilitates the comparison of strategies whether in a parametric case or in a semi-parametric one. The optimization program aims at finding an allocation that provides the maximal return’s expectancy alternately associated with a parametric VaR and a Cornish-Fisher VaR (CF VaR), the latter taking into account the Skewness and the Kurtosis of the distribution of returns.

In the following developments, we shall compare at each stage the results that would come from a classical optimization program (using VaR) and our “non Gaussian” program (using Cornish Fisher VaR).
- Optimization program for the parametric VaR

This optimization program is the following:

\[
\begin{align*}
\min_w & \left( \text{VaR}_\alpha (w) \right) \\
\text{s.t.} & \quad w^* \mu = \mu_p \\
\text{with:} & \quad \sum_{i=1}^{3} w_i = 1 \\
& \quad 0 \leq w_i \leq 1
\end{align*}
\]

With:
\[
\text{VaR}_\alpha = w^* \mu + z_\alpha * \sqrt{w^* \Sigma w}
\]

In this case, the program is a nonsmooth optimization problem (NSP). Nonsmooth optimization refers to the more general problem of minimizing functions that are typically not differentiable at their minimizers.

In order to solve it, we directly use the gradient sampling (GS) algorithm developed by Burke and alii (2005) and used by Curtis and Que (2012). This GS is a variable metric minimization program (VMM) and is more efficient than the previous programs: with the notation “O” after Landau, the algorithm by Curtis simplifies the linear complexity problem O(n) into a constant complexity O(1), which is data size independent.

The optimization program is:

\[
\begin{align*}
\min_w & \left( \text{VaR}_\alpha (w) \right) \\
\text{s.t.} & \quad w^* \mu = \mu_p \\
\text{with:} & \quad \sum_{i=1}^{3} w_i = 1 \\
& \quad 0 \leq w_i \leq 1
\end{align*}
\]

With:
\[
\text{VaR}_\alpha = w^* \mu + z_\alpha * \Sigma
\]

- Optimization program for the Cornish-Fisher VaR

In this case, the program is a nonsmooth optimization problem (NSP). Nonsmooth optimization refers to the more general problem of minimizing functions that are typically not differentiable at their minimizers.

In order to solve it, we directly use the gradient sampling (GS) algorithm developed by Burke and alii (2005) and used by Curtis and Que (2012). This GS is a variable metric minimization program (VMM) and is more efficient than the previous programs: with the notation “O” after Landau, the algorithm by Curtis simplifies the linear complexity problem O(n) into a constant complexity O(1), which is data size independent.

The optimization program is:

\[
\begin{align*}
\min_w & \left( \text{VaR}_\alpha (w) \right) \\
\text{s.t.} & \quad w^* \mu = \mu_p \\
\text{with:} & \quad \sum_{i=1}^{3} w_i = 1 \\
& \quad 0 \leq w_i \leq 1
\end{align*}
\]

With:
\[
\text{VaR}_\alpha = w^* \mu + z_\alpha \text{ Cornish–Fisher} * \sqrt{w^* \Sigma w}
\]

\[
\text{Z}_\alpha \text{ Cornish–Fisher} \approx Z_\alpha + \frac{1}{6} (Z_\alpha^2 - 1) * S + \frac{1}{24} (Z_\alpha^3 - 3Z_\alpha) * K - \frac{1}{36} (2Z_\alpha^3 - 5Z_\alpha) * S^2
\]

\[
S(X) = \frac{\mu_3}{\sigma^3} : \text{Skewness}
\]
\[ \hat{R}(X) = \frac{\mu_t}{\sigma_t^4} : \text{Kurtosis} \]

\[ K = \hat{R} - 3 : \text{peakedness} \]

\( \mu \): vector of average returns

\( w_i \): asset i weight in the portfolio \( \left( \sum w_i = 1 \right) \)

**Results**

- With a parametric 3% VaR

In the case when the risk is approximated by the returns’ variance, the risk reduction is simulated from a Gaussian. The standard deviation of total return over a holding period decreases with the square root of the length of the period and does not vary between the asset classes. The frontier is obtained minimizing the VaR for a given return. In the various scenarios, the risk which is taken by an investor is represented by an accepted capital loss. The investor will search a portfolio structure \( w_i \) on each efficient VaR frontier corresponding to a remaining holding period. Each portfolio must respect his loss target. Because the risk decreases with the holding period length, the higher the disinvesting horizon is far away, the higher the ratio of risky assets in the portfolio.

If the expected allocation does not correspond to any efficiency frontier, the resulting allocation is the one that minimizes the smallest VaR on the frontier. On the opposite, for long horizons, the VaR can be negative for all the portfolios i.e. correspond to gains. In these cases, the allocation that provides the greatest yield is chosen.

The asset allocation grid that results from the optimization, given a risk profile – i.e. a VaR -, is used to perform a back test from 1895 to 2011. As an example, we consider a 3% VaR with the confidence threshold of 99% for a 40 years period in the American case (the annex 1 presents a wilder comparison with different holding periods varying from 10 years to 40 years); the optimization problem searches on each frontier the allocation whose VaR is the closest to the chosen VaR. In this example (see charts 8a and b) the smallest VaR on the frontier for a one year period is 10%, i.e. presents a risks which is higher than the targeted one. Until we rich the 5th year, the allocations are much more risky than the chosen risk profile. After 25 years, i.e. with a remaining period of 15 years or less (on the left of chart 8b), investing in North American equities does not represent any loss probability; the investor is led to invest 100% of his wealth in “risky” assets.
The risk profile associated allocation allows a back test over 480 monthly overlapping rolling windows between 1895 and 2011. The back test shows that each of the 912 different allocations leads to a capital gain (i.e. a positive performance). The median yield associated with these allocations is 4.3% in real terms and the first percentile one reaches 2.3%. This excess performance, compared with the VaR criterion, is due to the rule that drives to adopt the maximum return allocation when all portfolios on an efficiency frontier have a negative VaR.

In order to provide another illustration of the same strategy, we estimated the replacement rates (ratios pension/revenue) resulting from the allocations. By lack of historical series, we took the average households’ disposable real income as a wages proxy. The individuals are supposed to save 6% of their income; when they are 65, the capital is annuitized according to the life expectancies provided by the United States Census Bureau. Graphs 9a and 9b provide the replacement rates; the median one reaches 25% of the households’ disposable income. These replacement rates have no vocation to be compared to a defined benefits pension fund replacement rate but are a concrete illustration of the exercise.

---

6 Source: The World Top Incomes Database (http://g-mond.parisschoolofeconomics.eu/topincomes/)
• With a semi parametric 3% VaR

In the non parametric case, the moments of the distribution are computed, for each holding horizon, using the assets historical returns. The volatility of stock returns tends to decline faster than in a Gaussian. Each frontier is built minimizing the CF VaR according to the 4 first moments on a moving period k*12 long, with k=1 to 40. This allows to have a holding horizon 40 years long with monthly data. Graphs 10a and 10b describe the mean-CF VaR efficient frontiers. An interesting aspect of these graphs is that stocks become less risky compared to Treasury bill or Bonds with the lengthening of holding periods. The slopes of the efficient frontiers are steeper in the semi-parametric model which determines a greater proportion of risky assets.

As in the previous case, the allocations are used for backtesting the non parametric strategy. With the American assets, the allocation strategies show that taking into account the high order (3 and 4) moments and the empirical reduction of risk is not neutral; with an accepted loss of 3%, the investor is totally invested in equities after 15 years i.e. for any horizon up to 25 years. The weight of bonds in the portfolio is then negligible.

Graphs 10a and 10b – Efficiency frontiers and allocations – semi parametric VaR - US

The returns’ distribution median value is higher than the value provided by the parametric program (5.30% compared to the previous 4.30%, Graphs 11a and 11b). The median replacement rate is also higher than the one which is estimated with the parametric allocation; it reaches 34.3% (to be compared with the previous 27.9%).
Notice that we get similar results if we consider the French series, in real terms and war periods excluded. Graphs 12a and 12b show that, for a given loss probability (a 3% VaR with a 99% confidence threshold in the present case), the optimal CF VaR frontiers lead to a higher share of “risky” assets than in the parametric case.

Graphs 12a and 12b – Portfolios allocations – parametric VaR vs semi parametric VaR - France

**Ranking the strategies**

In order to prioritize the performances of simulated portfolios, it is useful to use a criterion of stochastic dominance of first and second order with the first order stochastic dominance defined as follows: a random variable $F$ dominates $G$ if their cumulative distribution functions verify $F(x) \leq G(x)$, $\forall x \in [a,b]$. In other words, $G$ is more concentrated in weak returns. Simply speaking the first order dominance “pushes” the density function to the right.
In the ambiguous situations, the second order stochastic dominancy of $F$ on $G$ is translated by:

$$\forall x \in [a,b], T(x) = \int_a^x (G(t) - F(t)) dt \geq 0$$

Graphs 13a and 13d provide the portfolios’ returns cumulative distribution functions (CDF). We found that CF VaR strategy has first order stochastic dominance over parametric VaR strategy. Semi parametric model provides higher returns. The comparison of the replacement rates confirms without any equivocation the superiority of allocations computed with a non-parametric program compared to a parametric one.

Graphs 13a and 13b – Fonctions cumulatives de probabilités de rendements et de taux de remplacement – VaR paramétrique vs non paramétrique (CF VaR) – US

The verdict is the same for France (13c and 13d) and for United States: asset rebalancing strategies based on optimization procedures incorporating higher-order moments consistently dominate mean-variance optimization techniques.

Graphiques 13c et 13d – fonctions cumulatives de probabilités de rendements et de taux de remplacement – VaR paramétrique vs non paramétrique (CF VaR) – France
CONCLUSION

The statistical characteristics of various assets show that, in the long run, risks to stocks tend to diminish more rapidly than risks to bonds and T-Bill. More, the evidence of mean reversion in stock market prices calls for time diversification and life cycle allocation for default options retirement plans.

On the other hand, the non Gaussian distribution of equities implies that a CF VaR (non parametric VaR) instead of a VaR which relies on a Gaussian is used. A comparison of the result is striking: the semi parametric VaR leads to a portfolio that is much richer in equities than the portfolio stemming from the parametric VaR. The comparison between the replacement rates that come from both the strategies confirm without equivocation that the allocations computed with the non-parametric criterion are more profitable than the allocations coming from the parametric program.

This conclusion has another reach : it implies that considering that the equities returns follow a Gaussian is a simplistic hypothesis that leads to various errors such as stochastic simulations by Monte Carlo or bootstrapping. It also says that considering a simple VaR on one year, as in Solvency II, to build prudential rules for pensions funds will lead to suboptimal allocations.

REFERENCES


Male R., 2010, “Developing country business cycles: characterising the cycle”, Queen Mary working paper n°663, may, University of London


### Annex 1
Comparison of asset allocations with various holding periods

#### USA: LifeCycle Asset Allocation [1895-2011]

<table>
<thead>
<tr>
<th>Holding periods</th>
<th>CF VaR strategy</th>
<th>VaR strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 years</td>
<td>CFVaR=3%</td>
<td>VaR=3%</td>
</tr>
<tr>
<td>20 years</td>
<td>CFVaR=3%</td>
<td>VaR=3%</td>
</tr>
<tr>
<td>30 years</td>
<td>CFVaR=3%</td>
<td>VaR=3%</td>
</tr>
<tr>
<td>40 years</td>
<td>CFVaR=3%</td>
<td>VaR=3%</td>
</tr>
</tbody>
</table>

**Summary Statistics**

<table>
<thead>
<tr>
<th>Risk criteria</th>
<th>CF VaR strategy</th>
<th>VaR strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average return</td>
<td>1.7%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.5%</td>
<td>2.6%</td>
</tr>
<tr>
<td>Median Return</td>
<td>2.2%</td>
<td>2.1%</td>
</tr>
<tr>
<td>1st centile (return)</td>
<td>-3.8%</td>
<td>-4.3%</td>
</tr>
<tr>
<td>99th centile (return)</td>
<td>7.2%</td>
<td>7.1%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.15</td>
<td>-0.21</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.51</td>
<td>2.63</td>
</tr>
<tr>
<td>Average replacement rate</td>
<td>4.6%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Median replacement rate</td>
<td>4.4%</td>
<td>4.3%</td>
</tr>
<tr>
<td>1st centile (replacement rate)</td>
<td>3.1%</td>
<td>3.1%</td>
</tr>
<tr>
<td>99th centile (replacement rate)</td>
<td>8.8%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.07</td>
<td>-0.22</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.24</td>
<td>3.26</td>
</tr>
</tbody>
</table>

**Stochastic dominance tests**

\[ F_{CFVaR} \text{ vs } F_{VaR} \]

<table>
<thead>
<tr>
<th>Holding periods</th>
<th>10 years</th>
<th>20 years</th>
<th>30 years</th>
<th>40 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF VaR strategy</td>
<td>F_{CFVaR}</td>
<td>SOSD F_{VaR}</td>
<td>FOSD F_{VaR}</td>
<td>FOSD F_{VaR}</td>
</tr>
<tr>
<td>VaR strategy</td>
<td>F_{CFVaR}</td>
<td>SOSD F_{VaR}</td>
<td>FOSD F_{VaR}</td>
<td>FOSD F_{VaR}</td>
</tr>
</tbody>
</table>

FODS : first-order stochastically dominates – SOSD : second-order stochastically dominates


<table>
<thead>
<tr>
<th>Holding periods</th>
<th>CF VaR strategy</th>
<th>VaR strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 years</td>
<td>CFVaR=3%</td>
<td>VaR=3%</td>
</tr>
<tr>
<td>20 years</td>
<td>CFVaR=3%</td>
<td>VaR=3%</td>
</tr>
<tr>
<td>30 years</td>
<td>CFVaR=3%</td>
<td>VaR=3%</td>
</tr>
<tr>
<td>40 years</td>
<td>CFVaR=3%</td>
<td>VaR=3%</td>
</tr>
</tbody>
</table>

**Summary Statistics**

<table>
<thead>
<tr>
<th>Risk criteria</th>
<th>CF VaR strategy</th>
<th>VaR strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average return</td>
<td>1.8%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.2%</td>
<td>2.8%</td>
</tr>
<tr>
<td>Median Return</td>
<td>1.8%</td>
<td>1.1%</td>
</tr>
<tr>
<td>1st centile (return)</td>
<td>-6.4%</td>
<td>-5.7%</td>
</tr>
<tr>
<td>99th centile (return)</td>
<td>9.4%</td>
<td>7.2%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.07</td>
<td>-0.22</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.24</td>
<td>3.26</td>
</tr>
<tr>
<td>Average replacement rate</td>
<td>4.9%</td>
<td>4.9%</td>
</tr>
<tr>
<td>Median replacement rate</td>
<td>4.7%</td>
<td>4.7%</td>
</tr>
<tr>
<td>1st centile (replacement rate)</td>
<td>2.7%</td>
<td>2.8%</td>
</tr>
<tr>
<td>99th centile (replacement rate)</td>
<td>9.3%</td>
<td>8.9%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.07</td>
<td>-0.22</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.24</td>
<td>3.26</td>
</tr>
</tbody>
</table>

**Stochastic dominance tests**

\[ F_{CFVaR} \text{ vs } F_{VaR} \]

<table>
<thead>
<tr>
<th>Holding periods</th>
<th>10 years</th>
<th>20 years</th>
<th>30 years</th>
<th>40 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF VaR strategy</td>
<td>F_{CFVaR}</td>
<td>FOSD F_{VaR}</td>
<td>FOSD F_{VaR}</td>
<td>FOSD F_{VaR}</td>
</tr>
<tr>
<td>VaR strategy</td>
<td>F_{CFVaR}</td>
<td>FOSD F_{VaR}</td>
<td>FOSD F_{VaR}</td>
<td>FOSD F_{VaR}</td>
</tr>
</tbody>
</table>

Inconclusive

FODS : first-order stochastically dominates – SOSD : second-order stochastically dominates