imr Investment Management Review

Special Edition
Autumn 2017

CONFRONTING TRANSFORMATION IN THE INVESTMENT INDUSTRY

ADDRESSING INDIVIDUALS' RETIREMENT NEEDS

GROWING ACCEPTANCE OF THE FACTOR INVESTING AND SMART BETA PARADIGMS

ASSESSING RISK PREMIA IN BOND MARKETS

Published with the support of EDHEC-RISK Institute

FROM THE EDITOR

DANGEROUS TIMES FOR ASSET MANAGERS

Arjuna Sittampalam
Co-editor

Lionel Martellini
Co-editor

> … threats also offer opportunities that can make winners of industry players who seize them.

Transformative threats on several fronts are confronting the investment industry. But threats also offer opportunities that can make winners of industry players who seize them.

Already, four leading houses – Janus, Henderson, Standard Life and Aberdeen – have become just two through mergers ascribed widely to the shift towards passive investment, squeezing mainstream active fund managers.

However, the move towards passive management is not that straightforward. The emergence of so-called smart beta products is blurring the lines between active and traditional passive money management based on cap-weighted benchmarks. A three-way fight is underway between market cap-based passive, active and smart beta, with variants in each of these categories. Dynamic management of smart factor exposures is also expected to be a new source of added value.

At the core of active management, setting market prices and carrying out essential arbitraging is indispensable. Smart beta is not always clearly appreciated for what exactly it is because several types of products claim the label. Some are just active quantitative portfolio management tools. Others are genuine attempts to provide investors with well-diversified exposure to long-term rewarded risk factors.

The second problem receiving the most social and political attention concerns retail investors. The fund management industry stands accused of focusing on profitability at their expense, causing much distrust. However, the problems cannot be blamed entirely on fund managers. Hitherto, investment solutions addressing individuals' specific needs have not been available on an economic basis.

The third and the most dangerous threat is digital disruption. A leading consultant has stated that there are too many funds. Referring to the previous point, this has arisen from high distribution charges paid for by retail customers allowing their exploitation through an excessive proliferation of fund launches. Digitalisation, by accelerating change, should end this unsatisfactory situation. It will also endanger incumbents and foster new entrants.

Though the digital factor will reduce this exploitation, it does not produce investment solutions specific to each individual, but can facilitate their introduction through investor-centred, as opposed to performance-centred, dialogue. The fast-growing robo-advice sector enabled by digitalisation, already morphing into different forms, does not provide the answer either.

The goals-based framework and mass customisation proposed in this edition, in attending to the specific goals of investors, will potentially transform the individual sector by providing tailor-made solutions that are currently not available, except to the wealthiest.

The fixed income sector is vital to the security of anybody expecting retirement benefits from portfolios. Yet, compared with well-developed equity markets, it has more structural weaknesses embracing risk characteristics, indices and trading fragmentation.

Particularly, smart beta in fixed income is very much underdeveloped. In this edition, we observe the progress to date and look at what is still required to get on a sound footing in this area as well as in the related sphere of factor investing. We also cover how and why the new and important concept of risk allocation could become established thinking that replaces current asset allocation processes.

Most individuals are unlikely to be seriously interested in concepts such as bonds, shares and funds. Even if they invest their assets in these, many will depend on financial advisors for whom more education is vital.

The ideas of the goals-based framework, mass customisation, smart beta, factor investing and risk allocation have every chance of becoming embedded in individuals' investment practices. This will then require financial advisors to reach higher levels of sophistication and familiarity with these ideas.

A large proportion of this special edition is devoted to academically-grounded insights on the goals-based framework and related mass customisation challenges. If adopted, they should have a radical impact on individual investment management.

Lionel Martellini & Arjuna Sittampalam
Co-editors ■

A SNAPSHOT OF THIS ISSUE

The following five related articles are based on two academic papers that analyse a framework for meeting individuals' specific retirement goals currently lacking in the industry and how the challenge of mass customisation for the less wealthy can be solved. These will lead to revolutionary changes in the industry, if adopted, remedying the weaknesses of the investment process available to individual investors.

RETIREMENT GOAL-BASED INVESTING

This article, after addressing the deficiencies of existing retirement industry products, outlines the principles underlying the goals-based framework. Along with the challenges of mass customisation, a new bond retirement index series is also explained, designed to encourage the industry to come up with workable solutions for this framework *6*

A PATH-BREAKING SOLUTION FOR INDIVIDUAL INVESTORS' PROBLEMS

A non-academic resume of the paper 'Introducing a Comprehensive Investment Framework for Goals-Based Wealth Management' is an in-depth review of the goals-based framework aimed at providing a comprehensive understanding for everybody interested in the implementation of practical solutions. *9*

TAILOR-MADE FUNDS WITH ECONOMIES OF SCALE

A non-academic resume of the paper 'Mass Customisation versus Mass Production in Retirement Investment Management: Addressing a Tough Engineering Problem' explains an innovative way forward for meeting the challenge of mass customisation that allows large numbers of individuals to invest in just two types of portfolios. *19*

INDIVIDUAL INVESTMENT PROCESSES FOR THE 21ST CENTURY

While the way forward is clearly indicated by the two papers, it is the industry that needs to implement the new ideas in practice. This article points to the industry's serious difficulties in doing so, though rapidly changing conditions should encourage the radical transformation that is required. *26*

MULTI-ASSET PRODUCTS AND SOLUTIONS

This article puts the need for individual retirement solutions in the context of some of the latest thinking in Modern Portfolio Theory and approaches already used by institutional investors. *28*

A SNAPSHOT OF THIS ISSUE

EQUITY PORTFOLIOS WITH LIABILITY-HEDGING BENEFITS

Academic research has come up with a remarkable result of considerable interest to pension funds, using liability-driven approaches. Modern Portfolio Theory advocates that pension plans should invest in two portfolios, a liability-hedging portfolio consisting of fixed income securities and a performance-seeking portfolio exposed to riskier assets. However, the finding is that selecting certain types of equity in the risky portfolio can not only enhance the liability hedging benefits, but can also increase the overall performance. Alternatively, a higher allocation to equities becomes possible with commensurate extra performance with the same liability hedging benefits. *31*

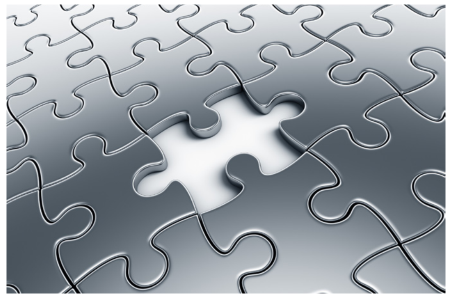
The following four related articles concern the relatively new and related paradigms of factor investing and smart beta, which are rapidly growing in acceptance and importance. The first two cover bond investing, while the other two pertain to the use of factors across and within assets classes and in equities.

BOND RISK PREMIA: THE NEW FRONTIER IN FACTOR INVESTING AND SMART BETA

The concepts of factor investing and smart beta are much less developed in bonds compared to equites. In the identification of factors, various difficulties arise. Existing bond indices have serious deficiencies which are partly related to factor exposures being unstable. A broad overview of current first generation smart beta approaches also highlights various weaknesses. *34*

FACTOR-BASED APPROACHES TO THE DESIGN OF SMART BOND PORTFOLIOS

The difficulties of measuring and identifying factors are referred to in this article. A review is carried out of the most consistently identified fixed income factors and it is pointed out that more research is required before bond risk factors can be reliably used in efficiently extracting risk premia. *37*



FACTOR INVESTING: EFFICIENT HARVESTING OF RISK PREMIA ACROSS AND WITHIN ASSET CLASSES

Asset allocation decisions have for years been based on weightings attached to different asset classes. In a new approach that is increasingly adopted, these decisions are based on exposure to risk factors cutting across asset classes. The theme of this article is the methodology used to implement this approach including the use of efficient factor indices as building blocks for each risk factor in the allocation process *41*

BE SERIOUS WITH EQUITY FACTOR INVESTING

In this article, the problems with equity factor methodologies used by some providers, when they depart from academically accepted practices, are analysed. These include the same provider using different definitions of the value factor at various times. Excessively concentrated bets without adequate diversification are also identified *44*

A SNAPSHOT OF THIS ISSUE

MEASURING VOLATILITY PUMPING BENEFITS IN EQUITY MARKETS

When portfolios are constructed on the basis of long-term weights, rebalancing is often carried out to restore the initial weights. This article investigates the conjecture that this rebalancing can be a source of extra investment performance. The results show that this is true for suitably selected categories of stocks. *47*

PREDICTING RISK PREMIA FOR TREASURY BONDS: THE EDHEC BOND RISK PREMIUM MONITOR

In general, the variation of bond yields along the yield curve reflects either expected future interest rates or the risk premium. Assessing the risk premium impact on the steepness of the yield curve is, thus, important to bond managers. This article describes the EDHEC Bond Risk Premium Monitor launched by EDHEC-Risk Institute as a tool to derive a state-of-the-art estimation of the risk premium using market and monetary policy information *50*

RETIREMENT GOAL-BASED INVESTING

Kevin Giron, Quantitative Research Engineer, *EDHEC-Risk Institute*

Lionel Martellini, Professor of Finance, *EDHEC Business School* and Director, *EDHEC-Risk Institute*

The need for new retirement solutions

With the need to supplement retirement savings via voluntary contributions, individuals will increasingly be responsible for their own saving and investment decisions. This global trend poses substantial challenges as individual investors not only suffer from behavioural limitations, but also typically lack the expertise needed to make educated investment decisions.

In response to these concerns, a number of new investment products have been proposed over the past few years by asset management companies. There are reasons to believe, however, that these products known as *target date funds* fall short of providing satisfactory solutions to the problems faced by individuals when approaching investment saving decisions.

First of all, *target date funds* offer a sole focus on the investment horizon without any protection of investors' minimum retirement needs. In particular, these products are not engineered to deliver replacement income in retirement, and do not achieve a proper hedging of the main risks related to the retirement investing decisions, namely investment risk, interest rate risk, inflation risk and longevity risk. Another important restriction is that most existing target-date funds do not allow for revisions of the asset allocation as a function of changes in market conditions. This is entirely inconsistent with academic prescriptions, and also common sense, which both suggest that the optimal strategy should also display an element of dependence on the state of the economy.

Towards improved forms of retirement investment solutions

Currently available investment options hardly provide a satisfying answer to the retirement investment challenge, and most individuals are left with an unsatisfying choice between, on the one hand, safe strategies with very limited upside potential, which will not allow them to generate the kind of *target* replacement income they need in retirement and, on the other hand, risky strategies offering no security with respect to *minimum* levels of replacement income.

This stands in contrast with a well-designed retirement solution that would allow individual investors to secure the kind of replacement income in retirement needed to meet their *essential* consumption goals, while generating a relatively high probability for them to achieve their *aspirational* consumption goals, with possible additional goals including healthcare, old age care and/or bequest.

“

…individuals will increasingly be responsible for their own saving and investment decisions. This global trend poses substantial challenges as individual investors not only suffer from behavioural limitations, but also typically lack the expertise needed to make educated investment decisions…

This recognition is leading to a new investment paradigm, which has been labelled *goal-based investing* (GBI) in individual money management – see Deguest et al. (2015) for a formal analysis. Within this paradigm, investors' problems can be fully characterised in terms of their lifetime meaningful goals, just as *liability-driven investing* (LDI) has become the relevant paradigm in institutional money management, where investors' problems are broadly summarised in terms of their liabilities.

The GBI retirement solutions are deeply rooted in simple and profound academic principles and involve:

- A dedicated and safe goal-hedging portfolio (GHP) that replicates risk factor exposures in investors' replacement income goals (dynamic replicating portfolio for inflation-linked deferred annuities);
- A common well-rewarded risky performance-seeking portfolio (PSP) that efficiently harvests risk premia in equity markets;
- A dynamic allocation to PSP versus GHP portfolios that (i) reacts to changes in market conditions and (ii) secures minimum replacement income levels while generating a high probability of achieving target replacement income levels.

As such, the framework builds upon a comprehensive and holistic integration of the three forms of risk management, namely hedging, diversification and insurance. This is in contrast to existing products or approaches used in institutional or individual money management, which are only based on selected risk management principles.

While each of these sources of value added is already used to some extent in different contexts, full integration of all these elements within a comprehensive disciplined investment management framework is required for the design of meaningful investment solutions.

EDHEC-Princeton retirement bond and retirement goal index series

The most natural way to frame an investor's retirement goal is in terms of how much lifetime guaranteed inflation-linked replacement income they will be able to afford to purchase upon retirement. Given that the biggest risk in retirement is the risk of outliving one's retirement assets, securing inflation-linked replacement income within the decumulation period can be achieved with inflation-linked annuities, which are the true risk-free assets for individuals preparing for retirement.

Annuity products, however, are cost inefficient, irreversible, and do not contribute to bequest objectives. These elements undoubtedly explain the low demand for annuities (i.e. the “annuity puzzle”, when annuitisation is not incentivised or mandatory). A good case can actually be made that annuitisation is a decision that is best taken close to retirement, if ever.

This point was made explicitly in Merton (2013): “Most people would not want to buy an actual life annuity during the accumulation period prior to retirement (…). If the individual buys an annuity on his life, it is reversible only at a high cost, which means if he dies two years from now, he loses everything. If his life circumstances change - get married, get divorced, start a second family - the lack of flexibility can be costly. The right time for him to determine the detail of post-retirement investment choices (…) is as close to retirement as possible when he has the most information about his health, his responsibilities, his opportunities, and his preferences.”

In the UK, the 2015 Pension Schemes Act, which has nullified the compulsory annuity purchase, creates a tremendous opportunity for asset managers to launch meaningful forms of retirement solutions. A key ingredient in these retirement solutions is a novel form of fixed-income portfolio, where the key focus should be on generating inflation-linked or cost-of-living-adjusted replacement income for a period of time roughly corresponding to the average life expectancy in retirement (say between 20 and 25 years after the retirement date). In parallel, late-life annuities can be purchased in decumulation to obtain protection against tail longevity risk.

In the context of these aforementioned changes in the retirement landscape, EDHEC-Risk Institute has teamed up with Princeton University's Operations Research and Financial Engineering (ORFE) department to launch a *retirement bond index series*. This initiative, supported by Merrill Lynch Wealth Management (MLWM) in the context of the “Risk Allocation Framework for Goal-Driven Investing Strategies” research chair at EDHEC-Risk Institute. More specifically, the so-called retirement goal bond index is designed to dynamically hedge interest rate risks and realised or expected inflation risks, thereby forming a dynamic proxy for a forward inflation-linked or cost-of-living adjusted bond ladder. The introduction of this retirement goal index series is consistent with the prescription of asset pricing theory, which has shown that T-Bills investment is a low volatility strategy from an absolute return perspective, but extremely risky for use in a retirement context since it leads to a highly volatile level of purchasing power expressed in terms of inflation-linked replacement income. In parallel, dynamic asset allocation *goal-based retirement indices* (which involve the use of retirement bond indices and equity indices) can also be proposed to allow investors to secure a minimum level of replacement income in retirement while generating attractive probabilities of achieving a target level of replacement income.

While these dynamic retirement goal-based indices are not investable and are meant to be used for education purposes only, our hope is that they will help inspire a new generation of asset and wealth managers design and distribute welfare-improving forms of retirement solutions that generate relatively high probabilities of individuals reaching their target levels of replacement income while securing their minimum levels of replacement income.

One key implementation challenge for these otherwise relatively straightforward concepts is the problem posed by the scalability requirements. Since it is hardly feasible to launch a customised dynamic allocation strategy for each investor, the challenge, which we discuss now, is to address the needs of a large number of investors through a limited number of funds.

Mass customisation versus mass production in retirement investing

Mass production (in terms of products) happened a long time ago within investment management, through the introduction of mutual funds and, more recently, exchange-traded funds. The new frontier in retirement investing is mass customisation (as in customised solutions), which by definition is a manufacturing and distribution technique that combines the flexibility and personalisation of “custom-made” solutions with the low unit costs associated with mass production. In other words, the true challenge is indeed to find a way to provide a large number of individual investors with meaningful dedicated investment solutions.

“

The most natural way to frame an investor's retirement goal is in terms of how much lifetime guaranteed inflation-linked replacement income they will be able to afford to purchase upon retirement.

In reality, different investors have different goals, as discussed above, and therefore the safe goal-hedging building blocks should be (mass) customised. Besides, the allocation to the safe rather than risky building blocks should also be engineered so as to secure each investor's *essential goals* (e.g. *minimum* levels of replacement income) while generating a relatively high probability of achieving their *aspirational* goals (e.g. *target* levels of replacement income).

That mass customisation is the key challenge that our industry is facing has long been recognised, but it is only recently that we have developed the actual capacity to provide such dedicated investment solutions to individuals. There are in fact two distinct dimensions of scalability, namely scalability with respect to the *cross-sectional dimension* (designing a dynamic strategy that can approximately accommodate the needs of different investors entering at the same point in time) and scalability with respect to the *time-series dimension* (designing a dynamic strategy that can approximately accommodate the needs of different investors entering at different points in time). The good news is that financial engineering can be used to meet these challenges – see Martellini and Milhau (2015) for a detailed analysis.

Addressing the mass customisation challenge will be facilitated by the convergence of powerful forces. On the one hand, production costs are strongly reduced, due to the emergence of *smart factor indices* as cost-efficient alternatives to active managers for risk premia harvesting. On the other hand, distribution costs are also bound to go down as the trend towards *disintermediation* is accelerating through the development of FinTech and robo-advisor initiatives.

Risk management, defined as the ability for investors, or asset and wealth managers acting on their behalf, to efficiently spend their dollar and risk budgets so as to enhance the probability of reaching their meaningful goals, will play a central role in an industrial revolution that will eventually lead to scalable, cost-efficient, investor-centric, welfare-improving retirement investment solutions.

This is the first of a group of related articles on retirement investing. Please see the next four on "A path-breaking solution for individual investors' problems", "Tailor-Made Funds with Economies of Scale", "Individual Investment Processes for the 21st Century" and "Multi-Asset Products and Solutions".

References

- Deguest, R., L. Martellini, A. Suri, V. Milhau and H. Wang. 2015. Introducing a Comprehensive Investment Framework for Goals-Based Wealth Management. EDHEC-Risk Institute Publication (March).
- Martellini, L., and V. Milhau. 2017. Mass Customisation versus Mass Production in Retirement Investment Management: Addressing a "Tough Engineering Problem". EDHEC-Risk Institute Publication (May).
- Merton, R. 2013. Observations on Financial Education and Consumer Financial Protection. CFA Institute Research Foundation.

A PATH-BREAKING SOLUTION FOR INDIVIDUAL INVESTORS' PROBLEMS

Arjuna Sittampalam, Editor, *IMR Magazine*

1. Introduction

Individual investors have different ambitions concerning their finances. Some of these can be very firm targets while others are more in the category of hopes. There are risks and uncertainties surrounding these definite and less solid aims, which are specific to each individual. Yet, by and large, what they get from financial advisors takes very little account of their individual goals but emphasises much more the risks attached to markets as a whole, an approach that is inadequate.

What is needed is a goals-based investment (GBI) approach that takes account of individuals' personal goals and preferences and the specific circumstances in relation to their future needs. A paper entitled, 'Introducing a Comprehensive Investment Framework for Goals-Based Wealth Management', by Romain Deguest, Lionel Martellini, Vincent Milhau, Anil Suri and Hungjen Wang puts forward an investment framework for catering to these specific needs and hopes. This article is based on the paper.

> **"What is needed is a goals-based investment (GBI) approach that takes account of individuals' personal goals and preferences and the specific circumstances in relation to their future needs.**

In some specific cases, goal-based thinking may not be called for. Take a high-flying 30-year-old earning $500,000 per annum and with wealth of just $25,000. This wealth is very small compared with the value of his large income which he expects to receive over many years in the future. In other words, this $25,000 is only a tiny proportion of his overall assets including the human capital represented by his massive earning power. Therefore, he might well invest his $25,000 with a risk of losing it all but also with the possibility of multiplying his investment many times in a few years. He would not have a particular goal in mind such as $100,000 or $1m.

However, take for example three 35-year old lawyers each with wealth of £200,000 and after-tax income of £30,000. A financial advisor might be tempted to put them in the same market risk category and persuade the client of this. Though a professional, the client could well be ignorant of finance. However, this would not be a good way of going about it.

The first lawyer might be solely concerned about receiving good replacement retirement income at age 65. The second lawyer might have some hopes of retiring early at 50 and then switching to an activity such as running a bookshop with much less income, in line with what he enjoys. The third might be much more focused on paying school fees for his children in the next 25 years and less concerned about his retirement standard of living. Each of these three need different portfolios to suit their separate objectives. Note that, unlike the first and third lawyers, the second has two explicit goals, not just one, namely retiring at 50 and running a bookshop. Firstly, he will need a minimum income of $15,000 per annum at age 50 to survive as an essential goal. Secondly, he will aim to have capital of $100,000 for the bookshop as an important, but not essential goal. This capital would not be essential as he might be able to obtain it through other means, such as a loan.

Let us take another example, this time, of a person with three different goals. A woman at 45 has $500,000 wealth and a regular income of $40,000 per annum. She might want a minimum retirement income at age 60 of $20,000 per annum which has to be achieved and therefore, an essential goal. In addition, she would like an extra $10,000 per annum, highly desirable but not vital and hence, this would be classified as an important goal, not essential one. Finally, she might also have vague hopes of having extra $200,000 for going on expensive holidays. She is keen on constructing an investment portfolio that might allow for the possibility of achieving this amount. This would be an example of an aspirational goal in that it cannot be afforded but is hoped for.

The above examples point to individuals' goals falling under three categories, essential, important and aspirational.

- An essential goal is what the investor has to meet as a sacrosanct target;
- An important goal is what he/she hopes to achieve, but would not be considered as a catastrophe if not attained;
- An aspirational goal is something that is desired, but the investor is realistic enough to understand that there might be a high probability of missing it.

The new GBI approach not only comes up with solutions for constructing the investor's portfolio, but also forms the basis for an ongoing dialogue between the individual and his/her advisor.

Differentiating between and allocating priorities to each of the three types of goals (essential, important and aspirational) is a core aspect of this framework. This categorisation makes it stand apart from investment practices in the retail industry to date.

2. The structure of the goals-based investment framework

2.1 Affordability

A key concept is that of affordability. Is the initial wealth sufficient to ensure that the essential goals can be secured with absolute certainty?

> **"The new GBI approach not only comes up with solutions for constructing the investor's portfolio, but also forms the basis for an ongoing dialogue between the individual and his/her advisor.**

If not, the investor has three options. Firstly, he/she can consider adding more money at the outset so that the initial wealth becomes adequate for the essential goal. Secondly, he/she can consider more money later. Finally, he/she has to be realistic and reduce the essential goal to a lower level. For instance, if an essential goal of $10,000 per annum is unaffordable, it can be reduced to $6,000 per annum which would make the goal certainly achievable.

It is not only essential goals that have to be affordable but also the important goals. The difference between the two types is that the essential goals have to be secured and guaranteed while the important ones, though also affordable, are not guaranteed.

Suppose that an individual at age 50 has capital of $400,000 and wants to have a minimum retirement income of $20,000 per annum at age 65. He/she would prefer to have $40,000 per annum at retirement but would not considered it essential. Assuming the necessary investment return from age 50 to age 65, his/her $400,000 capital would allow him/her to achieve not only the essential $20,000 goal per annum but also the higher $40,000 per annum. If he/she wants the latter to be guaranteed, it is possible that his/her entire wealth of $400,000 has to be tied up for this purpose. However, if he/she only wants the essential $20,000 goal to be guaranteed, then only half of the wealth is needed for this purpose and the other half can be invested in a riskier way to generate more wealth.

Why would one take this additional risk without securing the higher $40,000 income? Suppose also that at age 65 he/she wants to donate $250,000 to an only child or buy a second home in a holiday spot. This would be an aspirational goal that could not be afforded now. If he/she secures the entire $40,000, then this aspirational goal would not be achievable at all. But if he/she secures only the $20,000, the rest can be invested to give him/her a chance of reaching his/her aspirational goal while accepting that the important goal of $40,000 per annum may be put at risk.

In general, therefore, some risk is taken with important goals in order to increase the probability of achieving the aspirational ones. So, both the important and aspirational goals, while achievable, are exposed to risk.

2.2 Wealth and consumption

At a specific date in the future consider an individual aiming to withdraw money from his/her investment assets in order to finance specific consumption targets such as renovating a home, buying a yacht or an expensive car. These targets are consumption-based goals.

In contrast, wealth-based goals require the individual's wealth to reach particular levels at specific future dates.

In the end, all wealth is used for consumption including bequests. However, in a wealth-based goal the individual has no specific consumption usage in mind and is aiming for flexibility here.

The analysis in this article differs according to whether goals are wealth-based, consumption-based or both.

2.3 Three investment buckets

The investor's wealth and any future non-portfolio income are allocated to three different buckets.

- The personal bucket
- The market risk bucket
- The aspirational bucket

2.3.1 The personal bucket

The personal bucket would correspond to essential goals and these goals do not need to be precisely quantified in every case. For instance, a person might be living in an expensive home worth $1m and is determined to live in that home until death. This $1m would be a part of his/her wealth but his/her goal of staying in the home would be achieved by ringfencing this part as untouchable and the value of the home becomes irrelevant. Similar consideration could apply to art collections or comparable assets. At the other end of the economic spectrum, there may be a strong desire to avoid poverty or survive a possible stretch of unemployment, for which some cash cushion is put aside.

The essential goals that are quantifiable need to be catered for by the allocation of liquid assets to meet them. These assets have to comprise a separate portfolio that ensures that the essential goal is met. For instance, consider the above 45-year old woman from section 1 with the essential goal of $20,000 at retirement1. Her initial wealth should be more than sufficient to create enough assets at age 60 to assure this retirement income.

Assume that $300,000 is needed for this purpose. Then this $300,000 can be set aside in a portfolio of bonds that will provide the income with certainty. This is referred to as the goal-hedging portfolio (GHP).

In general, the essential goals are made certain by allocating enough assets to the GHP.

2.3.2 The market risk bucket

In the case of GBI, the assets assigned to the safe GHP fully secure essential goals, and therefore present no risk at all. As remarked in 2.1 above, the rest of the investor's assets are invested in riskier ways to aim for important and aspirational goals. A performance-seeking portfolio (PSP) is established for this purpose.

Such PSPs are not specific to any individual and can be used by anybody who is seeking to maximise their wealth by taking on risk. Modern Portfolio Theory (MPT) suggests choices that are discussed in a later section.

2.3.3 The aspirational bucket

Many aspire to a higher level of wealth and living standards in society. This aspiration is described as a desire for 'wealth mobility'. In the context of GBI, many investors aim for their aspirational goals through what is referred to as the wealth mobility portfolio (WMP) which falls within the aspirational bucket. This, however, does not mean that all aspirational goals target wealth mobility. For instance, consider again the 45-year old woman. She might decide to spend all the $200,000 in luxury expenditure in five years without gaining any mobility in her affluence status in society.

It is not essential that the aspirational bucket contains specific assets for aspirational goals. These can be targeted, in theory, through a leveraged allocation to the PSP, but in practice, many individuals are constrained from borrowing. Hence, the contents of an aspirational bucket often consist of the investor's personal choice of investments. These are typically privately held and have strong upside potential, possibly permitting the achievement of upward wealth mobility within or between affluent segments of the population. Examples of such assets are human capital, equity-based compensation, ownership in privately held companies with 100% majority or significant minority stakes.

3. The goal-hedging portfolio floor and the minimum-maximum principle

3.1 The floor

The three portfolios, GHP, PSP and WMP reflecting the three risk buckets above are the building blocks of the investor's GBI approach. The structure of the paper involves, firstly, the construction of each building block and, secondly, the allocation of the investor's initial wealth and any future income to each of these building blocks.

The GHP has been defined as the portfolio that secures the essential goal. The initial wealth (including the value at inception of any future income) has to be sufficient for this goal to be affordable and above a certain minimum. This minimum corresponds to the lowest possible value of the GHP that is consistent with the goal being attained and is referred to as the 'floor value'.

3.2 Suitable assets

While the minimum wealth has to be above this floor, it is not enough. This is just one of two conditions for affordability. The second condition is that there has to be a set of assets in which the minimum wealth can be invested so that it reaches the value of the essential goal at the latter's horizon. If there are no such assets that make this goal unattainable, then the goal is unaffordable even though the initial wealth is above the floor. But if an appropriate set of assets is available, the goal is reachable and is said to be replicable. It is assumed that all the goals outlined in section 4 are replicable.

Take the example of a 40-year-old woman with $100,000, requiring her wealth at age 60 to be $110,000 with adjustments for inflation in the 20 intervening years. To be certain of achieving this, she has to have access to a bond, the value of which increases with inflation over the years.

However, not all markets have inflation-linked bonds of the precise maturity of 20 years. In such cases the goal would not be affordable as there is the danger that high inflation well above expected rates will surpass any return from the available investments.

3.3 The minimum-maximum principle

Though the initial wealth can be above the minimum floor level as indicated earlier, what happens if more than the floor level of the initial wealth is invested in the GHP? In this case, less becomes available for the PSP reducing the chances of overall wealth increasing. Only by investing the minimum possible amount of the initial wealth in the GHP, subject to the floor being safeguarded, will the PSP performance prospects be maximised. This implies that the amount invested in the GHP should be exactly the floor value, no more no less, as far as practicable.

For instance, assume a man with $350,000 of initial wealth at age 55 wanting a retirement income of $10,000 per annum at age 65. To achieve this retirement income let us suppose that the minimum initial wealth has to be $150,000 for the GHP. In this case, this goal will be attained and $200,000 will be leftover for the PSP, allowing him to aim for other important goals. If he decides to invest, say, $300,000 in the GHP then only $50,000 will be available for the PSP reducing the prospects of his overall wealth increasing.

4. The goal-hedging portfolio in the absence of non-portfolio income

The various concepts defined and explained in the foregoing sections form the essential background to the central task in the GBI of constructing the three GHP, PSP and WMP building blocks, and the all-important process of allocating the investor's assets to each of these blocks.

Three types of cashflows are relevant to the investment framework:

- The initial wealth available for investment
- A possible consumption stream which the investor aims to withdraw
- A possible non-portfolio income stream which will add to the wealth

The non-portfolio income represents additional contributions from the investor as opposed to portfolio income (coupons, interests, etc. generated by the investments). In this section, it is assumed to be absent while the next section covers the presence of such income.

Note that the two conditions for affordability and the minimum GHP and maximum PSP principles outlined in section 3 above are intuitively clear in simple cases. But it is important to bear these in mind for the more complicated situations that are covered in this article.

Several different circumstances are:

- Wealth-based goal with a single horizon
- Wealth-based goal with multiple horizons
- Consumption-based goal
- Multiple wealth-based goals
- Multiple consumption-based goals
- Wealth-based and consumption-based goals

In this section and section 5, we focus on the GHP while the PSP and the WMP are explained in section 6.

4.1 The minimum wealth condition for affordability

The minimum wealth required for constructing the GHPs is derived using elementary financial theory (the principles of compound interest and related discount rates) as well as advanced probability techniques (involving stochastic calculus used in the celebrated Black-Scholes formula for option pricing). Important results and observations are presented here with minimal mathematics and the interested reader should read the original paper for the technical detail.

Under the goals-based framework, investors introduce initial wealth at inception and possibly more infusions of cash later, which are allocated to the GHP and other buckets. What is put into the GHP requires particular attention as these cashflows are aimed at meeting the essential goals.

The general approach followed through this section is best introduced and explained by a simple example of an actual real-life goal.

Consider somebody starting with an initial wealth of $10,000 at inception. In the general case, this initial wealth will be referred to as A_0.

Then assume that he/she has a goal of producing $20,000 at year 10. In the general case, year 10 becomes year T, the time horizon of the goal and the goal of $20,000 becomes G_T.

The central problem is to check affordability. In other words, whether A_0, the initial wealth ($10,000 in this example) at inception is sufficient to produce G_T, the essential goal ($20,000 in this example) at year 10.

For this purpose, G_T is discounted back to its present value G_0 at inception. This discounting is based not only on elementary compound interest techniques but also draws on stochastic theory, given that investment returns and market conditions which might enable G_0 to increase to G_T in ten years are uncertain. Hence, G_0 is the minimum amount that has to be available at inception to reach G_T at year 10.

The test is then whether A_0, the initial wealth ($10,000 in this case) is greater than or equal to the minimum amount needed, G_0. If it is, the goal is affordable with the initial wealth. Otherwise, it is not.

The simple example above was based on just a single goal at year 10. In general, however, there may be more than one goal (multiple goals) at times T1, T2…..T_r which could be wealth-based, consumption-based or both. While the above reasoning for the single goal remains applicable, complications are introduced with the multiple goals but the general approach stays valid.

4.2 Wealth-based goal with a single horizon

A wealth-based goal at a particular time horizon T requires the wealth of the investor at that time to be equal to or above the value of the goal, G_T at the horizon after starting with initial wealth, A_0 at inception.

This simple case exactly corresponds to the example outlined in section 4.1 above and hence, the affordability criterion derived there is repeated here.

The present value of G_T, the terminal wealth at time T, is G_0. The test is then whether A_0, the initial wealth is greater than or equal to the minimum amount needed, G_0, the floor value. If it is, the goal is affordable with the initial wealth. Otherwise, it is not.

Remember that the test of affordability also requires the existence of a replicable set of assets that can achieve the goal starting with G_0. As mentioned in section 3 above, all the goals in sections 4 and 5 are assumed to be replicable. If there is no such set, then the goal is not affordable even if the initial wealth is above the floor.

4.3 Wealth-based goal with multiple horizons

Often a wealth-based goal would require a minimum level of wealth not just at one date in the future but at many different dates. For example, consider somebody who aims to have $200,000 at the end of year 10 and wants this wealth level to increase by inflation at the end of each of the following 15 years. This represents a wealth-based goal with multiple horizons at years 10, 11, 12…25.

In the general case, minimum wealth levels are required at various time horizons, T_1, $T_2 \ldots T_r$ where r is the number of horizons. In the above case, r is 16 and the horizons are T_1 (10), T_2 (11), T_3 (12)…T_{16} (25).

Though the general principle as in 4.1 above remains the same, complications are introduced by multiple horizons. The initial wealth has to be higher than the combined present value of the wealth required at all the time horizons T_1, $T_2 \ldots T_r$. However, it is not straightforward as the notion of present value is not well-defined for a multiple set of future pay-offs at different dates. A specially defined formula for present value is used in this case based on advanced financial techniques.

4.4 Consumption-based goal

The general principles, outlined in section 4.1, in evaluating the minimum wealth required for consumption-based goals are the same as for the wealth-based goals.

In general, a consumption-based goal is a stream of consumption at a series of dates T_1, $T_2 \ldots . T_r$ in the future. But sometimes it pertains only to a single consumption involving, say, the purchase of an expensive yacht. This is catered for by having the consumption value as zero at all but one date in the series.

Unlike in the case of the wealth-based goals with multiple horizons, the present value at year 0 of the consumption stream is easily derived and is equal to the sum of the present values of each consumption payment at T_1, $T_2, \ldots . T_r$. This present value is the floor value of the GHP.

Consider a consumption stream of $200 at the end of year 3, $400 at the end of year 5 and $100 at the end of year 8.

At year 0, assume that the present values of these three consumption payments at years 3, 5 and 8 are $180, $370 and £72 respectively.

Then the total present value of the three payments at year 0 is $622 ($180+$370+$72) which is the floor value of the GHP.

The condition of affordability is that the initial wealth is equal to or exceeds this floor value, G_0.

The set of replicable assets is, in general, a set of securities (often pure discount bonds) maturing on the consumption dates with pay-offs equal to the consumption payments.

4.5 Multiple goals and joint affordability

While the foregoing referred to just one goal, either wealth-based or consumption-based, in reality many individuals have more than one goal of each type.

Then the question of joint affordability arises. Whereas previously the criteria for the affordability of a goal on its own were established, in the multiple case conditions for the different goals to be jointly affordable need to be identified.

4.5.1 Multiple wealth-based goals

The general case of several single wealth-based goals each with multiple horizons is considered and analysed here for their joint affordability. On the face of it, multiple goals might seem to produce additional complications.

Consider that there are three goals A, B and C. The three goals are equivalent to a single goal D with the horizons combining the separate horizons. A numerical example will make this clearer. Suppose

goal A - $200 at year 2, $300 at year 6

goal B - $150 at year 1, $500 at year 9

goal C - $600 at year 7, $800 at year 8

Goal D is the combination of goals A, B and C.

goal D - $150 at year 1, $200 at year 2, $300 at year 6, $600 at year 7, $800 at year 8, $500 at year 9.

The three separate goals A, B and C have become equivalent to a single goal D. Similarly, multiple wealth-based goals can be treated as a single wealth-based goal with multiple horizons.

4.5.2 Multiple consumption-based goals

They are treated in the same way as multiple wealth-based goals and are equivalent to a single consumption-based goal combining the separate goals with pay-offs at different horizons. A major difference is that evaluating the sum of pay-offs rather than their maximum is involved. The floor value, which the initial wealth of the combined goal has to equal or exceed for joint affordability, is the sum of the individual floor values of each of the separate goals. This is simpler than computing the maximum as required for multiple wealth-based goals.

4.5.3 Wealth-based and consumption-based goals

This is a simple extension of the logic applying to multiple wealth-based goals in section 4.5.1 and multiple consumption-based goals in section 4.5.2 above. These goals can be all treated as a single wealth-based goal and a single consumption-based goal.

Following the principles in the sections above, the joint affordability of these goals is again based on a series of specially defined present values that identifies the floor value for the GHP.

5. The goal-hedging portfolio in the presence of non-portfolio income

The analysis in section 4 was based on a single infusion of capital, the initial wealth at inception with no other cashflows being brought in by the investor. There is an additional complexity in the presence of incoming cashflows (referred to as non-portfolio income or income distinct from portfolio income such as dividends and interest received), though the general principles in section 4 remain applicable.

There are alternative ways of financing the GHP differing in how much credit is taken for the later income flows and how much of the initial wealth is assigned to the GHP, with the rest being allocated to the PSP.

Clearly, as explained in section 3 above, the lower the amount of the initial wealth that is invested in the GHP to secure the essential goals, the more money is available for the PSP to increase the investor's wealth. It is, therefore, in the investor's interest to use the method that allocates the lowest amount of the initial wealth to the GHP.

This lowest possible initial wealth is referred to as the 'cheapest' GHP. Selecting the cheapest financing method involving the maximum use of future income and minimum use of initial wealth should be the way forward. In line with this, when consumption-based goals are involved, the initial wealth will be utilised only to finance those parts of the consumption pay-outs not catered for by available income. This principle is highlighted by the various techniques.

5.1 LIQ technique

Consider an initial wealth of $100 at year 0, an income of $40 at the end of year 1 and a consumption of $100 at the end of year 2. Taking no credit for the income in year 1, most of the initial wealth of $100 can be used to buy a discounted zero-coupon bond that pays out $100 in two years to meet the consumption outflow. Since the price of this discounted bond at year 0 will be close to $100 (allowing for the small amount of inbuilt interest accumulation), very little will be available for the PSP. This technique is referred to as LIQ.

Under this method, the initial wealth allocated to the GHP will be a little less than $100, say, $96 if we can assume that this $96 initial wealth will increase to the required consumption pay-off of $100 at the end of year 2 after allowing for bond interest of about $4 over the two years. In this case the amount available for the PSP at inception will be only about $4, the difference between the initial wealth and the allocation to the GHP.

5.2 INC-ZER technique

The second choice is to buy a discount bond at year 0 that pays out $60 at year 2, allowing for the $40 at year 1 being available at year 2. This method referred to as INC-ZER clearly leads to a cheaper GHP than the one via the LIQ method, with only about $60 of the initial wealth being used and leaving the rest for the PSP. But it assumes that the $40 at year 1 earns zero interest up to year 2.

With INC-ZER a little less than $60 of the initial wealth is allocated to the GHP relying on the $40 of income at the end of year 1 to enable the required $100 consumption pay-off at year 2. Here, the amount available to the PSP at year 0 is a little over $40, about ten times what is produced by the LIQ method.

5.3 INC-CMP technique

In the INC-ZER method, no allowance is made for the income of $40 at year 1 to earn interest. In another method, INC-CMP, this allowance is made. As a result, there is a reduction in the proportion of the $100 pay-out that needs to be financed from the initial wealth. So, INC-CMP leads to a GHP cheaper than what INC-ZER produces.

5.4 INC-FWD technique

Forward contracts do not always exist in a market but if they do, investing in such contracts can represent an even cheaper method of establishing the GHP, referred to as INC-FWD.

5.5 Other situations

Another case involves multiple income flows interspersed with consumption outflows. The approach to selecting the cheapest portfolio remains the same, although the process is a bit more complicated. At any particular date of a consumption pay-out, all the unused income up to that date should be allowed for before calculating how much of the initial wealth will be used to contribute to this pay-out. The overall principle that future income should be preferred to current wealth in securing the essential goals still applies. Hence, the maximum amount of the initial wealth is made available for the PSPs aimed at increasing the investor's wealth and helping to achieve non-essential goals.

The principle of preferring income to initial wealth in establishing the floor holds true only when income is certain. If the income flows are uncertain, it might be better to use liquid wealth rather than the uncertain income to achieve the essential goals with certainty.

In the general case, a formula is derived for the floor of the GHP, the minimum initial wealth for the goal to be affordable. The securing of this goal involves the use of exchange options. Based on this general solution, specific formulae are also derived for each of the above methods: LIQ, INC-ZER and INC-CMP. In some cases, bonds (including zero-coupon bonds) can be used instead of exchange options.

A special case involves retirement where there is a series of income flows before retirement (accumulation phase) and consumption outflows after retirement (decumulation phase). This case is analysed in detail in the paper.

6. Performance-seeking portfolio and wealth-mobility portfolio

In the two previous sections, the construction of the GHP corresponding to the personal bucket was explained in detail. In this section, the other two building blocks are covered: PSP corresponding to the market risk bucket and WMP corresponding to the aspirational bucket.

6.1 Performance-seeking portfolio

While the set of GHPs corresponds to the essential goals that are personal to every investor, the PSP seeks to maximise performance through a well-diversified portfolio investing in assets that are tradeable in financial markets. From this perspective, even hedge funds and private equity can be considered as they are relatively diversified attempts to exploit market opportunities.

MPT points to a type of portfolio known as the maximum Sharpe ratio (MSR) portfolio, which all investors should use to achieve optimal performance with an efficient risk-reward ratio. Typical investment practice is to identify investments in various asset classes and then to allocate the portfolio among these.

In MPT terms, the procedure involves the first step of generating investable proxies for MSR portfolios in each asset class which are usually taken to be market indices. In equities particularly there is a growing belief that market cap-weighted indices are not suitable and techniques such as smart beta are being increasingly adopted. Then the second step is to allocate the PSP assets to each of these separate asset classes in order to establish a multi-class PSP.

This second step is known in the industry as asset allocation referring to weights to be assigned to each asset class. Recently, however, this asset allocation process is being challenged strongly by a relatively new but important idea referred to as risk allocation based on identifying the underlying risk factors within each asset class.

The market risk bucket has the most liquid assets. The excess non-portfolio income is invested in this bucket which also finances withdrawal of funds for non-essential consumption.

6.2 Wealth mobility portfolio

The WMP, in attempting to meet aspirational needs, aims at producing high returns which are unlikely from the PSP. Typically, WMP has a very high degree of idiosyncratic risk. Often these assets are chosen by the investor himself/herself without consulting a professional advisor as stated in section 2.3.3

7. Hierarchical classification of goals

The three types of goals – essential, important and aspirational – are all goals that the investor hopes to achieve with different levels of probability. The essential goals demand certainty while the aspirational ones need the lowest probability. However, there might be serious problems when all three types of goals are combined in the investor's overall portfolio, with potential dangers to higher priority goals from lower priority ones.

The concept of affordability was described and the criteria for it were analysed in earlier sections. When different goals are combined in the same portfolio the concept of joint affordability becomes important. Two individual goals could be each separately affordable but the initial wealth may not be enough to afford both of them.

If, however, the initial wealth is sufficient to provide for both the goals together, then the two goals are said to be jointly affordable. The same applies to a set of larger number of goals if they are all affordable on a joint basis.

An individual goal may be affordable in its own right, but may no longer be so when combined with a higher priority goal or a set of higher priority goals and therefore, will not be affordable on a joint basis.

For the purpose of prioritisation, all essential goals need to be made secure and have to be affordable. If the complete set of essential goals is not jointly affordable but each goal within the set is individually so, then the investor has to either bring in more initial wealth to make them jointly affordable or make some of the lower priority essential goals merely important ones.

Within the set of important goals, as well as in the set of aspirational ones, there has to be orders of priority in terms of investor's preference.

To avoid conflicts of priorities, an explicit hierarchy within and across types of goals needs to be established at their outset. When cashflows relating to lower priority goals occur before higher priority ones, there is the danger that the use of these cashflows for the low priority goals can affect the achievement or the probability of the high priority goals. So, the rules are

- A consumption cashflow for an important or aspirational goal that occurs before a consumption for an essential goal should take place if, and only if, the use of this cashflow for the low priority goal does not make any of the essential goals aspirational (non-affordable).
- The same applies to a consumption outflow for an important goal that occurs before an outflow for a more important goal. It should be allowed if, and only if, it does not convert the more important goal to an aspirational status.
- A consumption cashflow for an aspirational goal that occurs before a consumption for an important goal should take place if, and only if, the use of this cashflow for the low priority goal does not make any of the important goals aspirational (non-affordable).
- A consumption cashflow for an aspirational goal that occurs before a consumption for some other essential goals should take place if, and only if, the use of this cashflow for the low priority goal does not decrease the probability of achieving any of these other goals.

8. Allocation between building blocks

8.1 Main principles

The allocation between the three buckets – personal (GHP), market risk (PSP) and aspirational (WMP) – is based on two principles:

- The strategy must secure all essential goals
- It should lead to the highest success probabilities for the non-essential goals

In MPT, deriving an optimal portfolio in the presence of goals introduces complications studied extensively in academic research.

The GBI strategy draws upon the mainstream academic theory of expected utility maximisation. The concept of utility is explained with a simple example. An extra $1,000 gained by a billionaire provides him/her with very low additional utility while the same amount has much higher utility if received by somebody earning just $1,000 a month. This concept is a cornerstone in MPT.

In the allocation process the WMP is ignored as it is pre-selected as the personal choice of the investor unless he/she is able and willing to liquidate this bucket and make the proceeds available for the PSP. Thus, the allocation process concerns for the most part just the division of liquid wealth between the PSP and the GHP.

8.2 The simplest approach: Buy and hold

The simplest and most natural strategy is to secure the GHP by investing the minimum wealth required in it and to allocate the rest to the PSP. Therefore, the amount invested in the PSP is the difference between the total wealth and the amount invested in the GHP.

The GHP is purchased at inception and the rest is invested in the PSP. This is described as the 'buy and hold' strategy because in the absence of additional income it remains unaltered. If the total wealth shrinks with the PSP falling in value, the gap between the total wealth and the GHP goes down close to zero. In fact, it is the amount invested in the PSP that approaches zero and in the extreme, the wealth is invested entirely in the GHP. This strategy is akin to the constant proportion portfolio insurance (CPPI) strategy where the risk disappears when the portfolio hits the floor.

8.3 The optimal approach

This buy and hold strategy, however, is not the most optimal. A detailed analysis of the optimal strategy, based on MPT principles, leads to more complex formulae which are simplified with the use of a multiplier m. Intuition also suggests the use of this multiplier through the extension of the buy and hold strategy. This strategy then becomes just a special case with m equal to 1.

In the light of the above, the amount invested in the PSP initially is not the difference between the total wealth and the initial GHP value. Instead, it takes a new value of m times this difference and what is invested in the GHP becomes correspondingly less, below the floor value F. The latter is seemingly put at risk, but not really as explained below.

In the buy and hold strategy of section 8.2, the amount invested in the PSP, P is equal to the difference between the total wealth, T and the amount invested in the GHP, G.

Remember that the GHP value has to be greater than or equal to the floor value as this is one of the conditions for essential goals' affordability. But it should not be greater based on the principle that the cheapest possible GHP will allow the maximum investment in the PSP increasing the probability of more wealth. So, G has to be set exactly at the floor value F, neither higher nor lower, and thus it coincides with the required floor. Then, the equation

$$P = T - G$$

becomes equivalent to

$$P = T - F$$

When m is introduced, P is multiplied m times and becomes P_{new}

$$P_{new} = m\,(T\text{-}F)$$

Then G becomes G_{new}, the difference between the total wealth and P_{new}

$$G_{new} = T - P_{new}$$

When m is greater than 1, P increases to P_{new} and correspondingly G which was the floor value falls to G_{new}. In this situation, G_{new} is below the floor value. It might then appear that the floor value is no longer secure because the new GHP value, G_{new} is below that value.

But this is not so. If the total wealth falls to the floor value, P_{new} becomes zero, G_{new} becomes just T, T becomes equal to F and thus G_{new} becomes equal to F. So, G_{new} reaches the floor value again.

This explains why the appearance of the floor value being insecure is not a real danger. As long as the total wealth stays above the floor value, it is always possible to secure it. The worst is when the total wealth falls to the floor value but as mentioned here, G_{new} automatically achieves the floor value, thus removing any danger and securing the essential goal.

When m is taken as 1, the formula becomes the same as in the buy and hold strategy. Generally, m can vary over time but in the GBI framework m is fixed at a particular value.

8.4 The different situations and rebalancing

The various cases (wealth-based goals, consumption-based goals and multiple goals) covered in section 4 lead to the same formula structure where the CPPI principle holds. As the total wealth approaches the floor value, the PSP allocation shrinks and when the total wealth reaches the floor value, the entire wealth is securely invested in the GHP. Conversely, if the total wealth increases, the allocation to the PSP rises with the probability of achieving more performance.

It was shown above that the introduction of the multiplier (greater than 1) poses no risk to the security of the floor. However, this assumes that rebalancing is continuous as the PSP drops in value. In reality, rebalancing and associated trading take place at discrete time intervals. If the PSP falls by a large amount between rebalancing dates, the floor itself might be breached before the next rebalancing occurs. This risk becomes more serious if the exposure to the PSP is large as a consequence of a high value for m.

8.5 Multiple goals

Multiple essential goals introduce another dimension. For instance, in the case of two multiple goals the floor is the higher of the two different floors. The GHP then as a safe asset corresponds to the higher floor, which is the one more likely to be breached as the wealth falls.

If the two floors are reasonably close to each other, each can replace the other as the higher floor from time to time. Which floor is higher, therefore, is re-evaluated at each rebalancing date and correspondingly the safe GHP has to be switched to the new higher floor if necessary.

The case of more than two multiple goals is a clear extension of this two-goal process.

8.6 Cap on wealth

An interesting situation arises when a cap is imposed on the maximum wealth level. In this event the probability of reaching high wealth levels, though below the cap-imposed maximum level, can be enhanced.

8.7 The presence of non-portfolio income

The presence of income flows requires decisions at each income date as to what to do with the incoming cash. The general principle remains valid that consumption expenses should be primarily financed with income and that liquid wealth is used only when income on its own cannot meet consumption needs.

Note that under the general principle the GHP floor takes into account, at the outset, expected income flows. So, as each income flow comes in, some of it is allocated to the GHP and the rest to the PSP. Thus, the GHP continues to secure essential consumption-based goals.

9. Case Studies

Individual investors are typically classified by life stage and affluence.

Three life stage clusters are:

- LS1: Accumulation (age less than 55 years)
- LS2: Transition (age between 55 and 65 years)
- LS3: Decumulation (age higher than 65 years)

The affluence dimension is usually grouped under:

- A1: Mass affluent ($250,000 to $1m)
- A2: Affluent/high net worth (HNW) ($1m to $5m)
- A3: Ultra high net worth (UHNW) (>$5m)

These nine clusters can be further grouped under three overall clusters:

- C1: Accumulation/transition < $5m (LS1/A1, LS1/A2, LS2/A1, LS2/A2)
- C2: Decumulation < $5m (LS3/A1, LS3/A2)
- C3: UHNW, whatever the life stage (LS1/A3, LS2/A3, LS3/A3)

Three case studies are analysed in detail relating to each of the C1, C2 and C3 clusters.

- Case study 1 - HNW/UHNW transition: A HNW/UHNW couple with substantial assets in the transition phase. This is a proxy for cluster C1.
- Case study 2 - HNW retiree: A HNW couple at the beginning of the decumulation/retirement phase. This is a proxy for cluster C2.
- Case study 3 - Affluent accumulator: An affluent young couple in the middle of the accumulation phase. This is a proxy for cluster C3.

The three case studies are not covered in detail here and only some aspects are commented on.

In each of these, the individuals' multiple goals and existing assets are listed. The latter are allocated under each bucket and the split between different asset types is given. For instance, in case study 1 the market risk bucket consists of $2.15m which is divided between $1.5m in equities, $0.6m in US fixed income and $0.05m in cash. The current allocations in each of the case studies imply an existing strategy. In comparison, GBI strategies are also developed.

The analysis in the case studies shows that the GBI approaches developed lead to superior results overall compared with the existing strategies in terms of probabilities of reaching the goals.

In general, some components of the personal bucket are left untouched without bringing them into the GBI. These include the individual's residence and any cash holdings held as protection against emergencies and unforeseen financial shocks. Any liquid wealth in the personal bucket that is not ringfenced in terms of a particular goal is aggregated with the entire wealth in the market risk bucket as a supply of liquid assets to structure the GBI.

In the case of the aspirational bucket, the investments in general are assumed to be illiquid and left untouched for GBI purposes. But in some cases how the situation might be improved is examined, should these assets be liquid. In such cases assets in this bucket are added to the other liquid wealth for the implementation of the GBI, the success rate of which is enhanced by this change.

In two of the case studies the introduction of a possible cap is investigated for potential improvements in increasing the probabilities of achieving the goals. Indeed, there are such improvements. The impact of taxes is also considered and it is found that they might affect some goals but the essential goals remain robust to them.

The use of the multiplier m leads to an increase in the allocation to the PSP and the prospects of higher performance. In theory, m should have no effect on the safe GHP, hence the essential goals. But as explained in section 8, because of rebalancing dates at discrete intervals, some risk can occur with high values of m.

In the case studies, the impact of m for various values has been analysed with the base value fixed at 5. It is found that the essential goals remain secure for m up to 6 and a risk appears only at 7. But even up to 10 the risk is small. The rebalancing frequency has a bigger impact for example when decreasing it from monthly to quarterly.

Overall, the three case studies provide examples of how the principles of constructing GHPs set out in sections 4 and 5 and the question of allocating assets to the three buckets in section 8 are implemented in practice. Readers who would like to have an enhanced grasp of the theoretical principles in the earlier sections are advised to read the case studies in detail.

10. Impact of taxes

The existence of taxes (on dividends, coupon and interest income as well as capital gains) introduces a frictional factor. In the case of income the approach is straightforward as it is a matter of just multiplying the relevant income by a factor allowing for the after-tax rate. However, these after tax-rates vary according to the investor's exposure to the various asset classes each of which have different yield levels.

Capital gains taxes are affected by rebalancing decisions leading to the realisation of gains or losses. Therefore, they can be avoided in theory by not having any rebalancing.

In practice, however, rebalancing has to take place in some situations. Even in the buy and hold strategy referred to in section 8.2, while there is no rebalancing between the GHP and the PSP, within the latter rebalancing is required in order to keep weights in line with a long-term allocation.

Furthermore, if the multiplier m outlined in section 8.3 is greater than 1, when the PSP falls in value, rebalancing has to take place to reduce exposure to it. But in such cases usually the returns are negative with losses, though often leaving overall tax payments still positive.

A major problem in accounting for capital gains taxes is that they depend on the volume of transactions within a year and therefore, it is difficult to allow for it in advance. Several ad-hoc adjustments are proposed.

In the presence of taxes having wealth just equal to the present value of the goal is not sufficient as future taxes can depress wealth to take it below the present value. Ideally, the floor should be increased by an amount equal to the present value of the year-end tax payment. But as this is dependent on future rebalancing, a better option is to raise the floor by the amount of taxes already generated since the beginning of the year.

Taxes could render a previously affordable goal unaffordable but every case has to be looked at on its own. For example, in case study 1 of the previous section, it is shown that the GBI strategy is relatively unaffected by taxes in securing the essential goal, except for high values of the tax rate and the multiplier m.

Hence, the conclusion is that a lower value of the m should be chosen to secure the essential goals.

11. Inputs, outputs and mass customisation

The aim of the GBI framework is to identify and classify goals under the three risk buckets based on investor preferences and status of the funding at any given time, with the aim of developing strategies that secure essential goals and ensure substantial probabilities for non-essential ones. Various inputs and outputs are needed to keep asset allocation under review.

11.1 Inputs

11.1.1 Subjective inputs

The starting point are obtaining investor preferences regarding goals and checking their affordability relative to the funding available. This has to be subject to periodic revisions as the value of the overall portfolio and market developments and conditions change. Also, the investor's goals might be altered for events such as a substantial upward or downward shock in financial circumstances or a birth in the family triggering an educational goal. The revisions demand regular meetings with a financial advisor.

11.1.2 Objective inputs

In the process of establishing the overall GBI strategy, the present values of assets at various times in the future need to be evaluated. This requires in some cases theoretical models and assumptions about past data.

For example, consider zero-coupon bonds. In general, a zero-coupon bond maturing at every relevant date in the future is not available and special techniques are needed to fill the gaps. Furthermore, to estimate future performance of bonds various models are utilised.

> “
> The aim of the GBI framework is to identify and classify goals under the three risk buckets based on investor preferences and status of the funding at any given time…

The simplified forms of the GBI analysed in the paper are not subject to the risk of using an inappropriate model. Essential goals are secured in any event and so is the substantial access to upside potential, though not quantified, of PSP assets. But what is subject to models and not robust are the actual quantitative assessments relating to important and aspirational goals.

11.2 Outputs

Outputs are needed for both monitoring and adjusting the GBI strategy and also generating important information for a vital dialogue at regular intervals between the financial advisor and the investor. As with inputs, there are subjective as well as objective outputs relating to the probabilities of reaching goals, any expected shortfalls and asset allocation recommendations.

Various measures of success are available for the assessment of strategies whether GBI or others such as the investor's existing approach. Examples are:

- Probability of success of the goal

- The expected maximum shortfall and variations on this theme – the worst possible shortfall under all scenarios, not just the average expected

These measures relate to the different types of goals such as wealth-based and consumption-based and all the relevant horizons. They are computed by simulating future portfolio performance that depends on assumptions regarding the specific returns on various assets and the evaluation of relevant risk factors such as interest rates and inflation. These return and risk parameters can be re-estimated on every occasion when the portfolio is reviewed.

An important output relates to the opportunity cost of playing safe with any given essential goal. If this goal was not essential, additional assets would be available for the market risk bucket leading to a higher probability of more wealth. This extra wealth is foregone if the goal remains essential and represents the opportunity cost of securing the essential goal, a measure that the investor could find useful in his/her allocation to various risk buckets. This process could be widened to allow the investor to estimate the opportunity cost of staying with the current strategy rather than going for the more optimal GBI strategy.

11.3 Allocation advice

It is emphasised that essential goals must have a 100% probability of achievement irrespective of models and parameter values chosen. An example of this principle being violated lies in an equity allocation. If a stock index is assumed to generate a sufficiently high return in the future with a long enough time horizon, then a strategy based on a very high weighting in the stock index can suggest that an essential goal is certain to be achieved. But this forecast return is highly unreliable and in truth, the essential goal would not be secured.

In contrast, as explained in section 8, referring to the allocation process between the GHP and the PSP, when the total wealth falls to the GHP level, the wealth becomes invested entirely in the GHP, thus securing the essential goal. This result is independent of models.

11.4 Mass customisation

A substantial problem arises in real life in the implementation of a GBI solution. Investors, in general, share a financial advisor with many others. This severely limits the time that the advisor can give any particular client unless they are very wealthy. Thus, it is difficult for the advisor to offer a high degree of customisation needed for the GBI solution. It is, therefore, necessary to group individual investors into clusters according to their characteristics. The question then is whether the tailored benefits of the framework remain available with such groupings.

The answer depends on which of the three building blocks – the GHP, the PSP and the WMP – are being considered as well as on the final process of allocating between these three.

11.4.1 The goal-hedging portfolio

These are specific to every investor and demand a high degree of customisation. In the case of consumption flows, a solution is to invest in bonds, the maturities of which match consumption cashflows such as pension payments. The reinvestment of coupons issued by these bonds also presents a problem. It is suggested that if this exact cashflow matching of bond maturities with consumption payment dates is difficult, advanced techniques referred to as immunisation or convexity adjustments and various interest rate management tools can be used. Thus, it might be possible for financial advisors to achieve a reasonably acceptable implementation of GBI strategies if there is available a series of bond portfolios, both real and nominal, which can be used as proxies for GHPs.

There is a trade-off between the range of portfolios the advisors can use and the probabilities of achieving the essential goals. If this range is too narrow, these probabilities can fall unacceptably below 100%.

11.4.2 The performance-seeking portfolio

The PSP represents the easiest. In theory, all investors can have the same PSP with their different expected performance targets by varying the proportions allocated to the optimal maximum MSR portfolio and cash. Leverage can be used in attempting to exceed the return available on the MSR portfolio.

In practice, however, there are some obstacles such as constraints on leverage, short sales and foreign investments as well as exclusion of particular asset types. This would mean that different investors might need different PSPs, but even this is manageable with a relatively small number of PSPs.

11.4.3 Allocation between building blocks

The biggest obstacle to mass customisation is the all-important process of allocating the investor's wealth between the building blocks, particularly between the GHP and the PSP as set out in section 8 (and ignoring the WMP for the reasons given there). It is, for the most part, impossible to offer the same allocation strategy to several investors even if their goals are similar to each other. The allocation to the various blocks depends on the specific situation of each investor including current wealth levels and the extent of progress towards each goal.

The big question still remains as to whether a limited number of portfolios (encompassing different proportions in the GHP and the PSP) can do the job of mass customisation. It is suggested that model portfolios used by financial advisors with equity allocations of 20%, 40%, 60%, 70% and 80% and the rest of each allocation in bonds and cash can be interpreted as crude approximations for GBI strategies. The idea is that equities represent risks, bonds represent long-term consumption needs and cash provides the protection required by GHPs against losses. However, proper mass-scale customisation needs a dedication allocation to a limited number of building blocks requiring a satisfactory information technology system.

12. Summary and conclusion

The innovative and revolutionary goals-based investment (GBI) framework starts with concepts such as 'affordable' that are already used in everyday language. Note that it is quantification that converts investment decision-making from a nebulous gut feeling-based function to a scientific process. The expression of loose ideas in precise quantifiable terms forms the bedrock of the new process.

'Affordability' is given an exact quantifiable definition. The test for affordability, a key aspect of the approach, is partly based on the identification of a floor, another word in ordinary use that people can relate to but expressed here in terms of the minimum wealth needed.

Essential, important and aspirational are yet again popular words but are used in a precise way that converts vague feelings that individual investors already have about their financial ambitions into more accurately pinpointed investment goals which play a critical role in the framework.

In addition to the minimum wealth floor, it is also important to note that a goal, to be affordable, requires the existence of a suitable set of assets that can ensure its achievement.

To date, the concept of risk among individuals and their advisors is just that of the markets. The retail sector focuses on deciding whether they want high risk, medium risk or low risk for their portfolios, hoping that the outcome from the chosen risks will satisfy in some way what they hope to achieve in terms of wealth without any precise target in mind. On the other hand, the new approach deals with the risks of not achieving individuals' specific goals. Market risks are only a subset of these.

The three investment buckets – personal, market risk and aspirational – classify risks in a different way and bring a new understanding to the risk-reward thinking of most individuals and their advisors in terms of goals. In pursuit of these goals, investors' assets are allocated to three investment portfolios – the goal-hedging portfolio (GHP), the performance-seeking portfolio (PSP) and the wealth mobility portfolio (WMP).

The GHP is the one that is subject to a minimum floor in order to ensure affordability and security of essential goals. The PSP is established for the purpose of trying to create extra wealth through investment performance. The WMP, very specific to each investor, typically comprises assets selected personally with a potential high return.

The construction of the GHP takes account of various types of multiple goals that reflect the real-life behaviour of many investors, thus enhancing the practical value of the investment solution prescribed. The PSP is based on Modern Portfolio Theory principles and common to all investors. Setting up the WMP does not arise as it is already preselected.

In the face of multiple goals it is important that the investor ranks them in a clear hierarchy even within the essential, important and aspirational classes, so that implementation of the framework does not violate the hierarchy of preferences.

Taxes present a problem. Income taxes are less of a difficulty but capital gains taxes are unpredictable, depending on the level of transactions and are a source of friction that detracts a little from the smooth workings of the GBI approach. However, the effect is not too serious.

Finally, it is important to keep this investment approach under continual review in terms of progress towards goals. Various inputs and outputs, part of the monitoring process, need to be incorporated into an ongoing dialogue between investors and their advisors. Such a dialogue takes the interaction between these two parties to a new more sophisticated level.

A major concern is that, by and large, financial advisors might find the advocated solution that needs to be tailored for each client uneconomic. Hence, the question of whether mass customisation is possible comes up. But

> "
> **…such customisation in a proper sense might be difficult until an information technology system is made available that can cope with the inputs and outputs demanded by the framework.**

such customisation in a proper sense might be difficult until an information technology system is made available that can cope with the inputs and outputs demanded by the framework.

The GBI framework is innovative not just in catering for individuals' specific goals. It also breaks new ground in several of its component features requiring investors and their advisors to bring more precision into what they want. Their more concrete financial desires can thus be brought closer to fulfilment than is currently the case.

Overall, widespread adoption of the new process will enable the individual sector to access and reach the level of professional thinking that is taken for granted by leading institutional investors.

This is the second of a group of related articles on retirement investing. Please see the previous one on "Retirement Goal-Based Investing" and the next three on "Tailor-Made Funds with Economies of Scale", "Individual Investment Processes for the 21st Century" and "Multi-Asset Products and Solutions".

References

• Deguest, R., L. Martellini, V. Milhau, A. Suri and H. Wang. 2015. Introducing a Comprehensive Investment Framework for Goals-Based Wealth Management. EDHEC-Risk Institute Publication (March).

TAILOR-MADE FUNDS WITH ECONOMIES OF SCALE

Arjuna Sittampalam, Editor,
IMR Magazine

1. Introduction

The asset management industry has to change radically in the next decade or so. How it will change is uncertain but change it must given the wall of problems it faces, particularly in the retail sector where its offerings are widely not considered fit for purpose.

A new academic paper 'Mass customisation versus mass production in retirement investment management: Addressing a tough engineering problem' by Lionel Martellini and Vincent Milhau points a way forward for the retail industry, a road that should revolutionise the industry for the better. This article is based on the paper.

What is wrong in retail? The most pressing problem lies in the growing shift from defined benefit to defined contribution pension schemes, wherein individual beneficiaries are landed with responsibility for investment decision-making which they are not equipped for, given widespread illiteracy. The imperative for the solution to a potentially serious crisis on this front becomes ever more pressing year by year with fast-growing ranks of individuals under the threat of poor retirement living standards.

A brief review of what is available to prospective individual retirees helps to explain the deficiencies. The retail industry currently offers funds on a mass production basis that takes no account of the individuals' specific needs. A relatively recent innovation, target-date funds, goes in this direction in a small way. Investors are classified broadly by age such as 35, 40 or 50 and separate versions of these funds are offered to each group based on these entry ages. In all cases the funds close at the retirement age, typically 65.

However, these funds focus merely on the terminal value which the individuals receive at retirement. They do not deal with the problems of ensuring adequate retirement income and the risks attached to it including investment returns, interest rates, inflation and longevity. Even in the run-up to retirement they do not cater for the specific goals of individuals.

In the past, there was little choice. Mass production of mutual funds and more recently exchange-traded funds were the vehicles available in the retail industry. It was obviously not economical to produce a customised fund specific to the needs and preferences of each individual unless extremely rich.

Nobel laureate Robert Merton stated as far back as 2003 that the uniformity in advice given to retail customers need not continue as the tools and theories are available for more customised solutions. He saw mass customisation that combines the economies of scale and low costs attached to mass production with the flexibility of solutions tailored for individuals as a tough engineering problem and the challenge was how to solve it.

2. An individual's goals

The results in the paper are derived using elementary financial theory (the principles of compound interest and related discount rates) as well as advanced probability techniques (involving stochastic calculus

> "
> **The most pressing problem lies in the growing shift from defined benefit to defined contribution pension schemes, wherein individual beneficiaries are landed with responsibility for investment decision-making which they are not equipped for…**

used in the celebrated Black-Scholes formula for option pricing). Important results and observations are presented here with minimal mathematics and the interested reader should read the original paper for the technical detail.

In line with the above, an important principle that needs to be understood is that of compounding and discounting. When a person deposits, say, $1,000 in a savings account and keeps it for 10 years without withdrawals, at the end of the period his/her money increases to over $1,219 assuming that the compounded interest rate is constant at 2%. Thus the $1000 is multiplied by a 10 year compounding factor to arrive at $1219. In a reverse approach, a sum of $1,219 in 10 years has a present value of $1,000 at the start, year 0. This value is derived from the terminal value of $1,219 by discounting for the interest received. In this case, the $1219 is multiplied by a 10 year discounting factor to produce $1000.

In general, interest rates are not known for certain in advance and investment returns in riskier assets are not predictable. Estimated probabilities of these variables reaching different levels are brought into play with stochastic techniques but the general idea remains the same.

2.1 The two stages of mass customisation

Individuals, in general, have an idea of the minimum standards of living they want in retirement below which they would feel unhappy. In addition, they might aspire to a much higher standard that would be desirable, but not the end of the world if not attained. In terms of retirement income, these two standards are referred to as essential and aspirational goals respectively.

The challenge of mass customisation is met with a solution that encompasses two stages. The first stage concerns a single individual whose specific needs are the two goals, essential and aspirational. At this stage, a portfolio is constructed that enables his/her essential income to be secured while providing a good chance of achieving the desired higher aspirational income.

The second stage addresses the problem of scalability where this portfolio structure can be used for a large number of individuals with just a limited number of funds that can be set up economically by the industry. The achievement of this would be the fulfilment of the tough engineering problem mentioned by Merton. This second stage is covered from section 5 onwards.

2.2 Two portfolios: Goal-hedging and performance-seeking

The first task is to identify the maximum possible retirement income, $ri_{max,0}$ that can be purchased with an initial wealth. Consider a man at time 0 with an initial wealth of A_0. The maximum income he can receive at time T, the retirement date, will be based on the wealth, A_T available at that time. A_T can be derived from the initial wealth, A_0 by multiplying it by a compounding factor which reflects the investment return between the times 0 and T. A_T is then used to buy an annuity that will pay $ri_{max,0}$ per annum for the rest of his life.

The investor's entire initial wealth has to be used to receive this maximum income. Therefore, nothing is left to aim at the possibility of getting something higher. In other words, $ri_{max,0}$ is the highest level that can be afforded with certainty.

However, the above investor can decide that it is not vital for him to receive $ri_{max,0}$ entirely and that a lower income would be acceptable. For this, he will clearly need only part of his initial wealth.

This lower essential income can be expressed as $\delta_{ess} ri_{max,0}$ where δ_{ess} is less than 1 and therefore, $\delta_{ess} ri_{max,0}$ is less than $ri_{max,0}$. He is then able to try for his aspirational higher income, $\delta_{asp} ri_{max,0}$ where δ_{asp} is greater than 1. For this purpose, he can use the remaining part of his initial wealth that was not used to secure his essential goal.

This part of the initial wealth is then invested in a risky portfolio which offers a good chance of reaching his aspirational retirement income, $\delta_{asp} ri_{max,0}$, instead of $ri_{max,0}$.

We now have two portfolios, the goal-hedging portfolio (GHP) that secures the essential income goal, $\delta_{ess} ri_{max,0}$ with certainty and the performance-seeking portfolio (PSP) which offers the chance of achieving the aspirational income goal, $\delta_{asp} ri_{max,0}$.

Before the PSP is introduced, A_0 at time t=0 will increase to A_1 at time t=1, A_2 at time t=2 and so on until it reaches A_T at time T through the compounding process mentioned above. The maximum retirement income available at time T will remain the same throughout.

When the PSP is brought in, however, the overall wealth levels at all times after the start will no longer be the same because of the performance of the PSP. So, at the times t=1, 2, 3…T the wealth levels will be not A_1, A_2, A_3…A_T, but will have values of B_1, B_2, B_3…B_T. Each of these will produce not a retirement income of $ri_{max,0}$, but different retirement incomes of $ri_{max,1}$, $ri_{max,2}$, $ri_{max,3}$…$ri_{max,T}$ instead.

For example, suppose that the PSP performs so strongly between the times 4 and 5 as to double the wealth. Then there will be a huge jump in the maximum retirement income that can be secured, because of this performance. Thus, at any time t between 0 and T there will be not an A_t leading to $ri_{max,0}$, but instead a different B_t which will produce $ri_{max,t}$.

Note that an individual may not have just one aspirational goal. He/She may have several values in mind for δ_{asp}, each of which corresponds to a different level of aspirational income. But this is no great matter because the PSP, in principle, can achieve any reasonable level with some probability.

Typically, δ_{ess} and δ_{asp} are chosen to be 80% and 130% respectively of the maximum affordable income. So, δ_{ess} is 0.8 and δ_{asp} is 1.3. In other words, under this typical scenario 20% of the maximum income is sacrificed at the outset to secure the essential goal and this 20% is released for the PSP to achieve an extra 30% above the maximum.

2.3 Level of funding

$ri_{max,t}$ at time compared with $ri_{max,0}$ at time 0 is referred to as the funded ratio, R_t. It represents the level of funding of the retirement income possible at time t relative to that at time 0.

$$R_t = \frac{ri_{max,t}}{ri_{max,0}}$$

When t equals 0 the funded ratio is 1. R_t should be greater than or equal to δ_{ess}. If this ratio reaches a value of δ_{asp}, it means that $ri_{max,t}$ is $\delta_{asp} ri_{max,0}$ and has achieved the level of the aspirational goal even before retirement, a very positive development.

The implications of the above will be explained in detail in later section.

2.4 Presence of contributions

The foregoing assumes only initial wealth at the beginning and no subsequent contributions from the individual into the retirement pot. In general, however, individuals may also have regular contributions each year until retirement. This will add to the wealth at retirement and thus increase the maximum affordable retirement income.

But the problem is that at inception the conditions, under which these contributions can be invested, are uncertain and it makes the task of calculating the maximum affordable retirement income more difficult.

Suppose that at inception bond yields are around 5%. The initial GHP can be set up on the basis of this 5% bond rate. 10 years later what can be done with the contributions coming in? At that time bond yields might have fallen to 2% or risen to 10%, a huge difference in circumstances affecting the future returns on these contributions.

This problem is avoided by selling a bond at the outset referred to as an 'accumulation bond'. The underlying concept is easily explained by thinking of this bond as a borrowing or a loan taken out.

Consider the parallel of a woman having $300,000 of liquid wealth for investment and a house worth $200,000. She is also able to save $10,000 per annum out of her employment income which she intends to add to her liquid wealth over the next 25 years. However, instead, she takes out a loan, for which she is assumed to be creditworthy.

She calculates the amount of the loan that she can afford and it is equal to the present value of her savings of $10,000 in the next 25 years so that the loan can be paid off with these savings.

Assume that this present value is $150,000. So, the loan is set at $150,000 which she then adds to her liquid wealth. She has now a liquid wealth of $450,000 and a loan of $150,000. She pays off this loan in instalments from her savings of $10,000 in the next 25 years. By taking out the loan she can have access to a bigger pot of wealth of $450,000 at the very outset instead of having to wait for her savings in the future years.

Note that while her pot of wealth for investment at the outset has increased to $450,000, her net financial wealth has not changed as it is calculated by deducting the loan of $150,000 from $450,000 leading to $300,000 as before.

For the retirement scheme the same principle applies. The loan taken out is equal to the discounted present value, V at time 0 of the future annual contributions until retirement. This capital V is then added to the existing initial wealth, W to form a bigger pool of wealth, X equal to V+W allowing a higher maximum retirement income.

As explained above, the loan is merely one method of borrowing and a similar borrowing is done through an accumulation bond (a securitised form of the loan).

Remember that when governments and corporates issue bonds, all they are doing is borrowing money from the public, though they are doing it by issuing(selling) a tradable security. The public are the lenders and the government is the borrower. When the bond matures the government repays the amount of the bond. In technical language, somebody buying a bond is taking a long position in the bond and somebody selling the bond is referred to as taking a short position or short selling the bond. The buyer, in general, during the lifetime of the bond receives coupons. Conversely, the seller pays out coupons.

The value of this bond at the outset, as in the case of the loan above, is equal to the present value, V of the stream of contributions. Coupons are used on the same basis as the loan instalments and correspondingly, just as the loan was paid off, the accumulation bond is fully extinguished along with the last annual contribution. Therefore, this bond unlike most of the traded variety has no capital repayment at the end.

Under these circumstances the total portfolio has now three parts, namely, GHP, PSP and a negative position (a short position) in the accumulation bond which is being paid for by the contributions coming in.

Remember that while the total liquid wealth of the investor at the outset increases to X, his/her net wealth is unaltered because the value of the accumulation bond, V has to be deducted from the total wealth to arrive at the net financial wealth, W.

While at the beginning of the period the value of the accumulation bond is the present value of the stream of the contributions. at a later time t, its value decreases to become the present value of the contributions still outstanding from t to the retirement date, T. When the retirement date is reached, no further contributions will be paid and therefore, the value of the accumulation bond with no outstanding contributions due will be zero.

3. Optimal investment policy

The essential goal, $\delta_{ess} ri_{max,0}$ has to be secured by retirement date, T. The minimum wealth that is required at age T to achieve this essential goal is referred to as G_T. $\delta_{ess} ri_{max,0}$, when multiplied by K ($\delta_{asp}/\delta_{ess}$), leads to $\delta_{asp} ri_{max,0}$, the aspirational goal.

On the same basis, this minimum wealth, G_T for the essential goal when multiplied by K produces the wealth, KG_T needed at retirement time T for the aspirational goal.

So, the optimal investment policy involves making sure that;

1. the wealth, W_T at time T equals or exceeds G_T
2. the probability of achieving the aspirational wealth level, KG_T at time T is maximised.

The optimal investment policy with these two objectives is derived using Modern Portfolio Theory (MPT) and the solution leads to the optimal strategy which has the following components:

a) the GHP that replicates the annuity financing the essential goal

b) the PSP based on the maximum Sharpe ratio (MSR) portfolio and cash

c) the accumulation bond

This optimal policy involving dynamic trading between the GHP and the PSP leads to a terminal wealth that only produces two values, G_T (necessary for the essential goal) and KG_T (desirable for the aspirational goal), though the latter is not certain. Under this policy nothing in-between can be achieved. So, this points to several problems with the optimal strategy.

Firstly, with a fixed KG_T and δ_{asp} scalability becomes difficult as the value of δ_{asp} differs among individuals. Also, remember that many individuals do not have a fixed aspirational goal, but several. Therefore, assuming a particular value for δ_{asp} will certainly not suit them.

Secondly, implementing this optimal strategy is difficult because dynamic hedging between the building blocks is complicated by the difficult task of predicting expected returns. Moreover, massive unacceptable levels of leveraging might be required.

Finally, investors might be unhappy with only two possible outcomes, G_T and KG_T and nothing in-between which they might consider a gamble.

4. Building blocks and allocation

4.1 Allocation between the building blocks

To make sure that essential goals are achieved, not only does the investor's wealth have to be above the minimum value, G_T at the retirement date, T but also in the event of exiting the fund any time before that date, his/her wealth at the time needs to be sufficient to secure his/her essential retirement income. This requires his/her wealth at any time t between the start date and T to be above a floor level, G_t.

Since this applies to all individuals, it means that the value of the entire fund, X_t has to exceed a floor value, F_t. This floor value is computed by using a formula that differs according to how the fund is made scalable and less customised, as explained in later sections. For the purpose of this section, it is sufficient to note that there is a well-defined floor value above which the total wealth must stay at all times.

It was explained in section 3 that the optimal investment policy consisted of investing in a combination of a GHP and a PSP. A feature of this strategy is that the allocation to the PSP falls to zero when the total wealth falls to the level of the GHP. Then the fund becomes fully invested in the GHP. However, the strategy has the weaknesses listed earlier and is not implementable. The same feature of the PSP falling to zero is captured by making the allocation to the PSP a multiple m of the difference between the fund value, X_t and the floor value of the GHP, F_t. Thus, financial theory shows that the amount invested in the PSP, $A_{PSP,t}$ is

$$A_{PSP,t} = m\,(X_t - F_t)$$

The amount invested in the GHP, $A_{GHP,t}$ is the difference between total fund value and the amount invested in the PSP.

$$A_{GHP,t} = X_t + Vt - A_{PSP,t}$$

where V_t is the value of the outstanding contribution stream and also the value of the accumulation bond

Finally, the amount in the accumulation bond, $A_{ACC,t}$ is

$$A_{ACC,t} = - V_t$$

V_t, the value of the accumulation bond is negative because it is a short position as explained in section 2.4 and equal to the present value of the outstanding contributions at time t.

Since the amount invested in the PSP is not the difference between the wealth and the floor value but is a higher level of m times this difference, what is invested in the GHP becomes correspondingly less, below the floor value. However, this value is not in danger because as Xt falls to the floor value, the amount invested in the PSP becomes zero and the total wealth is entirely invested in the GHP. Therefore, the floor is secure.

The multiplier m in theory can be varied over time but is fixed here.

4.2 Funded ratio developments

In implementing the strategy based on the multiplier m, the funded ratio, $ri_{max,t}/ri_{max,0}$ referred to in section 2.3 needs to be monitored.

At no time should this ratio be allowed to fall below δ_{ess} because this would imply that the essential goal has become unaffordable. To avert this happening, enough PSP assets need to be transferred to the GHP in time.

Should the ratio reach of δ_{asp}, the aspirational goal becomes affordable but needs to be made secured. Remember that this development has taken place with a rise in the risky portfolio PSP. If no action is taken, the PSP can fall again with the aspirational achievement lost.

To avoid this, as soon as the aspirational goal is reached, the investor is advised to secure it by liquidating the PSP and investing the proceeds in an annuity that preserves the retirement income $\delta_{asp} ri_{max,0}$ so that it is fully secured.

4.3 Construction of the building blocks – practical issues

The following practical considerations apply to the three building blocks.

4.3.1 Performance-seeking portfolio

In theory, while the PSP is based on a proxy of a MSR portfolio, in practice, it is expected that any well-diversified portfolio which aims to produce returns with a good risk-reward ratio will be used.

4.3.2 Goal-hedging portfolio

The GHP involves the use of annuities both at retirement and at inception when deferred annuities are used for the establishment of the GHP. The value of these annuities are not based on market prices. The reasons include market quotes differing by provider and the fund itself being unable to invest in annuities as they are only sold to individuals. So, the value of the annuities are modelled by estimating inputs such as the cashflows, inflation rates and survival probabilities.

The deferred annuities are subject to interest rate risk which is catered for duration hedging techniques.

4.3.3 Accumulation bond

The value of the accumulation bond, being equal to the present value of the outstanding contributions, (see section 2.4) is based on discount rates in the calculation the present values and therefore, is exposed to interest rate risk. Hence, as in the case of deferred annuities, duration hedging techniques are brought into play.

5. Scability and different floors

5.1 Parameters for mass customisation

In general, individuals differ according to the following variables:

D1 - Gender

D2 - Retirement age

D3 - Retirement date

D4 - Start date of the accumulation phase

D5 – The initial contribution at the entry date and the annual scheduled contribution

D6 - Essential (D6a) and aspirational (D6b) goals expressed as fractions of the initially affordable income

At one extreme of scalability, it will be good to have a single portfolio that accommodates all individuals who vary by each of the above variables. This is an unattainable ideal, though it will represent full scalability.

At the other extreme, every individual with different characteristics will have his/her own portfolio, the fully customised solution. This is also unattainable for the practical reason that fund management companies cannot economically provide this.

What is feasible is to have a solution in between the two extremes with scalability with respect to some of the variables and customisation with regard to the rest. The overall aim is to cater to all the different types of individuals with a minimal number of different funds that can be established.

In the portfolio solution outlined in section 4, it was shown that no account is taken off a particular individual's aspirational goal. It is left for him/her to decide when his/her goal is reached and take action accordingly. So, this D6b can dropped from consideration for scalability.

It is assumed that D2, the retirement age is 65 as is the case in various countries

The age and gender differentiation is catered for by having five entry ages 35, 40, 45, 50 and 55 and having different funds for each age group and gender. Individuals' falling between these ages are assigned to the nearest relevant age group in this list. In such cases, because their age will differ a bit from the age assumed in the design of the portfolio, particularly in the GHP, they will not enjoy the same complete guarantees as the ones with the same exact age do.

D6a, the essential goal indicated by δ_{ess}, the relevant fraction of the maximum affordable retirement income at the start, is fixed as mentioned earlier at 80% for everybody as a base level, though in some circumstances for reasons explained in a later section a slightly higher ratio can be used.

Hence, the scalability process has to now only deal with D4, investors having different entry dates and D5, different levels of initial contribution (wealth) at the start date and the regular annual contribution at subsequent years.

In section 4 above, the allocation between the GHP and the PSP was based on a floor level, details of which was not given because it depended on the level of scalability in different versions of the portfolio. Section 5.2 below deals with the case of a portfolio without scalability with respect to the variables, D4 and D5.

In section 5.3, the different shape of the formula for the floor of a portfolio open to entry by investors at any date after the inception of the fund is explained, thus scaling for D4. The floors in 5.2 and 5.3 are different but both can be used in the implementation of the building block allocation as outlined in section 4. Thus, the portfolios in sections 5.2 and 5.3 are practicable.

The derivations of the floors in 5.2 and 5.3 involve some very basic mathematical reasoning on an intuitive rather than rigorous level. Readers content with accepting that the floors can be derived and implemented in the portfolio strategy outlined in section 4, can skip 5.2 and 5.3, merely appreciating that scalability with respect to entry dates can be achieved just by introducing a new formula for the floor level. Scalability with respect to contribution levels remains and is dealt with in section 6.

5.2 Basic floor for different contribution levels and entry dates

As explained in section 4, a critical aspect of the allocation strategy between the three building blocks the GHP, the PSP and the accumulation bond is the floor level of the portfolio which has to be protected at all times and this floor level is that of the GHP.

In this subsection, the floor of a fund scalable in neither D4 or D5 is discussed.

All investors share the same entry date and the same ratio of annual contribution to initial wealth.

In general, investors entering at the start will bring initial wealth (contributions), and commit to scheduled annual contributions. They may also make unscheduled contributions from time to time. Though individuals' contributions differ, it is not the size of the wealth or contribution that is important but the ratio y/W0 of the annual contribution y to initial wealth W0 which is assumed to be the same for all investors.

If unscheduled contributions occur, the reasoning below would not hold. However, because these cannot be anticipated in advance, they cannot be allowed for in evaluating the floor for this fund. Therefore, for the purpose of the establishing the floor, these unscheduled contributions need to be ignored and assumed to be zero.

Because all the individuals have the contribution ratio, it can be shown mathematically that this ratio is also equals the overall fund's ratio, C/X_0 where C is the total amount of regular contributions and X_0 is the total of initial contributions invested and hence the initial value of the fund. The number of shares of the fund held by each investor remains constant.

So, if investor A has 10 times the initial wealth (contribution) of investor B, then his regular contribution will also be 10 times than the other's because the ratio is the same and the number of shares of the fund held by each initial individual remains constant. Investor A will have 10 times the shares of the fund as investor B but each will have their share constant.

Consider an example of a fund where the lowest possible investment is an initial wealth of $1000 and a commitment to an annual contribution of $100. No unscheduled contributions are allowed. This 2-pronged purchase price of a share is correspondingly referred to as (1000, 100). All investors can purchase only either this minimum or a multiple of this minimum. Then suppose that another investor brings in (5000, 500). His investment is 5 times the minimum. Equivalently, imagine that this investor is replaced by 5 different investors each bringing in (1000,100).

From the perspective of the fund, the impact is the same whether one investor invests five times the minimum or 5 investors invest the minimum. So, the entire fund can be assumed to be populated by investors with the minimum purchase only, all having one share each.

From another perspective, also suppose the entire fund consists of 8000 shares of (1000,100) each. Again, it makes no difference to the fund if all 8000 shares are held by just one rich individual who brings in (8,000,000, 800,000). These examples don't hold if unscheduled contributions take place but the absence of these was explained to be a necessary assumption for deriving the floor.

The maximum affordable retirement income, $ri_{max,0}$ at the retirement date, T is determined by E_0, the sum of the investor's initial wealth, W_0 and the present value of his/her regular contributions, V_0 outstanding at time 0 between the times 0 and T.

So, E_0 is $E_0 = W_0+V_0$

The essential retirement income, $\delta_{ess} ri_{max,0}$ is determined by G_0 which is equal to $\delta_{ess} E_0$.

$G_0=\delta_{ess}(W_0 +V_0\_)$

This G_0, because of the compounding factor, will become G_1, G_2 and so on and reach G_T at time T which will ensure $\delta_{ess} ri_{max,0}$ at retirement. Correspondingly, G_0 will reach G_t at time t before T.

As explained in section 4, the investor exiting at any time and still having his/her essential goal secured has to be allowed for. So, at time t his/her wealth, has to be at least G_t less the value of outstanding contributions which corresponds to the floor value at that time.

The initial value of the fund, X_0 is the sum of the all individuals' initial wealth and the fund itself will have a contribution stream which is the sum C of the all individuals' contribution with a present value, Q_0.

Then H_0, the sum of X_0 and Q_0 at time 0, will secure the maximum retirement income of all the individuals at time T.

$$F_0 = \delta_{ess} H_0$$
$$= \delta_{ess}(X_0 + Q_0 0)$$

If we imagine the whole fund to be held by one individual as in the example above, then X_0 is his initial contribution and Q_0 is the present value of his annual regular contribution stream C and the floor exactly corresponds to the single individual bringing in W_0, referred to earlier.

Thus, F_t will be derived from F_0 which will progress to F_t at any time t. F_t is then expressed as a formula that is based on $\delta_{ess}(X_0 + Q_0)$ and the present value of the outstanding contributions as it progresses to F_T at T.

5.3 Allowing for different entry dates

Also note that any time t. X_t will secure retirement goals different from what F_t will secure, a point that is key to the next section.

What if somebody enters the fund outlined in the section above, say, 5 years, after the start date with all the other parameters including his/her contribution ratio being the same?

When the new investor enters at time t after the start date of the fund, the value of the floor at the time he/she enters will be F_t and the value of the fund will be X_t.

This floor value, F_t is the minimum level required at time t to secure the essential goal of all the investors who entered at time 0. But when the new entrant comes in, his/her essential goal at retirement will be based on the fund value, X_t at the time of his/her entry. This essential goal could be higher than the original essential goal set at time 0 for the fund, and if so, only that is secured and his/her higher replacement goal will not be secured.

For example, consider a fund with an original floor F_0 of $500,000 at time 0 which will secure an essential retirement income goal of $40,000 per annum. Suppose at time t=5 this F_0 has become F5 equal to, $550,000 through the compounding factor. The original essential goal of $40,000 pa will remain secured.

Also suppose that the fund value X5 at year 5 has jumped to $1,100,000 because of the PSP portfolio's performance allowing an essential retirement goal of $80,000 to be affordable. Because only the original retirement income set at year zero is protected though the floor F_t at year 5, the new entrant's retirement goal of $80,000 will not be fully protected.

This problem is dealt with by ensuring that the floor is set a value higher than the maximum replacement goal that can be afforded by the fund value at any time between the start of the fund and up to and including the time the new entrant comes in so that all who entered even at earlier times but later than the start have their essential goals protected by this maximum floor.

This higher floor, referred to as the relative max drawdown floor, ensures that the essential goal of an investor who enters at any time after the start date is secured. A fund with this floor is then available for all entrants irrespective of entry dates and therefore, this fund is scalable for the entry date and the only parameter that still needs to be scalable is the contribution ratio, C.

6. Scability for different contribution levels

6.1 Two funds for everybody

Now that a method for catering to different entry dates in the same fund has been established, the only parameter that still needs scalability is the contribution ratio, c, equal to the regular annual contribution divided by the initial wealth. For every value of c separate fund has to be established referred to as (1,c).

Without this scalability, the situation is still not satisfactory for fund providers as individuals vary widely in this ratio and it is not economical to provide virtually a different fund for each client corresponding to an exact c value.

There will be thus no benefit from scalability unless different individuals with different c's can be accommodated in a small number of funds.

The big breakthrough achieved in the paper is to show that scalability with respect to widely varying contribution levels can be achieved by everybody investing in just two funds. Their vastly different contribution levels are allowed for by each individual allocating different proportions of their initial wealth and future contributions to each of these two funds.

The two funds are referred to as (1,0) and (1,c) but in this case c takes a fixed value, typically, 1/50, 1/25, 1/20, 1/10, 1/5 and 1.

In these funds 1 in (1,0) and (1,c) stands for an initial wealth of $1 and in the (1,0) fund regular contributions are zero. The same terminology is used irrespective of the actual wealth in the fund.

Suppose in a the (1,c) fund where for example c is fixed at the value 1/10 and the initial contributions is set at $1000, then, the regular contribution is equal to 1/10 of $1,000, $100 per annum.

Consider another person whose initial wealth is $500 which means that the income is 1/10 of $500, $50 per annum. The ratio of initial wealth to regular contribution is 10 times in both cases, reflecting c fixed at 1/10. So, both individuals would invest in an identical way except that the first one would have double the assets of the other.

The choice of fixed c is that of the fund management company. A different company may choose a different c for these two funds.

People with different ratios of initial contribution to final contribution then invest different proportions the two funds, (1,c) and (1,0) as shown below

6.2 Allocation between (1,0) and (1,c) funds

So the allocation of an investor's wealth where the initial wealth is C_o and their regular annual contribution C in subsequent years is split up between the two funds, (1,0) and (1,c) as follows.

There are two cases considered:

a) When the initial wealth C_o is equal to or larger than the annual regular contribution C in subsequent years. Typically, this is the case.

b) When the initial wealth C_o is smaller than the annual regular contribution.

The first case a is the most straightforward one. At date 0, (C_o -C) is invested in the (1,0) fund and C is invested in the (1,c) fund

Then C in subsequent years is invested in the (1,c) fund.

If there are irregular unscheduled contributions this will be invested in the (1,0) fund. When the investor makes exactly the same annual contribution every year, the number of shares held in each fund by him remain constant from year to year. Thus, this strategy is described as a buy and hold strategy where the allocation between the two funds is fixed.

In general, if stated above there are additional unscheduled contributions then these are invested in the (1,0) fund and his number of shares in this fund increases but the corresponding number of shares in the (1,c) does not change.

An important issue is that the investor never reallocates between the two funds and therefore, does not sell any part of either fund and thus not incur transaction fees or taxes.

It is also shown that the investor through this strategy enjoys the same guarantee as he/she would enjoy had he/she invested in a hypothetical (1,c) fund where c is not fixed but equals the investor's personal contribution ratio.

Thus, his/her essential goal is secured and therefore, this strategy of investing in the two fixed funds is an effective strategy for achieving scalability for individuals with differing contribution levels, D5.

Case b is more difficult than case a. In this case, C_o less than C and investing C_o-C as in case a in the (1,0) fund is not possible since C_o-C. Therefore, the entire contribution at the beginning is invested in the (1,c) fund.

In subsequent years only a proportion of the regular contribution is invested in the (1,c) fund and the rest is in the (1,0). In general, the essential goal in this case is not fully secured.

Under these conditions only part of the regular contributions that can be treated as essential contributions can be secured in terms of the essential goal. However, if the fixed c is sufficiently large e.g., 1, then it might be secured.

7. Resets and new funds

7.1 Risk budgets and resets

It will be recalled that in the cases δ_{ess} of the maximum goal is essential level and the rest is free for investments for taking risks in order to reach the aspirational goal.

Typically, δ_{ess} starts at 80% so that the risk budget amounts to 20% of the initial wealth.

It is true that if the multiplier m is greater than 1, initially more than the risk budget can be invested in the PSP but even here if the PSP falls sufficiently, the risk budget of 20% come into play.

This is all well and good if everybody enters at year 0. But new entrants coming in at later dates will have a higher floor level, the relative max drawdown floor, explained in section 5.3.

The relative drawdown maximum floor never goes down because of the running maximum stipulation. Thus, new investors will come in at a level higher than 80% and they have limited upside of say, 12%, should the floor has become for example 88% of the fund value. It is then the case that the upside for the new investors is now smaller not the difference between 100 and 80, but the difference between 100 and 88. In general, it is possible that the floor for new investors has increased even closer to 100. In these cases, they will have very little upside to achieve aspirational goals.

To remedy this situation, the initial floor is set a higher level which is more than δ_{ess}. Then if the floor value as a proportion of the fund value reaches a very high level such as 95%, it is reset to a lower level, subject to being higher than δ_{ess}, which improves aspirational chances for new entrants.

The number of resets allowed is unlimited as long as the reset level remains above δ_{ess} and when it reaches δ_{ess} further resets are not allowed,. Resets can be implemented for all (1,c) funds. Typically, resets occur after poor relative performance of the PSP relative to annuities, causing a rise in γ_t.

7.2 New funds

New (1,c) are launched on two types of occasions. Firstly, when younger cohorts are coming for whom the existing funds don't fit in terms of the age criterion.

The second situation occur when old funds are either sterilised or quasi-sterilised. What sterilisation means is that when the fund value reaches the floor the PSP allocation becomes zero and the fund is fully invested in the GHP. Though the essential goal is safe, aspirational goals are no longer possible justifying the use of 'sterilisation'.

In the case of quasi-sterilised funds, the fund value has not quite reached the floor but almost there and hence, the potential for good performance is very limited.

Typically, even after the reset only a large outperformance of the PSP can save the day as the gap is still small. Then a new fund becomes necessary with the same gender, maturity date and age date as the existing fund but with a later start date.

A mechanical rule such the maximum floor value not exceeding 95% can point to the need for a new fund is indicated but in practice, it also depends on the fund provider's assessment of other criteria such as the demand for new funds.

8. Numerical analysis

A major part of the analysis consists of examining how (1,0) and (1,c) performed under simulation for different values of m in various age groups, 35, 45 and 55. Two sets of probabilities are evaluated. Also, separate simulations are carried out based on start dates of 1996 and 2016. These two dates are distinguished by the prevalence of very high interest rates at the earlier time and exceptionally lower ones more recently.

It is found that essential goals were secure up to m=3 in all cases but slightly less so for higher values of m when 2016 interest data were used. But the risks became higher with the 1996 data. This is understandable because high interest rates made it more difficult for the stocks in the PSP to outperform the annuities of the GHP. In many cases, while the risk to the essential goals were low, there were much higher chances of achieving the aspirational goals up to δ_{asp} 130.

It is shown the optimal value of m is 3 based on the simulations from the perspective of keeping the floor secure but maximising the aspirational achievements and that m=1 is never the best choice.

A very interesting conclusion is that for older people nearing retirement, there is a case for increasing the value of m and hence the risk taken in order to increase the probability of achieving aspirational goals. This result goes counter to conventional wisdom that people should take less risk as they get older.

In the analysis, strategies with different levels of customisation in terms of the probabilities of achieving essential and aspirational goals are compared.

In addition, customised strategies are compared with many currently available retirement funds such as balanced funds and target date funds. It is found that while many of them performed rather well on the upside, it is the PSP that achieves the best success probabilities for high aspirational goals. On the downside, many of these funds fail to secure 80% of the maximum retirement income in a reliable way.

The effect of resets was investigated and it is found that the reset process delays the creation of a new fund in many cases. This development is much to be desired given fund providers need to limit the expenses of launching new funds.

For a fuller appreciation of the principles and techniques outlined above, reading of the numerical analysis chapter of the original paper is recommended though the technical detail given there can be taken as read.

9. Summary and conclusion

The objective of the mass customisation solution is to establish a reasonably small number of portfolios which will cater for individuals with different entry ages, retirement dates, genders, entry dates into the fund, contribution levels and their essential and aspirational goals defined as proportions of δ_{ess} and δ_{asp} of their maximum affordable retirement income.

In terms of gender and age the individuals can be grouped in clusters of 35, 40, 45, 50 and 55 and separate gender portfolios. δ_{ess} is assumed to be 80% and δ_{asp} in some solutions has to be fixed and is usually 130%. But in the actual portfolio strategy outlined in the paper, δ_{asp} does not have to be fixed and thus does not need scalability. Therefore, the mass customisation solution focuses on scaling just the entry date and contributions levels.

The overall portfolio consisted of three parts, the goal-hedging portfolio (GHP) securing the essential goal, the performance-seeking portfolio (PSP) and the accumulation bond. The latter is introduced so that the expected regular contributions at the outset can be invested along with the initial wealth by effectively borrowing in advance the value of the initial contributions. The borrowing takes place through the accumulation bond which is paid off with the contributions coming in.

People entering at different dates are easily allowed for by having a different floor that is distinct from the case where all investors have the same entry date. The question allowing for investors having different contribution levels is solved by the remarkable step of everybody investing in different combinations of two basic portfolios, the combinations varying according to their contribution levels.

As funds get older, the risk budget that is the amount that can be invested in the PSP can decrease leading new comers at later dates having less upside risk. The cost of providing them with new funds is significantly reduced by allowing the floor level of the fund to be reset but new funds are still needed when reset possibilities are exhausted and also for catering to new younger members.

The prescription in the paper promises to make a big difference to the lot of future retirees, if adopted widely. The tough engineering problem – in the title of the paper, words used earlier by the distinguished Robert Merton appears to have been cracked with the ingenious introduction of just two fund types in which masses can invest while allowing for the individual needs.

This is the third of a group of related articles on retirement investing. Please see the previous two on "Retirement Goal-Based Investing" and "A Path-Breaking Solution for Individual Investors' Problems", and the next two on "Individual Investment Processes for the 21st Century" and "Multi-Asset Products and Solutions".

"The tough engineering problem…appears to have been cracked with the ingenious introduction of just two fund types in which masses can invest while allowing for the individual needs.

Reference

• Martellini, L. and V. Milhau. 2017. Mass Customisation versus Mass Production in Retirement Investment Management: Addressing a Tough Engineering Problem. EDHEC-Risk Institute Publication (May).

INDIVIDUAL INVESTMENT PROCESSES FOR THE 21st CENTURY

> The advisory sector has to recognise and accept the superiority of the new solutions compared with what they are accustomed to.

Arjuna Sittampalam, Editor, *IMR Magazine*

For the most part, private individuals have come a poor second in terms of the attention paid to their problems by the asset management industry. The solutions put forward in the two in-depth academic studies conducted by EDHEC-Risk Institute, summarised in the preceding two articles, go a long way to redressing the balance. These solutions, if adopted by the industry as is needed, will put individuals closer to par with leading institutional investors, with respect to the techniques and advice available for their specific needs.

The two papers on the goals-based framework and mass customisation complement each other very well. Together they are applicable to the entire wealth spectrum of individuals with the former covering the mass affluent upwards and the latter addressing the needs of the more average investor. These solutions could trigger some much-needed revolutionary changes in asset management.

The goals-based framework, while introducing path-breaking precision to the implementation of individuals' goals, will make the dialogue between an individual and his/her advisor more sophisticated. The ingenious solution to the mass customisation problem of allowing large numbers to address their specific needs just by investing in two funds will vastly improve the lot of future retirees. These people, coping with investment decisions that they are not qualified for and suffering from inadequate advice, are shooting in the dark when applying their hard-earned savings towards achieving a reasonable retirement income.

However, the fact that a solution is available and should be adopted does not mean that it will be. A powerful push is required from the industry to start with. Three groups have to adapt. Fund management companies have to not only provide new types of funds but also be proactive in propagating the new processes in a committed manner.

The advisory sector has to recognise and accept the superiority of the new solutions compared with what they are accustomed to. Previously, they had nothing better available and had to resort to fairly broad groupings of funds but the future is a different matter.

Finally, end-investors have to have a basic understanding of why the investment solution is much better for them than any vehicle they used before.

The industry will not find it easy to change direction. Both asset managers and advisors have to consider costs of training staff, new computer systems and the time needed to educate clients. Time and money is already invested in existing products which most institutions would be reluctant to abandon without strong motivation.

In fact, any commercial organisation would be reluctant to change its business model and product portfolio unless confident that the appropriate market exists. This is the classic chicken and egg argument. The market cannot arise without the supply and in turn, suppliers need to be assured of some demand.

In the past, innovative funds were introduced, classic examples of which are index funds, exchange-traded funds and hedge funds; but these are all funds. The new solutions represent investment processes, not just products and there has to be collaboration between asset managers and the intermediaries from the word 'go'. Perhaps the solution is for a few enlightened asset management companies to take the lead and undertake the training and supporting of advisors who are willing to participate.

Fortunately, some of the current trends in the industry militate in favour of change. The most important impetus comes from the downward pressure on fees with active managers being squeezed by passive and smart beta vehicles. Industry consolidation is widely predicted. The economies of scale offered by mass customisation could encourage some of the survivors to kickstart the process. Note that BlackRock has announced its intention to shift its investment decision-making to more computerised practices, relying less on human involvement.

In spite of Trump's presidency in the US being unenthusiastic about the new rule mandating fiduciary advice, it looks as if that is the way of the future with the momentum being unstoppable. If the new processes can be shown to be obligatory in a fiduciary 'best advice' sense, then the pressure for advisors to adopt it could be intense and the rest of the world could follow.

Hitherto, the industry has been able to sell products with high distribution charges which many advisors had every incentive to recommend without much regard for the customer's best interests. It will be different in the future. Digital disruption reinforcing the pressure for higher fiduciary standards will help with the propagation of a much-needed focus towards goals-based solutions.

If the new solutions are included in academic curricula, much progress could be made. Widespread industry applications of Modern Portfolio Theory took off with the vast increase in the number of alumni learning about it at universities. This year, the CFA Institute announced the introduction of fintech and robo advisory elements into their syllabus, remarkably within a few short years of their widespread usage. Should the CFA and other organisations incorporate the new approaches in their examinations, a major breakthrough becomes more likely.

Capital markets can also make important contributions. Some of the finer detail in the goals-based and mass customisation solutions ideally need sophisticated instruments such as zero-coupon bonds, indexed vehicles and particular types of options. Even many developed markets in various countries do not have access to these and the two papers certainly have limited applicability in many emerging financial markets worldwide.

The new approaches should have universal applicability and financial markets should help to ensure this. If market impediments are removed and innovation is encouraged by governments, even developing markets could leapfrog many of their developed counterparts in using the available sophisticated tools. This can already be seen in Asia with respect to derivatives.

With ageing societies in many countries, inadequate pensions potentially pose social problems. Furthermore, it is widely accepted that long-term savings are in the interests of society. Individual savings and retirement provisions, therefore, are of vital concern to society as a whole and governments need to play a part.

The impact of taxes was pinpointed as a problem, though minor in effect, in the goals-based framework paper and governments can and should do something about this. Pension schemes are already exempt from taxes. Perhaps the same should apply to other long-term savings.

The new robo advisory trend is fast growing. At this stage, it is difficult to predict whether the new proposals will compete with them or come in at the top end as the cream of partly-automated systems for individuals.

The fund management industry is not in very good odour, as many leaders admit. Keith Skeoch, the joint CEO of the new Standard Life Aberdeen group, has written in FTfm that the industry needs to rebuild trust. If the asset managers pick up the baton of implementation following the new ideas, then it could go a long way in repairing its image in society.

This is the fourth of a group of related articles on retirement investing. Please see the previous three on "Retirement Goal-Based Investing", "A Path-Breaking Solution for Individual Investors' Problems" and "Tailor-Made Funds with Economies of Scale" and the next one on "Multi-Asset Products and Solutions".

> If the new solutions are included in academic curricula, much progress could be made… Should the CFA and other organisations incorporate the new approaches in their examinations, a major breakthrough becomes more likely.

References

- Deguest, R., L. Martellini, V. Milhau, A. Suri and H. Wang. 2015. Introducing a Comprehensive Investment Framework for Goals-Based Wealth Management. EDHEC-Risk Institute Publication (March).
- Martellini, L. and V. Milhau. 2017. Mass Customisation versus Mass Production in Retirement Investment Management: Addressing a Tough Engineering Problem. EDHEC-Risk Institute Publication (May).
- Rutter Pooley, C. 2017. CFA revamps examinations to cover big data and robo-advice. Financial Times (11 May).
- Skeoch, K. 2017. The Investment Industry Needs to Rebuild Trust. FTfm. (03 July).

MULTI-ASSET PRODUCTS AND SOLUTIONS

Vincent Milhau, Research Director, *EDHEC-Risk Institute*

Within the old paradigm, the asset management industry had long focused on the design of products with attractive performance, in a context where the distinction between active management with high fees and the passive replication of a benchmark at more moderate fees was a structuring frontier. At least since the turn of the millennium, profound changes have been taking place, which have led to the value proposal for investment management being entirely rethought. The first of the changes is the increasing interest from investors for "smart beta" products, which give them access to investable risk premia at much lower costs than actively managed funds. The factor investing approach is an important aspect of this trend, because these premia can be interpreted as compensation for bearing exposure to undiversifiable risk factors. The second paradigm change is perhaps even more important and consists in the recognition that investors, be they individuals or big institutions, need dedicated (customised or mass-customised) *solutions* to achieve their *goals*, as opposed to off-the-shelf products.

> … since the turn of the millennium, profound changes have been taking place, which have led to the value proposal for investment management being entirely rethought

Modern portfolio theory can serve as a useful guide towards the design of meaningful solutions, by teaching us what the right building blocks are, and how they should be combined over time to improve investor welfare. In practice, it will be needed, of course, to deviate from theoretical optimal portfolios to satisfy implementation constraints, but at least the key desirable properties of the building blocks and the allocation rule can be accounted for when defining implementable solutions. In particular, fund separation theorems state that optimal strategies combine at least two blocks: a performance-seeking portfolio with the highest possible Sharpe ratio, and a liability or goal-hedging portfolio that should track the present value of an investor's liabilities or goal. This distinction has important implications when it comes to a better understanding of multi-asset solutions, since it implies that good risk-adjusted performance is only one possible objective – admittedly, one that is far from obvious to reach – for such multi-asset products, and the focus of a multi-asset solution should extend beyond this simple objective.

Scientific Diversification in Performance-Seeking Portfolios

The maximum Sharpe ratio (MSR) portfolio, defined as the portfolio that maximises the expected excess return over the risk-free rate per unit of risk taken, is a fundamental tenet of modern portfolio theory. It appears in (i) the two-fund separation theorem of Sharpe's (1964) Capital Asset Pricing Model, (ii) dynamic portfolio strategies designed to strike the best balance between expected return and risk (Merton, 1973), (iii) optimal liability-driven investing strategies for pension funds or insurance companies (Martellini and Milhau, 2012), and (iv) goal-based strategies that maximise the probability of reaching a target wealth level (Browne, 1999).

While conceptually straightforward, the prescription to hold the MSR portfolio has proved extremely difficult to implement in practice, because of the need to estimate expected returns and covariances and the dramatic impact of estimation errors on out-of-sample performance. Estimation risk is so large that it can offset the expected benefits of scientific diversification, and the "naive" diversification rule that weights all constituents equally proves to be a not-so-easy-to-outperform benchmark (DeMiguel, Garlappi and Uppal, 2009). In view of this problem, it is tempting to use a portfolio construction technique that does not rely on expected returns, which are more the most difficult parameters to estimate.

The global minimum variance (GMV) portfolio is attractive from that perspective because it is the only point of the efficient frontier that depends solely on covariances, and it can be justified as an MSR portfolio under the agnostic assumption that all expected returns are equal. Moreover, in asset allocation, the number of constituents is typically small, so the sample size is sufficiently greater than the universe size to make sample risk in the covariance matrix limited (Kan and Zhou, 2007). Were constituents too numerous with respect to available samples, efficient estimation procedures, such as statistical shrinkage, are available to obtain a more robust parameter estimate. However, it should be noted that even with a perfectly estimated covariance matrix, a GMV portfolio has no reason to deliver an attractive Sharpe ratio because its objective is to minimise risk, regardless of the opportunity cost in terms of performance. Moreover, it tends to be concentrated in the least volatile asset classes, like Government bonds, so it does not accurately reflect an intuitive idea of what a "well-diversified portfolio" should look like.

Risk and Factor Allocation

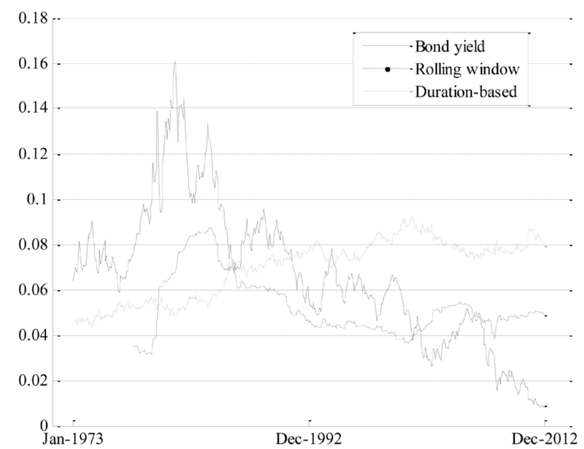
To construct a well-diversified portfolio, one may go back to conventional wisdom, which advocates to "spread eggs across many baskets". But what are eggs and baskets in asset allocation? The equal-weighting rule answers that eggs are dollars and baskets are constituents. But it is well known that a 50%-50% stock-bond portfolio, while being perfectly diversified in terms of dollar contributions, is poorly diversified in terms of risk contributions because its volatility is largely dominated by equity risk, which is by far the most volatile asset class. Risk parity has become a popular approach to correct for such imbalances since its goal is to ensure that all constituents (baskets) have the same contributions to risk (eggs) in the portfolio. In particular, this avoids excessive concentration in assets with low volatilities, even though such constituents still tend to have the largest weights.

This methodology, however, has its own limits. The first is inherent to the use of volatility as a risk measure: while volatility is symmetric and equally penalises upside potential and downside risk, investors are clearly more averse to downside risk than to uncertainty on the upside. As a result, standard forms of multi-class risk parity portfolios tend to overweight bonds, even in an environment where interest rates are expected to go up from historically low values to higher levels and bonds have a high downside risk. The adaptation of risk parity portfolios to interest rate conditions is the focus of a research paper by Martellini, Milhau and Tarelli (2015), supported by Lyxor in the context of the "Risk Allocation Solutions" research chair at EDHEC-Risk Institute. In this paper, we argue that a "duration-based volatility" measure for bonds, equal to current duration multiplied by interest rate volatility, more adequately captures increasing risk as interest rates decrease (see Exhibit 1). Extensions of the risk parity methodology to non-symmetric risk measures, such as semi-variance (the variance restricted to negative returns), value-at-risk or expected shortfall, are also discussed in this paper and in a related effort by Roncalli (2013). Users of these latter methods should be aware that they bear the risk of estimation errors in expected returns (required for downside risk estimates), but not as much as by attempting to estimate the MSR portfolio, because under suitable mathematical conditions given by Roncalli (2013), the extended risk parity portfolio is long-only, unlike the MSR portfolio, which can completely rule out assets with perceived unattractive risk-return profiles.

The second limitation of the risk parity approach is that by focusing on the contributions of constituents, it completely disregards the underlying factors that explain the risk of constituents, which can lead to a misleading picture for correlated constituents with a strong common factor. For instance, an equally-weighted portfolio of two bonds with similar durations spreads dollars evenly, and is almost at risk parity, because the two volatilities are close to one another. Yet, the risk of this portfolio is related to a single factor bet, which is a bet on the level of interest rates. To better assess the extent of diversification in a multi-class portfolio, Carli, Deguest and Martellini (2014) propose to look through the constituents and to consider the risk contributions of underlying systematic factors. Implicit factors, which are extracted by analysing the returns of the securities to explain, are well suited for this purpose because they are uncorrelated, and they exhaust by definition all uncertainty over asset returns, in contrast to macroeconomic factors, which generally have low explanatory power. The *effective number of uncorrelated bets* that the authors use was introduced in Deguest, Martellini and Meucci (2014) to quantify the degree of diversification as a quantitative measure of the dispersion of risk contributions (eggs) across factors (baskets). In the previous example, it would be close to 1 because portfolio risk is almost entirely explained by the interest rate level factor.

The factor perspective in the construction of well-diversified portfolios is gaining popularity because it is theoretically supported by asset pricing models, and because it makes intuitive sense. As Ang (2014, p. 195) puts it, "factors are to assets what nutrients are to food", so it is important for investors to be informed about what an appropriate diet is before deciding how to collect these nutrients. Moreover, factors go a long way towards explaining returns on long-only actively managed portfolios, as Ang, Brandt and Denison (2014) show in their study of the Norwegian sovereign wealth fund. Smart beta products, which have been largely developed in the equity class, are investment vehicles that allow factor premia

Exhibit 1: Yield-to-redemption and volatility estimates for a bond index (January 1973 – December 2012).



Source: Martellini, Milhau and Tarelli (2015).

to be collected at low costs and have made the distinction between active and passive management less clear (see the contribution on smart beta in the equity space in this special issue).

From Investment Products to Investment Solutions

A well-diversified "risky" performance-seeking portfolio is only one piece of the asset allocation process, albeit an important one. A multi-asset multi-factor solution should also involve a "safe" building block, as well as a suitably chosen allocation strategy to the risky and safe building blocks. The ultimate objective is to offer a high probability of reaching the investor's aspirational goals while securing their essential goals, defined as minimum levels of wealth or consumption. Clearly, such a value proposition is very different from delivering a high alpha or outperforming a given benchmark, and this is precisely why investment solutions cannot be based on diversification principles alone.

The second important component for a meaningful solution is a safe liability- or goal-hedging portfolio (LHP/GHP), the role of which is to replicate the present value of a process that characterises an investor's objectives which are summarised in terms of their liabilities in an institutional money management context, or their goals in an individual money management context. For an individual saving for retirement, this process is given by the discounted value of the future expected replacement income cash flows. To construct a LHP/GHP, a factor approach is also helpful, provided the risk factors that impact the liability value are identified. Interest rates are almost always a major source of risk due to the discounting mechanism, but inflation risk is also often present in situations where the target wealth level or income stream are inflation-linked. Thus, real bonds or nominal bonds are typically important ingredients of the LHP/GHP. Duration and convexity hedging, which are traditionally employed to hedge interest rate risk, are examples of factor-matching techniques used in this context.

Other building blocks may be necessary, as posited by Merton's (1973) fund separation theorem: when investment opportunities (expected returns and volatilities) vary over time, "hedging portfolios" are required, which enjoy higher returns when market conditions get worse. For instance, bonds perform well when interest rates decrease, and stocks tend to have higher returns when the prospective equity premium is lower.

The final step in the design of investment solutions is the suitable allocation to the blocks. If the investor's objective is to achieve the best trade-off between expected return and risk without any explicit target wealth or consumption level, Merton (1973) shows that the optimal allocation not only depends on the investor's horizon and risk aversion, but also on the current values of risk premia. Target-date funds with a deterministic glide path are not a satisfactory implementation of these prescriptions, if only because they do not react to changes in market conditions (see the contribution on improved forms of target date funds in this special issue). Research has shown that enhanced long-term investing strategies based on these principles and satisfying realistic implementation constraints can be designed, which lead to higher welfare (Martellini and Milhau, 2010).

When the investor explicitly states wealth or consumption goals, meaningful investment solutions should maximise the probabilities of reaching them in favourable scenarios, and/or minimise the expected shortfall in those scenarios where market conditions do not make it possible for the goals to be reached. This multi-asset investment framework requires the use of risk-budgeting techniques, in which the allocation to the various building blocks is a function of the current distance to the goals. These methods have long been employed in portfolio

> **The ultimate objective is to offer a high probability of reaching the investor's aspirational goals while securing their essential goals…**

insurance, but they should also be applied in a goal-based investing context. An extensive review of these strategies is done by Deguest et al. (2015) in a research paper supported by Merrill Lynch Wealth Management, and the contribution by Kévin Giron and Lionel Martellini on retirement goal indices in this issue describes an application to the retirement investing problem.

As a conclusion, investors do need multi-factor products designed to efficiently harvest risk premia across and within asset classes to generate attractive risk-adjusted performance. That said, they also need hedging building blocks and dynamic investment solutions products that efficiently allocate to the various blocks in such a way as to deliver high probabilities of reaching their meaningful goals.

This is the fifth of a group of related articles on retirement investing. Please see the previous four on "Retirement Goal-Based Investing", "A Path-Breaking Solution for Individual Investors' problems", "Tailor-Made Funds with Economies of Scale", and "Individual Investment Processes for the 21st Century".

References

- Ang, A. 2014. Asset Management: A Systematic Approach to Factor Investing. Oxford University Press.
- Ang, A., M. Brandt and D. Denison. 2014. Review of the Active Management of the Norwegian Government Pension Fund Global. Technical Report.
- Browne, S. 1999. Reaching Goals by a Deadline: Digital Options and Continuous-Time Active Portfolio Management. *Advances in Applied Probability* 31(2): 551-577.
- Carli, T., R. Deguest and L. Martellini. 2014. Improved Risk Reporting with Factor-Based Diversification Measures. EDHEC-Risk Institute Publication.
- Deguest, R., L. Martellini and A. Meucci. 2014. Risk Parity and Beyond – From Asset Allocation to Risk Allocation Decisions. EDHEC-Risk Institute Working Paper.
- Deguest, R., L. Martellini, V. Milhau, A. Suri and H. Wang. 2015. Introducing a Comprehensive Risk Allocation Framework for Goals-Based Wealth Management. EDHEC-Risk Institute Publication.
- DeMiguel, V., L. Garlappi and R. Uppal. 2009. Optimal versus naive diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies* 22(5): 1915-1953.
- Kan, R. and G. Zhou. 2007. Optimal Portfolio Choice with Parameter Uncertainty. *Journal of Financial and Quantitative Analysis* 42(3): 621-655.
- Martellini, L. and V. Milhau. 2012. Dynamic Allocation Decisions in the Presence of Funding Ratio Constraints. *Journal of Pension Economics and Finance* 11(4): 549-580.
- Martellini, L. and V. Milhau. 2010. Towards the Design of Improved Forms of Target-Date Funds. Bankers, *Markets and Investors* 109: 6-24.
- Martellini, L., V. Milhau and A. Tarelli. 2015. Towards Conditional Risk Parity – Improving Risk Budgeting Techniques in Changing Economic Environments. *Journal of Alternative Investments* 24(3): 5-29.
- Merton, R. 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica* 41(5): 867-887.
- Roncalli, T. 2013. Introducing Expected Returns in Risk Parity Portfolios: A New Framework for Strategic and Tactical Asset Allocation. SSRN working paper.
- Sharpe, W. F. 1964. Capital Asset Prices – A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance* 19(3): 425-442.

EQUITY PORTFOLIOS WITH LIABILITY-HEDGING BENEFITS

Lionel Martellini, Professor of Finance, *EDHEC Business School* and Director, *EDHEC-Risk Institute*

Vincent Milhau, Research Director, *EDHEC-Risk Institute*

Liability-Driven Investing and Beyond: Fund Separation versus Fund Interaction Theorems

Asset-liability management (ALM) for pension funds has become relatively straightforward, in principle, within the paradigm known as liability-driven investing (LDI). In a nutshell, when extended to ALM, modern portfolio theory and the *fund separation theorem* unambiguously advocate that pension plans should implement the suitable combination of a liability-hedging portfolio (LHP) invested in fixed-income securities and aiming to match the risk factors impacting the value of their liabilities as well as possible, and a performance-seeking portfolio (PSP) aiming to efficiently harvest risk premia across and within risky asset classes, and most importantly in equity markets around the globe.

When a pension fund is underfunded, pension assets are by definition insufficient to cover the liabilities, but the pension fund may in principle optimally borrow the required amount to make up for the gap between pension assets and pension liabilities and also maintain a levered investment in performance-seeking assets which may contribute to solving the funding problem without requiring exceedingly high levels of additional contributions.

While this clear separation between the search for performance and the desire to hedge liabilities is perfectly intuitive and sensible in theory, it suffers from a number of limitations in terms of real-world implementation. The main limitation is undoubtedly the presence of leverage constraints, which implies that most underfunded pension funds cannot use as much leverage as would be required to fully hedge their liabilities. In practice, pension funds end up investing all their assets in a zero- or low-leverage portfolio mostly containing stocks and bonds, with a key trade-off between a dominant allocation to equities (say a 60:40 stock-bond split), which generates attractive levels of expected returns but also implies high levels of funding ratio volatility, or a more moderate equity allocation (say a 40:60 stock-bond split), which requires lower ALM risk budgets but correspondingly also generates lower upside potential.

In this context, the question arises whether it would make sense for a pension fund to hold a customised equity portfolio engineered to exhibit enhanced liability-hedging properties versus holding an off-the-shelf broad equity index. Intuition indeed suggests that a better alignment of the PSP with respect to the liabilities would lead to an increased allocation to stocks for the same level of volatility of the funding ratio, which in turn would generate higher access to the equity risk premium.

In a recent paper (Coqueret, Milhau and Martellini, 2016)[1], we analyse whether LDI solutions can be enhanced by the design of performance-seeking equity benchmarks with improved liability-hedging properties. We confirm this intuition and show that improving hedging characteristics of the performance portfolio generates welfare gains unless this improvement comes at an exceedingly large opportunity cost in terms of performance, a result that we call the *fund interaction theorem.* While two competing effects exist in principle (a better alignment of the equity portfolio with the liabilities leads to a higher allocation to equities for the same ALM risk budget due to enhanced liability-friendliness, but it may also lead to a lower reward per dollar invested compared to a pure focus on performance), our empirical analysis actually suggests that the selection of stocks with above-average liability-hedging properties leads to both a higher degree of liability-friendliness (as expected) and also to better performance due to increased exposure to rewarded factor tilts.

In this context, we find that very substantial increases in investor welfare would result from switching from a standard off-the-shelf cap-weighted (CW) equity benchmark to an equity benchmark designed to exhibit above-average liability-hedging properties. For inflation-linked liabilities, we find that the

1 - Coqueret, G., L. Martellini and V. Milhau. 2017. Equity Portfolios with Improved Liability-Hedging Benefits. *Journal of Portfolio Management* 43(2): 37-49.

use of a minimum variance equity benchmark based on a double-sort procedure of stocks according to (high) dividend yield and (low) volatility would have generated, over the period 1999-2012, an annualised excess return reaching 270 basis points for the same funding ratio volatility, as well as a lower funding ratio drawdown, compared to what is obtained with the use of the standard cap-weighted S&P 500 index as a benchmark.

Equity Benchmarks with Improved Liability-Friendliness

We consider two alternative approaches to the definition of liability-friendliness. The first one is based on *cash-flow matching* capability: under this definition, liability hedging aims to find securities with dividend payments that match the pension payments as closely as possible. The stocks which are expected to display above-average liability-friendliness in terms of cash-flow matching capacity are those that generate large and stable dividend yields.

The second definition is based on *factor exposure matching*. Since perfect cash-flow replication is typically difficult to achieve in practice, investors who need to hedge liabilities may instead choose to match the risk factor exposures of their assets with those of their liabilities. The objective pursued in this case is to immunise the funding ratio against variations in the risk factors that impact liabilities, and the success is measured in terms of tracking error with the liability proxy.

In this setting with a focus on risk factor matching, a stock will be said to be liability-friendly if the tracking error of the stock returns with respect to the returns on the liability proxy is low. Given the decomposition of the tracking error into two components (one that is related to the portfolio volatility and one that is related to the portfolio correlation with the liability proxy), a low tracking error can be achieved if the volatility of the stock is low and/or if the correlation between the stock and the liability proxy is high.

Using data from the CRSP database from 1975-2012, we construct portfolios with stocks originating from the S&P 500 universe. We cast the analysis at the individual stock level, as opposed to the sector level, given the expected presence of very substantial levels of cross-sectional dispersion in interest rate hedging benefits across individual stocks. The portfolios are rebalanced every year in March. In the analysis, the liability proxy is computed as a constant maturity bond and its returns are computed using 15Y US Treasury yields. The second step of the procedure establishes the weights that are assigned to each stock. We start by considering equal weights for all stocks (no selection EW) to assess the benefits of the selection stage, and we additionally provide the results for the cap-weighted portfolio of all stocks (no selection CW), which is the commonly used benchmark.

We find that the various selection procedures indeed deliver what they are designed for. In particular, the equally-weighted (EW) portfolio of the 20% of stocks with the lowest volatilities has a tracking error of 14.6% with respect to our liability proxy over the sample period, while the EW portfolio of the 20% of stocks with the highest volatilities is almost twice as large at 27.8%. This spectacular improvement in tracking error does not only emanate from a lower portfolio volatility; it is also linked so a strong increase in correlation with the liabilities. Hence, the selection of low volatility stocks generates a positive 7.7% correlation with the liability proxy, while a selection of high volatility stocks generates a negative correlation of -6.7%. Intuitively, this improvement can be traced down to the fact that low volatility stocks, which tend to be low dividend uncertainty stocks, are the stocks that tend to be the closest approximations to fixed-income securities, and as a result, the best approximation to bond-like liabilities. In terms of correlations, the high correlation selection ranks only second (although close to first), with a large turnover, suggesting that empirical correlations are highly unstable. We further observe that all selections increase the Sharpe ratio as well as the turnover, compared to both the EW and CW benchmarks, and the increased liability-friendliness of the portfolios is therefore not penalised by lower risk-adjusted performance. We also confirm that the selection on dividend yields generates a statistically and economically significant increase in this dimension with respect to the use of the standard S&P 500 index as a benchmark.

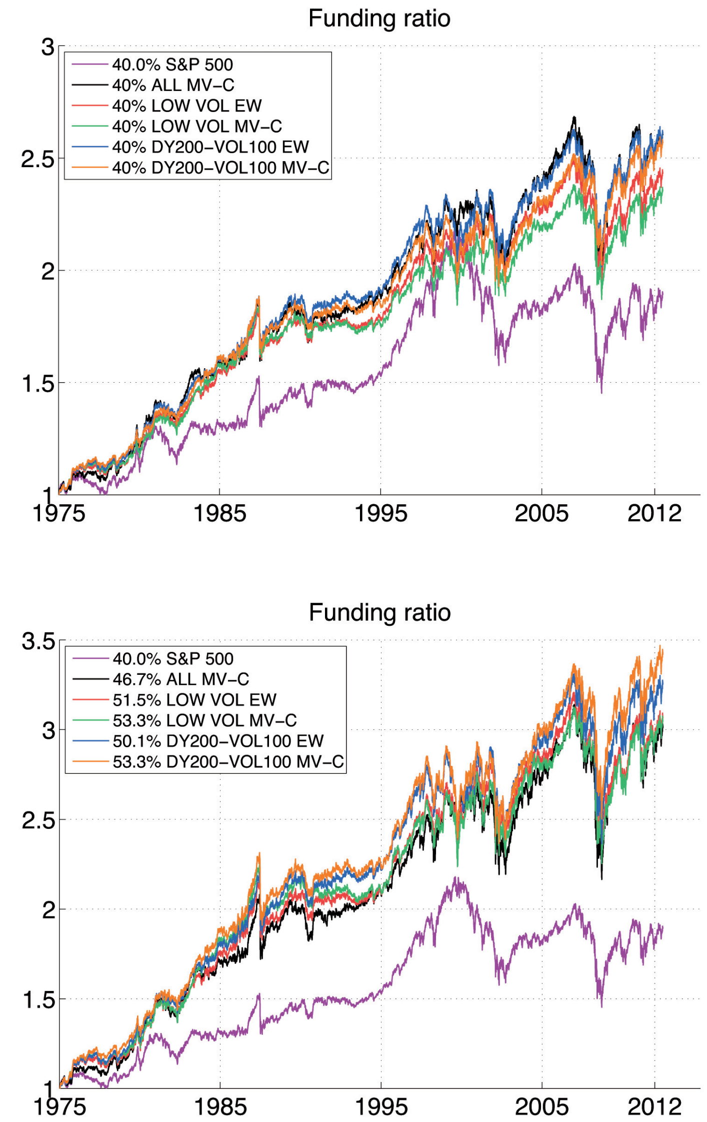
Addressing the focus on liability hedging through a double-sort procedure, starting with the 200 highest dividend yield (DY) stocks, selecting the 100 lowest volatility stocks amongst them, and subsequently performing a minimum variance optimisation, is found to lead to further improvements in the liability-friendliness of the selected portfolios.

Hence, combining the double-sort selection procedure with the minimum-variance weighting scheme with norm constraints (MV-C), leads to improvements in all indicators with respect to the base case results, and reaches the following attractive levels – 14.1% tracking error, 12.5% volatility, 8.2% correlation and 5.4% average dividend yield. Overall, double sorts starting with DY and then low volatility generate comparable levels of factor matching liability-friendliness (tracking error at 14.1%) with improved cash-flow matching properties (average DY at 5.40% compared to selection purely based on volatility. The Sharpe ratio further increases to 0.79, even though this comes at the cost of a higher turnover, which should deserve dedicated attention at the implementation stage.

Measuring the Impact on Investor Welfare

Due to the resulting improvement in liability-hedging benefits, liability-driven investors can allocate a higher fraction of their portfolios to equities without a corresponding increase in funding ratio volatility. For example, we find that a pension fund allocating 40% to equities on the basis of a cap-weighted equity benchmark can allocate as much as 53.3% to a minimum variance portfolio of selected stocks from the aforementioned double-sort procedure for the same volatility of the funding ratio. This substantial increase in equity allocation without a corresponding

Exhibit 1 – Historical Trajectories for the Funding Ratio



increase in ALM risk budgets confirms that the aforementioned improvements obtained in terms of improved liability-friendliness are economically significant.

The resulting increase in equity allocation for the same ALM risk budget, combined with an improved risk-adjusted performance of the dedicated equity benchmark with respect to the S&P500 index, leads to an improvement in performance, reaching 158 basis points annualised over the 1975-2012 sample period. This improvement can be decomposed into a contribution purely emanating from the increase in equity allocation assuming no impact on performance (39 basis points) and a contribution purely emanating from the improved performance of the equity benchmark assuming no increase in allocation (119 basis points).

In terms of historical trajectories, we plot the evolution of the funding ratio over the sample period assuming an initial funding ratio normalised at 100%, and compute the corresponding risk and return indicators.

We find that LDI strategies based on all improved liability-friendly portfolios strongly outperform LDI strategies based on the S&P 500 over the sample period, and Exhibit 1 shows that the outperformance is even more spectacular when the allocation to the improved equity benchmark is adjusted to generate the same volatility of the funding ratio as when investing 40% in the S&P 500 index. Moreover, the volatility of the funding ratio lies between 1.1% and 1.9% lower in absolute value when the equity benchmark is the liability-friendly portfolio compared to the use of the S&P 500.

With the exception of the MV-C portfolio of all stocks, the reduction of the maximum drawdown reaches at least 10% in absolute value, or 30% in relative value.

Furthermore, we observe that even after controlling for the volatility LDI strategies with liability friendly portfolios dominate those with S&P500 in terms of extreme risk measures (funding ratio maximum drawdown).

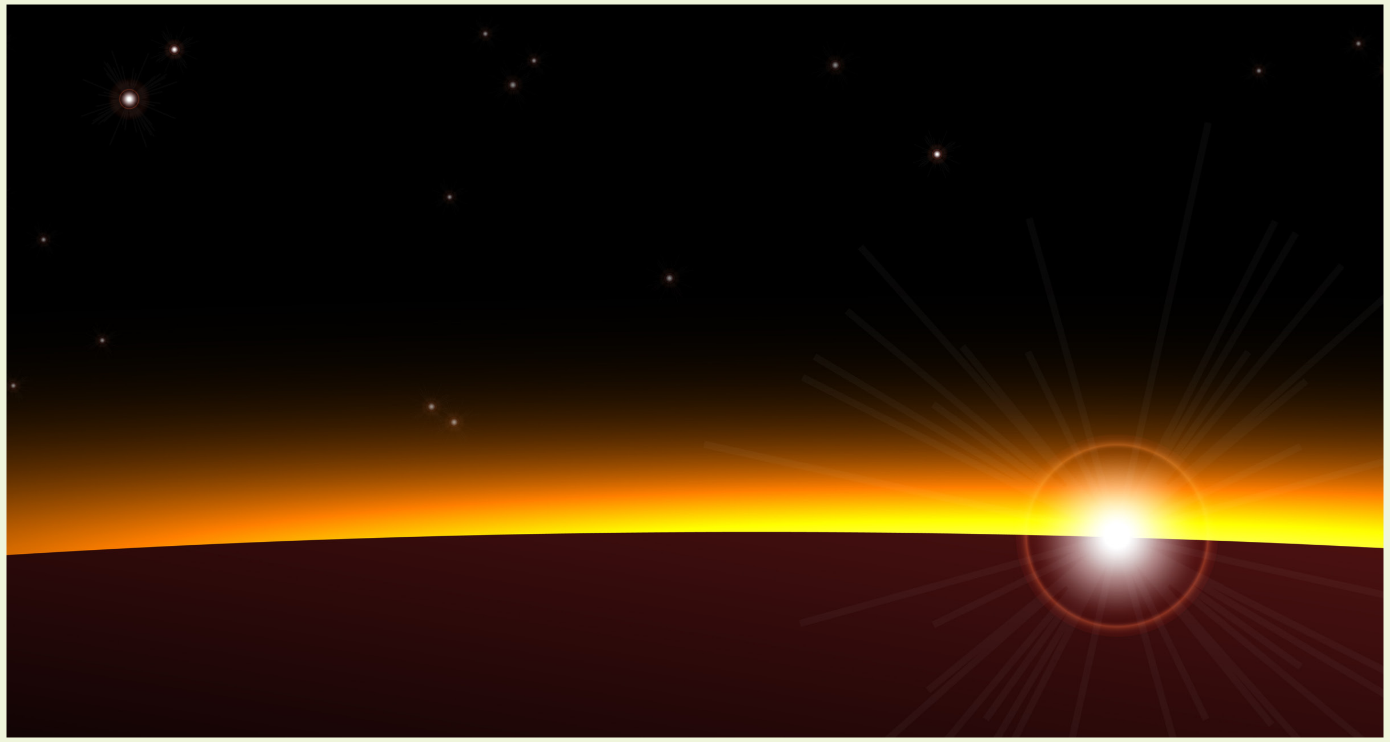
Increasing Equity Allocation without a Corresponding Increase in Risk Budgets

The LDI paradigm advocates that pension plans divide their investments between fixed-income securities matching the investor's liabilities, and a riskier performance-seeking portfolio that is heavily invested in global equities. However, asset managers can actually provide more value to LDI clients by incorporating low volatility and high-yielding equities as part of a broader LDI solution.

While separating performance from liability matching makes theoretical sense, in practice, pension funds cannot use as much leverage as necessary to fully hedge their liabilities. If a manager places the majority of a pension fund's money into equities, the portfolio could run the risk of high levels of short-term funding volatility. On the other hand, a more moderate equity allocation would require lower risk budgets but less upside potential. For underfunded pension funds, this type of allocation would have a lower chance of solving the funding problem without substantial levels of additional contributions.

We have found that it is possible to construct a customised "liability-friendly" equity portfolio with better liability-hedging properties than an off-the-shelf broad equity index. Customisation can allow for an increased allocation to stocks for the same level of funding ratio volatility, which could result in higher overall performance, barring the improvement in liability-hedging benefits coming at an exceedingly large opportunity cost in terms of risk-adjusted performance.

> **"…it is possible to construct a customised "liability-friendly" equity portfolio with better liability-hedging properties than an off-the-shelf broad equity index.**



BOND RISK PREMIA: THE NEW FRONTIER IN FACTOR INVESTING AND SMART BETA

Frank Fabozzi, Professor of Finance, *EDHEC Business School*

Lionel Martellini, Professor of Finance, *EDHEC Business School* and Director, *EDHEC-Risk Institute*

Riccardo Rebonato, Professor of Finance, *EDHEC Business School*

Bond Risk Premia: The New Frontier in Factor Investing and Smart Beta

The abundance of theoretical and empirical research on factor investing in the equity universe stands in sharp contrast to the relative scarcity of research about how to efficiently harvest risk premia in bond markets. That relatively little is known about the out-of-sample performance of factor-based bond portfolio optimisation models is perhaps surprising given the importance of fixed-income investments within the portfolios of institutional and private investors.

From the investment practice standpoint, a similar contrast actually exists between factor investing in the equity space (which is a relatively mature subject) and factor investing in bond markets (which is still is in its infancy). It is indeed fair to say that the so-called smart beta approach is now firmly grounded in equity investment practices, and the key question for an increasing majority of institutional investors is not whether one should use smart beta, but rather which and how much smart beta to use. In parallel, interest in smart beta equity products is rapidly growing in retail and private wealth management. In contrast, the concept of smart beta in the fixed-income space is still relatively less mature, despite the obvious importance and relevance of the subject.

> “… the concept of smart beta in the fixed-income space is still relatively less mature, despite the obvious importance and relevance of the subject.

One first possible explanation behind the scarcity of the research on risk premia harvesting in bond markets and related investable solutions is that bonds are often held as part of investors' hedging portfolios, where the focus is on matching interest rate risk factor exposures on the asset side to interest rate risk factor exposures on the liability side, as opposed to risk/reward ratio maximisation. This is, however, not a sufficient reason for ignoring the need to generate attractive risk-adjusted performance. After all, there are infinitely many bond portfolios with a given target duration, and selecting the one with the highest risk-adjusted performance should intuitively improve investor welfare.

Another possible explanation is that bond portfolio construction models pose a number of technical and implementation challenges that do not exist in the equity universe. In particular, individual bonds, unlike stocks or constant maturity bond indices, have a finite maturity, which imposes a very specific structure on the covariance matrix and expected return vector. Moreover, no-arbitrage relationships exist between bonds of various maturities and these relationships impose a set of constraints on risk/return and factor exposure parameter estimators, for which there is no equivalent in equity portfolios.

Additional implementation hurdles include the lack of reliable market prices for bonds

> “… individual bonds, unlike stocks or constant maturity bond indices, have a finite maturity, which imposes a very specific structure on the covariance matrix and expected return vector.

that typically trade in thin markets and the concern over high implementation costs.

In what follows, we provide a brief review of the theoretical, empirical and practical challenges related to factor investing in the fixed-income space. The modern approach to factor investing actually suggests that we should first identify robust and economically-motivated sources of risk in fixed-income markets, select securities on the basis of the desired factor tilts, and then apply an efficient weighting scheme aiming to enforce the highest level of diversification to eliminate as much unrewarded risk as possible from investor portfolios. In this context, our ambition is to discuss each one of these aspects, with applicability to various segments of fixed-income markets, most importantly sovereign and corporate bond markets.[1]

Limits of Traditional Bond Benchmarks

Over recent years, a number of concerns have been expressed about the (ir)relevance of existing forms of corporate and sovereign bond indices offered by index providers, many of which are simply not investable in the first place. Beyond the lack of investability, there are two major problems with existing bond benchmarks, namely (1) an excessive concentration and lack of proper diversification and (2) a lack of stability and control over the underlying set of factor exposures.

Lack of Diversification in Sovereign and Corporate Bond Indices

Lack of diversification is the first major problem with bond indices that simply weight the debt issues by their market value. Given the large share of the total debt market accounted for by issuers with large amounts of outstanding debt, market-value-weighted corporate bond indices will have a tendency to overweight bonds with large amounts of outstanding debt. This weighting scheme will inevitably lead to excessive concentration and lack of diversification. The issuer concentration in government-bond indices is very undesirable from the perspective of the construction of a well-diversified portfolio: Staal et al. (2015), for instance, document that 55% of the Barclays Treasury Index (which is supposed to refer to 38 different countries) is made up of US and Japanese bonds, and that the same two bonds account for over 65% of the total risk of the index.

While the claim that over-represented issuers may give rise to adverse selection and poor diversification is compelling, simplistic rules to correct for issue size are likely to be ineffective, since it is not size in itself that is 'dangerous', but size relative to ability to repay and to refinance debt (see the next section for a more detailed discussion of *fundamentally-weighted* indices). So, for example, a correction of the market-capitalisation benchmarking rule based on the naive comparison of the size of debt outstanding for US Treasuries or Greek government debt would do little to rebalance an over-representation of supposedly riskier bonds; nor would looking at the amount of debt outstanding as a fraction of GDP do much to redress re-financing risk in the case of the government debt of Italy and Japan, two markets which are predominantly financed by foreign and domestic investors, respectively.

Moreover, it is often argued that such indices not only give too much weight to *some* assets but also that they give too much weight to *riskier* assets, a problem that has been dubbed the "bums' problem" (Siegel, 2003). This concern may be particularly relevant in the corporate bond market if, during periods of sector-specific speculative bubbles, some industry sectors (such as tech stocks in the run-up to 2001 or financial stocks in the run-up to 2006) found it 'excessively' easy to issue cheap debt, thereby becoming over-represented in a market-capitalisation benchmark.

While it is actually debatable whether debt weighting really leads to the most risky securities being over-weighted[2], it is clear that market-value debt-weighted indices leads to concentrated portfolios that are in opposition with investors' needs for efficient risk premia harvesting, which involves holding well-diversified portfolios. In a nutshell, a good case can be made that existing bond indices tend to be poorly diversified portfolios, regardless of whether or not the over-weighting applies to the wrong constituents.

1 - In principle, some of the insights discussed in this paper can also be of relevance to other fixed-income markets such as MBS markets, swap markets, etc.

2 - A higher weight for an issuer with a high market value of debt does not necessarily mean that the index is over-weighting issuers with a high face value of debt. An issuer with a high amount of par value debt outstanding will only get a high weight if the market value is relatively close to par value, which implies that the issuer is not perceived to be very risky. It is therefore not clear why the market value-weighted index should become riskier. In addition, loading onto riskier issuers should not necessarily be a problem if this risk is rewarded by higher expected returns.

Lack of Stability and Control in Factor Exposures

In addition to the problem of concentration, fluctuations in risks' exposure (such as duration or credit risk in existing indices) are another source of concern – see Campani and Goltz (2011) for more detail. Such uncontrolled time variation in risk exposures is incompatible with the requirements of investors that these risk exposures be relatively stable so that allocation decisions are not compromised by implicit choices made by an unstable index. For example, an asset-liability mismatch would be generated by changes in the duration of the bond index if the latter were used as a benchmark for a pension fund bond portfolio.

More generally, it appears that existing bond indices can be regarded as more "issuer-friendly" than "investor-friendly", in the sense that these bond indices passively reflect the collective decisions of issuers regarding the maturity and size of bond issues. Bond issuers make a decision on the duration of the bonds that they issue, the decision being based on minimising funding costs for the issuer. When investors make duration decisions, they are based on optimising an investor's investment objectives. There is no reason to believe that there will be any consistency between the cost minimisation objective of issuers and the investment optimisation objective of investors in the determination of market duration. As a result, the duration that is obtained from an index can be considered as a "historical accident" (Siegel, 2003).

> … existing bond indices can be regarded as more "issuer-friendly" than "investor-friendly", in the sense that these bond indices passively reflect the collective decisions of issuers regarding the maturity and size of bond issues.

As is the case for equity indices, the lack of diversification of unrewarded specific risk exposures, and the lack of explicit control over systematic rewarded risk exposures, are two distinct major flaws in bond indices if they are to be used as benchmarks for investors' portfolios. A number of index providers or asset managers have launched new forms of smart beta bond portfolio indices to try to address some of the challenges with traditional weighting schemes based on the market value of debt.

Review of First Generation Smart Beta Initiatives in Bond Markets

Here, we provide a broad overview of first generation initiatives, which can be broadly classified in two different categories that have historically focused on fixing the concentration problem – fundamental approaches and diversification approaches. We then move on to an analysis of more recent factor-based approaches, which focus on stability in factor exposure problem.

Fundamental indexing in the bond market is a direct transfer of methodologies originally developed for equities. In the methodology used in Arnott et al. (2010) for corporate bond indices, the following five factors are used to assign a score to each (investment grade and high yield) corporate bond: book value of assets, total dividends, total cash flow, sales and face value of the debt issue. First, weights are computed for each corporate and with respect to each factor, by using the trailing five-year average of each of the above metrics over the aggregate five-year average across all corporations. While it might seem unclear why it would be desirable to use a 5-year trailing value as opposed to the current value for the fundamentals, one practical advantage of this procedure is that the implied smoothing leads to a reduction in turnover. Then, the composite measure is obtained by equally weighting four of the measures: assets, dividends, cash flow and sales.

For emerging market sovereign bonds, the approach developed by Arnott et al. (2010) is based on the following factors: total population, square root of land area (as a crude approximation for resources), total gross domestic product, energy consumption and face value of the debt issue. Similarly to the case of corporate bonds, weights are first computed for each country and with respect to each factor, by using smoothed five-year averages of the above-mentioned metrics over the aggregate metric across all countries. Then, a country's aggregate weight is the equally-weighted average of its score on the individual factors.

The fundamental approach for constructing a bond index raises several concerns. First, the methodology used does not address the concern over the stability of factor exposure. Moreover, the problem of concentration is approached with a purely ad-hoc methodology, and better-diversified portfolios could be constructed on the basis of standard risk models. More importantly, it is unclear why some backward-looking trailing average of some arbitrarily selected variables should contain more useful information than say bond ratings, which for all their flaws are based on a much richer information set.

By contrast, the diversification approaches rely on risk and/or return parameter estimates as well as risk-based portfolio optimisation models to design well-diversified bond portfolios. These include minimum concentration (MC) portfolios (which correspond to the closest approximation of an equally-weighted strategy subject to constraints such as duration or weight constraints), global minimum variance (GMV) portfolios and diversified risk parity (DRP) portfolios, also known as factor risk parity portfolios.

One may also use expected return estimates based on economically motivated priors for Sharpe ratio maximisation exercises as in Deguest et al. (2013).

This is the first of a group of related articles on smart beta and factor investing. Please see the next three on "Factor-Based Approaches to the Design of Smart Bond Portfolios", "Factor Investing: Efficient Harvesting of Risk Premia Across and Within Asset Classes" and "Be Serious with Equity Factor Investing!".

References

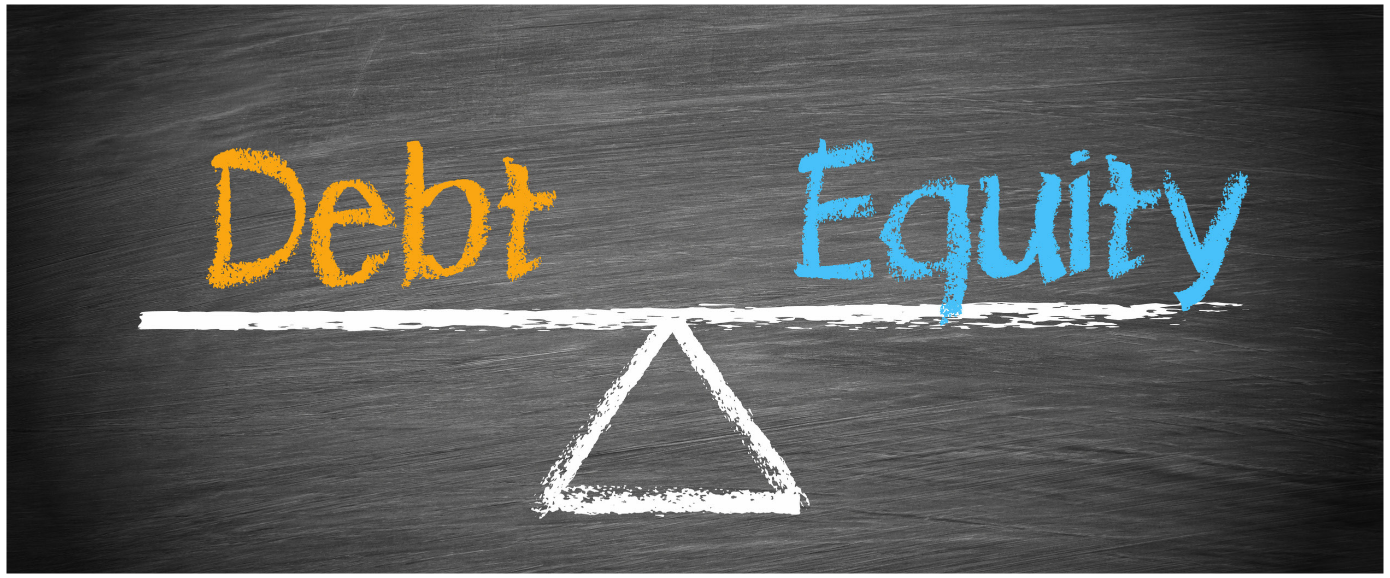
• Arnott, R., J. Hsu, L. Feifei and S. Shane. 2010. Valuation Indifferent Weighting for Bonds. *The Journal of Portfolio Management* 36(3).

• Campani, C. and F. Goltz. 2011. A Review of Corporate Bond Indices: Construction Principles, Return Heterogeneity, and Fluctuations in Risk Exposures. EDHEC-Risk Institute Publication.

• Deguest, R., F. Fabozzi, L. Martellini and V. Milhau. 2013. Bond Portfolio Optimisation in the Presence of Duration Constraint. EDHEC-Risk Institute Publication.

• Siegel, L. 2003. Benchmarks and Investment Management. The Research Foundation of the Association for Investment Management and Research. Charlottesville, Virginia.

• Staal, L., M. Corsi, S. Shores and C. Woida. 2015. A Factor Approach to Smart Beta Development in Smart Beta. *Journal of Index Investing* 6(1).



FACTOR-BASED APPROACHES TO THE DESIGN OF SMART BOND PORTFOLIOS

Frank Fabozzi, Professor of Finance, *EDHEC Business School*

Lionel Martellini, Professor of Finance, *EDHEC Business School* and Director, *EDHEC-Risk Institute*

Riccardo Rebonato, Professor of Finance, *EDHEC Business School*

The modern approach to factor investing (see for example Amenc, Goltz and Martellini (2013) for an application to the equity space) suggests that before applying a weighting scheme, we should first identify robust and economically-motivated sources of differential cross-sectional returns, which can be regarded as 'factors', with rational or behavioural interpretation, or 'frictions'. In this context, more work is required both in academia and in the industry to start addressing these challenges in a careful way, before we are able to see the emergence of improved bond benchmarks that will provide investable answers to investors' needs.

An economic motivation is not just an academic add-on. Understanding the source and origin of the cross-section differential returns matters a lot from the point of view of a robust benchmark creation:

• If the origin of the excess returns can be traced to a source of systematic risk, then the attending compensation (the corresponding 'market price of risk') will not disappear by discovering it, but may decrease or increase over time with variations in the stochastic discount factor (the investors' risk aversion)[1];

• If the excess returns are due to a behavioural finance 'irrationality' it could, in principle, be arbitraged away by rational investors, and revealing the behavioural anomaly could therefore be the first step towards its disappearance. However, its persistence or otherwise may be linked to the availability or scarcity of 'arbitrage capital'[2];

• If the excess returns are due to institutional frictions, they can be an 'easy' source of profitability for investors who are not affected the regulatory or institutional constraints. However, they can disappear at the stroke of a regulatory pen;

• If, finally, the excess returns are truly due to an 'anomaly', then they are likely to disappear after discovery as the anomaly gets exploited (for instance, one of the explanations for the 'disappearance' of the equity size factor is that it was really an anomaly which was readily arbitraged away once identified).

For the sake of brevity, we therefore refer in this article to all these sources of differential cross-sectional returns as 'generalised factors', but it should be kept in mind that, from the perspective of the creation and long-term profitability of investable factor benchmarks, the underlying differences can be very important. It turns out that the recent discovery about generalised factors that used to 'work well' for equities also seem to be effective in the fixed-income space does create an embarrassing explanation problem. As Asness, Moskowitz and Pedersen (2013) eloquently put it, "[t]he strong correlation structure among value and momentum strategies across such diverse asset classes is difficult to reconcile under existing behavioural theories, while the high Sharpe Ratios of a global […] diversified portfolio presents an even more daunting hurdle for rational-risk-based models." So, we seem to find more and more factors and to observe that the 'old' factors seem to work even where they were not expected (or even supposed) to; however, we are farther and farther away from a unified understanding of why they do. As argued above, unfortunately this imperfect understanding does matter when it comes to building robust and stable investable portfolios.

1 - Cochrane (2010) stresses the recent shift in emphasis in asset pricing research from variations in expectations to variation in discount rates.

2 - This clear distinction between 'irrationality'-based and institutional-based source of differential cross-sectional returns can easily become blurred: the availability of the speculative capital that should arbitrage away irrationalities may, for instance, have become greatly reduced because of regulatory initiatives such as the Volker rule in the States, or the Liikanen proposal in the EUR area.

Modern 'generalised factor discovery' in the fixed-income area is also made more difficult by the problem of proxies. Very often, a candidate factor (such as liquidity or value) is difficult to measure, or even to define, precisely. Furthermore, as recognised as early as 1993 by Fama and French (1993), a straightforward transposition of the factors that are most popular in the equity space (such as, for instance, value, momentum or low volatility) to the fixed-income arena is not straightforward.[3]

This has led to the proliferation of proxies, and, sometimes, to the creation of proxies of proxies. For instance, Asness, Moskowitz and Pedersen (2013) investigate value and momentum in fixed income, and circumvent the problem of defining value for bonds (a concept that, according to Fama and French (1993) 'has no obvious meaning for […] bonds' – see footnote 5) by arguing that "individual stock portfolios formed from the negative of past 5-year returns are highly correlated with those formed on *BE/ME* ratios in our sample. […] Hence, using past 5-year returns to measure value seems reasonable. …" The logic here is to use a proxy (the 5-year returns) to stand in for another proxy (value) for an unspecified latent factor. The choice may well be reasonable, but the link to the desired proxy (value), let alone to the latent factor, is neither transparent nor unique: indeed, in the literature a number of additional measures of value for bonds have been proposed, such as, the 5-year yield change in nominal yields, the difference in the 10-year yield to the 5-year inflation forecast (for real bonds), the 10-year yield spread to the short rate, not to mention a composite average of all three measures.

In addition, this concern is of relevance for the construction of robust benchmarks. As we have seen, despite the suggestive labels attached to the factors, what has actually been studied in the recent literature are often proxies more or less loosely associated to the additional fundamental quantities they 'stand in for'. This can create not only ambiguity, but also ample scope for data snooping and over-fitting.[4] Needless to say, the cost of over-fitting in-sample is poor performance out-of-sample. Apart from the dangers of over-fitting and data mining, discovering more and more factor proxies (just in the equity space, Fama recently counted more than 350!) is not necessarily desirable from the point of view of the construction of a truly diversified portfolio.[5] Indeed, as Cochrane (2010) reports, "in 2007 and 2008, hedge funds found to their dismay that portfolios they had constructed to exploit multiple "signals" all fell at the same time". This is exactly suggestive of a single source of "risk" corresponding to multiple signals of "return". A multitude of putative proxies can therefore engender a false sense of high portfolio diversifiability, and obscure the fact that a very small number of underlying latent economic factors may manifest themselves through a variety of highly correlated measurable proxies. After all, if a factor proxy is associated with a true risk premium, the attending 'excess' return is simply a compensation for receiving good or bad pay-offs in periods of high or low consumption, respectively. High or low consumption, in turn, can be parsed in terms of a relatively small number of (often highly correlated) economic configurations, such as low growth, or high unemployment. A principled and parsimonious approach to proxy analysis is therefore essential, especially in the nascent field of fixed-income 'smart beta', but has arguably not received sufficient attention.

3 - … explanatory variables like size and book-to-market equity have no obvious meaning for government and corporate bonds…', page 4.

4 - On this point, for direct a discussion of investment implications see Abu-Mostafa, Magdon-Ismail and Lin (2012) and and Bailey et al. (2014).

5 - Bender, Briand, Nielsen and Stefek (2010) and Ilmanen and Kizer (2012) emphasise the diversification benefits of factor-based portfolio building. Their point is well-taken, as long as the underlying commonality of many factors, and the attending reduction in effective diversifiability, are recognised.

6 - The training period is defined the period over which the relative out- or under-performance of bonds of different maturities is monitored and used to construct the long/short portfolio. The investment period is the period over which the zero-cash portfolio is held.

All these qualifications should be clearly kept in mind as we review in what follows the generalised fixed-income factors that have been identified most consistently in the modern literature. Among the best documented fixed-income generalised factors, one finds the following:

1. Low risk has been found both in corporate bonds and in Treasuries – see, for example, de Carvalho et al. (2014), Ilmanen et al. (2004), Frazzini and Pedersen (2014) and Howeling and van Zundert (2017). In the case of Treasuries, it is well known (Naik et al., 2016) from time-series studies that low maturity bonds offer a higher Sharpe ratio in virtually every economic environment. While these results are statistically very robust, their economic significance and portfolio exploitability is not obvious. This is especially true for Treasuries and high-yield corporate bonds, given that the recent very-low-yield environment has made the attainment of equity-like returns only obtainable with unfeasibly high leverage – de Carvalho et al. (2014) document a required leverage of 50 for the lowest-risk (and highest Sharpe ratio!) bonds. The existence of leverage-constrained investors also suggests a (regulatory/institutional) reason for the low risk generalised factor.

2. Momentum has been documented in a variety of asset classes, including fixed-income – as can be seen from the very title of the well-known paper of Asness, Moskowitz and Pedersen (2013), "*Value and Momentum Everywhere*". It should be remembered, however, that momentum is not truly a generalised factor (i.e. an attribute of a security – such as its volatility – or of an issuer – such as the debt-to-equity ratio), but rather the result of a trading strategy. The origin of this well-documented feature is still hotly debated in academic circles. When it comes to fixed-income instruments, the profitability of momentum is also still debated. For investment-grade bonds, Khang and King (2004) and Gebhardt, Hvidkjaer and Swaminathan (2005) find reversal, while Jostova et al. (2013) find no momentum. In preliminary work, Rebonato (2017) uses the Treasury-derivate discount bond data by Gurkaynak, Sack and Wright (2007) to create self-financing (zero-cash) portfolios that are long 'winners' and short 'losers'. With this strategy, it is found that momentum in US Treasuries is profitable over a relatively wide combination of training and investment periods.[6] Popisil and Zhang (2010) and Jostova et al. (2013) find momentum in the corporate high-yield market.

3. Value was identified as a factor for corporate bonds by L'Hoir and Boulhabel (2010) and by Asness, Moskowitz and Pedersen (2013) for US Treasuries. Above, we mentioned the difficulties in ascribing a precise meaning to 'value' in the fixed-income context, and the particularly 'creative' adoption of proxies and second-order proxies that this difficulty has generated. To solve these problems, current work (Rebonato and Hong, 2017) attempts to make use of affine structural models or of time-series return-predicting factors (see Cochrane and Piazzesi, 2005; Cielsak and Povala, 2010), to establish cross-sectional differentiations in the returns of Treasury bonds. While initial results are encouraging, these early studies still require further analysis and confirmation.

4. Liquidity is a particularly promising rewarded 'true' factor in the fixed-income space, as poorer returns due to a reduction in liquidity are often associated with adverse economic conditions (and hence with periods of low consumption). Recently Acharya and Pedersen (2005) have put liquidity on a firm asset-pricing footing by embedding its treatment in a CAPM-like framework, and identifying three liquidity-related 'betas' in addition to the traditional market beta.[7] The problem with their approach is that liquidity is a latent quantity, and the choice of a suitable proxy is fraught with difficulties – see Adler (2012) for a discussion of the different aspects and 'types of' of liquidity.[8] Indeed, away from the equity arena, direct measures of liquidity gleaned from transaction-level data (such as bid-offer spreads) are both difficult to obtain, and of dubious reliability.[9]

> "… more research is clearly needed before risk premia can be as efficiently extracted in the fixed-income space as in the equity space…

7 - In addition to a (liquidity-corrected) market beta, additional terms appear in the pricing equation that are linked to the covariances between the asset illiquidity and the market return, the asset return and the market illiquidity, and the asset illiquidity and the market illiquidity.

8 - The recent literature differentiates between market liquidity (linked to the ability to transact in market size without significantly affecting the price), and funding liquidity (linked to the availability of capital to fund positions in risky assets.)

9 - For instance, as liquidity begins to deteriorate, dealers tend to reduce to a 'market size' in which they are willing to transact before changing the bid-offer spread.

10 - For a justification of Duration-Times-Yield as a measure of bond risk, see Fisher (2006).

11 - See Exhibits 17 to 19 in the above-mentioned paper for more detail.

It is probably for this reason that very few studies document the link between excess returns and liquidity in the fixed-income area. In an attempt to remedy this situation, current work by Rebonato and Hong (2017) focuses on the establishment of a robust liquidity proxy for fixed income, constructed from publicly available price-based time series by modifying a methodology pioneered by Ludvigson and Ng (2009). They show that this measure of liquidity is closely linked to measures such as 'liquidity and noise' (Hu, Pan and Wang, 2013), and funding liquidity (Adrian, Etula and Muir, 2013), and argue that a large part of the yield premium provided by TIPS over US Nominal Treasuries can be explained as a compensation for liquidity risk factor. Rebonato and Naik (2016) also document the impact of liquidity on the pricing of TIPS relative to nominal Treasury bonds by using a variety of visible market proxies, but they do so without trying to combine them as in Rebonato and Hong (2017).

5. Carry can in general be defined as the funding of a riskier, longer-dated, higher-yielding strategy by shorting a lower-risk, shorter-dated and lower-yielding security. As in the case of liquidity, "carry" is likely to be linked to a true factor, in the sense that it rewards undiversifiable risk-taking likely to materialise in situations of market distress. This interpretation is indirectly confirmed by Lettau, Maggiori and Weber (2014), who apply the ideas behind the Downside-CAPM Asset pricing model, and show convincingly i) that the extra return afforded by a variety of carry strategies can be explained as compensation for poor pay-offs in periods of poor equity returns, and ii) that the well-known poor power of the CAPM model when it comes to explaining cross-sectional returns can be substantially increased once downside risk is taken into account. In the case of corporate bonds, carry is, of course, earned by investing in riskier, longer-dated securities while funding the long position in safer, shorter-dated bonds (perhaps Treasuries). By using yield-times-duration[10] as a measure of risk, de Carvalho et al. (2014) argue that, while longer-dated, lower-rated bonds offer the highest returns, low-risk, short-maturity bonds provide the best Sharpe ratios. Along similar, but not identical, lines, industry studies suggest that the best trade-off between extra yield and risk (the best Sharpe ratio) is offered just below the investment-grade rating.

6. Size as a generalised factor in fixed income has been recently analysed by Howeling and van Zundert (2017). They do so by looking at the total size of the company's debt, rather than the size of the individual issue. In their empirical analysis, every month they form a portfolio made up of the 10% of bonds with the smallest company index weights. Of course, since smaller companies tend to issue smaller bonds, and since smaller issue tend to be less liquid, the size-generalised factor inevitably picks up a liquidity contribution. They find size to contribute to portfolio profitability when combined with Low Risk, Value and Momentum. They find that the returns from this multi-factor portfolio are up to three times higher than the market, and that the excess return cannot be explained by risk or by equivalent equity factors. We should keep in mind the caveats expressed above about the possible convergence of different factors in situations of market distress.

7. Fallen angels, defined as corporate bonds downgraded to just below investment grade, seem to offer a premium when investing for a short period (up to 6 months). The effect has been analysed by Staal et al. (2015) for a comparison of the profitability of an unconditional strategy that always buys the Barclays US high-yield index with a fallen-angel strategy.[11] The same study reports a ratio of 4 to 3 in favour of the fallen-angel strategy for the relative Sharpe Ratios for the period from April 1998 to March 2015. Institutional frictions are the most likely explanation for the fallen-angel effect, which is often exploited by deploying downgrade-tolerant strategies, because a large number of mandates prohibit investment in non-investment-grade securities, and force liquidation after a downgrade.

New Frontiers in Smart Bond Portfolios

Overall, it seems that more research is clearly needed before risk premia can be as efficiently extracted in the fixed-income space as in the equity space, and this must be done with emphasis on the implementation constraints and data qualities issues, some of which are highly specific to bond markets. Now that the construction of smart beta equity portfolios has almost become a well-established technology, we see these challenges as an important new frontier in smart beta investing.

This is the second of a group of related articles on smart beta and factor investing. Please see the previous one on "Bond Risk Premia: The New Frontier in Factor Investing and Smart Beta" and the next two on "Factor Investing: Efficient Harvesting of Risk Premia Across and Within Asset Classes" and "Be Serious with Equity Factor Investing!".

References

• Abu-Mostafa, Y., M. Magdon-Ismail and H-T. Lin. 2012. Learning from Data – A Short Course. AMLbook.

• Acharya, V. and L. H. Pedersen. 2005. Asset Pricing with Liquidity Risk. *Journal of Financial Economics* 77: 375-410.

• Adler, D. 2012. The New Field of Liquidity and Financial Frictions. The Research Foundation of CFA Institute Literature Review: 1-37.

• Adrian T., E. Etula and T. Muir. 2013. Financial Intermediaries and the Cross-Section of Asset Returns. Federal Reserve Bank of New York Staff Reports, no. 464.

• Amenc, N., F. Goltz and L. Martellini. 2013. Smart Beta 2.0. EDHEC-Risk Institute Publication.

• Asness, C. S., T. J. Moskowitz and L. H. Pedersen. 2013. Value and Momentum Everywhere. *Journal of Finance* 68(3): 929-985.

• Bailey, D. H., J. Borwein, M. López de Prado and Q.J. Zhu. 2014. Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Over-fitting on Out-Of-Sample Performance. Notices of the AMS 61(5): 458.

• Bender J., R. Briand, F. Nielsen and D. Stefek. 2010. Portfolio Risk Premia: A New Approach to Diversification. *The Journal of Portfolio Management* 36(2): 17-25.

• Cieslak, A and P. Povala. 2010. Understanding Bond Risk Premia. Kellogg School of Management working paper. Wisconsin University and University of Lugano, available at https://www.gsb.stanford.edu/sites/default/files/documents/fin_01_11_CieslakAnna.pdf.

• Cochrane, J. H. 2010. Discount Rates. University of Chicago and NBER working paper.

• Cochrane, J. H. and M. Piazzesi. 2005. Bond Risk Premia. *American Economic Review* 95(1): 138-160. For technical details and the regression coefficients, consult the appendix at https://www.aeaweb.org/aer/data/mar05_app_cochrane.pdf.

• de Carvalho R. L., P. Dugnolle, X. Lu and P. Moulin. 2014. Low-Risk Anomalies in Global Fixed Income. *Journal of Fixed Income* 23(4): 51-70.

• Fama, E. F. and K. R. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33: 3-56.

• Frazzini, A. and L. H. Pedersen. 2014. Betting Against Beta. *Journal of Financial Economics* 111(1): 1-25.

• Fisher, L. 2006. Yield-Elasticity: A New Objective Measure of Interest-Rate Risk. *Journal of Applied Finance* 16: 161-173.

• Gebhardt, W. R., S. Hvidkjaer and B. Swaminathan. 2005. Stock and Bond Market Interaction: Does Momentum Spill Over? *Journal of Financial Economics* 75(3): 651-690.

• Gurkaynak, R. S., B. Sack, J.H. Wright. 2007. The US Treasury Yield Curve: 1961 to the Present. *Journal of Monetary Economics* 54: 2291-2304. Data available at http://www.federalreserve.gov/econresdata/researchdata/feds200805.xls and http://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html.

• Howeling, P. and J. van Zundert. 2017. Factor Investing in the Corporate Bond Market. *Financial Analysts Journal* – CFA Institute Publications 73(2): 100-115.

• Hu, G. X., J. Pan and J. Wang. 2013. Noise as Information for Illiquidity. *Journal of Finance* 68(6): 2341-2382.

• Ilmanen, A., R. Byrne, H. Gunasekera and R. Minikin. 2004. Which Risks Have Been Best Rewarded? *The Journal of Portfolio Management* 30(2): 53-57.

• Ilmanen, A. and J. Kizer. 2012. The Death of Diversification Has Been Greatly Exaggerated. *The Journal of Portfolio Management* 38(3): 15-27.

• Jostova G., S. Nikolova, S. Philipov and C.W. Stahel. 2013. Momentum in Corporate Bond Returns. *Review of Financial Studies* 6(7): 1649-1693.

• Khang, K. and T.H.D. King. 2004. Return Reversal in the Bond Market: Evidence and Causes. *Journal of Banking and Finance* 28(3): 569-593.

• Lettau, M., M. Maggiori and M. Weber. 2014. Conditional Risk Premia in Currency Markets and other Asset Classes. *Journal of Financial Economics* 114(2): 197-225.

• L'Hoir, M. and M. Boulhabel. 2010. A Bond Picking Model for Corporate Bond Allocation. Journal of Portfolio Management 33(2): 131-139.

• Ludvigson, S. C. and S. Ng. 2009. Macro Factors in Bond Risk Premia. *Review of Financial Studies* 22(12): 5027-5067.

• Naik, V., M. Devarajan, A. Nowobilski and S. Page. 2016. Factor Investing and Asset Allocation – A Business Cycle Perspective. CFA Institute Research Foundation Publication.

• Popisil, L. and J. Zhang. 2010. Momentum and Reversal Effects in Corporate Bond Prices and Credit Cycles. *The Journal of Fixed Income* 20(12): 101-115.

• Rebonato, R. 2017. Smart Beta in Fixed Income. Working paper presented at the L'AGEFI Day: Indexing, ETF & Smart Beta Summit, London (8 June).

• Rebonato, R. and S. Hong. 2017. A New Measure of Liquidity from Publicly Available Data. [To be submitted to *The Journal of Fixed Income*].

• Rebonato, R. and V. Naik. 2016. Can Liquidity Explain the Recent Fall in Break-Even Inflation? PIMCO working paper (March).

• Staal, L., M. Corsi, S. Shores and C. Woida. 2015. A Factor Approach to Smart Beta Development in Smart Beta. *Journal of Index Investing* 6(1).

FACTOR INVESTING: EFFICIENT HARVESTING OF RISK PREMIA ACROSS AND WITHIN ASSET CLASSES

Lionel Martellini, Professor of Finance, *EDHEC Business School* and Director, *EDHEC-Risk Institute*

Risk factors, defined as systematic underlying sources of risk that impact a large set of securities, have long been used in investment practice for analysing the risk and performance of actively or passively managed portfolios. More recently, a new approach has emerged, where factors have a more explicit role in the investment process. This approach, known as factor investing, recommends that allocation decisions be directly expressed in terms of risk factors, as opposed to standard asset class decompositions.

While the relevance of factor investing is now widely accepted amongst sophisticated institutional investors, an ambiguity remains, however, with respect to the exact role that risk factors are expected to play in the investment process. In fact, the term factor is used with many different meanings and this can be a source of confusion, and sometimes disappointment, about the benefits that can be expected from factor investing approaches. For factor investing to gain even wider acceptance, some clarification is needed with respect to the various definitions of factors and the benefits of factor investing in an institutional context.

> … factor investing within asset classes allows for a more efficient harvesting of risk premia compared to traditional approaches…

In a recent research paper (Martellini and Milhau, 2017) supported by Amundi in the context of the "ETF, Indexing and Smart Beta Investment Strategies" research chair at EDHEC-Risk Institute, we argue that there actually exist two main types of benefits that can be expected from factor investing. On the one hand, factor investing across asset classes allows for a better structuration of the investment process, both from an asset-only perspective and from an asset-liability management perspective. On the other hand, factor investing within asset classes allows for a more efficient harvesting of risk premia compared to traditional approaches that focus for example on sector decompositions.

Factor Investing and Risk Allocation Decisions: A More Efficient Structuration of the Investment Process

From an allocation perspective, factor investing is the process in which investors decide how much to allocate to each factor as opposed to each asset class. Even if factor allocation decisions must of course be eventually translated back into asset weights in order to be implemented, there are reasons to believe that this approach allows for a better structuration of the investment process, both from an asset-only perspective and from an asset-liability management perspective.

To provide some initial intuitive support for this claim, we should note that the first rationale for factor investing is that the focus on factors allows investors to gain a more holistic understanding of risk and performance by allowing investors to analyse the commonalities in the key drivers for the returns on seemingly disparate asset classes.

One of the key problems experienced by institutional investors during the sub-crime crisis and the severe market downturn that has followed is indeed that even a seemingly well-diversified allocation to multiple asset classes can hide an extremely concentrated set of factor exposures. In this context, the use of risk factors with relatively high explanatory power is critically important since it is only by framing the allocation exercise in the factor space as opposed to the asset space that investors will be able to understand how well or how poorly diversified their performance portfolios actually are.

As macro-economic factors have typically low explanatory power for asset returns, implicit factors appear to be the most natural option for representing underlying sources of risk and the more appropriate tools in risk budgeting exercises (Deguest et al., 2013). Implicit factors, extracted in a multi-asset universe via principal component analysis or minimum linear torsion techniques (Meucci et al., 2013) for a more robust outcome, can be used to build efficiently diversified portfolios. To better assess the contributions of underlying risk factors, Meucci (2009) and Deguest et al. (2013) propose to decompose the portfolio returns (which can be seen as combinations of correlated asset returns or correlated factor returns) as a combination of the contributions of uncorrelated implicit factors. The factor risk parity (FRP) portfolio is then defined as the portfolio that maximises the "effective number of uncorrelated bets" (ENUB), that is it ensures an equal contribution of the various uncorrelated implicit factors to the risk of the portfolio.

Explicit macro-economic factors like the GDP or inflation, on the other hand, can still be useful as state variables that characterise market conditions, or more precisely state variables that define various regimes of economic activity, and these regimes are relevant for asset allocation if expected performance and risk of assets and liabilities vary across these states of the economy. There are well-known advantages to identifying and employing economic regimes within a consistent asset allocation framework – see Ang and Bekaert (2002) or Mulvey and Liu (2016) for a recent reference.

One may for example identify four regimes based on whether inflation and real GDP fall above or below their median values. Formally, let us define regime 1 as a situation where inflation is below the median and growth is above average, regime 2 as a situation where growth and inflation are above average, regime 3 as a situation where both inflation and growth are below average, and regime 4 as a situation where inflation is above the median and growth is below average. It can then be shown (Martellini and Mulvey, 2017) that U.S. and international equities do best when inflation remains modest (regimes 1 and 3), and suffer when inflation is above the median, especially with growth below average (regime 4). Government bonds, on the other hand, do reasonably well except for when growth and inflation are above average (regime 2). Real estate has excellent returns under regime 3, and has under-performed during regime 4. Commodities outperform under regimes 1 and 2, but suffer when inflation is above average and growth is below average. Treasury inflation-linked bonds do best when both inflation and growth are below average (regime 3). These historical patterns, backed by sound economic theory explaining the patterns, confirm that the construction of a well-developed portfolio should take into account the behaviour of asset categories under varying economic regimes. These relationships also demonstrate the opportunities to take advantage of patterns to improve risk-adjusted returns at the asset allocation level, and even more so when liabilities are to be taken into account since liabilities for most pension plans are also sensitive to changes in economic growth and inflation.

Factor Investing and Benchmarking Decisions: A More Efficient Harvesting of Risk Premia

Factors are not only useful to understand the cross-sectional and time-series determinants of risk and returns across asset classes, but they can also be used as investable building blocks within asset classes. Individual securities earn their risk premium through exposure to rewarded factors, while the remaining risk is uncompensated for. The academic literature has identified a number of rewarded factors, the existence and persistence of which seem to be robust over different time periods and across different regions. Starting with the equity space where the concept of factor investing is most mature, the most commonly accepted risk premia include the value and the size factors (Fama and French, 1993), the momentum factor (Carhart, 1997), the low volatility factor (Ang et al., 2006), and the quality factor (Asness et al., 2013).

The outstanding question from an investor perspective is to determine the best way to harvest such multiple risk premia. This decision is in fact embedded within the choice of a benchmark, which is then used as a reference portfolio for passive or active mandates. While cap-weighted (CW) indices are typically used as default investment benchmarks by asset owners and asset managers, they have in fact been shown to suffer from two main shortcomings. On the one hand, CW indices are ill-suited investment benchmarks because they tend to be concentrated portfolios that contain an excessive amount of unrewarded risk. On the other hand, CW indices represent bundles of factor exposures that are highly unlikely to be optimal for any investor, if only because they have not been explicitly controlled for. For example, by construction CW equity indices show a large cap and a growth bias, while the academic literature has instead shown that small cap and value were the positively rewarded risk exposures. In practice, it has been shown that the use of investable forms of smart factor indices allows for substantial improvements in risk-adjusted performance compared to these traditional benchmarks (Amenc et al., 2014).

It is fair to say that the smart beta approach is now firmly grounded in equity investment practices, and the key question for an increasing majority of institutional investors is not whether one should use smart beta, but instead which and how much smart beta to use. In contrast, the concept of smart beta in the fixed-income space is still relatively less mature, despite the obvious importance and relevance of the subject. Over recent years, a number of concerns have been expressed, however, about the (ir)relevance of existing forms of corporate and sovereign bond indices offered by index providers, both in terms of lack of diversification and absence of control of the underlying factors exposures (Campani and Goltz, 2011). More generally, it appears that existing bond indices can be regarded as more "issuer-friendly" than "investor-friendly", in the sense that these bond indices passively reflect the collective decisions of issuers regarding the maturity and size of bond issues, with no control over risk factor exposures associated with such choices nor over the reward that investors should deserve from holding a well-diversified portfolio of such factor exposures. In this context, there is an increasing understanding of the need for improved bond benchmarks that will provide adequate answers to investors' needs through the construction of investable proxies for rewarded risk factors. This relates to factors that are suitably defined in fixed-income markets, namely interest rate risk (level of the yield curve, slope of the yield curve, curvature of the yield curve) and credit risk factors, liquidity risk factors, low risk factors, carry factors, value factors, momentum factors, etc. This issue is discussed further within this special issue, in an article by Riccardo Rebonato and myself.

“ … the key question for an increasing majority of institutional investors is not whether one should use smart beta, but instead which and how much smart beta to use…

Smart Beta and Beyond: Maximising the Benefits of Factor Investing

It is possible for institutional investors to use a comprehensive risk allocation framework to coherently formalise the factor investing process across and within asset classes. In a nutshell, the framework involves two steps. In a first step, the combined use of implicit factors that explain cross-sectional differences in risk and return parameters as well as explicit macro-economic factors that explain their changes over time allows asset owners (or asset managers servicing them) to implement more efficient allocation decisions. In a second step, when the allocation decisions are translated back into more traditional asset class decompositions, the use of smart factor indices as building blocks allows for a more efficient harvesting of risk premia.

While these two steps have been related in this discussion to a factor investing effort across and within asset classes, respectively, let us note in conclusion that the situation is in fact more subtle. One the one hand, FRP techniques can be applied to smart factor indices as opposed to broad asset classes so that implicit and macro-economic factors can and should impact the allocation to investable proxies for rewarded micro-economic factors within asset classes. On the other hand, as opposed to using investable proxies for risk premia in equity and fixed-income markets, one may seek to identify investable portfolios that replicate factors influencing returns in several classes, such as the value and momentum everywhere factors (Asness et al., 2013). These questions represent a fertile area of investigation both for academic research and investment practice, and we expect more innovation in the years ahead in terms of how factor investing can to lead to further developments in welfare-improving investment solutions for institutional and individual investors.

This is the third of a group of related articles on smart beta and factor investing. Please see the previous two on "Bond Risk Premia: The New Frontier in Factor Investing and Smart Beta" and "Factor-Based Approaches to the Design of Smart Bond Portfolios" and the next one entitled "Be Serious with Equity Factor Investing!".

References

- Amenc, N., R. Deguest, F. Goltz, A. Lodh and L. Martellini. 2014. Risk Allocation, Factor Investing and Smart Beta: Reconciling Innovations in Equity Portfolio Construction. EDHEC-Risk Institute Publication (July).
- Ang, A. and G. Bekaert. 2002. International Asset Allocation with Regime Shifts. Review of Financial Studies 15(4): 1137–1187.
- Ang, A., R. J. Hodrick, Y. Xing and X. Zhang. 2006. The Cross-Section of Volatility and Expected Returns. Journal of Finance 61(1): 259–299.
- Asness, C., A. Frazzini and L. H. Pedersen. 2013. Quality Minus Junk. Working Paper.
- Asness, C. S., T. J. Moskowitz and L. H. Pedersen. 2013. Value and Momentum Everywhere. Journal of Finance 68(3): 929-985.
- Campani, C. and F. Goltz. 2011. A Review of Corporate Bond Indices: Construction Principles, Return Heterogeneity, and Fluctuations in Risk Exposures. EDHEC-Risk Institute Publication (June).
- Carhart, M. 1997. On Persistence in Mutual Fund Performance. Journal of Finance 52(1): 57–82.
- Deguest, R., L. Martellini and A. Meucci. 2013. Risk Parity and Beyond – From Asset Allocation to Risk Allocation Decisions. EDHEC-Risk Institute Working Paper (August).
- Fama, E. F. and K. R. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. Journal of Financial Economics 33(1): 3–56.
- Martellini, L. and J. Mulvey. 2017.
- Introducing a Regime-Aware ALM system as a Coherent Factor-Based Investment Framework for DB Pension Plans. EDHEC-Risk Institute Working Paper (forthcoming).
- Martellini, L. and V. Milhau. 2017. Smart Beta and Beyond: Maximising the Benefits of Factor Investing. EDHEC-Risk Institute Working Paper (forthcoming).
- Meucci, A. 2009. Risk and Asset Allocation. Springer.
- Meucci, A., A. Santangelo and R. Deguest. 2013. Measuring Portfolio Diversification Based on Optimized Uncorrelated Factors. SSRN Working Paper.
- Mulvey, J. and H. Liu. 2016. Identifying Economic Regimes: Reducing Downside Risks for University Endowments and Foundations. The Journal of Portfolio Management 43(1): 100-108.

BE SERIOUS WITH EQUITY FACTOR INVESTING!

> … many factor indices show considerable divergence from academic definitions.

Felix Goltz, Research Director, *ERI Scientific Beta* and Head of Applied Research, *EDHEC-Risk Institute*

While sharing the same objectives, equity indices that aim to provide multiple factor exposures may opt for very different implementation methods, thus reflecting differences in the underlying beliefs about multi-factor investing. This article looks at the conceptual considerations involved in designing different approaches. The key issues that we discuss involve the robustness, consistency and diversification of different approaches when designing multi-factor indices.

Factor definitions

Product providers across the board put strong emphasis on the academic grounding of their factor indices. At the same time, they try to differentiate their products using proprietary elements in their strategy, often leading to the creation of products using new factors or novel strategy construction approaches that may or may not be consistent with the broad consensus on empirical asset pricing in the academic literature. As for factor definitions, many factor indices show considerable divergence from academic definitions.

For example, the Fama and French (2012, 2015) factor definitions, which are widely used in academic research, are based on straightforward stock selection criteria such as price-to-book value for example. However, for most factor or multi-factor offerings, product providers typically favour more complex factor definitions that may indeed reflect a stark disagreement with academic research. For example, some providers use industry or regional adjustments for certain variables within a given factor score while not using the same adjustments for other variables that make up the same factor score. Moreover, providers often use variables that are quite far removed from the original factor definition – for example, change in asset turnover in quality scores. In fact, most of the Quality indices on offer have more to do with the precepts of stock-picking gurus than with the academic literature, where profitability and investment have been identified as asset pricing factors.

While the definitions found in the reference academic research rely on straightforward variables and make a choice of transparently and simply selecting one key metric to come up with a factor score for each stock, the proprietary definitions from most providers use different sets of variables, as well as various adjustments, and often consist of complex combinations of several variables.

A mismatch with academic factor definitions creates two problems. The first, which we have already mentioned, is that it is difficult to refer to academic evidence to justify one's factor offering and at the same time distance oneself from the empirical framework used for that same research using factor definitions different from those used by the researchers cited. The second is that this complexification and/or creation of ad-hoc proprietary factors is a source of potential data-mining problems.

Selecting proprietary combinations or making proprietary tweaks to variable definitions offers the possibility of improving the performance of a factor index in a backtest. In general, proprietary factor definitions increase the amount of flexibility providers have in testing many variations of factors and thus pose a risk of data-mining. In fact, it appears that providers sometimes explicitly aim to select ad-hoc factor definitions that have performed well over short-term backtests.

The question is whether the improvement of the "enhanced" factor definition will also hold going forward, especially if there is no solid economic foundation for it. There is clearly a risk that one ends up with what academics have termed "lucky factors". Harvey and Liu (2015) show that by snooping through data on a large number of candidate factors and retaining those with the highest t-stat, one takes the risk of uncovering flukes, which will not repeat out of sample. Perhaps even more importantly, it is unclear what – if anything – factors with extensive proprietary tweaks still have in common with the factors from academic research. Therefore, the empirical evidence in favour of the academic factors and their economic grounding cannot be transposed to such new proprietary factors.

While the selection bias potentially exists for any strategy, there is an additional bias that is specific to so-called composite scoring approaches. These are factor definitions that draw on combinations of multiple variables. Novy-Marx (2015) analyses the bias inherent in backtests of composite scoring approaches. Novy-Marx argues that the use of composite variables in the design and testing of smart beta strategies yields a "particular pernicious form of data-snooping bias". He shows that creating a composite variable based on the in-sample performance of single variable strategies generates an over-fitting bias. The author concludes that, "combining signals that backtest positively can yield impressive backtested results, even when none of the signals employed to construct the composite signal has real power".

A simple reason for why composite scores may be more prone to generating biased results is that a composite variable requires more inputs and thus increases the number of possible choices. There seems to be wide-ranging awareness that composite strategies, by having more inputs, will lead to increased data-mining risk. Pedersen (2015) argues that, "we should discount backtests more if they have more inputs and have been tweaked or optimised more". Likewise, Ilmanen (2011) states that analysis involving "tweaks in indicator specification" is "even more vulnerable to data-mining than is identification of the basic regularities".

For investors conducting due diligence on commonly-offered smart beta strategies, it thus appears important to investigate not just the backtested performance but also the underlying data snooping risk, given that both selection bias and over-fitting bias may be present when proprietary composite scores are being used. Moreover, one can argue that backtests of strategies that do not employ complex proprietary scores are naturally more robust and the backtested performance of such strategies needs to be discounted less than that of complex proprietary factor definitions. In the next section, we further investigate the biases stemming from methodological choices, particularly looking at what happens when a consistent approach may be lacking in the index design.

Inconsistencies in methodologies

A major source of potential data-mining bias that may result in overstated backtested performance is the flexibility offered by the testing of many variations in search of the winning one. Such flexibility is obviously increased when a provider allows index methodologies to be inconsistent. On the contrary, a very effective mechanism to avoid data-mining is to establish a consistent framework for smart beta index creation. Such a framework can limit ad-hoc choices while providing the necessary flexibility needed for smart beta index construction. Surprisingly, while most major index providers argue that cap-weighted indices should employ a consistent set of rules across regions to avoid unintended investment outcomes, said consistency is often overlooked for factor indices.

Perhaps the most severe form of inconsistency is inconsistency among index offerings across time. For example, it is commonplace that two multi-factor indices launched at different points in time by the same provider use different definitions of the Value factor. This may be surprising, especially for the Value component, as Value seems to be among the most standard factors. Just like inconsistencies across factors open the room for a large number of variations in index design, it is clear that inconsistencies over time further increase such flexibility. Such inconsistency across time is, however, widely present among index offerings. Amenc et al. (2015) emphasise that inconsistency over time is all but day-to-day business for index providers.

> … a very effective mechanism to avoid data-mining is to establish a consistent framework for smart beta index creation.

Concentration issues

An important issue that can be easily neglected when constructing a multi-factor index is diversification. Positive exposure to rewarded factors is obviously a strong and useful contributor to expected returns. However, products that aim to capture explicit risk-factor tilts often neglect adequate diversification. This is a serious issue because diversification has been described as the only "free lunch" in finance. Diversification allows a given exposure to be captured with the lowest level of risk required. In contrast, gaining factor exposures exposes investors to additional types of risk, and therefore, such exposures do not constitute a "free lunch". They instead constitute compensation for risk in the form of systematic factor exposures. Such capturing of risk premia associated with systematic factors is attractive for investors who can accept the systematic risk exposure in return for commensurate compensation.

However, factor-tilted strategies, when they are very concentrated, may also take on other non-rewarded risks. Non-rewarded risks come in the form of idiosyncratic or firm-level risk, as well as potential risk for sector concentration, currency, sovereign or commodities risk exposure. Financial theory does not provide any reason why such risk should be rewarded. Therefore, a sensible approach to factor investing should not only look at obtaining a factor tilt, but also at achieving proper diversification within that factor tilt.

In fact, if the objective were to obtain the most pronounced value tilt, for example, the strategy that corresponds to this objective is to hold 100% in the single stock with the largest value exposure. This clearly shows that the objective of maximising the strength of a factor tilt is not reasonable. While practical implementations of concentrated factor-tilted indices will be less extreme than this example, we can expect problems with high levels of idiosyncratic risk and high levels of turnover whenever index construction focuses too much on concentration and pays too little attention to diversification.

One of the possible ways to construct a multi-factor index is to combine different single factor indices. Amenc et al. (2016) show that well-diversified factor indices which pursue a diversification objective through an alternative weighting scheme based on a relatively broad stock selection provide considerable benefits over more concentrated single factor indices. Their results suggest that well-diversified factor portfolios or indices outperform their highly-concentrated counterparts in terms of risk-adjusted performance, because concentrated factors may be highly exposed to unrewarded factors. In addition, they show that factor-tilted portfolios on narrow stock selections present implementation drawbacks such as higher turnover.

Concentration may arise in particular in indices that do not have such a diversification objective, especially in multi-factor indexing methodologies that, rather than combining single factor indices, actually build multi-factor indices from the stock level up.

In addition to concentration, stock level approaches contain further issues that we turn to now.

When using multi-factor scores in portfolio optimisation, it should not be forgotten that the score is ultimately used as a proxy for expected returns. It is well known for example that mean-variance optimisation that integrates expected returns can result in an "error maximisation exercise" since expected returns are hard to estimate at the individual stock level, and since mean-variance optimisers are very sensitive to estimation error for expected returns (Best and Grauer, 1991).

Achieving high absolute factor scores at the portfolio level by concentrating on picking champion stocks that score highly on all targeted factor dimensions is probably intuitively attractive but it is predicated on a high-precision relationship between factor scores and returns at the stock-level. There is no question that factor investing is motivated by an attempt to capture higher long-term returns through the right risk exposures. However, return estimation at the stock level is notoriously difficult. Black (1993) distinguishes between explaining returns, which is easy because it is really explaining variance, and predicting returns, which is hard. He contends that the accurate estimation of average expected return requires decades of data. For variance, he notes, "We can use daily (or more frequent) data to estimate covariances. Our estimates are accurate enough that we can see the covariances change through time". To estimate expected returns, on the other hand he writes, "Daily data hardly help at all." and "We need such a long period to estimate the average that we have little hope of seeing changes in expected return".

The search for champion stocks as measured by their factor scores is a stock-picking exercise that relies implicitly but heavily on the accuracy of expected return predictions. Attempting to improve stock-level return forecasts, even when this is done with the support of a factor model, is a largely futile exercise, reminiscent of traditional stock pickers. It may be useful to pause and remember that it is precisely the lack of persistent success in stock picking that has led to more institutional investors shifting toward passive strategies. If efforts are to be made to improve the risk-adjusted returns of factor investing, it is more on the risk dimension side, where we can rely on 60 years of progress in financial econometrics to estimate convergent estimators of volatilities and covariances.

> **"If efforts are to be made to improve the risk-adjusted returns of factor investing, it is more on the risk dimension side…**

When academics have tested standard factors, they have done so by running portfolio sorts, and assessing return differences at the portfolio level, not by assessing returns at the stock level. For example, they have observed that, on average, value stocks tend to have higher returns than growth stocks over the long-term. If one now tries to design strategies based on very fine distinctions at the stock level, such relations may be drowned in noise. More generally, making very fine distinctions at the stock level is prone to capturing estimation error.

Thus, any stock-level approach needs to be handled with care and one needs to assess whether suitable mechanisms have been built in to achieve robustness.

Conclusions

The offerings in the area of multi-factor indices are multiplying rapidly and investors have to assess how such indices match their investment needs. Given that most products have been launched recently, analysis of risk and performance is mostly limited to backtested data. Therefore, the methodological principles behind index construction should become a key area of attention in the assessment of these indices. Analysing robustness requires an assessment of index design principles and the conceptual considerations underlying index design. Our brief review of offerings aims to shed light on several issues such as complex proprietary factor definitions, potential inconsistencies in methodologies, and concentration issues.

In principle, multi-factor indices aim at a common goal – outperforming cap-weighted benchmarks by providing exposure to multiple rewarded factors. As discussed here, the ways to do this are nonetheless quite diverse. A key consideration for investors is how robust the performance presented in backtests is expected to be. Highly parameterised approaches naturally contain higher risks of overstated backtest performance than more parsimonious index design methods. In particular, since the bottom-up approach is more flexible, it can more easily fall prey to data-mining. It is always possible to find a combination of factor definitions, multi-factor scoring and a weighting scheme that will select the right stocks in sample. In-sample over-fitting, however, would lead to disappointing out-of-sample performance. In terms of due diligence, the bar on innovative bottom-up methods should be set higher than for classic top-down approaches, and investors would be well advised to ask for live track records of a significant length when a provider shows a lot of creativity.

There is no doubt that more elaboration on factor definitions and the use of more granular stock-level information allow the data to be fitted better and help to produce backtests that suggest superior performance, but the ultimate question investors should ask is that of the robustness of the advertised index performance in live conditions.

This is the fourth of a group of related articles on smart beta and factor investing. Please see the previous three articles on "Bond Risk Premia: The New Frontier in Factor Investing and Smart Beta", "Factor-Based Approaches to the Design of Smart Bond Portfolios" and "Factor Investing: Efficient Harvesting of Risk Premia Across and Within Asset Classes".

References

• Amenc, N., F. Ducoulombier, F. Goltz, A. Lodh and S. Sivasubramanian. 2016. Diversified or Concentrated Factor Tilts? *Journal of Portfolio Management* (42): 64-76.

• Amenc, N., F. Goltz, A. Lodh and S. Sivasubramanian. 2015. Robustness of Smart Beta Strategies. *Journal of Index Investing* 6(1): 17-38.

• Best, M. J., and R. R. Grauer. 1991. Sensitivity Analysis for Mean-Variance Portfolio Problems. *Management Science* 37(8): 980-989.

• Black, F. 1993. Beta and Return. *The Journal of Portfolio Management* 20(1): 8-18.

• Fama, E. F. and K. R. French. 2012. Size, Value, and Momentum in International Stock Returns. *Journal of Financial Economics* 105(3): 457-472.

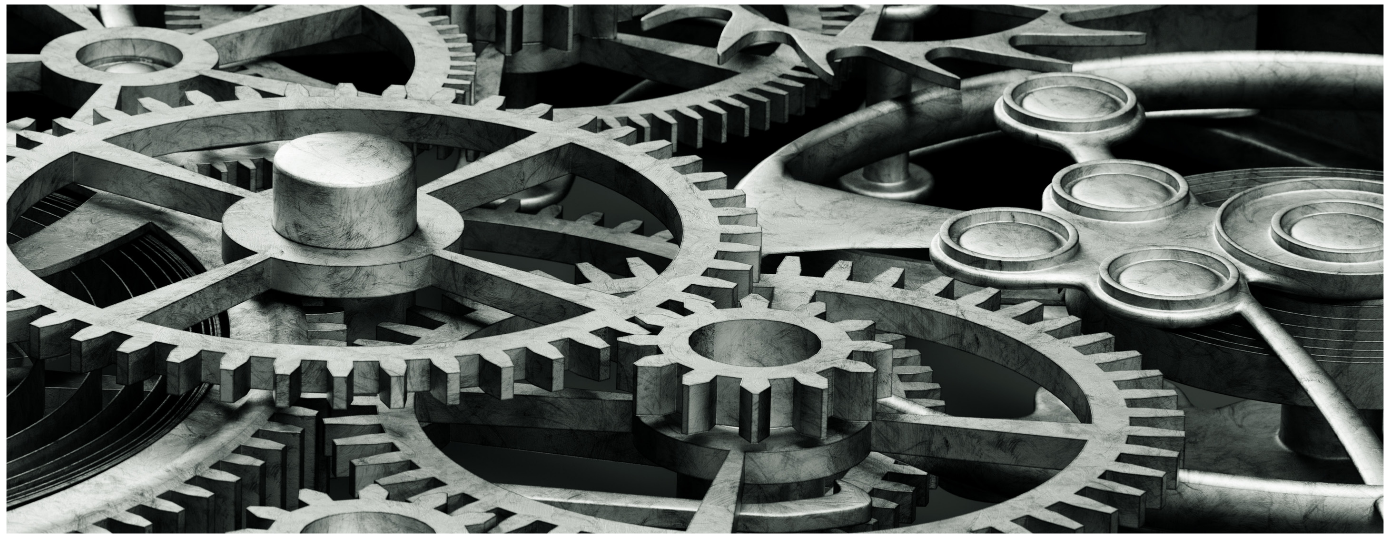
• Fama, E. F. and K. R. French. 2015. A Five-Factor Asset Pricing Model. *Journal of Financial Economics* 116(1): 1-22.

• Harvey, C. R. and Y. Liu. 2015. Lucky Factors. SSRN Working Paper.

• Ilmanen, A., 2011. Expected Returns: An Investor's Guide to Harvesting Market Rewards. John Wiley & Sons.

• Novy-Marx, R. 2015. Backtesting Strategies Based on Multiple Signals. National Bureau of Economic Research Working Paper.

• Pedersen, L. H. 2015. Efficiently Inefficient: How Smart Money Invests & Market Prices Are Determined. Princeton University Press.



MEASURING VOLATILITY PUMPING BENEFITS IN EQUITY MARKETS

This research is supported by BdF Gestion.

Jean-Michel Maeso, Quantitative Research Engineer, *EDHEC-Risk Institute*

Lionel Martellini, Professor of Finance, *EDHEC Business School* and Director, *EDHEC-Risk Institute*

Introduction

It has been argued that portfolio rebalancing, defined as the simple act of resetting portfolio weights back to the original weights, can be a source of additional performance. This additional performance is known as the rebalancing premium, sometimes also referred to as the *volatility pumping* effect or the *diversification bonus*, since volatility and diversification turn out to be key components of the rebalancing premium. The rebalancing premium, intrinsically linked to long-term investing, is typically defined as the difference between the expected *growth rate* of a rebalancing strategy and the expected *growth rate* of the corresponding buy-and-hold strategy, where the portfolio *growth rate* is the compounded geometric mean return of the portfolio – a meaningful measure of performance in a multi-period setting.

The growth rate of a portfolio $G_p(0,T)$ on the period $[0;T]$ is defined as:

$$G_P(0,T) = \frac{1}{T}\ln\left[\frac{P_T}{P_0}\right]$$

If we now consider a fixed-weight portfolio P^{reb} and a buy-and-hold portfolio $P^{b\&h}$, then the *rebalancing premium* $RP(0,t)$ over the period $[0; t]$ is simply defined by:

$$RP(0,t) = \mathbb{E}\left[G_{P^{reb}}(0,t)\right] - \mathbb{E}\left[G_{P^{b\&h}}(0,t)\right] = \mathbb{E}\left[\frac{1}{t}\ln\frac{P_t^{reb}}{P_t^{b\&h}}\right]$$

Empirical Analysis of the Rebalancing Premium

The base universe of our empirical study consists of the 132 stocks extracted from CRSP, which have continuously been in the S&P 500 index from November 1985 to December 2015. While this obviously implies the presence of a survivorship bias in equity portfolio performance, there is no reason to assume that it will impact the main comparative results of our analysis. We use a resampling procedure and build one set of 30 randomly selected equally-weighted portfolios and another set with the 30 randomly selected corresponding buy-and-hold portfolios. Then we average the (historical) rebalancing premium for the two sets of 30 randomly selected portfolios. More precisely, the rebalancing premium that we use to compare the rebalanced portfolios with the buy-and-hold portfolios is reported in this setting as the average of the rebalancing premium across the 30 random portfolios. This particular procedure mitigates the impact of stock selection biases. We consider the following set of values for the number of stocks in each randomly selected universe: 2, 10, 30, 50 and 132. In the case where we take $N = 132$, which is equal to the total size of the universe under analysis, we of course obtain a single portfolio, as opposed to 30 different portfolios.

We make the following assumptions: (i) the initial weight invested in each asset is $w_i = 1/N$ and (ii) the rebalancing frequency is 1 month. Firstly, in Exhibit 1 we focus on the historical rebalancing premium as a function of time horizon (ranging from 1 month to 30 years) for different numbers of stocks in the portfolios. For a 5-year time horizon and a number $N = 30$, the historical rebalancing premium is reasonably high at 85 bps. In a different configuration, with a 10-year time horizon and $N = 50$ risky assets, the historical rebalancing premium is 113 bps. We assess that the number of stocks considered has almost no influence as long as it exceeds a minimum value of around 10. This first perspective shows that the historical rebalancing premium from our S&P 500 base universe is higher than 50 bps if the number of stocks N is higher than or equal to 2 and the time horizon is at least 5 years. If we take a number of stocks N higher than or equal to 10, then the historical rebalancing premium is higher than 50 bps for time horizons higher than or equal to 2 years.

Exhibit 2 shows the distribution of the difference in growth rates for a 5-year time horizon and also displays the evolution of that difference over time on the period November 1985-December 2010. Each date on this chart corresponds to a 5-year period starting date. The analysis of the 5-year realised growth rates difference allows us to have a more precise view on all the 5-year historical scenarios and not only their average. The average growth rate difference (i.e. the historical 5-year rebalancing premium) is 86 bps. The growth rate difference achieves the highest value (higher than 100 bps) for the starting dates in the period January 1996-January 2000 and the period August 2004-October 2008, and the lowest value (lower than -50 bps) for the starting dates in the period July 2002-March 2004. We note that among the 302 historical 5-year scenarios, 36% of them display a growth rate difference higher than 100 bps, 61% display a growth rate difference higher than 50 bps and 16% display a negative growth rate difference. Overall, these results suggest that the rebalancing premium can be substantial in equity markets.

Rebalancing Premium and Stock Characteristics

The objective of this section is to determine whether the rebalancing premium differs for various groups of stocks. To see this, we test for the empirical relationship between the (out-of-sample) historical rebalancing premium and standard characteristics such as market capitalisation, book-to- market ratio, past performance, volatility and serial correlation. We are also interested in the persistence of the criteria used in the stock selection process since it is only if the characteristic is persistent that investors could benefit from tilting their portfolio towards that particular characteristic in an attempt to increase the rebalancing premium.

We still consider the 132 stocks which were in the S&P 500 over the period November 1985- December 2015 as our base universe and take five possible time horizons: 1, 2, 3, 4 and 5 years. We do not consider longer horizons for a persistence criterion.

For a given characteristic (market capitalisation for instance) we build at each initial (end-of-month) date t0 two sets (1 "high" and 1 "low") of two portfolios (1 equally-weighted and 1 buy-and-hold):

1. The first set of portfolios ("high") is made of the 30 best performing stocks of the base universe according to the characteristic at the initial time;

2. The second set of portfolios ("low") is made of the 30 worst performing stocks of the base universe according to the characteristic at the initial time.

The investment universe for each portfolio is held constant over the corresponding time-horizon, which allows us to analyse the influence of the characteristic on the volatility pumping effect. We then compute, for each characteristic and each time horizon, the average rebalancing premium of the "high" and "low" sets. We compare the "high" and "low" sets to a set of portfolios built randomly with 30 stocks from the base universe. This approach allows us to see if the characteristic has an impact on the rebalancing premium.

In Exhibit 3, we assess that "market capitalisation", "book-to-market ratio", "past performance" and "volatility" are sorting characteristics that have a significant impact on the rebalancing premium.

Small cap, growth, past loser and high volatility portfolios display respective out-of-sample 5-year rebalancing premia of 138, 107, 131 and 125 bps when random portfolios (N=30) display, on average, a 5-year rebalancing premium of 85 bps. The 5-year rebalancing premium can be enhanced by more than 40 bps if stocks are selected according to a characteristic such as market capitalisation rather than randomly. On the other hand, "serial correlation" as a sorting characteristic does not lead to a substantially higher rebalancing premium for time horizons higher than 1 year: for instance, the 5-year rebalancing premium for portfolios of stocks with negative serial correlation is 67 bps. This result suggests the presence of a relative lack of persistence of serial correlations at the individual stock level.

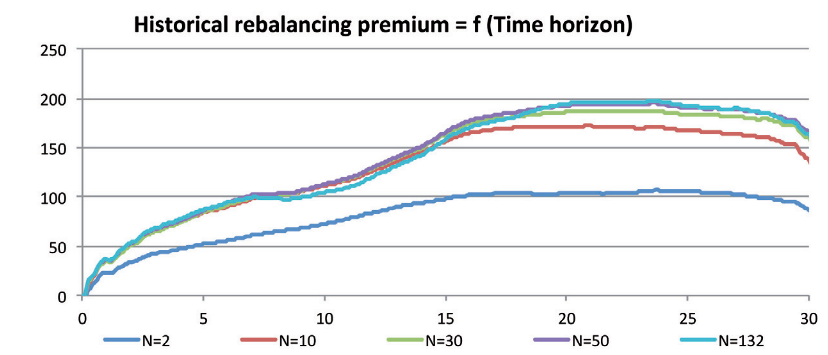
> "The analysis of the volatility pumping effect may also be transported beyond the equity universe, either in a bond portfolio context or in a multi-asset context

Conclusion

Using a selection of stocks from the S&P 500 universe we find an average historical rebalancing premium of almost 90 bps (in the absence of transaction costs) for a 5-year time horizon. Our analysis on individual stocks' characteristics highlights that size, value, momentum and volatility are sorting characteristics that have a significant out-of-sample impact on the rebalancing premium. In particular, the selection of small cap, low book-to-market, past loser and high volatility stocks generates a higher out-of-sample rebalancing premium compared to random portfolios for time horizons from 1 year to 5 years. Taken together, these results suggest that a substantial rebalancing premium can be harvested in equity markets over reasonably long horizons for suitably selected types of stocks.

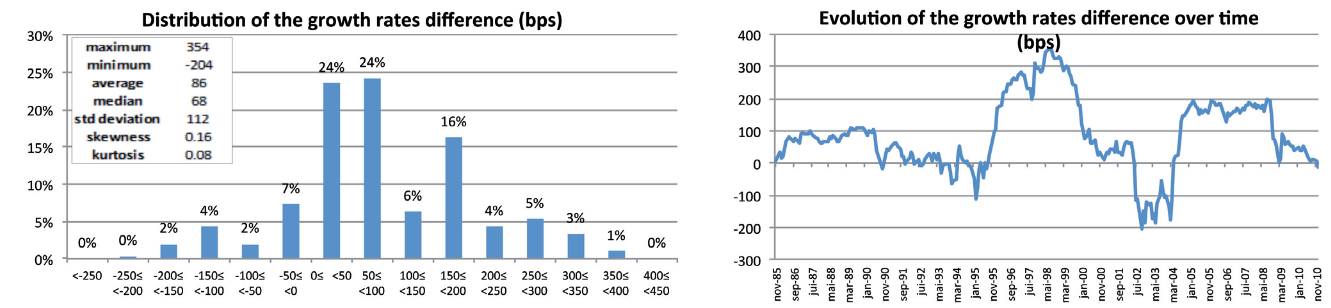
While our analysis has focused on an individual stock universe, it could be usefully applied to various equity benchmark portfolios such as style, sector, factor or country indices. The analysis of the volatility pumping effect may also be transported beyond the equity universe, either in a bond portfolio context or in a multi-asset context. Once a deep understanding of how to most efficiently harvest the rebalancing premium has been obtained, we could also focus on how to transport these benefits in a portfolio context. In particular, one would like to analyse the conditional performance of the rebalancing premium harvested within and across asset classes so as to better assess its diversification benefits.

Exhibit 1 – Historical rebalancing premium (bps)



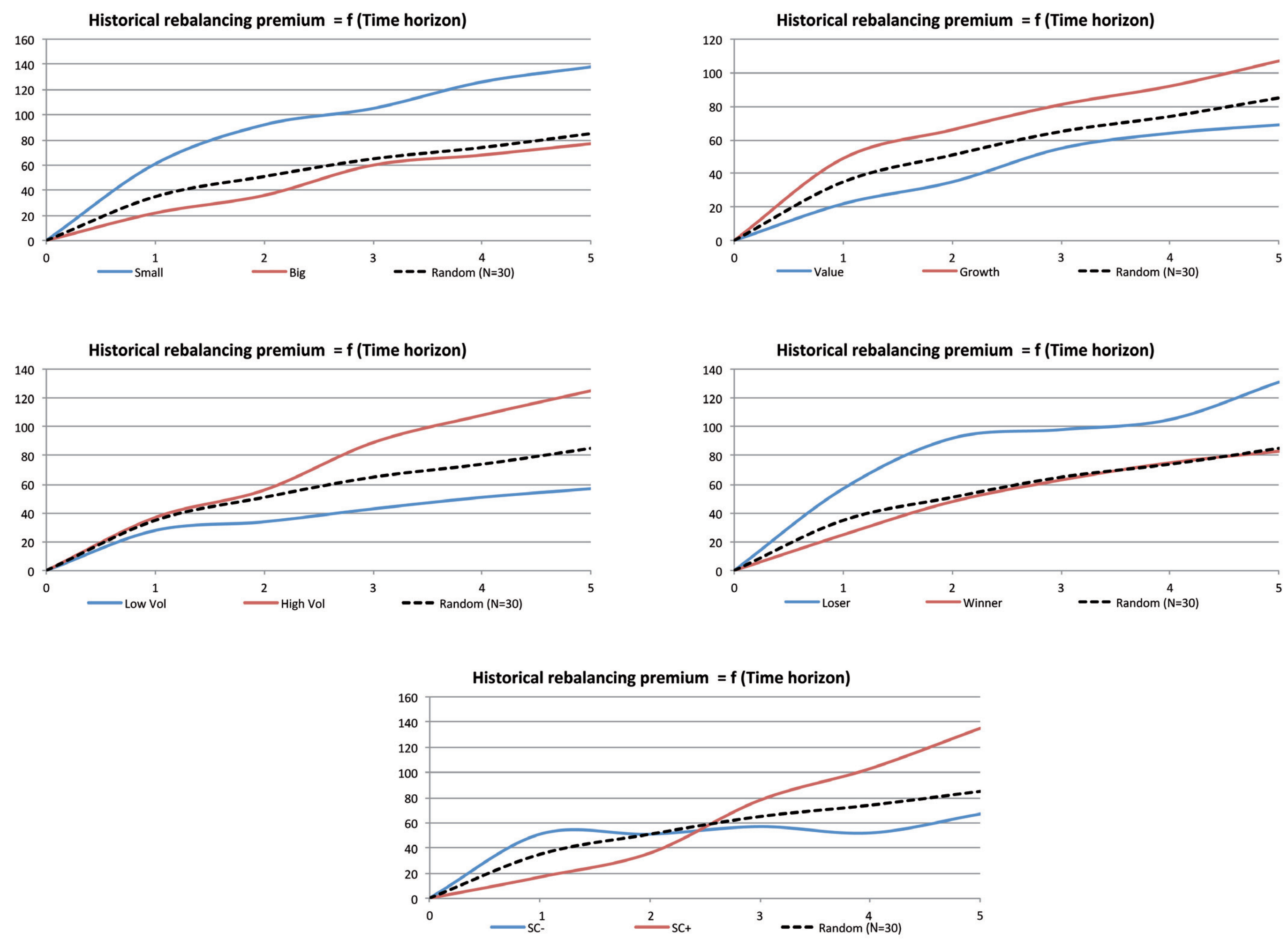
This figure displays the historical rebalancing premium (in bps) as a function of the time horizon (in years) considered for different number of stocks in the portfolios.

Exhibit 2 – Historical distribution of the growth rates difference (bps)



This figure displays the distribution of the average (across random portfolios) difference between the realised growth rate of the rebalanced portfolios and that of the corresponding buy-and-hold portfolios. The time horizon considered is 5 years and the number of stocks in the portfolios is $N = 50$. It also displays the evolution of the 5-year average (across random portfolios) growth rates difference over time.

Exhibit 3 – Historical rebalancing premium (bps) with market capitalisation, book-to-market ratio, volatility, past performance and serial correlation as sorting characteristics



This figure displays, for different characteristics, the historical rebalancing premium (in bps) as a function of the time horizon (in years) when portfolios are sorted by corresponding characteristic of the stocks (blue lines and red lines). The characteristics tested are market capitalisation, book-to-market ratio, volatility, past performance and serial correlation. We also display the rebalancing premium (black dashed lines) when the $N = 30$ stocks are randomly selected from the base S&P 500 universe. The time horizons considered are 1, 2, 3, 4 and 5 years.

PREDICTING RISK PREMIA FOR TREASURY BONDS: THE EDHEC BOND RISK PREMIUM MONITOR

Riccardo Rebonato, Professor of Finance, *EDHEC Business School*

> Deciding the relative portfolio weights among the different risk factors hinges in great part on the time-varying compensation attached to these different factors.

1. Why Risk Premia Matter

Investors in the Treasury market often observe an upward-sloping yield curve.[1] This means that, by assuming 'duration risk', they can very often invest at a higher yield than their funding cost. Yet, if the steepness of the yield curve purely reflected expectations of future rising rates, no money could on average be made from this strategy. This prompts the obvious question: When does the steepness of the yield curve simply reflects expectations of rising rates, and when does it embed a substantial risk premium?

The investment relevance of being able to answer these questions is clear. Take, for instance, a bond manager whose performance is assessed against a Treasury benchmark. His or her main strategic investment choices boil down to deciding whether to be long or short duration with respect to the benchmark. Knowing how well s/he is compensated for taking this duration risk is key to long-term performance. Or take a multi-asset portfolio manager. Deciding the relative portfolio weights among the different risk factors hinges in great part on the time-varying compensation attached to these different factors.

In all these cases, and in many more, being able to estimate in a reliable and robust manner the risk premium attached to yields is key to successful investing. It is for this reason that the EDHEC-Risk Institute is launching the EDHEC Bond Risk Premium Monitor: a robust tool to derive a state-of-the art estimation of the risk premium using market and monetary-policy information. This article goes on to explain how this task is achieved, and the theoretical underpinnings of the analytical tools used for the task.

2. Predicting Excess Returns

What predicts excess returns in Treasury bonds? And how much can one explain?

Until recently, the answers to both questions used to be: 'the slope', and 'rather little', respectively. States of the world characterised by a steep upwardsloping yield curve used to be considered indicators of positive expected excess returns. However, the degree of predictability was modest (with R^2 of the regression of the predicted and realised excess returns never exceeding 20%).

To understand why the slope was deemed to be a good predictor of excess returns consider Tab (1).

Now, recessionary periods are associated with the monetary authorities cutting rates and therefore engineering an upward-sloping yield curve. It is also natural to assume that investors should become more risk averse in the troubled recession periods. It is therefore plausible to deduce that the yield-curve slope should explain excess returns (see, eg, Fama (1986), Stambaugh (1988), Fama and French (1989), Dahlquist and Hasseltoft (2016)).

Starting from the mid-2000s, several results have questioned this received wisdom[2] : these more recent investigations suggest that different return-predicting factors may be far more complex than the simple slope[3] ; and their predictions of excess returns sometimes produce much higher R^2. Why is this the case?And what is the economic significance of the new, more complex, factors?

The motivation of the question can be readily understood by looking at Figures (1) and (2), which focus on the predictions made by the old- and new generation factors.

More precisely, Figure (1) shows the realised average excess returns, and the excess returns predicted by the slope and other 'new generation' return predicting factors. While all these predictions are all strongly correlated it is clear that the new-generation factors add a substantial twist to the slope story.

1 - Since 1971, the yield curve has been upward sloping (with the 10-year yield above the 1-year yield) for almost 84% of the time. Unless investors repeatedly and erroneously expected rates to rise almost all of the time, this is prima facie evidence of the existence of a risk premium.

2 - Some reference papers for the new wave of excess-return studies are Cochrane and Piazzesi (2005), Cieslak and Povala (2010a, b), Hellerstein(2011), Rebonato (2015), Dai, Singleton and Yang, (2004), Cochrane (2015).

3 - For instance, the return-predicting factor of Cochrane-Piazzesi (2005) is usually referred to as a 'tent', and is built by giving weights of different sign and magnitude to five forward rates. In general, the common feature of the new-generation factors is that they require (implicitly or explicitly) much hihger principal components than the second — sometimes as high as the fifth.

Tab 1: Sharpe ratios for the excess return 'carry' strategy applied to US Treasuries during the 1955-2014 period, subdivided i) into different chronological sub-periods, ii) into periods of recessions or expansions, and iii) during tightening cycles. Data adapted from Naik et al. (2016). (1)

	2-year	5-year	10-year
Full Sample	0.20	0.20	0.16
1955-1986	0.04	−0.01	−0.07
1987-2014	0.59	0.56	0.49
Recession	0.82	0.72	0.59
Expansion	0.01	0.06	0.05
1st half Expansion	0.52	0.50	0.45
2nd half Expansion	−0.61	−0.50	−0.48
Tightening Cycles			
1979:Q3-1981:Q2	−1.06	−1.13	−1.23
1993:Q3-1995:Q1	−0.79	−0.86	−0.86
2004:Q2-2006:Q2	−1.52	−0.90	−0.50

Figure 1: Average excess returns from the invest-long/fund-short strategy described in the text, and the excess returns predicted by the slope and by 'new generation' return-predicitng factors (in-sample analysis, US data).

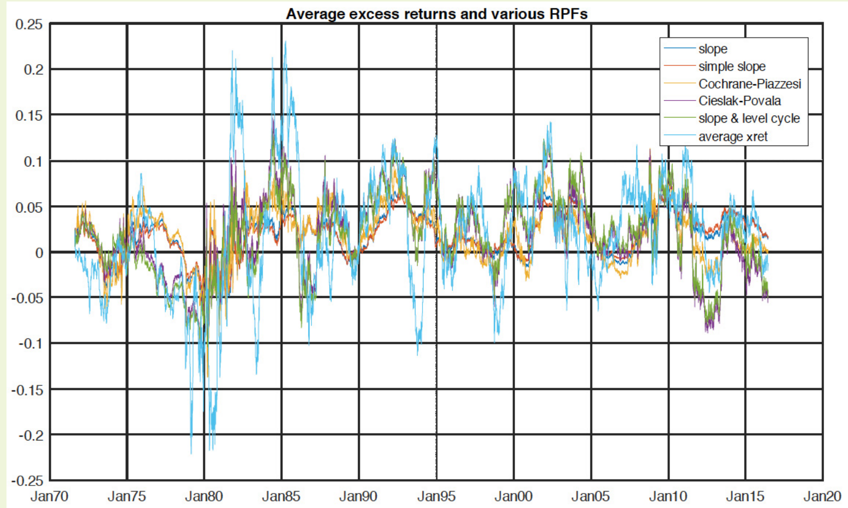


Figure 2: The differences between the prediction of average excess returns produced by the slope, and the predictions produced by the new-generation factors (in-sample analysis, US data.)

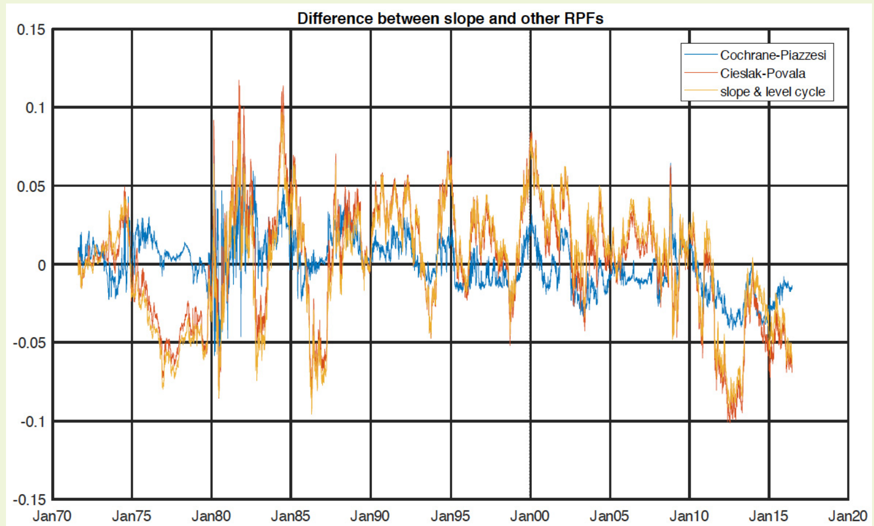


Figure 3: Normalised frequency components of the power spectrum for the 2- 6- and 9-year returns. Frequencies in years^{-1}.

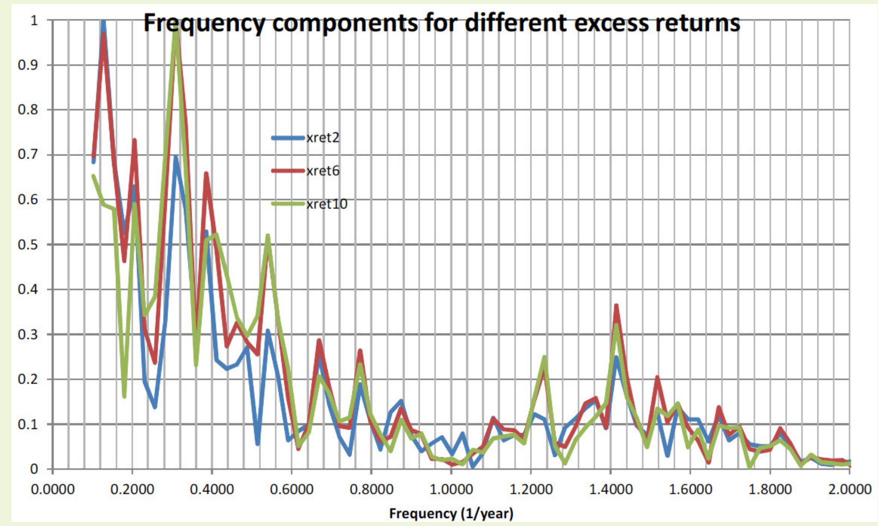


Figure (2) makes this intuition clearer by showing the differences between the prediction produced by the slope, and the predictions produced by the new generation factors (in-sample analysis). Despite the fact that the new return predicting factors are constructed following very different prescriptions, what is added on top of the slope predictions is remarkably similar.

This qualitative analysis therefore prompts the following questions:

1. Are these 'extra predictions' informative, or, as Bauer and Hamilton (2015) argue, are they just a result of over-fitting?
2. Why do such apparently different return-predicting factors produce such similar incremental predictions (with respect to the slope predictions)?
3. What is their financial and economic interpretation?

A full answer would take too long a detour (see, eg, Rebonato (2018)). We can however summarise the main findings as follows.

The first insight is linked to the power spectrum of excess returns: one can clearly see both low-frequency (business-cycle) components (well captured by the 'old' slope factor), but also a much higher frequency contribution, that requires higher principal components to be captured. It is (in part) because of its ability to capture these high-frequency components that a factor such as the Cochrane-Piazzesi fares better than the slope by itself.

This is shown in Figs (3) (4) and (5). The first figure shows that at all investment horizons there are important contributions from both low- ('business cycle') and high-frequencies components.

When we look at Figs (4) and (5), which show the frequency spectrum of the slope factor and of a 'new-generation' factor, we note how the slope recovers the low frequency peaks of the excess returns, but completely misses the medium and high-frequency components. Contrast this with the power spectrum of the 5-year excess returns and one of modern factors, which displays a remarkable match across all frequencies.

The second 'modern' insight alluded to above suggests that a large fraction of return predictability comes from detecting the cyclical straying of yields from a long-term fundamental trend. Once an effective decomposition of the yield dynamics into trend and cycle is carried out, one finds that the different degrees of mean reversion of the various return-predicting factors explain the different degrees of excess returns predictability very well.

Why do these two different 'types of' factors help the prediction of excess returns?

We propose that two distinct financial mechanisms can explain excess returns: the first, associated with low-frequency changes in

Figure 4: Frequency power spectrum of the 2-year excess returns and of the slope return-predicting factor.

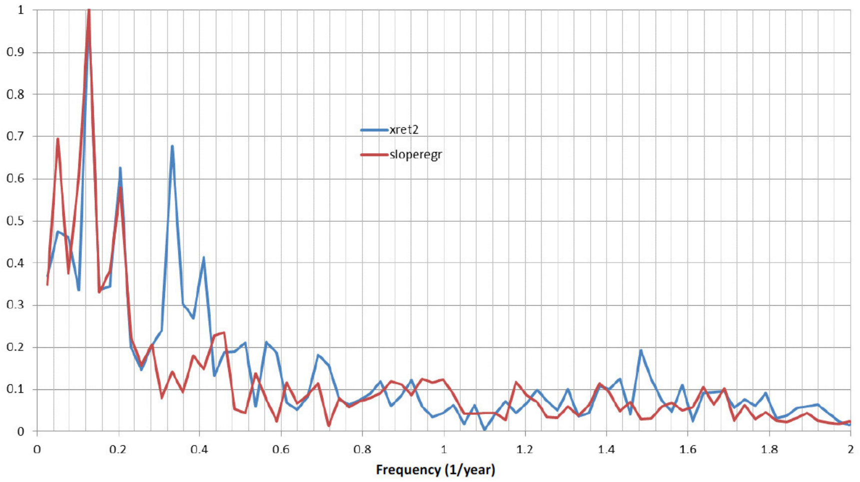
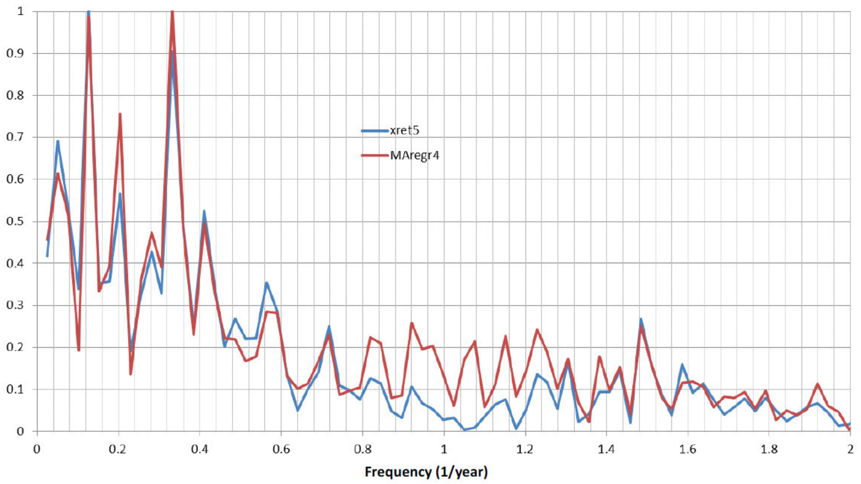


Figure 5: Frequency power spectrum of the 5-year excess returns and of the return-predicting factor built using the slope and the cycle to the 4-year moving average of the level of yields.



excess returns, is linked with changes in risk aversion with business-cycle periodicity. As for the second financial mechanism, associated with higher-frequency cycles, we suggest that it comes from the actions of pseudo-arbitrageurs who bring the level and slope of the yield curve back in line with fundamentals. These deviations have a much quicker mean-reversion, and are therefore associated with the higher-frequency components of the excess return spectrum.

The full picture is more complex, but one can see the two key insights are that the frequency components of excess returns and their mean reverting properties give us a very effective procedure to construct powerful and very parsimonious return-predicting factors: in order to predict excess returns we need a good frequency match (across high and low frequencies), and a good match of the speed of mean reversion. When these two conditions are satisfied, a number of similarly (and highly) effective and robust factors can be built almost by inspection. The new factors are parsimonious (they only require one slope-like component and one cycle-like component), intuitively understandable (thanks to the financial interpretation offered above) and highly effective (both in-sample and out-of-sample they predict as well as, and often better than, the Cochrane-Piazzesi or the Cieslak-Povala Factors).

3. Regularising the Statistical Information

As interesting as these results are, all predictions about risk premia gleaned from purely statistical studies suffer from two main shortcomings:

1. there is no guarantee that the risk premia thus estimated will be consistent with the absence of arbitrage;

2. no use is made of any information about the level of market yields: clearly, an estimate of, say, a −3% term premium has a different degree of ex ante plausibility depending on whether the corresponding market yield is, say, at 6% or 2%.

Traditionally, the 'other' route to estimating risk premia has been via the use of arbitrage-free affine term-structure models. Unfortunately, affine models do incorporate information about the level of market yields, and do ensure absence of arbitrage, but rarely do they have the flexibility to capture the rich information conveyed by the statistical analysis.[4] Both approaches are useful, but neither tells the whole truth.

The EDHEC Bond Risk Premium Monitor exploits the relative strengths of the two approaches and tries to overcome their weaknesses. It does so by complementing the predictions from the statistical estimate with the assessment of the risk premium coming from a member of the family of affine models described in Rebonato (2017). As state variables the chosen model uses the short rate, its own stochastic reversion level and the market price of risk:[5]

$$dr_t^{\mathbb{Q}} = \kappa_r^{\mathbb{P}} \left[\theta_t - r_t\right] dt + \sigma_r dz_t^r \qquad (2)$$

$$d\theta_t^{\mathbb{Q}} = \kappa_\theta^{P} \left[\widehat{\theta}_t^{\mathbb{P}} - \theta_t\right] dt + \lambda_t \sigma_\theta dt + \sigma_\theta dz_t^\theta \qquad (3)$$

$$d\lambda_t = \kappa_\lambda \left[\widehat{\lambda}_t - \lambda_t\right] dt + \sigma_\lambda dz_t^\lambda \ , \qquad (4)$$

where r_t, θ_t and λ_t are the time-t value of the short rate, of its instantaneous reversion level (the 'target rate') and the market price of risk, respectively; σ_r, σ_θ and σ_λ are the associated volatilities: $\widehat{\theta}_t^{\mathbb{P}}$ and $\widehat{\lambda}_t$ are the reversion levels of the 'target rate' and of the market price of risk, respectively; and the increments dz_t^r, dz_t^θ and dz_t^λ suitably correlated. The model is fully specified once the initial state, r_0, θ_0 and λ_0 is given.

The reader is referred to Rebonato (2017) for the financial motivation of the model, and for a detailed description of its performance. For our purposes, the important observation is that the information about the $\mathbb{P}$-measure path of the Fed funds (the 'short rate') comes from the forward guidance (the 'blue dots') provided quarterly by the Fed. See Figure (6).

When the two sources of information are combined, we obtain for the 10-year term premium the composite robust estimates shown in Figs (7) and (8).

As Fig (7) shows, the correlation among the statistical and the model-based estimate are above 90% for all the models. This is remarkable, considering how different the approaches and the sources of information are. This congruence gives us confidence about the robustness and the reliability of the combined approach.

4. Conclusions

In this article, we have given a glimpse of the latest and most exciting research strands carried out in the academic world in general, and at EDHEC-Risk Institute in particular, about the robust estimation of the yield risk premia. The predictions about the term premia for various yield maturities of the US Treasuries will be regularly provided, together with more formal research papers on these and related topics. Much work remains to be done, for instance by looking at different currencies, and at related asset classes. However, we believe that the present offering can already be of real practical use and interest for practitioners and for academics.

Figure 6: The path traced by the Fed 'blue dots' and the most similar path for the expectation of the $\mathbb{P}$-measure path of the Fed funds.

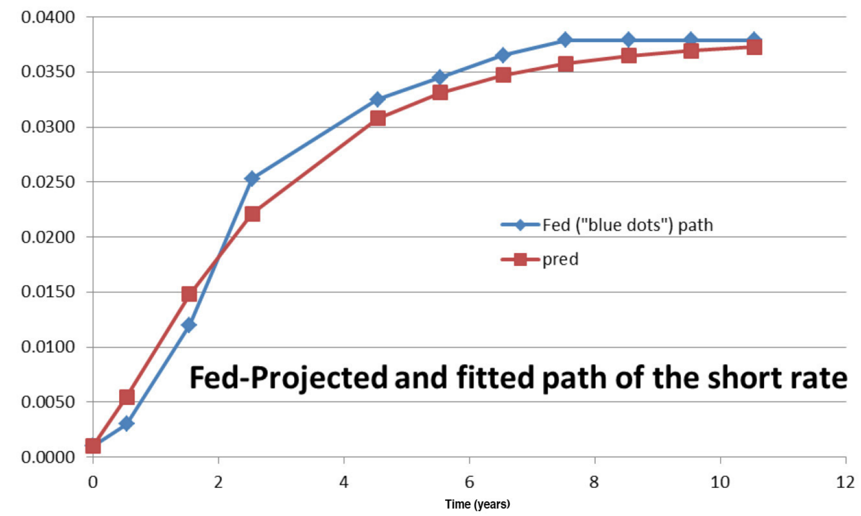


Figure 7: Estimates from the statistcial models and the affine model, rescaled to have the same volatility. The correlation among all the estimate is over 90%.

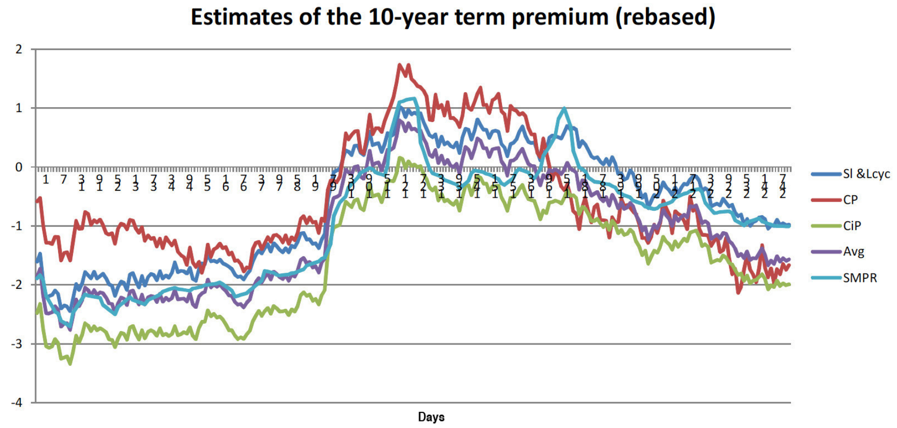
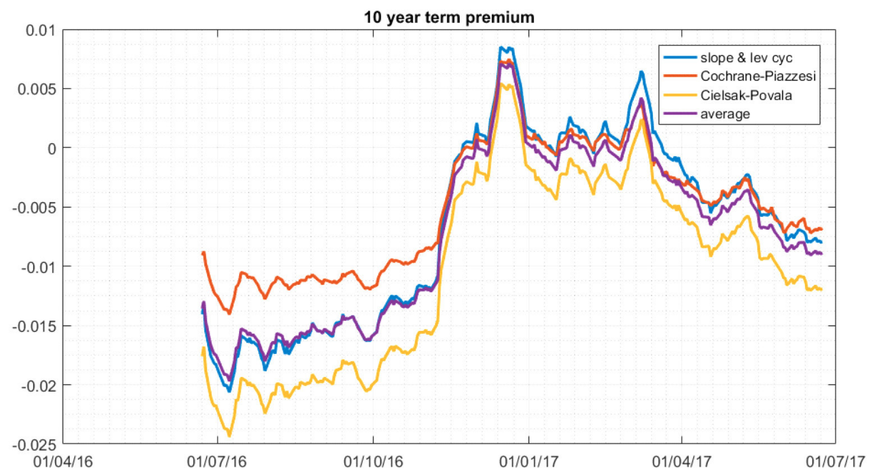


Figure 8: The result of combining the statistical estimates of the 10-year term premium from the models in the legend combined with the estimate from the affine model.



4 - This is usually because the affine dependence of the market price of risk on the state variables (required to retain tractability) is too stylised to be quantitatively useful.

5 - To simplify the analysis, and in line with standard findings (see, eg, Cochrane & Piazzesi (2005, 2008), Adrian, Crump & Moench (2013)), the model assumes that investors only seek compensation for the uncertainty about the level of rates, which we proxy in our approach as the long-term reversion level of the reversion level.

References

• Adrian, T., R. K. Crump and E. Moench. 2013. Pricing the Term Structure with Linear Regressions. Federal Reserve Bank of New York. Staff Report No 340, August 2008, Revised 2013. Available at http://www.newyorkfed.org/research/staff_reports/sr340.pdf.

• Bauer, M. D. and J. D. Hamilton. 2015. Robust Risk Premia. Federal Reserve Bank of San Francisco. Working Paper. Available at http://econweb.ucsd.edu/~jhamilto/bh.

• Cieslak, A. and P. Povala. 2010a. Understanding Bond Risk Premia. Working Paper – Kellogg School of Management, Wisconsin University and University of Lugano. Available at https://www.gsb.stanford.edu/sites/default/files/documents/fin_01_11_CieslakAnna.pdf.

• Cieslak, A. and P. Povala. 2010b. Expected Returns in Treasury Bonds. Working Paper – Northwestern University and Birbeck College, subsequently published in *The Review of Financial Studies* 28(10): 2859–2901.

• Cochrane, J. H. 2015. Comments on "Robust Risk Premia" by Michael Bauer and Jim Hamilton. Working Paper – University of Chicago. Available at http://faculty.chicagobooth.edu/john.cochrane/research/Papers/bauer_hamilton_comments.p.

• Cochrane, J. H. and M. Piazzesi. 2005. Bond Risk Premia. *American Economic Review* 95(1): 138-160. For technical details and the regression coefficients, see also https://www.aeaweb.org/aer/data/mar05_app_cochrane.pdf.

• Cochrane, J. H. and M. Piazzesi. 2008. Decomposing the Yield Curve. Working Paper – University Chicago and NBR.

• Dahlquist, M. and H. Hasseltoft. 2016. International Bond Risk Premia, in Handbook of Fixed Income Securities (Editor – P. Veronesi). Chapter 9: 171-190.

• Dai, Q., K. J. Singleton and W. Yang. 2004. Predictability of Bond Risk Premia and Affine Term Structure Models. Working Paper – Stern School of Business.

• Fama, E. F. 1986. Term Premiums and Default Premiums in Money Markets. *Journal of Financial Economics* 17: 175-196.

• Fama, E. F. and K. R. French. 1989. Business Conditions and Expected Returns Stocks and Bonds. *Journal of Financial Economics* 25: 23-49.

• Hellerstein, R. 2011. Global Bond Risk Premiums, Federal Reserve Bank of New York Staff Report no 499, June 2011.

• Naik, V., M. Devarajan, A. Nowobilski and S. Page. 2016. Factor Investing and Asset Allocation – A Business Cycle Perspective. CFA Institute Research Foundation Publication.

• Rebonato, R. 2015. Return-Predicting Factors for US Treasuries: On the Similarity of 'Tents' and 'Bats'. *International Journal of Theoretical and Applied Finance* 18(4): 1-14.

• Rebonato, R. 2017. Affine Models with Stochastic Market Price of Risk. *International Journal of Theoretical and Applied Finance* 20(4): 1750027—1-1750027—38.

• Rebonato, R. 2018. Bond Pricing and Yield Curve Modelling — A Structural Approach, Cambridge University Press, Cambridge, UK (forthcoming).

• Stambaug, R. F. 1988. The Information in Forward Rates: Implications for Models of the Term Structure. *Journal of Financial Economics* 22: 613-652.

Investment Management Review

Co-editors: Dr Arjuna Sittampalam, Lionel Martellini, Professor of Finance at EDHEC Business School and Director of EDHEC-Risk Institute

Publisher: Angus I Milne
aim@sageandhermes.com

Subscriptions:
aim@sageandhermes.com
+44 (0) 1202 419 408

Published by
Sage & Hermes Limited
44 Worthington Road,
Surbiton, Surrey KT6 7RX, UK

+44 (0) 1202 419 408
enquiries@imrmagazine.com
www.imrmagazine.com