Stochastic modeling applied to portfolio optimization problems

Jean-Michel Maeso

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES FONDAMENTALES ET APPLIQUÉES

THÈSE DE DOCTORAT

Modélisation stochastique appliquée à des problèmes d'optimisation de portefeuille

Jean-Michel MAESO

Laboratoire Jean Alexandre Dieudonné (LJAD)

Présentée en vue de l'obtention du grade de docteur en Mathématiques
d'Université Côte d'Azur
Dirigée par : Frédéric Patras
Co-encadrée par : Vincent Milhau
Soutenue le : 1er décembre 2022

Devant le jury, composé de :
Christophette Blanchet-Scalliet, Maîtresse de conférences, Ecole Centrale de Lyon.
Mireille Bossy, Directrice de recherche, INRIA.
Guillaume Coqueret, Associate Professor, EM Lyon.
Francine Diener, Professeur des universités, Université Côte d'Azur.
Stéphane Loisel, Professeur des universités, ISFA-Université Lyon 1.
Lionel Martellini, Full Professor, Edhec.
Vincent Milhau, Directeur de recherche, Edhec-Risk Institute.
Frédéric Patras, Directeur de recherche, CNRS.
Sylvain Rubenthaler, Maître de conférences, Université Côte d'Azur.

Thèse

pour obtenir le titre de

Docteur en Sciences

Discipline : Mathématiques

présentée et soutenue par

Jean-Michel MAESO

Modélisation stochastique appliquée à des problèmes d'optimisation de portefeuille

Stochastic modeling applied to portfolio optimization problems

---

Thèse dirigée par

Frédéric PATRAS

et

Vincent MILHAU

soutenue le 1 décembre 2022.

Jury :

M.	Christophette BLANCHET-SCALLIET	*Examinatrice*
M.	Mireille BOSSY	*Examinatrice*
M.	Guillaume COQUERET	*Rapporteur*
M.	Francine DIENER	*Présidente du jury*
M.	Stéphane LOISEL	*Rapporteur*
M.	Lionel MARTELLINI	*Membre invité*
M.	Vincent MILHAU	*Co-Directeur de thèse*
M.	Frédéric PATRAS	*Co-Directeur de thèse*
M.	Sylvain RUBENTHALER	*Examinateur*

Resumé. Cette thèse est constituée de trois parties indépendantes. Le premier chapitre s'appuie sur la théorie de portefeuille stochastique développée par Robert Fernholz et propose une analyse théorique puis empirique de l'optimisation de l' "excess growth rate" d'un portefeuille d'actions. L' "excess growth rate" est défini comme la différence entre l'espérance du rendement logarithmique du portefeuille et la somme pondérée des espérances des rendements logarithmiques de ses constituants. Cette quantité peut s'exprimer comme une fonction des termes de la matrice de covariance des actifs du portefeuille. Le deuxième chapitre étudie l'impact en termes de performance du rebalancement des poids des constituents d'un portefeuille vers ses valeurs initiales. L'étude est centrée sur la "rebalancing premium", définie comme la différence de performance logarithmique entre deux portefeuilles avec les mêmes poids initiaux (et donc initialement identiques), l'un étant rebalancé à une certaine fréquence vers ses poids initiaux alors que l'autre non. L'objet de ce chapitre est de fournir une analyse théorique, numérique et empirique approfondie de la "rebalancing premium" pour des portefeuilles d'actions et d'examiner les conditions dans lesquelles elle peut être optimisée. Le troisième chapitre développe un cadre théorique complet et flexible d'investissement pour un individu à la retraite (phase de décumulation) qui modélise les risques de marché, de longévité et le risque de dépenses imprévues liées aux besoins en soins de longue durée. L'univers d'investissement va au-delà des actions et des obligations et inclut des balanced funds, des target date funds ainsi que des annuités. Ce cadre théorique a été implémenté avec le logiciel R afin de pouvoir résoudre ensuite numériquement un problème d'optimisation de portefeuille.

Mots clés: Calcul stochastique, optimisation, simulations de Monte Carlo, construction de portefeuille.

Abstract. This thesis consists of three independent parts. The first chapter is based on the stochastic portfolio theory developed by Robert Fernholz and proposes a theoretical and empirical analysis of the optimization of the "excess growth rate" of an equity portfolio. The "excess growth rate" is defined as the difference between the expected logarithmic return of the portfolio and the weighted sum of the expected logarithmic returns of its components. This quantity can be expressed as a function of the terms of the portfolio's asset covariance matrix. The second chapter studies the impact in terms of performance of the rebalancing of the weights of the constituents of a portfolio towards its initial values. The study focuses on the "rebalancing premium", defined as the difference of the logarithmic performance between two portfolios with the same initial weights (and therefore initially identical), one being rebalanced at a certain frequency towards its initial weights while the other is not. The purpose of this chapter is to provide a thorough theoretical, numerical and empirical analysis of the rebalancing premium for equity portfolios and to examine the conditions under which it can be optimized. The third chapter develops a comprehensive and flexible investment framework for an individual in retirement (decumulation phase) that models market risk, longevity risk, and the risk of unexpected expenses related to long-term care needs. The investment universe goes beyond stocks and bonds and includes balanced funds, target date funds and annuities. This theoretical framework was implemented with R software in order to be able to solve a portfolio optimization problem numerically.

Keywords: Stochastic calculus, optimization, Monte Carlo simulations, portfolio construction.

Remerciements

Ceux qui me connaissent bien savent que j'ai deux passions depuis l'adolescence : le tennis et les mathématiques. Bien que j'aurais pu écrire quelques pages sur ce sport qui me fait tant vibrer, le sujet de cette thèse porte uniquement sur les mathématiques (je n'ai pas trouvé de sujet de thèse sur les applications du calcul stochastique au tennis), ce qui signifie que la plupart de mes amis et ma famille ne dépasseront probablement pas cette page de remerciements. Je ne leur en veux pas.

La réalisation de cette thèse a été une expérience exceptionnelle sur le plan intellectuel et sur le plan humain. Je ne m'attendais pas au niveau de difficulté que j'aurais à surmonter pour produire ce travail et je suis reconnaissant envers mes deux co-directeurs Frédéric Patras et Vincent Milhau pour leur encadrement, leur bienveillance et pour m'avoir fait profiter de leurs nombreuses connaissances et idées originales. Je tiens à remercier tous les membres de mon jury pour leur temps précieux et leurs commentaires qui ont contribué à améliorer la qualité de cette thèse.

Je remercie mon professeur de mathématiques de classes préparatoires Michel Alessandri, un professeur exceptionnel et un homme formidable, pour m'avoir transmis sa passion des mathématiques et m'avoir inspiré dans mon parcours professionnel et mes choix personnels.

Vincent, merci pour tout ce que tu m'as apporté, c'est un plaisir et une chance de travailler avec un chercheur de ton niveau. Je remercie également Lionel Martellini avec qui j'ai pris plaisir à travailler et qui m'a soutenu dans mon projet de réaliser une thèse de mathématiques sur mon temps libre : votre rencontre a été un tournant dans ma vie professionnelle et une source d'inspiration. Merci à Riccardo Rebonato et à Dominic O'Kane pour tout ce qu'ils m'ont appris en travaillant à leurs côtés.

Enfin, je tiens à remercier mes amis et ma famille pour leur amour et leur soutien constants. Merci à Vincent (Jacquot), Valentine, Fabien, Hugo, Caroline, Mohamed, Nicolas, Alexis, Benjamin, Serge, Alexandre, Benoît, Arnaud, Henri-Pierre, Zulma, Pierre-Anthony et Véronique pour tous les bons moments partagés lors de ces trois dernières années, que ce soit sur un court de tennis, dans une salle de sport ou ailleurs. Merci à mes parents, Patricia et Georges, qui m'ont beaucoup soutenu à leur manière bien au-delà de ces trois dernières années.

Contents

General Introduction

This general introduction aims to present the thematic of the thesis and its structure composed of three independent contributions.

Since the second half of the twentieth century, sophisticated mathematical tools from probability theory have been used to model and solve problems in finance such as portfolio selection and pricing of derivatives. Stochastic calculus, in particular, has been widely applied in finance to model asset prices and uncertainty in financial markets. The mathematical object at the basis of stochastic calculus is the Brownian motion, whose name is a tribute to the botanist Brown (1828) who reported his observations of the disordered motion of plant pollen particles when they were observed under a microscope. Bachelier (1900), in his doctoral thesis, was the first to introduce the Brownian motion in finance, and that process has become the building block of most pricing models in finance, notably the Black-Scholes formula (Black and Scholes (1973)). In this thesis, which deals with three research topics that are motivated by future practical applications in finance, we will extensively apply stochastic modeling on risk estimation and portfolio selection problems and call upon notions belonging to statistics and econometrics.

Academic research on portfolio selection was pioneered by Markowitz (1952), who developed the mean-variance single-period framework and introduced the concept of diversification defined as the risk management technique that aims to earn the highest expected return possible for the amount of risk taken. He also defined the efficient frontier as the set of portfolios that have a maximum expected return for a given level of risk (i.e. volatility). In a general setting involving n risky assets with expected return vector $\mu = (\mu_i)_{1\leq i\leq n}$, with covariance matrix $\Sigma = (\sigma_{i,j})_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}$ and weights $\boldsymbol{w} = (w_i)_{1\leq i\leq n}$ the mean-variance investor optimization problem is formulated as:

$$\boldsymbol{w}^\star = \arg\max_{\boldsymbol{w}} \mu_P(w) \quad \text{u.c.} \begin{cases} \sigma_P(\mathbf{w}) \leq \sigma^\star \\ \mathbf{1}^\top \boldsymbol{w} = 1 \end{cases}$$

where $\mu_P = \sum_{i=1}^{n} w_i \mu_i$ is the portfolio expected return, $\sigma_P = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{i,j}}$ is the portfolio volatility and $\sigma^\star$ is the maximum level of risk.

A few years later, Markowitz (1956) demonstrated how to calculate efficient frontier portfolios using quadratic programming:

$$\boldsymbol{w}^\star = \arg\max_{\boldsymbol{w}} \left[\mu_P(w) - \frac{\varphi}{2}\sigma_P^2(\mathbf{w})\right] \quad \text{u.c. } \mathbf{1}^\top \boldsymbol{w} = 1$$

where φ is the risk-aversion parameter.

Later, Tobin (1958), Sharpe (1964) and Lintner (1965b,a) extended the mean-variance framework by introducing a risk-free asset in the investment universe and developed the capital asset pricing model (CAPM). Merton (1969) extended the portfolio selection approach of Markowitz in a continuous-time setting allowing for rebalancing.

The main limitation of the Markowitz approach is the high sensitivity of optimal portfolios to risky asset expected return estimation (see for instance Merton (1980) and DeMiguel et al. (2009) for more details). Then, different alternative methods of equity portfolio construction that

only require a covariance matrix estimation and no expected return have been developed and investigated:

- Capitalization Weighted (CW), which weights stocks by their capitalization.
- Equally-Weighted (EW), which attributes the same weight to each constituent in the investment universe.
- Global Minimum Variance (GMV), which aims to minimize the ex-ante volatility by using the covariance matrix structure.
- Equal Risk Contribution (ERC), which equalizes the contribution of each constituent to the portfolio volatility (see Maillard et al. (2010) for more details).
- Maximum Diversification (MD), which aims at maximizing the diversification ratio defined as the weighted average of the volatilities of the constituents over the portfolio volatility (see Choueifaty and Coignard (2008) for more details).

Fernholz and Shay (1982) and Fernholz (2002) introduced a continuous-time mathematical framework labeled as stochastic portfolio theory involving a standard Ito-process model for a financial market composed of n risky stocks with correlation matrix $C_t = (\rho_{i,j}(t))_{\substack{1\leq i\leq n\\1\leq j\leq n}}$, volatility vector $V_t = (\sigma_i(t))_{1\leq i\leq n}$, covariance matrix $\Sigma_t = (\sigma_{i,j}(t))_{\substack{1\leq i\leq n\\1\leq j\leq n}} = (\rho_{i,j}(t)\sigma_i(t)\sigma_j(t))_{\substack{1\leq i\leq n\\1\leq j\leq n}}$, and weight vector $w_t = (w_i(t))_{1\leq i\leq n}$. Stock prices $S_i(t)$ are modeled as processes verifying the following stochastic differential equations:

$$dS_i(t) = S_i(t)\left(\mu_i(t)dt + \sum_{\nu=1}^{m}\beta_{i,\nu}(t)dW_\nu(t)\right),\ i = 1,\dots,n$$

where $(W_1(.),\dots,W_m(.))$ is a vector of m independent Brownian motions ($m \geq n$), $\mu(.) = (\mu_1(.),\dots,\mu_n(.))$ is the vector of stocks' expected return and $\beta(.) = \left(\beta_{i,\nu}(.)\right)_{\substack{1\leq i\leq n\\1\leq \nu\leq m}}$ is the $(n\times m)$ matrix of volatilities representing the sensitivity of each stock to the m sources of uncertainty.

It can be shown that the portfolio value $P(t)$ verifies:

$$\begin{aligned} d\left(\ln P(t)\right) &= \sum_{i=1}^{n} w_i(t)d\left(\ln S_i(t)\right) + g(t)dt - \sum_{i=1}^{n} w_i(t)g_i(t)dt \\ &= \sum_{i=1}^{n} w_i(t)d\left(\ln S_i(t)\right) + g^\star(t)dt \end{aligned}$$

where $g^\star(t)$ is labeled as the portfolio excess growth rate and defined as the expected growth rate of the portfolio $g(t)$ minus the weighted-average of the growth rates $g_i(t)$ of its constituents. The excess growth rate can be rewritten as a function of the covariance matrix and the portfolio weights:

$$\begin{aligned} g^\star(t) &= g(t) - \sum_{i=1}^{n} w_i(t)g_i(t) \\ &= \sum_{i=1}^{n} w_i(t)\mu_i(t) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i(t)w_j(t)\sigma_{i,j}(t) - \left(\sum_{i=1}^{n} w_i(t)\mu_i(t) - \frac{1}{2}\sum_{i=1}^{n} w_i(t)\sigma_{i,i}(t)\right) \\ &= \frac{1}{2}\left(\sum_{i=1}^{n} w_i(t)\sigma_{i,i}(t) - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i(t)w_j(t)\sigma_{i,j}(t)\right) \end{aligned}$$

Finally the excess growth rate can be written as half of the difference between the weighted-average variance of the assets in the portfolio and the portfolio variance:

$$g^{\star}(t) = \frac{1}{2}\left(w_t^{\top}\left(V_t \odot V_t\right) - w_t^{\top}\Sigma_t w_t\right)$$

We note that if all assets have the same expected growth rate, then maximizing the portfolio excess growth rate is equivalent to maximizing the portfolio expected growth rate, which in turn is well-known to be equivalent to maximizing expected utility for an investor endowed with logarithmic preferences, i.e. who wants to maximize his expected log-wealth (Latané (1959)). Interestingly, the excess growth rate maximizing portfolio, just like the equally-weighted portfolio (EW), global minimum variance (GMV), equal risk contribution (ERC) or maximum diversification portfolios described above, only requires estimates for the covariance matrix, and not for the vector of expected returns. In an empirical context, one might wonder whether adopting a portfolio strategy that periodically maximizes the ex-ante portfolio excess growth rate would allow one to obtain a risk-adjusted performance similar to well-known portfolio strategies such as equally-weighted or equal risk contribution. We formally define the Max $g^{\star}$ portfolio strategy as the weighting scheme that maximizes the ex-ante excess growth rate of the portfolio in the absence of a risk-free asset, and the MVR[1] strategy as the risky asset portfolio, when combined with the risk-free asset, maximizes the ex-ante excess growth rate in a universe with a risk-free asset.

The first chapter will be dedicated to a detailed theoretical analysis of the properties of the excess growth rate maximizing Max $g^{\star}$ portfolio, its link with the MVR portfolio as well as its relationship with respect to other popular weighting schemes. It will also provide a thorough empirical analysis of the benefits of maximizing an equity portfolio excess growth rate in a context where the constituents are individual stocks, by comparing the Max $g^{\star}$ and MVR portfolios to benchmark portfolios such as cap-weighted portfolios, equally-weighted portfolios, minimum variance portfolios, equal risk contribution portfolios, and maximum diversification portfolios. At the end of this first chapter, we will be able to answer the following questions:

- What are the analytical expressions of the Max $g^{\star}$ and MVR portfolios when weights are not constrained?
- How does a strategy that periodically[2] maximizes the ex-ante portfolio excess growth rate performs compared to other portfolio strategies in different individual stock universes (regional universes, industry universes and factor-tilted universes)?

In the second chapter, we keep the continuous-time framework introduced by Fernholz and Shay (1982) and Fernholz (2002) to study the benefits of portfolio rebalancing, defined as the simple act of resetting portfolio weights back to the original weights. Portfolio rebalancing has been documented as a potential source of additional performance known as the *rebalancing premium*, also sometimes referred as the *volatility pumping* effect (see Luenberger (1997)) or *diversification bonus* (see Booth and Fama (1992), Erb and Harvey (2006), Willenbrock (2011) and Qian (2012)) 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 (see Bernstein and Wilkinson (1997), Hallerbach (2014)), where the portfolio *growth rate* is the compounded geometric mean return of the portfolio (see Kelly (1956) and Latané (1959)), a meaningful measure of performance in a multi-period setting.[3] In the context of a general stochastic multi-period framework with time horizon T, where P_0 denotes the portfolio initial value and P_T its final value, we define the growth rate of

[1] MVR stands for Maximum Volatility Return and our definition is derived from Mantilla-Garcia (2016) portfolio excess growth rate optimization problem.

[2] typically with a 3-month rebalancing frequency

[3] The rebalanced portfolio and the buy-and-hold portfolio have the same initial weights.

a portfolio $G_P(0,T)$ on the period $[0;T]$ as:

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

Note that this quantity is stochastic as seen from the initial date since the final portfolio value P_T is unknown as seen from the initial date. If we now consider a fixed-weight portfolio P^{reb} and a buy-and-hold portfolio $P^{b\&h}$ based on holding n assets with the same initial weights $(w_1,\ldots,w_n)$ (this weights are held constant in the rebalanced portfolio, while they drift away in the buy-and-hold portfolio), then the *rebalancing premium* $RP(0,t)$ over the period $[0;t]$ is simply defined by:

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

The first analysis of the benefits of rebalancing in a continuous-time setting can be traced back to the seminal contribution by Merton (1971), who establishes that a continuously rebalanced fixed-weight portfolio strategy is optimal for a risk-averse investor in a multi-period setting with a constant opportunity set. This result suggests that if market conditions never change and if relative risk aversion is independent of wealth, there is no reason to change portfolio weights, so a fixed-mix is optimal. Following this first analysis, a large number of papers, including notably Fernholz and Shay (1982), Wise (1996), Fernholz (2002) and Gabay and Herlemont (2007) compare a fixed-weight portfolio with a buy-and-hold portfolio and analyse the probability that the continuously rebalanced portfolio outperforms the buy-and-hold portfolio. In practice, portfolios cannot be rebalanced continuously, risk and return parameters are time-varying, time horizon is finite and transaction costs can have a strong impact on the performance of a rebalanced strategy. Several authors have subsequently pointed out the limitations of an analysis of the rebalancing premium in such a stylised analytical setting (see Chambers and Zdanowicz (2014), Cuthbertson et al. (2015)) and others have analysed whether the alleged benefits of rebalancing subsist empirically when taking into account the impact of transaction costs, initial weighting schemes, investment universe, rebalancing frequency, time-varying correlations and volatilities and the performance metrics used to compare the performance and/or risk-adjusted performance of the portfolios. The conclusions of these empirical analysis of the rebalancing premium in realistic settings have been mostly positive. In particular, Arnott and Lovell (1993), Tsai (2001) and Harjoto and Jones (2006) find that a rebalanced strategy outperforms the corresponding buy-and-hold strategy considering a two-asset (equity and bond) universe.[4] A common limitation of these papers, however, is that their empirical protocol is based on a fixed historic period, which raises questions about the robustness of the reported findings.

Another key limitation in the related literature is that it does not provide a clear understanding of the actual origins of the rebalancing premium. In particular, no attempt is made to control for differences in exposure of the rebalanced portfolio and the corresponding buy-and-hold portfolio with respect to rewarded risk factors. These differences are expected to have an impact on the relative performance of the portfolios, an impact which is not directly related to the volatility pumping effect. For example, in an equity context and using the Fama-French-Carhart factor model to measure risk exposures, the intuition suggests that the buy-and-hold strategy would lead to growth, large cap and momentum biases relative to the corresponding fixed-mix strategy while the fixed-mix strategy would lead to value, small cap and contrarian biases relative to the corresponding buy-and-hold strategy. The main contribution of the second chapter is precisely to measure and analyse the sources of the rebalancing premium effect by disentangling the isolated impact of the various aforementioned components. In this second chapter, we aim at determining the main drivers of the rebalancing premium in a theoretical and numerical framework and then empirically analysing the difference of performance between a rebalanced portfolio and a buy-and-hold portfolio from a risk-adjusted performance perspective.

[4] Jaconetti et al. (2010) assert that there is no optimal frequency when considering a rebalancing strategy.

In a stochastic portfolio theory setting where stock prices $S_i(t)$ follow geometric Brownian motions and the rebalanced portfolio is continuously rebalanced, the rebalancing premium can be analytically decomposed as:

$$RP(0,T) = \mathbb{E}\left[\frac{1}{T}\ln\left(\frac{P_T^{reb}}{P_t^{b\&h}}\right)\right] = g^* + \mathbb{E}\left[\frac{1}{T}\left(\ln\frac{\prod_{i=1}^{n} S_i(T)^{w_i}}{\sum_{i=1}^{n} w_i S_i(T)}\right)\right]$$

where $g^\star$ is the excess growth rate defined and studied in the first chapter.

A second part of the chapter is dedicated to an empirical analysis of the rebalancing premium in realistic settings. To control for differences in factor exposures, we regress the excess monthly returns of the rebalanced portfolio over the buy-and-hold portfolio on the market, size, value and momentum factors, and interpret the alpha of this regression as the marginal value added by a pure volatility pumping effect, defined as the fraction of the rebalancing premium that is not related to differences in factor exposures. As such, this analysis is related to Plyakha et al. (2012), who consider stocks in the major US equity indices over forty years, compare the performance of equally-weighted portfolios with price-weighted and value-weighted portfolios and analyse the difference of performance using inter alia Fama-French-Carhart four-factor linear regression. In a second step, we analyse how this pure volatility pumping effect depends upon the covariance matrix of the underlying assets, the serial correlation in asset returns, the rebalancing frequency, the stocks' characteristics and the portfolio weights.

At the end of this second chapter, we will be able to answer the following questions:

- What are the main drivers of the rebalancing premium?
- What are the individual stock characteristics that have a significant out-of-sample impact on the rebalancing premium?
- Is it possible to reap a substantial rebalancing premium in equity markets in practice?

A common thread to the first two chapters was the excess growth rate term and a constant time horizon. In the third chapter, we tackle a portfolio optimization problem in a retirement context where the time horizon is stochastic. The main goal of this third chapter will be to present a formal analysis of efficient investment strategies for individuals facing long-term care risk in the decumulation phase of their life-cycle. Our main contribution is to develop a comprehensive and flexible framework for providing personalized advice to retirement investment decisions in the presence of life event risk, with a relatively rich menu of investment opportunities that includes balanced funds, target date funds, but also various types of annuity products for which we use realistic market quotes.

The investment problem for individuals and households in the decumulation phase can be broadly defined as a combination of consumption and bequest goals, subject to a dollar budget defined in terms of initial wealth. One of the key challenges for financial advisers is to provide personalized advice to individuals as far as their retirement investment decisions are concerned. Indeed, with the need to supplement public and private retirement benefits via voluntary contributions, the so-called third pillar of pension systems, individuals are becoming increasingly responsible for their own retirement savings and investment decisions. This global trend poses substantial challenges to individuals, who often lack the expertise required to make such complex financial decisions in the face of multiple sources of uncertainty that include longevity risk, market and inflation risks as well as risk of unforeseen expenses due to life events such as long-term care needs.

In principle, annuity products are attractive since they can be used to generate a target level of replacement income throughout retirement and can therefore solve the problem of longevity,

market and inflation risks. This intuition has been confirmed by many papers, including Pfau (2013), Huang et al. (2012), Milevsky (2013), Milevsky and Macqueen (2015), Vernon (2014, 2015) and Pang and Warshawsky (2012) inter alia, which have demonstrated the value of annuity contracts in a retirement investment context. In practice, however, the demand for such products is extremely low despite their risk-free nature in a retirement investment context. Using the RAND Health and Retirement Study dataset for the cohort aged 65-75 in 1998, Pashchenko (2013) for example reports that only 5% of individuals receive income from annuities, with a peak at 12.2% among the highest income quintile and a low at 0.4% for the lowest quintile. Common explanations of the "annuity puzzle" are related to the fact that annuities involve counterparty risk and high levels of fees, and also that they do not contribute to bequest objectives. One additional key drawback of annuity products is their severe lack of flexibility. Indeed, annuitization is an almost irreversible decision, unless one is willing to bear the costs of high surrender charges, which can amount to several percentage points of the invested capital.

One of the main goals of this third chapter will be to assess whether the opportunity cost of such substantial rigidity may partly explain the annuity puzzle, and to provide a quantitative measure of the associated impact on the optimal demand for annuities. To perform this analysis, we introduce a simulation framework that incorporates a market simulation engine, involving Monte-Carlo simulations coupled with realistic long-term capital market assumptions (CMAs), a product simulation engine incorporating scenarios for stocks, bonds as well as annuity products, and an individual investor simulation engine incorporating mortality risk scenarios, as well as target levels of replacement income cash-flows including random shocks to replacement income needs due to life events such as long-term care needs. To sum up, the main tasks of the framework will be:

1. To model market risks (stocks, bonds, short term) via Monte-Carlo simulations,
2. To model longevity risk, life event risk and to define a target level of replacement income for retirement including the risk of life event occurrence,
3. To define a broad investment universe including: equities, bonds, balanced funds, target date funds and an exhaustive range of annuities,
4. To determine which is the optimal portfolio among the different possible investment portfolios for retirement with respect to an objective function to be specified.

The framework main parameters will be the following: age, sex, initial wealth, initial withdrawal rate (as a percentage of initial wealth), cost of living adjustment (COLA in short), risk aversion of the individual, market risk model parameters, mortality distribution, life event occurence age distribution, nursing home cost in case of life event occurence, investment universe, number of Monte Carlo simulations, objective function to maximize and step of the grid of weights on which the objective function is calculated.

Several authors have examined different frameworks with annuities to determine the optimal strategy for retirees, based on trade-offs supporting minimum spending needs and preserving financial assets. Pfau (2013) considers an asset allocation among stocks, bonds, inflation-adjusted single-premium immediate annuities (SPIAs), fixed SPIAs and variable annuities with guaranteed living benefit riders (VA/GLWBs). He finds that an optimal retirement income portfolio generally consists of stocks and fixed SPIAs. In the same vein, Huang et al. (2012) introduce the Efficient Income Frontier (EIF) concept based on the trade-off between income risk and legacy potential with asset allocation among traditional investments, SPIAs and variable annuities with guaranteed living benefit riders (VA/GLWBs). They suggest that the inclusion of annuities in a retiree's portfolio reduces income risk, increases legacy potential or increases the amount of income that can likely be sustained throughout retirement. Milevsky (1998) complements this body of research by analyzing the optimal age at which to annuitize and concluding that the decision to do so should come relatively late in retirement.

Milevsky (2013) and Milevsky and Macqueen (2015) develop the concept of product allocation, a technique that seeks to hedge against key retirement risks by allocating savings across a variety of investment and insurance-based products. They design their optimal asset allocation with traditional investments, income annuities (immediate or deferred income) and deferred annuities with a guaranteed withdrawal benefit rider and find that the trade-off between income sustainability and financial legacy is at the heart of the product allocation strategy.

Vernon (2014, 2015) develops an analytical framework to determine the optimal retirement income generators (RIGs) for retirees that include: drawing from Social Security, a Systematic Withdrawal Plan, buying a guaranteed lifetime annuity, a reverse mortgage, interest and dividend income and real estate rental income. He observes that (1) no single retirement income generator delivers on all potential retirement goals, (2) clients will need to prioritize and make trade-offs between various retirement goals, and (3) annuities can be used strategically to improve retirement security. Pang and Warshawsky (2012) compared the following six strategies: (1) systematic withdrawal from mutual funds, (2) fixed payout immediate life annuity, (3) immediate variable annuity for life, (4) variable annuity (VA in short) plus guaranteed minimum withdrawal benefit (GMWB in short), (5) systematic withdrawal from mutual funds plus fixed payout immediate life annuity, and (6) systematic withdrawal from mutual funds plus fixed payout life annuity, with gradual annuitization over time. They find that none of the six strategies considered dominates the others and that the best advice may be to segment wealth to establish minimum necessary income and hedge longevity risk, then focus on growth.

Overall these authors emphasize the added value of the annuity in a retirement portfolio context in different frameworks. In this chapter our modulable framework will allow us to test the potential benefits of annuities in the decumulation phase documented in the literature and also the robustness of these benefits when life events are taken into account.

Chapter

1 Maximizing an Equity Portfolio Excess Growth Rate

Abstract

It has been claimed that, for dynamic investment strategies, the simple act of rebalancing a portfolio can be a source of additional performance, sometimes referred to as the volatility pumping effect or the diversification bonus because volatility and diversification turn out to be key drivers of the portfolio performance. Stochastic portfolio theory suggests that the portfolio excess growth rate, defined as the difference between the portfolio expected growth rate and the weighted-average expected growth rate of the assets in the portfolio, is an important component of this additional performance (see Fernholz and Shay (1982), Fernholz (2002) and Gabay and Herlemont (2007) inter alia). In this context, one might wonder whether maximizing a portfolio excess growth rate would lead to an improvement in the portfolio performance or risk-adjusted performance. This chapter provides a thorough empirical analysis of the maximization of an equity portfolio excess growth rate in a portfolio construction context for individual stocks. In out-of-sample empirical tests conducted on individual stocks from 4 different regions (US, UK, Eurozone and Japan), we find that portfolios that maximize the excess growth rate are characterized by a strong negative exposure to the low volatility factor and a higher than 1 exposure to the market factor, implying that such portfolios are attractive alternatives to competing smart portfolios in markets where the low volatility anomaly does not hold (e.g., in the UK, or in rising interest rate scenarios) or in bull market environments. In parallel, our empirical analysis of US factor-tilted universes shows an outperformance in terms of mean return, growth rate and Sharpe ratio for portfolios that maximize the excess growth rate with respect to equally-weighted portfolios for 4 stock selections, namely past winners, past losers, high vol and high investment stocks. These results suggest in particular that maximizing a portfolio excess growth rate is a welfare improving approach for momentum portfolios.

This chapter is the fruit of a collaboration with Lionel Martellini [1]. A part of this chapter led to the publication of an article in *Quantitative Finance* in July 2020.

1.1 Introduction

It has been claimed that, for dynamic investment strategies, the simple act of rebalancing a portfolio can be a source of additional performance, sometimes referred to as the volatility pumping effect or the diversification bonus because volatility and diversification turn out to be key drivers of the portfolio performance. Stochastic portfolio theory suggests that the portfolio excess expected growth rate (or excess growth rate in short), denoted by $g^\star$ and defined as the difference between the portfolio expected growth rate and the weighted-average expected growth rate of the assets in the portfolio, is an important component of this additional performance (see Fernholz and Shay (1982), Fernholz (2002) and Gabay and Herlemont (2007) inter alia). In this context, one might wonder whether maximizing a portfolio excess growth rate would allow one to substantially increase its risk-adjusted performance. Besides, if all assets have the same expected growth rate, then maximizing the portfolio excess growth rate is equivalent to maximizing the portfolio expected growth rate, which in turn is well-known to be equivalent to maximizing expected utility for an investor endowed with logarithmic preferences. Interestingly, the excess growth rate maximizing portfolio, just like the minimum variance, equal risk contribution or maximum diversification portfolios, only requires estimates for the covariance matrix, and not for the vector of expected returns.

The main purpose of this chapter is to provide a thorough empirical analysis of the benefits of maximizing an equity portfolio excess growth rate in a context where the constituents are individual stocks. We formally define the Max $g^\star$ portfolio strategy as the weighting scheme that

[1] Professor of finance at EDHEC Business School and director of EDHEC-Risk Institute

maximizes the ex-ante excess growth rate of the portfolio in the absence of a risk-free asset, and the MVR[2] strategy as the risky asset portfolio, which combined with the risk-free asset, maximizes the ex-ante excess growth rate in a universe with a risk-free asset. We also compare the Max $g^\star$ and MVR portfolios to benchmark portfolios such as cap-weighted portfolios, equally-weighted portfolios, minimum variance portfolios, equal risk contribution portfolios, and maximum diversification portfolios. Our analysis relies upon an improved estimate of the covariance matrix based on an implicit factor model, with factors extracted from a principal component analysis and random matrix theory used in the robust selection of statistically relevant factors.

In out-of-sample empirical tests conducted on individual stocks from four different regions (US, UK, Eurozone and Japan), we find that portfolios that maximize the excess growth rate are characterized by a strong negative exposure to the low volatility factor and a higher than 1 exposure to the market factor, implying that such portfolios are attractive alternatives to competing smart portfolios in markets where the low volatility anomaly does not hold (e.g., in the UK, or in rising interest rate scenarios) or in bull market environments.

The rest of the chapter is organized as follows. In Section 1.2, we provide a detailed theoretical analysis of the properties of the excess growth rate maximizing Max $g^\star$ portfolio, its link with the MVR portfolio as well as its relationship with respect to other popular weighting schemes. In Section 1.3, we perform a series of thorough empirical tests on the basis of individual stocks in four different regions. Finally our conclusions and suggestions for further research can be found in Section 1.4.

1.2 Theoretical Analysis of the Max Excess Growth Rate Portfolio

We first introduce a general continuous-time framework, before presenting a detailed theoretical analysis of the portfolio excess growth rate. All the stochastic processes below are defined under the physical measure $\mathbb{P}$ and under a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, F_t)$ where Ω is a set of outcomes, $\mathcal{F}$ is a sigma-algebra that represents the set of all measurable events, $\mathbb{P}$ is a probability measure and F_t is a filtration.

1.2.1 General Theoretical Framework

As in Fernholz (2002) we define a standard Ito-process model for a financial market composed of n risky stocks with correlation matrix $C_t = (\rho_{i,j}(t))_{\substack{1\le i\le n\\ 1\le j\le n}}$, volatility vector $V_t = (\sigma_i(t))_{1\le i\le n}$, covariance matrix $\Sigma_t = (\sigma_{i,j}(t))_{\substack{1\le i\le n\\ 1\le j\le n}} = (\rho_{i,j}(t)\sigma_i(t)\sigma_j(t))_{\substack{1\le i\le n\\ 1\le j\le n}}$, and weight vector $w_t = (w_i(t))_{1\le i\le n}$. We consider stock prices $S_i(t)$ as processes verifying the following stochastic differential equations:

$$dS_i(t) = S_i(t)\left(\mu_i(t)dt + \sum_{\nu=1}^{m}\beta_{i,\nu}(t)dW_\nu(t)\right),\ i = 1,\dots,n \tag{1.1}$$

where $(W_1(t),\dots,W_m(t))$ is a vector of m-independent Brownian motions ($m \ge n$), $\mu(t) = (\mu_1(t),\dots,\mu_n(t))$ is the vector of stocks' expected return and $\beta(t) = \left(\beta_{i,\nu}(t)\right)_{\substack{1\le i\le n\\ 1\le \nu\le m}}$ is the $(n\times m)$ matrix of volatilities representing the sensitivity of each stock to the m sources of uncertainty.

The Ito lemma applied to stock prices shows that (1.1) is equivalent to

$$d\left(\ln S_i(t)\right) = g_i(t)dt + \sum_{\nu=1}^{m}\beta_{i,\nu}(t)dW_\nu(t),\ i = 1,\dots,n \tag{1.2}$$

[2]MVR stands for Maximum Volatility Return and our definition is derived from Mantilla-Garcia (2016) portfolio excess growth rate optimization problem.

where $g_i(t) = \mu_i(t) - \frac{1}{2}\sigma_{i,i}(t)$ are the individual expected growth rates of the stocks. The covariance matrix Σ_t is linked to the $\beta(t)$ matrix as follows:

$$\sigma_{i,j}(t) = \sum_{\nu=1}^{m} \beta_{i,\nu}(t)\beta_{j,\nu}(t) \tag{1.3}$$

In this setup, we define a strategy at time t where the investor put a fraction $w_i(t)$ of his wealth in asset i. The value of the investor portfolio at time t is $P(t) = \sum_{i=1}^{n} w_i(t)S_i(t)$ with $\sum_{i=1}^{n} w_i(t) = 1$. Fernholz (2002) establishes that the value of the portfolio $P(t)$ follows the process:

$$dP(t) = P(t)\left(\sum_{i=1}^{n} w_i(t).\frac{dS_i(t)}{S_i(t)}\right) = P(t)\left(\mu(t)dt + \sum_{\nu=1}^{m} \beta_\nu(t)dW_\nu(t)\right) \tag{1.4}$$

with $\mu(t) = \sum_{i=1}^{n} w_i(t)\mu_i(t)$ denoting the expected return of the portfolio, and $\beta_\nu(t) = \sum_{i=1}^{n} w_i(t)\beta_{i,\nu}(t)$ denoting the volatility coefficients of the portfolio. The variance of the portfolio is given by:

$$\sigma^2(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i(t)w_j(t)\sigma_{i,j}(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i(t)w_j(t)\sum_{\gamma=1}^{m} \beta_{i,\gamma}(t)\beta_{j,\gamma}(t) = \sum_{\gamma=1}^{m} \beta_\gamma^2(t) \tag{1.5}$$

Where $\beta_\gamma = \beta_{\gamma,\gamma}$.

The Ito lemma applied to the portfolio value shows that (1.4) is equivalent to:

$$d\left(\ln P(t)\right) = \left(\mu(t) - \frac{1}{2}\sum_{\gamma=1}^{m} \beta_\gamma^2(t)\right)dt + \sum_{\nu=1}^{m} \beta_\nu(t)dW_\nu(t) = g(t)dt + \sum_{\nu=1}^{m} \beta_\nu(t)dW_\nu(t) \tag{1.6}$$

where $g(t) = \mu(t) - \frac{1}{2}\sum_{\gamma=1}^{m} \beta_\gamma^2(t) = \mu(t) - \frac{1}{2}\sigma^2(t)$ is the expected growth rate of the portfolio.

We have from Equation (1.2) that $\sum_{i=1}^{n} w_i(t)d\left(\ln S_i(t)\right) = \sum_{i=1}^{n} w_i(t)g_i(t)dt + \sum_{i=1}^{n}\sum_{\gamma=1}^{m} w_i(t)\beta_{i,\gamma}(t)dW_\gamma(t)$ and then we can rewrite Equation (1.6) as:

$$\begin{aligned} d\left(\ln P(t)\right) &= \sum_{i=1}^{n} w_i(t)d\left(\ln S_i(t)\right) + g(t)dt - \sum_{i=1}^{n} w_i(t)g_i(t)dt \\ &= \sum_{i=1}^{n} w_i(t)d\left(\ln S_i(t)\right) + g^\star(t)dt \end{aligned} \tag{1.7}$$

where $g^\star(t)$ is labeled as the portfolio excess growth rate and defined as the expected growth rate of the portfolio $g(t)$ minus the weighted-average of the growth rates of its constituents. We show below that the excess growth rate can be rewritten as a function of the covariance matrix

and the portfolio weights:

$$\begin{aligned} g^\star(t) &= g(t) - \sum_{i=1}^{n} w_i(t) g_i(t) \\ &= \sum_{i=1}^{n} w_i(t)\mu_i(t) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i(t)w_j(t)\sigma_{i,j}(t) - \left(\sum_{i=1}^{n} w_i(t)\mu_i(t) - \frac{1}{2}\sum_{i=1}^{n} w_i(t)\sigma_{i,i}(t)\right) \\ &= \frac{1}{2}\left(\sum_{i=1}^{n} w_i(t)\sigma_{i,i}(t) - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i(t)w_j(t)\sigma_{i,j}(t)\right) \end{aligned} \tag{1.8}$$

Finally the excess growth rate can be written as half of the difference between the weighted-average variance of the assets in the portfolio and the portfolio variance:

$$g^\star(t) = \frac{1}{2}\left(w_t^\top (V_t \odot V_t) - w_t^\top \Sigma_t w_t\right) \tag{1.9}$$

$\odot$ refers to the Hadamard product: for two matrices M and N of the same dimension $(m \times n)$, $M \odot N$ is a matrix of the same dimension as M and N with elements defined as follows: $\forall (i,j) \in [\![1,m]\!] \times [\![1,n]\!], (M \odot N)_{i,j} = M_{i,j} \times N_{i,j}$.

Besides, if all assets have the same expected growth rate, then maximizing the portfolio excess growth rate is equivalent to maximizing the portfolio expected growth rate, which in turn is equivalent to maximizing expected utility for an investor endowed with logarithmic preferences, i.e. who wants to maximize his expected log-wealth (Latané (1959)). In an empirical context, one might wonder whether adopting a portfolio strategy that periodically maximizes the ex-ante portfolio excess growth rate would allow one to obtain a risk-adjusted performance similar to well-known portfolio strategies such as equally-weighted or equal risk contribution.

1.2.2 Maximization of the Excess Growth Rate Term

We give a proof of the positivity of the excess growth rate for a long-only portfolio in Appendix 1.A for a general setting involving n_t risky assets with correlation matrix $C_t = (\rho_{i,j}(t))_{\substack{1\le i\le n_t \\ 1\le j\le n_t}}$, volatility vector $V_t = (\sigma_i(t))_{1\le i\le n_t}$, covariance matrix $\Sigma_t = (\sigma_{i,j}(t))_{1\le i\le n_t 1\le j\le n_t} = (\rho_{i,j}(t)\sigma_i(t)\sigma_j(t))_{\substack{1\le i\le n_t \\ 1\le j\le n_t}}$ and weights $w_t = (w_i(t))_{1\le i\le n_t}$.

In the particular case of a portfolio made of two risky assets, for each time t the portfolio with the highest excess growth rate is the equally-weighted portfolio. In our more general setting with n_t risky assets, if we assume that all the volatilities at time t are identical (equal to σ_t), the excess growth rate simplifies as:

$$g_t^\star = \frac{\sigma_t^2}{2}\left(1 - \frac{1}{ENC(w_t)} - 2\sum_{i=1}^{n_t}\sum_{\substack{j=1 \\ j>i}}^{n_t} w_i(t)w_j(t)\rho_{i,j}(t)\right) \tag{1.10}$$

where the effective number of constituents of the portfolio $ENC(w_t)$ is defined as follows[3]:

$$ENC(w_t) = \frac{1}{\|w_t\|^2} \tag{1.11}$$

In this last case, the excess growth rate increases with the volatility σ_t, increases with the $ENC(w_t)$ and decreases with the pairwise correlations $\rho_{i,j}(t)$, suggesting that the portfolio with the highest excess growth rate is a well diversified (at least with a relatively high effective number

[3] $\|.\|$ denotes the Euclidian norm.

of constituents) portfolio with highly volatile and weakly correlated assets. Under the specific assumption of identical volatility for all the assets, the Max $g^\star$ portfolio is equivalent to other well-known 'smart' weighting schemes such as the Global Minimum Variance and the Maximum Diversification portfolios. We notice that if we also assume that all the assets have the same pairwise correlation at time t, then the portfolio with the highest excess growth rate is the equally-weighted portfolio, which is also the portfolio with the highest effective number of constituents.

In the general setting defined above, the excess growth rate $g_t^\star$ can be written as:

$$g_t^\star = \frac{1}{2}\left(w_t^\top (V_t \odot V_t) - w_t^\top \Sigma_t w_t\right) \tag{1.12}$$

and the optimization problem of the excess growth rate maximization of a portfolio $w_t^\star$ can be formulated as:

$$\begin{aligned} w_t^\star = \arg\max_{w_t} & \left(w_t^\top (V_t \odot V_t) - w_t^\top \Sigma_t w_t\right) \\ u.c.\ & w_t^\top \mathbb{1}_{n_t} = 1 \end{aligned}$$

The Lagrange function of the optimization problem is:

$$\mathcal{L}(w_t;\lambda) = \frac{1}{2}\left(w_t^\top (V_t \odot V_t) - w_t^\top \Sigma_t w_t\right) - \lambda(w_t^\top \mathbb{1}_n - 1) \tag{1.13}$$

The solution $w_t^\star$ satisfies the following first-order conditions:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial w_t}(w_t^\star;\lambda) = \frac{1}{2}(V_t \odot V_t) - \Sigma_t w_t^\star - \lambda \mathbb{1}_{n_t} = \mathbb{0}_{n_t} \\ \frac{\partial \mathcal{L}}{\partial \lambda}(w_t^\star;\lambda) = w_t^{\star\top} \mathbb{1}_{n_t} - 1 = 0 \end{cases} \tag{1.14}$$

From the first equation of (1.14), we have:

$$\begin{aligned} \lambda \mathbb{1}_{n_t} &= \frac{1}{2}(V_t \odot V_t) - \Sigma_t w_t^\star \\ \lambda\left(\mathbb{1}_{n_t}^\top \Sigma_t^{-1} \mathbb{1}_{n_t}\right) &= \frac{1}{2}\mathbb{1}_{n_t}^\top \Sigma_t^{-1}(V_t \odot V_t) - \mathbb{1}_{n_t}^\top w_t^\star \end{aligned}$$

From the second equation of (1.14), we have:

$$w_t^{\star\top} \mathbb{1}_{n_t} = \mathbb{1}_{n_t}^\top w_t^\star = 1 \tag{1.15}$$

We retrieve from above:

$$\lambda = \frac{1}{2}\frac{\mathbb{1}_{n_t}^\top \Sigma_t^{-1}(V_t \odot V_t) - 2}{\mathbb{1}_{n_t}^\top \Sigma_t^{-1} \mathbb{1}_{n_t}} \tag{1.16}$$

We reinject the value of λ in the first equation of (1.14):

$$\Sigma_t w_t^\star = \frac{1}{2}(V_t \odot V_t) - \lambda \mathbb{1}_{n_t} \tag{1.17}$$

Finally, in the stylized setting with n_t risky assets, no risk-free asset and no assumptions on the covariance matrix structure, the weight vector $w_t^\star = (w_1(t), ..., w_n(t))$ that maximizes the excess growth rate $g_t^\star$ is:

$$w_t^\star = \frac{1}{2}\left(\Sigma_t^{-1}(V_t \odot V_t) - \frac{\mathbb{1}_{n_t}^\top \Sigma_t^{-1}(V_t \odot V_t) - 2}{\mathbb{1}_{n_t}^\top \Sigma_t^{-1} \mathbb{1}_{n_t}} \Sigma_t^{-1} \mathbb{1}_{n_t}\right) \tag{1.18}$$

where Σ_t is the $(n_t \times n_t)$ covariance matrix, V_t the $(n_t \times 1)$ volatility vector and $\mathbb{1}_{n_t}$ a $(n_t \times 1)$ vector of ones. This analytical solution can lead to possibly negative weights, so in our long-only framework we have to compute the solution numerically.

If we now assume that every asset has the same volatility σ_t and the same pairwise correlation ρ_t, we have $\Sigma_t = \sigma_t^2 C_t = \rho_t J_{n_t} + (1-\rho_t) I_{n_t}$ with I_{n_t} the $(n_t \times n_t)$ identity matrix and $J_{n_t} = \mathbb{1}_{n_t}^\top \mathbb{1}_{n_t}$. If ρ_t is different from 1, the inverse of C_t is:

$$C_t^{-1} = \frac{1}{\sigma_t^2}\Sigma_t^{-1} = \frac{1}{1-\rho_t}\left(I_{n_t} - \frac{\rho_t}{\rho_t(n_t-1)+1} J_{n_t}\right) \tag{1.19}$$

From Equation (1.18), after having noticed that $V_t \odot V_t = \sigma_t^2 \mathbb{1}_{n_t}$ and $\Sigma_t^{-1} = \frac{1}{\sigma_t^2} C_t^{-1}$, we have:

$$\begin{aligned} w_t^\star &= \frac{1}{2}\left(\Sigma_t^{-1}(V_t \odot V_t) - \frac{\mathbb{1}_{n_t}^\top \Sigma_t^{-1}(V_t \odot V_t) - 2}{\mathbb{1}_{n_t}^\top \Sigma_t^{-1} \mathbb{1}_{n_t}} \Sigma_t^{-1} \mathbb{1}_{n_t}\right) \\ &= \frac{1}{2}\left(C_t^{-1}\mathbb{1}_{n_t} - \frac{\mathbb{1}_{n_t}^\top C_t^{-1}\mathbb{1}_{n_t} - 2}{\mathbb{1}_{n_t}^\top C_t^{-1}\mathbb{1}_{n_t}} C_t^{-1}\mathbb{1}_{n_t}\right) \\ &= \frac{1}{\mathbb{1}_n^\top C_t^{-1}\mathbb{1}_{n_t}} C_t^{-1}\mathbb{1}_n \end{aligned}$$

We use Equation (1.19) to simplify the following quantities:

$$\begin{aligned} C_t^{-1}\mathbb{1}_{n_t} &= \frac{1}{1-\rho_t}\mathbb{1}_{n_t} - \frac{n_t\rho_t}{(1-\rho_t)(\rho_t(n_t-1)+1)}\mathbb{1}_{n_t} \\ &= \frac{1}{\rho_t(n_t-1)+1}\mathbb{1}_{n_t} \\ \mathbb{1}_{n_t}^\top C_t^{-1}\mathbb{1}_{n_t} &= \frac{n_t}{\rho_t(n_t-1)+1} \end{aligned} \tag{1.20}$$

We finally obtain:

$$\begin{aligned} w_t &= \frac{\frac{1}{\rho_t(n_t-1)+1}}{\frac{n_t}{\rho_t(n_t-1)+1}}\mathbb{1}_{n_t} \\ &= \frac{1}{n_t}\mathbb{1}_{n_t} \end{aligned}$$

In the stylized setting made of n_t risky assets with the same volatility σ_t and pairwise correlation ρ_t, the portfolio displaying the highest excess growth rate is the equally-weighted portfolio.

If we now introduce a risk-free asset in the general setting with n_t risky assets as in Mantilla-Garcia (2016) then we can show that the portfolio that maximizes the excess growth rate is a mixed allocation between the risk-free asset and a risky asset portfolio. We define this risky asset portfolio as the MVR portfolio and give a detailed solution of the MVR optimization problem in Appendix 1.B. The weights of the MVR portfolio verify[4]:

$$w_t^{MVR} = \frac{\Sigma_t^{-1}(V_t \odot V_t)}{\mathbb{1}_{n_t}^\top \Sigma_t^{-1}(V_t \odot V_t)} \tag{1.21}$$

If we only assume a constant and equal pairwise correlation amongst all the risky assets but possibly different volatilities, then the MVR portfolio is a linear combination of an Equally-Weighted portfolio and an Equal Risk Contribution portfolio (see Mantilla-Garcia (2016) for further details). It is known that the analytical formula of the risky portfolio that maximizes the Sharpe ratio (i.e., the ratio of the expected excess return of the portfolio $\mu(t) - r_f(t)$ over

[4] As for the Max $g^\star$ portfolio, the analytical solution of the MVR portfolio can lead to possibly negative weights, so in our long-only framework we have to compute the solution numerically.

the volatility of the portfolio $\sigma(t)$) is given by $\frac{\Sigma_t^{-1}\tilde{\mu}_{n_t}}{\mathbb{1}_{n_t}^\top \Sigma_t^{-1}\tilde{\mu}_{n_t}}$ where $\tilde{\mu}_{n_t}$ is the vector of expected excess returns and Mantilla-Garcia (2016) put forward that the MVR portfolio coincides with the maximum Sharpe ratio portfolio if expected returns are proportional to return variances (i.e. if the vectors $\tilde{\mu}_{n_t}$ and $V_t \odot V_t$ are collinear).

It is crucial to note that the MVR and the Max $g^\star$ portfolios differ since they do not correspond to the same optimization problem. The Max $g^\star$ portfolio is formally defined as the weighting scheme that maximizes the ex-ante excess growth rate of the portfolio in a universe of n_t risky assets. In the augmented universe with a risk-free asset, the MVR portfolio is not the portfolio with the highest ex-ante excess growth rate: it is a building block of the portfolio that maximizes the ex-ante excess growth rate, the other building block being the risk-free asset (see Appendix 1.B for more details).

1.2.3 Overview of Other Weighting Schemes, Optimization Constraints and Reporting Indicators

1.2.3.1 Other Weighting Schemes

In addition to the Max $g^\star$ and MVR portfolios, we also present results for the following weighting schemes used for benchmarking purposes:

- Capitalization Weighted (CW), which weights stocks by their capitalization.
- Equally-Weighted (EW), which attributes the same weight to each constituent in the investment universe.
- Global Minimum Variance (GMV), which aims to minimize the ex-ante volatility by using the covariance matrix structure.
- Equal Risk Contribution (ERC), which equalizes the contribution of each constituent to the portfolio volatility (see Maillard et al. (2010) for details).
- Maximum Diversification (MD), which aims at maximizing the diversification ratio defined as the weighted average of the volatilities of the constituents over the portfolio volatility (see Choueifaty and Coignard (2008) for details).

Each portfolio is rebalanced at quarterly intervals in March, June, September and December. The rebalancing dates coincide with the third Friday of the corresponding month. The CW portfolio is computed by extracting the stocks' capitalization. The EW portfolio computation is straightforward. On the other hand the covariance matrix needs to be estimated for the ERC, GMV, MD, MVR, and Max $g^\star$ portfolios. The main challenge in covariance matrix estimation is the curse of dimensionality: because the number of stocks in the investment universe can be large (up to 500 stocks) and exceeds the number of observations, the historically-estimated sample covariance matrix will be non-invertible ; this is particularly disturbing since weighting schemes such as the minimum variance GMV or Max $g^\star$ investor's unconstrained portfolios are functions of the inverse of the covariance matrix. We will therefore estimate the covariance matrix at each step by using a statistical factor model where each factor is modeled as a linear combination of returns of the stocks. We use principal component analysis to extract the factors and consider the past two-year weekly returns (104 data points): we thus extract the factors from the data rather than using explicit factors. This choice is prompted by the desire to avoid imposing any view on which factors matter most in explaining the variability of stock returns. We then limit the number of statistical factors using a criterion from Random Matrix Theory in order to achieve parsimony and robustness (see Plerou et al. (1999) and Amenc et al. (2011) for more details).

1.2.3.2 Optimization Constraints

We constrain the ERC portfolio weights to be positive. We also constrain the GMV, MD, MVR and Max $g^\star$ portfolio weights to be positive and to ensure that their effective number of constituents (ENC in short) is greater than or equal to $\frac{n_t}{3}$, where

$$ENC(w_t) = \frac{1}{\|w_t\|^2} \tag{1.22}$$

and where $\|.\|$ denotes the euclidian norm. An application of Cauchy-Schwarz inequality shows that $ENC(w_t) \leq n_t$, and that equality holds if, and only if, all weights are equal.[5] This second constraint is meant to ensure a minimum level of maive diversification in the portfolios. The benefits of weight constraints have been widely documented in the litterature (see Jagannathan and Ma (2003) for hard minimum and maximum weight constraints and DeMiguel et al. (2009) for flexible constraints applying to the norm of the weight vector).

1.2.3.3 Excess Growth Rate and Turnover

We compute the annualized out-of-sample excess growth rate between two quarterly rebalancing dates $g^\star_{oos}(t \to t+1)$ of each strategy for the different universes as follows:

$$g^\star_{oos}(t \to t+1) = 4 \times \left(g_{strat,t \to t+1} - \sum_{i=1}^{n_t} w_i(t) g_{i,t \to t+1} \right) \tag{1.23}$$

$g_{strat,t \to t+1}$ denotes the realized growth rate of the strategy between two rebalancing dates, n_t the number of stocks in the portfolio at time t, $w_i(t)$ the weight of stock i in the portfolio at time t and $g_{i,t \to t+1}$ the growth rate of asset i between two rebalancing dates.

We denote with $g_{wa}(t \to t+1)$ the annualized weighted average of the realized growth rates of the constituents for a given strategy between two rebalancing dates:

$$g_{wa}(t \to t+1) = 4 \times \sum_{i=1}^{n_t} w_i(t) g_{i,t \to t+1} \tag{1.24}$$

We finally report for every strategy, the averages across all periods of the out-of-sample excess growth rate and of the (cross-sectional) weighted average of the constituent realized growth rates. These indicators will allow us to assess if the Max $g^\star$ strategy, which by construction is the strategy with the highest ex-ante excess growth rate, has the highest average ex-post excess growth rate, to compare the magnitude of the out-of-sample excess growth rate of the Max $g^\star$ strategy with the out-of-sample excess growth rates of the MVR and the other strategies and to assess if raising the out-of-sample excess growth rate is achieved at the expense of lowering the out-of-sample (cross-sectional) average asset growth rate.

We will also report the annualized one-way turnover of all the strategies. The annualized one-way turnover is computed as:

$$\theta_t = \frac{1}{2} \left(\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N_t} |w_{i,t} - w_{i,t-}| \right) \times 4 \tag{1.25}$$

where $w_{i,t}$ is the weight of stock i on the rebalancing date t, $w_{i,t-}$ the effective weight just before the quarterly rebalancing date t, T the number of quarterly observations and N_t the number of available stocks before and after rebalancing date t. We will estimate the annual impact of the transaction costs impact on portfolio performance as twice the one-way annualized turnover times 50 basis points as in Balduzzi and Lynch (1999) and DeMiguel et al. (2009).

[5] Cauchy-Schwarz inequality states that for any two vectors $(x_1, ..., x_{n_t})$ and $(y_1, ..., y_{n_t})$, we have $\left(\sum_i x_i y_i\right)^2 \leq \left(\sum_i x_i^2\right) \times \left(\sum_i y_i^2\right)$. Equality is achieved if, and only if, one of the two vectors is zero or the two vectors are parallel. Here, we take $y_i = 1$.

1.3 Empirical Analysis of the Max Excess Growth Rate Portfolio with Individual Stocks

We consider in this section several universes of individual stocks from four developed regions: United States, Eurozone, UK and Japan. The individual stock returns data used in our study emanates from the Scientific Beta databases.[6]

1.3.1 Long Term US Equity Universe

We consider the ERI Scientific Beta Long-Term Track Record (LTTR in short) US Universe as our base case investment universe over the period June 1970 - December 2015. This universe is composed every quarter of the 500 US stocks with the largest market capitalization. The strategies studied are rebalanced quarterly and a calibration period of two years of weekly returns is used for the covariance matrix estimation when needed.

Panel (A) of Table 1.1 reports risk, performance and diversification indicators relative to the CW, EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes applied to this universe over the period June 1970 - December 2015. The Max $g^\star$ portfolio has the worst risk-adjusted performance of all the portfolios with a Sharpe ratio of 0.29 (versus 0.39 for the CW portfolio and 0.54 for the ERC portfolio) and displays the highest extreme risk of all the portfolios with a maximum drawdown of 77.7 %. The MVR portfolio exhibits a risk-adjusted performance (Sharpe ratio of 0.45) close to that of the EW portfolio (Sharpe ratio of 0.46). The growth rates of the MVR and Max $g^\star$ portfolios are lower than that of the EW portfolio (11.0% for MVR and 9.0% for the Max $g^\star$ versus 11.4% for the EW strategy). We acknowledge the domination of the MVR and Max $g^\star$ portfolios by the EW portfolio in terms of growth rate and risk-adjusted return: the heterogeneity of the universe or a high negative exposure to one or many rewarded risk factors could be possible explanations of these relative underperformances. We notice that the Max $g^\star$ allocation is the one with the highest out-of-sample average annualized excess growth rate (7.8%), followed by the MVR strategy (6.0%), which is coherent with their corresponding ex-ante optimization functions. The CW and EW strategies have out-of-sample average annualized growth rates that are significantly lower, at respectively 3.1% and 4.2%. We also note that the high out-of-sample average excess growth rates of the Max $g^\star$ and MVR strategies contrast with the low out-of-sample weighted average asset growth rate of these strategies (1.2% for Max $g^\star$, 4.9% for MVR, 7.0% for CW and 8.1% for ERC). The MD strategy has the highest annualized one-way turnover (77.2%) while the Max $g^\star$, MVR, GMV and EW strategies display 60.4%, 74.7%, 66.9% and 25.4% annualized one-way turnovers. The impact of transaction costs on the Max $g^\star$ portfolio mean performance will be $60.4 \times 2 \times 50 = 0.604\%$ while it will respectively be 0.254% and 0.051% for the EW and CW portfolios.

[6]The data for the US region is extracted from CRSP.

Table 1.1: Risk, Performance and Diversification Statistics on the US stock universe

(A) Time Period: June 1970 - December 2015

	CW	**EW**	**GMV**	**ERC**	**MD**	**MVR**	**Max $g^\star$**
Annualised Mean Total Return (%)	11.5	12.7	12.3	12.7	12.7	12.4	11.5
Annualised Volatility (%)	16.8	16.7	11.9	14.4	13.6	16.5	22.3
Annualised Growth Rate (%)	10.1	11.4	11.6	11.7	11.8	11.0	9.0
Sharpe Ratio	0.39	0.46	0.62	0.54	0.57	0.45	0.29
Average ENC (%)	26	100	33	68	33	33	33
Maxdrawdown (%)	54.3	58.3	43.9	53.0	50.1	59.6	77.7
Daily VaR 1% (%)	−2.8	−2.8	−2.0	−2.4	−2.4	−2.8	−3.9
An. excess growth rate (%)	3.1	4.2	2.7	3.5	4.2	6.0	7.8
Av. growth rates (%)	7.0	7.1	8.8	8.1	7.6	4.9	1.2
Turnover (%)	5.1	25.4	66.9	35.7	77.2	74.7	60.4

(B) Time Period: June 2002 - December 2015

	CW	**EW**	**GMV**	**ERC**	**MD**	**MVR**	**Max $g^\star$**
Annualised Mean Total Return (%)	9.5	11.3	11.0	11.2	12.2	12.4	12.3
Annualised Volatility (%)	19.7	21.3	15.3	18.7	17.4	21.1	26.1
Annualised Growth Rate (%)	7.5	9.1	9.9	9.5	10.7	10.2	8.9
Sharpe Ratio	0.41	0.47	0.63	0.53	0.63	0.53	0.42
Average ENC (%)	23	100	33	73	33	33	33
Maxdrawdown (%)	54.3	58.3	43.9	53.0	50.1	59.6	67.8
Daily VaR 1% (%)	−3.5	−3.8	−2.7	−3.3	−3.0	−3.9	−4.7
An. excess growth rate (%)	2.8	3.8	2.3	3.2	4.0	5.9	7.2
Av. growth rates (%)	5.8	6.7	8.6	7.5	8.0	5.8	3.8
Turnover (%)	5.2	24.9	63.5	33.9	74.0	71.5	59.2

This table reports the risk, performance and diversification statistics for the CW, EW, GMV, ERC, MD, MVR and Max $g^\star$ portfolios for the US universe over the periods June 1970 - December 2015 (Panel (A)) and June 2002 - December 2015 (Panel (B)). Av. growth rates denotes the annualized out-of-sample (cross-sectional) weighted average of the constituent growth rates averaged across all the period. The data comes from ERI Scientific Beta databases. The risk-free rate is the 3-month Tbill rate. We use daily total return series in US dollar currency.

In what follows, we examine two alternative rebalancing frequencies, monthly and annual, for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ strategies on the US base universe over the period June 1970 - December 2015. Panel (A) of Table 1.2 shows risk, performance and diversification indicators relative to the EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes applied to this universe for a monthly rebalancing frequency over the period June 1970 - December 2015. We assess that the annualized mean returns are close to those obtained when assuming a quarterly rebalancing frequency (12.9% versus 12.7% for the EW, 12.4% versus 12.3% for the GMV and 11.4% versus 11.5% for the Max $g^\star$). We also report the same statistics for an annual rebalancing frequency in Panel (B) of Table 1.2: the Sharpe ratios of the EW, GMV and Max $g^\star$ weighting schemes are respectively 0.47, 0.65 and 0.35 versus 0.46, 0.62 and 0.29 when considering a quarterly rebalancing frequency. As expected, a given strategy displays a higher turnover when rebalanced more frequently: for instance the annualized one-way turnovers of the Max $g^\star$ strategy respectively are 88.6%, 60.4% and 35.8% for monthly, quarterly and annual rebalancing frequencies. The impact of the transaction costs on the Max $g^\star$ mean performance will respectively be: $88.6 \times 2 \times 50 = 0.886\%$, 0.604% and 0.358%. Overall this additional analysis suggest that the results are robust with respect to changes in rebalancing frequency. A quarterly rebalancing frequency seems to be an acceptable trade-off between deviations to target weights and transaction costs.

Table 1.2: Risk, Performance and Diversification Statistics on the US Long Term Stock Universe with Monthly and Annual Rebalancing Frequencies

(A) Monthly Rebalancing Frequency

	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max $g^\star$**
Annualised Mean Total Return (%)	12.9	12.4	12.7	12.8	12.4	11.4
Annualised Volatility (%)	16.9	11.7	14.4	13.6	16.7	22.7
Annualised Growth Rate (%)	11.5	11.8	11.7	11.9	11.0	8.8
Sharpe Ratio	0.47	0.64	0.54	0.58	0.44	0.28
Average ENC (%)	100	33	66	33	33	33
Maxdrawdown (%)	57.9	44.7	53.1	50.9	59.1	79.0
Daily VaR 1% (%)	−2.9	−2.0	−2.5	−2.4	−2.8	−3.9
An. excess growth rate (%)	4.3	2.8	3.7	4.5	6.5	8.2
Av. growth rates (%)	7.1	9.0	8.1	7.4	4.5	0.6
Turnover (%)	40.6	106.3	56.7	123.5	115.9	88.6

(B) Annual Rebalancing Frequency

	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max $g^\star$**
Annualised Mean Total Return (%)	12.6	12.5	12.5	12.9	12.6	12.3
Annualised Volatility (%)	16.3	11.8	14.1	13.3	15.7	21.0
Annualised Growth Rate (%)	11.2	11.8	11.5	12.0	11.4	10.1
Sharpe Ratio	0.47	0.65	0.53	0.60	0.48	0.35
Average ENC (%)	100	33	65	33	33	33
Maxdrawdown (%)	56.7	43.5	52.0	47.0	55.0	70.2
Daily VaR 1% (%)	−2.8	−2.1	−2.4	−2.3	−2.7	−3.7
An. excess growth rate (%)	4.0	2.6	3.4	4.0	5.6	7.4
Av. growth rates (%)	7.1	9.1	8.0	8.0	5.7	2.6
Turnover (%)	14.0	36.3	19.5	41.6	41.4	35.8

This table reports the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and Max $g^\star$ portfolios for the US universe over the periods June 1970 - December 2015 with a monthly rebalancing frequency (Panel (A)) and an annual rebalancing frequency (Panel (B)). The data comes from ERI Scientific Beta databases. The risk-free rate is the 3-month Tbill rate. We use daily total return series in US dollar currency.

We now investigate the factor exposures of the different portfolios to see if the differences in performance are explained by different factor exposures. We focus on the well-academically documented following factors: market, size, momentum, value and low volatility. To observe the factor exposures of the portfolios, we first perform a five-factor OLS linear regression of the portfolios' excess returns. We use the yield on Secondary Market US Treasury Bills (3 months) as a proxy for the risk-free rate. The market factor is proxied by the CW strategy excess returns time series. The size, value, momentum and low volatility factors are proxied by returns time series of long/short equity portfolios. We borrow the series from the long-term ERI Scientific Beta database. Each long/short factor has a long leg, which is formed with the top or bottom 30% in terms of characteristic, and a short leg, which is the other bottom or top 30% group. The choice of the characteristic and the definition of the long leg is the following:

- Size: stocks are sorted on their market capitalization and the long leg is the 30% lowest market cap stocks group.
- Value: stocks are sorted on their book-to-market (BE/ME) ratio and the long leg is the 30% highest BE/ME group.
- Momentum: stocks are sorted on their past 52-week returns, skipping the last four weeks and the long leg is the 30% highest past winners group.
- Low volatility: stocks are sorted on their past 104-week volatility and the long leg is the 30% low volatility group.

The factor exposures of the different weighting schemes are reported in Table 1.3 (Panel (A) for the June 1970 - December 2015 period and Panel (B) for the June 2002 - December 2015 period). Over the June 1970 - December 2015 period, the Max $g^\star$ portfolio has a null alpha and negative exposures to the value, momentum and low volatility factors, which are positively rewarded in the long run (see Table 1.4 for the positivity and the magnitude of the risk premia). The MVR portfolio displays a positive annualized alpha of 0.76% but lower than that of the EW portfolio (0.97%) over the same period. As expected, a common pattern of the MVR and Max $g^\star$ portfolios is their negative exposures to the low volatility risk factor. Given that the low volatility risk factor is positively rewarded with a risk premium of 0.5% (see Table 1.4), the high negative exposure of the Max $g^\star$ strategy to this risk factor (-0.3760) besides its negative exposure to the positively rewarded value risk factor (-0.0476) mostly explain the low risk-adjusted performance of this strategy (Sharpe ratio of 0.29) over the June 1970 - December 2015 period.

Table 1.3: Factor Exposures on the US stock universe

(A) Time Period: June 1970 - December 2015

	CW	**EW**	**GMV**	**ERC**	**MD**	**MVR**	**Max $g^\star$**
Alpha (%)	0.00	0.00	0.01	0.00	0.01	0.00	0.00
		(2.85)	(2.85)	(2.64)	(2.64)	(1.05)	(0.00)
Market	1.0000	1.0085	0.8009	0.9273	0.8655	0.9592	1.0443
		(579.99)	(289.81)	(408.15)	(286.88)	(256.88)	(258.98)
Value	0.0000	0.1013	0.1016	0.1167	0.0821	0.0602	−0.0476
		(41.95)	(26.48)	(37.00)	(19.61)	(11.62)	(−8.50)
Size	0.0000	0.2514	0.2379	0.2661	0.2951	0.3334	0.3064
		(102.87)	(61.23)	(83.31)	(69.59)	(63.52)	(54.05)
Momentum	0.0000	−0.0378	0.0147	−0.0065	0.0151	−0.0166	−0.0876
		(−19.46)	(4.77)	(−2.57)	(4.46)	(−3.98)	(−19.44)
Low Vol	0.0000	−0.0118	0.1770	0.0997	0.0942	−0.0277	−0.3760
		(−7.31)	(69.07)	(47.33)	(33.66)	(−7.99)	(−100.64)
Adj- R2 (%)	100.0	98.1	90.7	95.7	91.5	91.2	94.4
Annualised alpha (%)	0.00	0.97	1.54	1.17	1.55	0.76	0.00

(B) Time Period: June 2002 - December 2015

	CW	**EW**	**GMV**	**ERC**	**MD**	**MVR**	**Max $g^\star$**
Alpha (%)	0.00	0.00	0.01	0.01	0.01	0.01	0.00
		(1.40)	(3.13)	(2.62)	(2.67)	(1.49)	(0.39)
Market	1.0000	0.9946	0.8348	0.9321	0.8518	0.8890	1.0260
		(408.88)	(227.39)	(356.88)	(197.08)	(172.10)	(173.00)
Value	0.0000	0.0117	0.0338	0.0379	−0.0050	0.0085	−0.0966
		(2.76)	(5.27)	(8.32)	(−0.66)	(0.95)	(−9.69)
Size	0.0000	0.1625	0.0826	0.1422	0.1143	0.1455	0.1642
		(47.21)	(15.90)	(38.47)	(18.69)	(19.91)	(20.30)
Momentum	0.0000	0.0099	0.0569	0.0334	0.0684	0.0412	−0.0034
		(3.33)	(12.65)	(10.45)	(12.92)	(6.52)	(−0.48)
Low Vol	0.0000	−0.0340	0.1936	0.0695	0.0456	−0.1240	−0.3773
		(−11.64)	(43.97)	(22.19)	(8.80)	(−20.02)	(−55.06)
Adj- R2 (%)	100.0	98.9	95.1	98.4	94.7	94.7	95.9
Annualised alpha (%)	0.00	0.78	2.62	1.56	2.63	1.76	0.51

This table reports the results of the regression of the daily excess returns time series of the CW, EW, GMV, ERC, MD, MVR and Max $g^\star$ weighting schemes over the market, value, size, momentum and low volatility factors. The universe considered is the US universe over two periods (June 1970 - December 2015 and June 2002 - December 2015).

Table 1.4: Risk Premia

	Market	Value	Size	Momentum	Low Vol
US June 1970 - December 2015 (%)	6.5	1.2	1.2	4.7	0.5
US June 2002 - December 2015 (%)	7.6	0.8	2.3	−0.4	−2.6
Eurozone (%)	6.0	2.6	2.1	1.5	1.0
UK (%)	6.6	-0.1	3.3	3.0	-3.2
Japan (%)	6.9	10.0	4.5	-0.9	0.8

This table displays the risk premia, estimated by the annualized mean return, of the market, value, size, momentum and low volatility risk factors for different universes. The universes considered are the US universe over the periods June 1970-December 2015 and June 2002 - December 2015; and the Eurozone, UK and Japan universes over the period June 2002-February 2017.

If we keep the same investment universe but change the investment period to June 2002-December 2015 we notice that the low volatility risk factor is negatively rewarded (-2.6%, see Table 1.4). The annualized mean return, the annualized volatility, the annualized growth rate and the Sharpe ratio statistics relative to the EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes for the US universe over this shorter investment period are given in Panel (B) of Table 1.1. We assess that the MVR and the Max $g^\star$ strategies display better growth rates (respectively 10.2% and 8.9% versus 7.5%) and Sharpe ratio (respectively 0.53 and 0.42 versus 0.41) than the CW portfolio, suggesting that the performances of these weighting schemes are enhanced when the low volatility risk premium is negative.

We will thereafter compare these results to those obtained in other regions where the low volatility risk factor (see Ang et al. (2009) for a low volatility premium study outside the United States) can be negatively rewarded.

1.3.2 Empirical Tests on Regional Universes

We apply the same experimental protocol as above to the following universes for the period June 2002 - February 2017:

- Eurozone stocks (300 stocks).
- UK stocks (100 stocks).
- Japanese stocks (500 stocks from June 2002 to June 2016 and 300 stocks from June 2016 on).

Table 1.5 shows the risk, performance and diversification indicators relative to the CW, EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes for the Eurozone (Panel (A)), UK (Panel (B)) and Japan (Panel (C)) universes over the period June 2002 - February 2017.

For the UK region, the MVR and the Max $g^\star$ portfolios, with respective Sharpe ratios of 0.50 and 0.41, have better risk-adjusted performances than the CW portfolio (Sharpe ratio of 0.35). They also exhibit higher growth rates (respectively 10.0% and 9.0%) than that of the EW portfolio (8.7%): these results, favorable to the MVR and the Max $g^\star$ strategies, discernibly contrast with the results obtained in the Eurozone universe. Indeed for the Eurozone the MVR and Max $g^\star$ portfolios, with respective Sharpe ratios of 0.20 and 0.15, lower growth rates of respectively 3.7% and 2.6%, and higher maximum drawdowns of 60.0% and 68.1%, both display lower risk-adjusted performances than the CW portfolio (Sharpe ratio of 0.29, growth rate of 5.5% and Maximum Drawdown of 58.3%). The growth rates of the MVR and Max $g^\star$ portfolios are also lower than that of the EW portfolio (6.8%). In the previous US universe over the June 1970 - December 2015 period, the risk-adjusted performance and the growth rate of the Max $g^\star$ strategy were sharply below those of the EW and even the CW strategies; while the MVR strategy displays risk and performance statistics close to those of the EW strategy. In the Japanese universe, the Max $g^\star$ portfolio exhibits a risk-adjusted performance (Sharpe ratio of 0.32) close to that of the

CW portfolio (Sharpe ratio of 0.31). The MVR portfolio has a risk-adjusted performance (Sharpe ratio of 0.42) equal to that of the EW portfolio. In that same universe, the MVR and Max $g^\star$ also have lower growth rates (respectively 6.4% and 4.8%) than that of the EW portfolio (6.8%). As for the US region, we note that for each region the high out-of-sample average excess growth rates of the Max $g^\star$ and MVR strategies contrast with the low out-of-sample weighted average asset growth rate of these strategies. The level of turnover obtained for the different strategies for a given region are consistent with the results of the US regions: for instance the strategy with the highest turnover is always the MD portfolio and the Max $g^\star$ portfolio turnover is always lower than those of the MVR and GMV portfolios.

Table 1.5: Risk, Performance and Diversification Statistics on the Eurozone, UK and Japan stock universes over the period June 2002 - February 2017

(A) Eurozone Universe (300 Stocks)

	CW	**EW**	**GMV**	**ERC**	**MD**	**MVR**	**Max $g^\star$**
Annualised Mean Total Return (%)	7.7	8.7	9.2	8.6	7.5	5.2	5.2
Annualised Volatility (%)	21.1	19.5	12.9	16.5	14.3	17.6	22.7
Annualised Growth Rate (%)	5.5	6.8	8.4	7.2	6.5	3.7	2.6
Sharpe Ratio	0.29	0.36	0.58	0.42	0.20	0.20	0.15
Average ENC (%)	31	100	33	63	33	33	33
Maxdrawdown (%)	58.3	61.9	50.1	59.0	53.2	60.0	68.1
Daily VaR 1% (%)	−3.7	−3.5	−2.4	−3.0	−2.7	−3.2	−4.0
An. excess growth rate (%)	3.0	4.2	3.2	3.8	5.2	7.0	8.1
Av. growth rates (%)	2.5	2.6	5.1	3.4	1.2	−3.4	−5.5
Turnover (%)	4.6	25.3	61.6	34.4	77.7	71.2	57.0

(B) UK Universe (100 Stocks)

	CW	**EW**	**GMV**	**ERC**	**MD**	**MVR**	**Max $g^\star$**
Annualised Mean Total Return (%)	9.0	10.5	10.4	10.5	10.7	11.8	11.8
Annualised Volatility (%)	19.1	18.9	14.4	17.0	15.6	19.4	23.4
Annualised Growth Rate (%)	7.2	8.7	9.3	9.1	9.5	10.0	9.0
Sharpe Ratio	0.35	0.44	0.56	0.49	0.54	0.50	0.41
Average ENC (%)	32	100	33	83	33	33	33
Maxdrawdown (%)	44.4	50.1	33.0	46.0	40.9	52.2	59.6
Daily VaR 1% (%)	−3.4	−3.4	−2.7	−3.1	−2.8	−3.4	−4.1
An. excess growth rate (%)	2.5	3.1	2.0	2.7	3.1	4.9	5.7
Av. growth rates (%)	4.6	5.6	7.3	6.3	6.4	5.0	3.3
Turnover (%)	3.9	21.9	62.2	34.1	75.5	70.5	58.2

(C) Japan (500 and then 300 stocks)

	CW	**EW**	**GMV**	**ERC**	**MD**	**MVR**	**Max $g^\star$**
Annualised Mean Total Return (%)	7.0	9.0	9.6	9.1	9.0	8.3	7.6
Annualised Volatility (%)	22.2	21.2	15.6	18.8	17.0	19.2	23.8
Annualised Growth Rate (%)	4.6	6.8	8.4	7.4	7.5	6.4	4.8
Sharpe Ratio	0.31	0.42	0.61	0.48	0.52	0.42	0.32
Average ENC (%)	24	100	33	73	33	33	33
Maxdrawdown (%)	60.1	57.8	41.6	51.8	48.0	58.8	71.1
Daily VaR 1% (%)	−3.7	−3.6	−2.8	−3.3	−3.2	−3.6	−4.2
An. excess growth rate (%)	2.9	3.7	2.7	3.3	4.0	5.6	6.5
Av. growth rates (%)	1.7	3.1	5.7	4.0	3.5	0.8	−1.7
Turnover (%)	6.0	28.8	65.7	36.4	78.8	74.1	61.6

This table reports the risk, performance and diversification statistics for the CW, EW, GMV, ERC, MD, MVR and Max $g^\star$ portfolios built on three different regional universes: Eurozone (Panel (A)), UK (Panel (B)) and Japan (Panel (C)). The data comes from ERI Scientific Beta databases. The risk-free rate is the 3-month Tbill rate. We use daily total return series in local currencies. The period considered is June 2002 - February 2017.

Panel (A) of Table 1.6 displays the factor exposures of the different weighting schemes for the Eurozone universe. The MVR and Max $g^\star$ portfolios have strong negative annualized alphas (respectively -3.35% and -3.79%) and negative exposures to the low volatility factor (respectively -0.0448 and -0.2776). The positivity of the low volatility risk premium (1.0%, see Table 1.4) is a source of explanation of the underperformance of the MVR and Max $g^\star$ portfolios with respect to the other weighting schemes. The factor exposures of the different weighting schemes for the UK universe are listed in Panel (B) of Table 1.6. The MVR and Max $g^\star$ portfolios have higher annualized alphas (respectively 2.30% and 0.96%) than the EW portfolio (0.37%) and display negative exposures to the low volatility factor (respectively -0.1340 and -0.2937). Unlike for the US (June 1970 - December 2015 period) and the Eurozone universes, the low volatility risk premium is negative (-3.2%, see Table 1.4). That particular configuration where high volatility stocks outperform low volatility stocks on the long run can explain the good performances of the MVR and Max $g^\star$ portfolios relatively to the other weighting schemes. The factor exposures of the different weighting schemes for the Japanese universe are displayed in Panel (C) of Table 1.6. The MVR and Max $g^\star$ portfolios have lower annualized alphas (respectively 0.18% and -0.87%) than the EW portfolio (0.32%). The low volatility risk premium is positive (0.8%, see Table 1.4) and the negative exposure of the Max $g^\star$ to this factor (-0.2120) can justify its relatively weak risk-adjusted performance and growth rate.

Table 1.6: Factor Exposures on the Eurozone, UK and Japan stock universes

(A) Eurozone Universe (300 Stocks)

	CW	EW	GMV	ERC	MD	MVR	Max $g^\star$
Alpha (%)	0.00	0.00	0.01	0.00	0.00	−0.01	−0.01
		(0.19)	(1.80)	(0.25)	(−0.38)	(−2.24)	(−2.50)
Market	1.0000	1.0378	0.8204	0.9538	0.8600	0.9480	1.0816
		(274.18)	(153.69)	(201.82)	(131.23)	(110.39)	(124.75)
Value	0.0000	−0.0035	−0.0161	0.0230	−0.0051	0.0307	−0.0377
		(−0.80)	(−2.57)	(4.16)	(−0.67)	(3.06)	(−3.72)
Size	0.0000	0.3225	0.2949	0.3414	0.3588	0.4194	0.4272
		(68.45)	(44.39)	(58.04)	(44.00)	(39.24)	(39.59)
Momentum	0.0000	−0.0047	0.0512	0.0169	0.0536	0.0470	−0.0038
		(−1.38)	(10.77)	(4.01)	(9.19)	(6.15)	(−0.49)
Low Vol	0.0000	−0.0472	0.1655	0.0681	0.0755	−0.0448	−0.2776
		(−11.48)	(28.54)	(13.27)	(10.61)	(−4.80)	(−29.47)
Adj- R2(%)	100.0	98.3	92.5	96.4	90.8	89.5	93.6
Annualised alpha (%)	0.00	0.12	1.68	0.20	−0.44	−3.35	−3.79

(B) UK Universe (100 Stocks)

	CW	EW	GMV	ERC	MD	MVR	Max $g^\star$
Alpha (%)	0.00	0.00	0.01	0.00	0.01	0.01	0.00
		(0.40)	(1.86)	(1.13)	(1.57)	(1.36)	(0.55)
Market	1.0000	1.0389	0.8648	0.9808	0.8611	0.9118	1.0038
		(236.09)	(146.61)	(224.32)	(129.58)	(114.15)	(121.19)
Value	0.0000	−0.0057	−0.0370	−0.0136	−0.0358	−0.0191	−0.0098
		(−1.12)	(−5.39)	(−2.68)	(−4.63)	(−2.11)	(−1.02)
Size	0.0000	0.2780	0.1377	0.2350	0.1641	0.2094	0.2547
		(55.80)	(20.62)	(47.48)	(21.82)	(23.16)	(27.16)
Momentum	0.0000	−0.0231	0.0216	−0.0108	−0.0114	−0.0565	−0.0555
		(−6.08)	(4.24)	(−2.85)	(−1.98)	(−8.19)	(−7.76)
Low Vol	0.0000	−0.0136	0.1692	0.0603	0.0546	−0.1340	−0.2937
		(−3.84)	(35.58)	(17.12)	(10.20)	(−20.81)	(−43.99)
Adj- R2 (%)	100.0	96.5	89.1	95.7	88.2	89.0	91.9
Annualised alpha (%)	0.00	0.37	2.33	1.05	2.21	2.30	0.96

(C) Japan Universe (500 and then 300 Stocks)

	CW	EW	GMV	ERC	MD	MVR	Max $g^\star$
Alpha (%)	0.00	0.00	0.01	0.00	0.01	0.00	0.00
		(0.47)	(2.37)	(1.12)	(1.24)	(0.14)	(−0.67)
Market	1.0000	1.0267	0.8734	0.9722	0.9036	0.9539	1.0362
		(366.72)	(211.60)	(324.75)	(189.55)	(181.89)	(197.52)
Value	0.0000	0.0318	−0.0099	0.0264	−0.0327	−0.0381	−0.0636
		(6.89)	(−1.45)	(5.36)	(−4.16)	(−4.41)	(−7.35)
Size	0.0000	0.2522	0.2357	0.2422	0.3048	0.3754	0.4229
		(61.89)	(39.23)	(55.58)	(43.93)	(49.18)	(55.39)
Momentum	0.0000	0.0064	0.0461	0.0089	0.0352	0.0334	0.0054
		(1.78)	(8.68)	(2.30)	(5.74)	(4.95)	(0.80)
Low Vol	0.0000	−0.0306	0.3045	0.1523	0.1882	0.0427	−0.2120
		(−7.31)	(69.07)	(47.33)	(33.66)	(7.08)	(−35.09)
Adj- R2 (%)	100.0	98.5	93.9	97.8	93.1	93.5	95.8
Annualised alpha (%)	0.00	0.32	2.41	0.82	1.45	0.18	−0.87

This table reports the results of the regression of the daily excess returns time series of the CW, EW, GMV, ERC, MD, MVR and Max $g^\star$ weighting schemes over the market, value, size, momentum and low volatility factors. The universes considered are the Eurozone, UK and Japan universes over the period June 2002 - February 2017.

Table 1.7 summarizes the main results on the US, Eurozone, UK and Japanese universes. We notice that, while for the Eurozone universe the MVR and the Max $g^\star$ strategies are outperformed by the CW strategy, the MVR strategy outperforms the CW strategy in terms of mean return, Sharpe ratio and growth rate in all the other universes. The Max $g^\star$ strategy outperforms the CW strategy in terms of mean return, Sharpe ratio and growth rate in the UK and Japanese universes but not in the US universe. We also assess that apart from the UK universe where the low volatility risk premia is negative, the MVR and the Max $g^\star$ portfolios display poor or mixed performances compared with the other "diversified" weighting schemes such as the EW strategy. The MVR strategy, and to a greater extent the Max $g^\star$ strategy, exhibit higher volatilities than the other well-known "alternative" weighting schemes (for instance GMV and ERC). Another characteristic that deserves to be underlined is the strong negative exposure of the Max $g^\star$ strategy to the low volatility risk factor and a high exposure to the market factor (higher than 1) for the four universes studied. The MVR displays to a lesser extent the same pattern with a negative exposure to the low volatility risk factor in 3 out of the 4 universes considered (the exception being the Japanese universe) and an exposure to the size risk factor higher than those of the EW, GMV and MD strategies for the three out of the four universes (the exception being the UK universe where the size exposure of the EW is higher than that of the MVR strategy). We also notice that the exposure to the market factor of the MVR strategy for the four universes is higher than that of the GMV and MD strategies.

Table 1.7: Summary

	US 1970-2015	**US 2002-2015**	**Eurozone**	**UK**	**Japan**
Max $g^\star$ outperforms CW in term of Sharpe ratio		X		X	X
MVR outperforms CW in term of Sharpe ratio	X	X		X	X
	US 1970-2015	**US 2002-2015**	**Eurozone**	**UK**	**Japan**
Max $g^\star$ outperforms CW in term of growth rate		X		X	X
MVR outperforms CW in term of growth rate	X	X		X	X
	US 1970-2015	**US 2002-2015**	**Eurozone**	**UK**	**Japan**
Max $g^\star$ outperforms CW in term mean return		X		X	X
MVR outperforms CW in term of mean return	X	X		X	X
	US 1970-2015	**US 2002-2015**	**Eurozone**	**UK**	**Japan**
Max $g^\star$ outperforms EW in term of Sharpe ratio					
MVR outperforms EW in term of Sharpe ratio		X		X	
	US 1970-2015	**US 2002-2015**	**Eurozone**	**UK**	**Japan**
Max $g^\star$ outperforms EW in term of growth rate				X	
MVR outperforms EW in term of growth rate		X		X	
	US 1970-2015	**US 2002-2015**	**Eurozone**	**UK**	**Japan**
Max $g^\star$ outperforms EW in term mean return		X		X	
MVR outperforms EW in term of mean return		X		X	

This table compares the performances of the MVR and Max $g^\star$ portfolios to that of the corresponding EW and CW portfolios for the different regions (US, Eurozone, UK and Japan) in terms of mean return, growth rate and Sharpe ratio.

1.3.3 Empirical Tests on Industry Universes

In this section, we test within the US universe, homogeneous selections of stocks based on the industry classification. The data emanates from the ERI Scientific Beta Live database and the universe is divided into ten industries following the Thomson Reuters classification: Energy, Basic Materials, Industrials, Cyclical Consumer, Non-Cyclical Consumer, Financials, Healthcare, Technology, Telecoms and Utilities. For a given industry, the EW weighting scheme is a natural benchmark for the different weighting schemes. Tables 1.8, 1.9, 1.10 and 1.11 present for each of the aforementioned industry, risk, performance and diversification statistics relative to the EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes. The results of the MVR and Max $g^\star$ portfolios are heterogeneous amongst the different universes: for instance in the Basic Materials and Technology industries these portfolios outperform the EW portfolio in terms of mean return, growth rate and Sharpe ratio while in the Energy and Industrials sectors they are outperformed by the EW.

Table 1.8: Risk, Performance and Diversification Statistics on US Industry Universe (Part1)

Universe:	**S0-Energy**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	13.7%	12.7%	13.4%	11.9%	13.4%	13.1%
An. Volatility	30.8%	22.6%	28.0%	25.4%	34.3%	34.2%
An.growth rate	9.0%	10.2%	9.4%	8.7%	7.5%	7.3%
Sharpe ratio	0.41	0.51	0.43	0.42	0.36	0.35
Average ENC	100%	33%	85%	33%	33%	33%
Maxdrawdown	65.3%	50.4%	60.0%	53.7%	69.1%	70.9%
Daily VaR 1%	-5.3%	-3.9%	-5.0%	-4.6%	-6.1%	-6.1%
An.excess growth rate	3.1%	2.3%	3.0%	3.7%	4.7%	4.7%
Universe:	**S1-Basic Materials**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	13.0%	10.9%	12.3%	12.4%	18.8%	17.7%
An. Volatility	24.4%	17.7%	21.6%	19.3%	27.0%	29.1%
An.growth rate	10.0%	9.4%	10.0%	10.6%	15.1%	13.5%
Sharpe ratio	0.48	0.55	0.51	0.58	0.65	0.57
Average ENC	100%	33%	84%	34%	33%	33%
Maxdrawdown	59.7%	41.9%	52.4%	41.4%	53.5%	63.5%
Daily VaR 1%	-4.3%	-2.9%	-3.8%	-3.2%	-4.8%	-5.1%
An.excess growth rate	3.4%	2.0%	3.1%	3.2%	4.7%	5.0%
Universe:	**S2-Industrials**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	13.0%	10.2%	12.2%	9.8%	11.6%	12.3%
An. Volatility	21.5%	17.1%	20.1%	19.1%	24.4%	26.2%
An.growth rate	10.7%	8.8%	10.2%	7.9%	8.6%	8.8%
Sharpe ratio	0.55	0.53	0.54	0.45	0.43	0.42
Average ENC	100%	33%	88%	33%	33%	33%
Maxdrawdown	59.5%	55.5%	58.0%	61.6%	69.3%	70.4%
Daily VaR 1%	-3.8%	-3.1%	-3.6%	-3.5%	-4.5%	-4.6%
An.excess growth rate	2.1%	1.3%	1.9%	2.2%	3.6%	3.8%

This table details the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of Scientific Beta US Live universe stocks sorted on 4 sectors (Energy, Basic Materials, Industrials and Cyclical Consumer) based on Thomson Reuters classification over the period June 2002-February 2017. The yield on Secondary US Treasury Bills is used as a proxy for the risk-free rate. We use daily total return series in US dollars. All statistics are annualized.

Table 1.9: Risk, Performance and Diversification Statistics on US Industry Universe (Part2)

Universe:	**S3-Cyclical Consumers**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	10.8%	10.6%	10.9%	12.0%	12.8%	11.9%
An. Volatility	22.1%	17.9%	20.7%	19.9%	24.4%	26.0%
An.growth rate	8.4%	9.0%	8.7%	10.0%	9.9%	8.6%
Sharpe ratio	0.43	0.53	0.47	0.54	0.48	0.41
Average ENC	100%	33%	87%	33%	33%	33%
Maxdrawdown	67.1%	55.5%	64.4%	64.2%	75.9%	78.0%
Daily VaR 1%	-3.7%	-3.1%	-3.5%	-3.4%	-4.4%	-4.6%
An.excess growth rate	3.2%	2.0%	2.9%	3.2%	4.8%	5.1%
Universe:	**S4-Non-cyclical Consumers**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	11.0%	11.1%	11.1%	10.8%	10.6%	10.2%
An. Volatility	14.3%	13.1%	13.8%	15.4%	18.5%	19.3%
An.growth rate	10.0%	10.3%	10.1%	9.7%	8.9%	8.4%
Sharpe ratio	0.68	0.75	0.71	0.62	0.51	0.46
Average ENC	100%	34%	95%	37%	33%	33%
Maxdrawdown	35.3%	23.8%	32.8%	27.4%	37.5%	40.3%
Daily VaR 1%	-2.6%	-2.3%	-2.5%	-2.7%	-3.2%	-3.3%
An.excess growth rate	2.1%	1.1%	1.9%	3.0%	4.3%	4.5%
Universe:	**S5-Financials**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	10.6%	10.4%	10.4%	11.3%	11.8%	10.7%
An. Volatility	26.5%	18.4%	23.6%	19.8%	26.5%	29.4%
An.growth rate	7.1%	8.8%	7.6%	9.4%	8.3%	6.4%
Sharpe ratio	0.35	0.50	0.39	0.51	0.40	0.32
Average ENC	100%	33%	84%	33%	33%	33%
Maxdrawdown	72.1%	48.8%	66.2%	56.4%	73.3%	79.4%
Daily VaR 1%	-4.9%	-3.5%	-4.5%	-3.7%	-5.1%	-5.6%
An.excess growth rate	3.9%	1.6%	3.0%	2.3%	6.0%	7.8%

This table details the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of Scientific Beta US Live universe stocks sorted on 3 sectors (Cyclical Consumer, Non-Cyclical Consumer and Financials) based on Thomson Reuters classification over the period June 2002-February 2017. The yield on Secondary US Treasury Bills is used as a proxy for the risk-free rate. We use daily total return series in US dollars. All statistics are annualized.

Table 1.10: Risk, Performance and Diversification Statistics on US Industry Universe (Part3)

Universe:	**S6-Healthcare**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	13.4%	12.3%	13.3%	15.2%	14.3%	13.7%
An. Volatility	17.8%	15.6%	17.1%	18.1%	21.4%	22.8%
An.growth rate	11.9%	11.2%	11.9%	13.6%	12.0%	11.1%
Sharpe ratio	0.68	0.71	0.71	0.77	0.61	0.55
Average ENC	100%	33%	93%	34%	33%	33%
Maxdrawdown	41.1%	39.0%	40.4%	36.8%	37.0%	39.5%
Daily VaR 1%	-3.2%	-2.9%	-3.2%	-3.2%	-3.7%	-4.0%
An.excess growth rate	3.0%	2.2%	2.8%	3.6%	4.9%	5.1%
Universe:	**S7-Technology**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	13.4%	12.1%	12.8%	15.3%	17.1%	17.1%
An. Volatility	24.9%	19.5%	23.1%	21.7%	27.1%	29.3%
An.growth rate	10.4%	10.3%	12.9%	13.5%	13.5%	12.9%
Sharpe ratio	0.49	0.56	0.50	0.65	0.59	0.54
Average ENC	100%	33%	89%	33%	33%	33%
Maxdrawdown	56.4%	47.2%	53.8%	52.8%	63.4%	66.8%
Daily VaR 1%	-4.2%	-3.4%	-3.9%	-3.8%	-4.5%	-4.9%
An.excess growth rate	3.9%	2.2%	3.4%	4.0%	5.7%	6.0%
Universe:	**S8-Telecoms**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	12.9%	9.8%	11.5%	11.4%	16.3%	15.2%
An. Volatility	23.7%	20.3%	22.0%	24.7%	36.2%	36.1%
An.growth rate	10.1%	7.8%	8.4%	9.8%	9.8%	8.7%
Sharpe ratio	0.49	0.43	0.47	0.41	0.42	0.39
Average ENC	100%	41%	93%	66%	41%	42%
Maxdrawdown	61.5%	50.5%	58.5%	67.1%	81.8%	81.3%
Daily VaR 1%	-4.4%	-3.8%	-4.2%	-4.8%	-6.8%	-7.0%
An.excess growth rate	3.5%	0.9%	2.7%	3.5%	5.2%	5.7%

This table details the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of Scientific Beta US Live universe stocks sorted on 3 sectors (Healthcare, Technology and Telecoms) based on Thomson Reuters classification over the period June 2002-February 2017. The yield on Secondary US Treasury Bills is used as a proxy for the risk-free rate. We use daily total return series in US dollars. All statistics are annualized.

Table 1.11: Risk, Performance and Diversification Statistics on US Industry Universe (Part4)

Universe:	**S9-Utilities**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	11.0%	9.8%	10.8%	10.2%	10.3%	10.4%
An. Volatility	19.0%	16.4%	18.4%	19.8%	22.6%	22.8%
An.growth rate	9.2%	8.5%	8.3%	7.8%	7.8%	7.9%
Sharpe ratio	0.51	0.52	0.52	0.45	0.40	0.41
Average ENC	100%	33%	96%	33%	33%	33%
Maxdrawdown	50.7%	40.4%	48.8%	48.7%	58.6%	58.4%
Daily VaR 1%	-3.5%	-2.9%	-3.3%	-3.4%	-4.0%	-4.0%
An.excess growth rate	1.3%	0.7%	1.2%	1.8%	2.2%	2.2%

This table details the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of Scientific Beta US Live universe stocks sorted on 1 sector (Utilities) based on Thomson Reuters classification over the period June 2002-February 2017. The yield on Secondary US Treasury Bills is used as a proxy for the risk-free rate. We use daily total return series in US dollars. All statistics are annualized.

Table 1.12 examines the performances of the MVR and Max $g^\star$ portfolios in contrast of the corresponding EW portfolio for the different industries in terms of mean return, growth rate and Sharpe ratio. For five out of the ten sectors, the EW portfolio outperforms the Max $g^\star$ and MVR weighting schemes in terms of Sharpe Ratio and growth rate. The Max $g^\star$ and the MVR portfolios always display better statistics with respect to the EW for two sectors: Basic Materials and Technology. The MVR portfolio, but not the Max $g^\star$ portfolio, outperforms the EW portfolio in terms of mean return, growth rate and Sharpe ratio in two other industries: Cyclical Consumers and Financials. Overall, the MVR portfolio displays a better growth rate than the EW portfolio for 5 sectors out of 10 (versus 3 out of 10 for the Max $g^\star$ portfolio and a better risk-adjusted performance than the EW for 4 sectors out of 10 (vs. 2 out of 10 for the Max $g^\star$ portfolio). Overall, we cannot clearly explain the reasons why the Max $g^\star$ and the MVR strategies are more suitable for some industries than others.

Table 1.12: Industry Summary

	S0	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**
Max $g^\star$ outperforms EW in term of SR		X						X		
MVR outperforms EW in term of SR		X		X		X		X		
	S0	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**
Max $g^\star$ outperforms EW in term of growth rate		X		X				X		
MVR outperforms EW in term of growth rate		X		X		X	X	X		
	S0	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**
Max $g^\star$ outperforms EW in term mean return		X		X		X		X	X	
MVR outperforms EW in term of mean return		X		X		X	X	X	X	

This table compares the performances of the MVR and Max $g^\star$ portfolios to that of the corresponding EW portfolio for the different industries in terms of mean return, growth rate and Sharpe ratio.

1.3.4 Empirical Tests on Factor-Tilted Universes

In this section we test within the ERI Scientific Beta Live US universe, homogeneous selections of stocks based on fundamental characteristics relative to risk factors such as the book-to-market ratio (see Table 1.13), the market capitalization (see Table 1.14), the 12-month past performance skipping the last month (see Table 1.15), the past 104-week volatility (see Table 1.16), the growth in total assets (see Table 1.17) and the gross profitability ratio (GPR) (see Table 1.18). This complementary approach allows us (i) to analyze what sorting characteristics make the maximization of the excess growth rate relevant and (ii) to see if, with respect to a given characteristic, a relation exists between the risk-adjusted performance of the Max $g^\star$ portfolio and the level of the characteristic.

Table 1.13 displays risk, performance and diversification statistics relative to the EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes for the value and growth universes. The value (respectively growth) universe is built by selecting at each rebalancing date the 50 stocks with the highest (respectively lowest) book-to-market ratio. For the value universe, the MVR and Max $g^\star$ portfolios have risk-adjusted performances (respective Sharpe ratios of 0.52 and 0.39) lower than that of the EW portfolio (Sharpe ratio of 0.54) and annualized growth rates (respectively 11.8% and 8.6%) lower than that of the EW portfolio (11.9%). On the other hand, for the growth universe, the annualized growth rate of the MVR portfolio is higher than that of the EW portfolio (9.8% versus 8.6%) while their Sharpe ratio is identical (0.49). The Max $g^\star$ portfolio is outperformed by the EW portfolios in terms of Sharpe ratio (0.40 versus 0.49) and growth rate (8.5% versus 8.6%).

Table 1.13: US Live Universe - Value and Growth Selections - Risk, Performance and Diversification Statistics

Universe:	**Value**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	15.2%	16.4%	15.6%	17.3%	15.8%	13.7%
An. volatility	25.9%	18.4%	22.8%	21.6%	28.3%	31.9%
An.growth rate	11.9%	14.7%	13.0%	15.0%	11.8%	8.6%
Sharpe ratio	0.54	0.82	0.63	0.74	0.52	0.39
Average ENC	100%	33%	80%	33%	33%	33%
Maxdrawdown	63.0%	39.2%	55.9%	54.6%	72.4%	78.0%
Daily VaR 1%	-4.8%	-3.4%	-4.2%	-4.1%	-5.3%	-5.9%
An.excess growth rate	5.1%	2.2%	4.2%	4.0%	7.8%	9.4%
Universe:	**Growth**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	10.4%	10.1%	10.4%	12.1%	12.5%	12.2%
An. volatility	18.7%	13.9%	16.5%	16.4%	23.1%	27.2%
An.growth rate	8.6%	9.2%	9.1%	10.7%	9.8%	8.5%
Sharpe ratio	0.49	0.64	0.56	0.66	0.49	0.40
Average ENC	100%	33%	85%	33%	33%	33%
Maxdrawdown	54.7%	38.5%	46.9%	37.9%	66.3%	74.4%
Daily VaR 1%	-3.2%	-2.4%	-3.0%	-2.9%	-4.0%	-4.7%
An.excess growth rate	4.4%	1.7%	3.3%	3.8%	9.0%	10.1%

This table reports the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of two subsets of Scientific Beta US Live universe stocks over the period June 2002-February 2017. The first subset is defined at each rebalancing date as the 50 stocks with the highest book-to-market ratio (value stocks) and the second subset is defined at each rebalancing date as the 50 stocks with the lowest book-to-market ratio (growth stocks). The yield on Secondary US Treasury Bills is used as a proxy for the risk-free rate. We use daily total return series in US dollars. All statistics are annualized.

The risk, performance and diversification statistics relative to the EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes for the small cap and large cap universes are shown in Table 1.14 . The small cap (respectively large cap) universe is built by selecting at each rebalancing date the 50 stocks with the lowest (respectively highest) market capitalization. The MVR and Max $g^\star$ portfolios underperform the EW portfolio on the small cap universe with respective Sharpe ratios of 0.55 and 0.48 versus 0.58. However the growth rate of the MVR portfolio is higher than that of the EW portfolio (11.5% versus 11.2%). The results are different for the large cap universe where the MVR and Max $g^\star$ portfolios display a growth rate (respectively of 9.9% and 8.8%) higher than that of the EW portfolio (7.9%). The Sharpe ratio of the MVR portfolio on this universe is higher than that of the EW and Max $g^\star$ portfolios (0.52 versus 0.45 and 0.44).

Table 1.14: US Live Universe - Small Caps and Large Caps Selections - Risk, Performance and Diversification Statistics

Universe:	**Size-Small**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	13.4%	13.0%	13.4%	13.0%	14.3%	14.5%
An. volatility	21.1%	15.2%	18.6%	17.0%	23.6%	27.8%
An.growth rate	11.2%	11.9%	11.6%	11.6%	11.5%	10.7%
Sharpe ratio	0.58	0.78	0.65	0.70	0.55	0.48
Average ENC	100%	33%	100%	33%	33%	33%
Maxdrawdown	58.5%	44.9%	54.0%	50.8%	61.1%	66.8%
Daily VaR 1%	-3.8%	-2.5%	-3.1%	-2.9%	-4.2%	-4.7%
An.excess growth rate	4.1%	1.9%	3.4%	3.6%	6.6%	7.4%
Universe:	**Size-Big**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	9.5%	9.7%	9.8%	10.5%	12.0%	11.5%
An. volatility	18.5%	14.0%	16.8%	16.0%	20.6%	23.7%
An.growth rate	7.9%	8.8%	8.4%	9.3%	9.9%	8.8%
Sharpe ratio	0.45	0.61	0.51	0.58	0.52	0.44
Average ENC	100%	33%	88%	33%	33%	33%
Maxdrawdown	52.8%	41.0%	47.4%	44.7%	56.2%	64.5%
Daily VaR 1%	-3.3%	-2.5%	-3.0%	-2.9%	-3.8%	-4.1%
An.excess growth rate	2.2%	1.3%	1.9%	2.0%	3.4%	3.7%

This table reports the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of two subsets of Scientific Beta US Live universe stocks over the period June 2002-February 2017. The first subset is defined at each rebalancing date as the 50 stocks with the lowest market capitalization (small stocks) and the second subset is defined at each rebalancing date as the 50 stocks with the highest market capitalization (big stocks).

The risk, performance and diversification statistics relative to the EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes for the high momentum and low momentum universes are reported in Table 1.15. The high momentum (respectively low momentum) universe is built by selecting at each rebalancing date the 50 stocks with the highest (respectively lowest) 12-month past performance skipping the last month. The MVR and Max $g^\star$ portfolios perform significantly better than the EW portfolio on the high momentum and low momentum universes. For instance on the high momentum universe, the growth rates of the MVR and Max $g^\star$ portfolios are 11.2% and 9.7% versus 7.6% for the EW portfolios. The risk-adjusted performance of these two portfolios (respective Sharpe ratios of 0.55 and 0.45) is higher than that of the EW portfolio 0.39.

Table 1.15: US Live Universe - Winners and Losers Selections - Risk, Performance and Diversification Statistics

Universe:	**Momentum-Winners**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	10.4%	14.2%	11.5%	14.2%	13.9%	13.2%
An. volatility	23.7%	18.4%	21.4%	19.9%	23.1%	26.5%
An.growth rate	7.6%	12.6%	9.2%	12.3%	11.2%	9.7%
Sharpe ratio	0.39	0.71	0.48	0.66	0.55	0.45
Average ENC	100%	33%	86%	34%	33%	33%
Maxdrawdown	62.0%	44.3%	55.4%	50.3%	57.3%	63.7%
Daily VaR 1%	-4.0%	-3.2%	-3.8%	-3.5%	-4.1%	-4.6%
An.excess growth rate	4.1%	2.7%	3.7%	4.0%	5.6%	6.1%
Universe:	**Momentum-Losers**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	9.5%	9.6%	9.3%	11.3%	11.8%	11.6%
An. Volatility	29.1%	23.6%	26.9%	25.3%	30.5%	33.8%
An.growth rate	5.3%	6.9%	5.7%	8.1%	7.1%	5.9%
Sharpe ratio	0.29	0.36	0.30	0.40	0.35	0.31
Average ENC	100%	33%	85%	33%	33%	33%
Maxdrawdown	77.1%	66.8%	74.0%	66.3%	75.5%	81.2%
Daily VaR 1%	-5.4%	-4.4%	-4.9%	-4.5%	-5.6%	-6.0%
An.excess growth rate	7.8%	4.3%	6.6%	6.2%	10.6%	12.7%

This table reports the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of two subsets of Scientific Beta US Live universe stocks over the period June 2002-February 2017. The first subset is defined at each rebalancing date as the 50 stocks with the highest 12-month past returns skipping the last month (winners stocks) and the second subset is defined at each rebalancing date as the 50 stocks with the lowest 12-month past returns skipping the last month (losers stocks).

The risk, performance and diversification statistics relative to the EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes for the low volatility and high volatility universes are contained in Table 1.16. The low volatility (respectively high volatility) universe is built by selecting at each rebalancing date the 50 stocks with the lowest (respectively highest) past 104-week volatility. On the low volatility universe, except from the Max $g^\star$ which displays a growth rate of 9.5%, the growth rates of all the other weighting schemes are between 10.0% and 10.1%. The high volatility universe is more suited to the MVR and Max $g^\star$: these weighting schemes exhibit Sharpe ratios of 0.46 and 0.40 and growth rates of 10.5% and 9.1%, which are higher than that of the EW portfolio (Sharpe ratio of 0.38 and growth rate of 8.4%).

Table 1.16: US Live Universe - Low Vol and High Vol Selections - Risk, Performance and Diversification Statistics

Universe:	**Low Volatility**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	10.9%	10.7%	11.0%	10.9%	10.9%	10.4%
An. Volatility	13.5%	12.4%	13.2%	12.8%	13.3%	13.8%
An.growth rate	10.0%	10.0%	10.1%	10.1%	10.1%	9.5%
Sharpe ratio	0.72	0.76	0.74	0.76	0.73	0.66
Average ENC	100%	33%	94%	33%	33%	33%
Maxdrawdown	36.9%	35.9%	36.9%	36.9%	38.8%	39.4%
Daily VaR 1%	-2.5%	-2.2%	-2.4%	-2.3%	-2.4%	-2.4%
An.excess growth rate	0.9%	1.0%	1.0%	1.2%	1.3%	1.3%
Universe:	**High Volatility**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	13.6%	17.0%	13.7%	15.7%	15.0%	14.6%
An. Volatility	32.3%	27.0%	30.6%	27.6%	29.8%	33.2%
An.growth rate	8.4%	13.3%	9.1%	11.9%	10.5%	9.1%
Sharpe ratio	0.38	0.58	0.41	0.52	0.46	0.40
Average ENC	100%	33%	90%	33%	33%	33%
Maxdrawdown	75.3%	59.9%	72.6%	63.8%	67.9%	75.9%
Daily VaR 1%	-5.7%	-4.8%	-5.3%	-5.3%	-5.6%	-6.1%
An.excess growth rate	8.7%	7.3%	8.6%	9.2%	10.9%	12.6%

This table reports the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of two subsets of Scientific Beta US Live universe stocks over the period June 2002-February 2017. The first subset is defined at each rebalancing date as the 50 stocks with the lowest past 104-week volatility (low volatility stocks) and the second subset is defined at each rebalancing date as the 50 stocks with the highest past 104-week volatility (high volatility stocks).

Table 1.17 contains risk, performance and diversification statistics relative to the EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes for the low investment and high investment universes. The low investment (respectively high investment) universe is built by selecting at each rebalancing date the 50 stocks with the lowest (respectively highest) the growth in total assets. For the low investment universe, the MVR and Max $g^\star$ portfolios display a risk-adjusted performance (respective Sharpe ratios of 0.42 and 0.40) lower than that of the EW portfolio (Sharpe ratio of 0.54) and annualized growth rates (respectively 8.3% and 8.5%) lower than that of the EW portfolio (10.2%). On the other hand, for the high investment universe, the annualized growth rate of the MVR and Max $g^\star$ portfolios are higher than that of the EW portfolio (11.2% and 10.1% versus 8.6%) the same pattern applies for their Sharpe ratios (0.53 and 0.45 versus 0.43).

Table 1.17: US Live Universe - Low Investment and High Investment Selections - Risk, Performance and Diversification Statistics

Universe:	**Low Investment**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	12.3%	10.8%	11.8%	10.3%	11.0%	12.2%
An. Volatility	20.6%	15.4%	18.4%	17.8%	23.4%	27.2%
An.growth rate	10.2%	9.7%	10.1%	8.7%	8.3%	8.5%
Sharpe ratio	0.54	0.62	0.58	0.51	0.42	0.40
Average ENC	100%	33%	81%	33%	33%	33%
Maxdrawdown	58.7%	47.0%	55.7%	57.8%	70.1%	73.4%
Daily VaR 1%	-3.5%	-2.7%	-3.1%	-3.1%	-4.1%	-4.9%
An.excess growth rate	3.4%	2.2%	3.0%	3.8%	5.7%	5.9%
Universe:	**High Investment**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	11.3%	11.0%	10.7%	10.9%	14.1%	14.5%
An. Volatility	23.6%	16.9%	20.4%	18.6%	24.4%	29.6%
An.growth rate	8.6%	9.6%	8.7%	9.2%	11.2%	10.1%
Sharpe ratio	0.43	0.58	0.47	0.52	0.53	0.45
Average ENC	100%	33%	80%	34%	33%	33%
Maxdrawdown	60.7%	46.3%	55.1%	45.8%	55.5%	67.3%
Daily VaR 1%	-4.2%	-3.0%	-3.6%	-3.1%	-4.3%	-5.3%
An.excess growth rate	5.2%	2.8%	4.4%	4.5%	7.4%	8.2%

This table reports the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of two subsets of Scientific Beta US Live universe stocks over the period June 2002-February 2017. The first subset is defined at each rebalancing date as the 50 stocks with the lowest growth in total assets (low investment stocks) and the second subset is defined at each rebalancing date as the 50 stocks with the highest growth in total assets (high investment stocks).

Table 1.18 displays risk, performance and diversification statistics relative to the EW, GMV, ERC, MD, MVR and the Max $g^\star$ weighting schemes for the high profitability and low profitability universes. The high profitability (respectively low profitability) universe is built by selecting at each rebalancing date the 50 stocks with the highest (respectively lowest) gross profitability ratio. We observe on the high profitability and low profitability universes an underperformance of the MVR and Max $g^\star$ portfolios with respect to the EW portfolio in terms of Sharpe ratios and growth rate.

Table 1.18: US Live Universe - High Profitability and Low Profitability Selections - Risk, Performance and Diversification Statistics

Universe:	**High Profitability**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	11.5%	12.0%	11.8%	11.4%	11.8%	10.9%
An. Volatility	18.4%	14.4%	16.8%	16.0%	19.9%	22.6%
An.growth rate	9.8%	11.0%	10.4%	10.2%	9.8%	8.4%
Sharpe ratio	0.56	0.75	0.63	0.64	0.53	0.43
Average ENC	100%	33%	87%	34%	33%	33%
Maxdrawdown	46.3%	34.1%	40.3%	35.0%	43.3%	51.1%
Daily VaR 1%	-3.1%	-2.6%	-2.8%	-2.7%	-3.4%	-3.8%
An.excess growth rate	2.9%	1.9%	2.6%	3.3%	4.8%	5.2%
Universe:	**Low Profitability**					
Time period:	**21/6/02-28/2/17**					
	EW	**GMV**	**ERC**	**MD**	**MVR**	**Max g***
An.mean return	9.0%	7.8%	8.4%	6.4%	6.6%	6.9%
An. Volatility	29.6%	22.1%	26.8%	23.8%	31.0%	33.4%
An.growth rate	4.6%	5.4%	4.7%	3.5%	1.8%	1.3%
Sharpe ratio	0.26	0.30	0.27	0.22	0.17	0.17
Average ENC	100%	33%	81%	33%	33%	33%
Maxdrawdown	80.4%	63.8%	76.7%	70.5%	86.6%	88.7%
Daily VaR 1%	-5.4%	-4.1%	-5.1%	-4.3%	-5.8%	-6.2%
An.excess growth rate	5.2%	1.9%	4.0%	3.5%	9.7%	11.0%

This table reports the risk, performance and diversification statistics for the EW, GMV, ERC, MD, MVR and the Max $g^\star$ portfolios made of two subsets of Scientific Beta US Live universe stocks over the period June 2002-February 2017. The first subset is defined at each rebalancing date as the 50 stocks with the highest profitabiity ratio (high profitability stocks) and the second subset is defined at each rebalancing date as the 50 stocks with the lowest profitability ratio (low profitability stocks).

A more global picture is given in Table 1.19 about the risk and performance statistics of the MVR and Max $g^\star$ portfolios relatively to the EW portfolio for the different US Live tilted universes aforementioned. This table indicates for each tilted universe if the MVR and Max $g^\star$ portfolios outperform the EW portfolios in terms of Sharpe ratio, growth rate and mean return. Amongst the 12 tilted universes considered, the universes where the MVR and Max $g^\star$ portfolios outperform the EW portfolio both in terms of Sharpe ratio, growth rate and mean return are the following: high momentum, low momentum, high volatility and high investment.

Table 1.19: US Live Tilted Universes Summary

	Value	**Growth**	**Small**	**Big**	**Winners**	**Losers**	**Low Vol**	**High Vol**	**Low Inv**	**High Inv**	**High Prof**	**Low Prof**
Max $g^\star$ outperforms EW in term of SR					X	X		X		X		
MVR outperforms EW in term of SR				X	X	X	X	X		X		
Max $g^\star$ outperforms EW in term of growth rate				X	X	X		X		X		
MVR outperforms EW in term of growth rate		X	X	X	X	X	X	X		X		
Max $g^\star$ outperforms EW in term mean return		X	X	X	X	X		X		X		
MVR outperforms EW in term of mean return	X	X	X	X	X	X	X	X		X	X	

This table compares the performances of the MVR and Max $g^\star$ portfolios to that of the corresponding EW portfolio for the different tilts in terms of mean return, growth rate and Sharpe ratio. The region considered is the UK over the period June 2002 - February 2017.

We also conduct the same analysis on the UK universe in order to test the robustness of these results in a different region. Table 1.20 reports for each tilted universe if the MVR and Max $g^\star$ portfolios outperform the EW portfolios in terms of Sharpe ratio, growth rate and mean return for the UK region. The tilted universes are built, for a given tilt, by selecting the 10 stocks[7] with the highest characteristic according to this tilt. High momentum, low momentum and high investment are still tilted universes where the MVR and Max $g^\star$ portfolios outperform the EW portfolio both in terms of Sharpe ratio, growth rate and mean return. Unlike for the US region, the high volatility tilted universe displays a relative bad performance for the MVR and Max $g^\star$ portfolios with respect to the EW portfolio. Overall, it seems that the MVR and Max $g^\star$ portfolios outperform consistently the EW portfolio on universes tilted towards high momentum, low momentum and high investment. Since momentum is a rewarded risk factor, a practical application could be the development of bespoke MVR and Max $g^\star$ momentum indices.

[7]The UK universe is made of 100 stocks and, for the sake of consistency with the US universe, we select 10% of the stocks for the sorting procedure (i.e., 10 stocks for the UK universe).

Table 1.20: UK Live Tilted Universes Summary

	Value	**Growth**	**Small**	**Big**	**Winners**	**Losers**	**Low Vol**	**High Vol**	**Low Inv**	**High Inv**	**High Prof**	**Low Prof**
Max $g^\star$ outperforms EW in term of SR	X		X		X	X				X		
MVR outperforms EW in term of SR	X		X		X	X				X		
Max $g^\star$ outperforms EW in term of growth rate	X		X	X	X	X			X	X	X	
MVR outperforms EW in term of growth rate	X		X		X	X			X	X	X	
Max $g^\star$ outperforms EW in term of mean return	X		X	X	X	X			X	X	X	
MVR outperforms EW in term of mean return	X		X	X	X	X			X	X	X	X

This table compares the performances of the MVR and Max $g^\star$ portfolios to that of the corresponding EW portfolio for the different tilts in terms of mean return, growth rate and Sharpe ratio. The region considered is the United Kingdom over the period June 2002 - February 2017.

1.4 Conclusions and Extensions

Our empirical analysis shows that the MVR and Max $g^\star$ strategies achieve their stated objectives, namely generate portfolios with high out-of-sample excess growth rates. We argue that such portfolios can be meaningful complements to existing weighting schemes on specific universes where the low volatility risk premium is negative. Unlike other well-known allocation methodologies such as the GMV, ERC or MD weighting schemes, which by construction usually overweight less volatile stocks compared to the EW strategy and consequently lead to less volatile portfolios, the MVR and particularly the Max $g^\star$ portfolios have a higher volatility and higher exposure to the market factor. Our work can be extended in several dimensions. In an individual stock framework, it is difficult to construct ex-ante homogeneous universes, that is universes where the constituents have similar expected growth rates, which are theoretically favorable to the MVR and Max $g^\star$ portfolios. A possible extension of our analysis would consist in (1) implementing MVR and/or Max $g^\star$ strategies on various individual stock universes and (2) combining these improved benchmarks, and also potentially combining them with portfolios based on competing weighting schemes, so as to more efficiently diversify away unrewarded risk exposure in the resulting portfolio.

While we have focused on a broad set of universes to compare the risk and return characteristics of the MVR and Max $g^\star$ portfolio construction techniques without explicitly taking into account changes in market conditions, another possible extension could be to take into account the time-varying distribution of factor returns and use the characteristics of the MVR and Max $g^\star$ strategies such as their high exposures to the market risk factor or their negative exposures to the low volatility risk factor, to actively adjust the factor exposures of a portfolio with tactical tilts. More generally, one fruitful avenue of research would consist in providing a detailed analysis of the benefits associated with a dynamic allocation to risk factors while taking into account the time-varying properties of their risk premia.

1.A Positivity of the Excess Growth Rate

In a general setting involving n_t risky assets with correlation matrix $C_t = (\rho_{i,j}(t))_{\substack{1\leq i\leq n_t \\ 1\leq j\leq n_t}}$, volatility vector $V_t = (\sigma_i(t))_{1\leq i\leq n_t}$, covariance matrix $\Sigma_t = (\sigma_{i,j}(t))_{1\leq i\leq n_t 1\leq j\leq n_t} = (\rho_{i,j}(t)\sigma_i(t)\sigma_j(t))_{\substack{1\leq i\leq n_t \\ 1\leq j\leq n_t}}$ and positive weights $w_t = (w_i(t))_{1\leq i\leq n_t}$ the excess growth rate $g_t^\star$ verifies:

$$
\begin{aligned}
g_t^\star &= \frac{1}{2}\left(\sum_{i=1}^{n_t} w_i(t)\sigma_i^2(t) - \sum_{i=1}^{n_t}\sum_{j=1}^{n_t} w_i(t)w_j(t)\sigma_{i,j}(t)\right) \\
&= \frac{1}{2}\left(\sum_{i=1}^{n_t} w_i(t)\sigma_i^2(t) - \sum_{i=1}^{n}\sum_{j=1}^{n_t} w_i(t)w_j(t)\rho_{i,j}(t)\sigma_i(t)\sigma_j(t)\right) \\
&\geq \frac{1}{2}\left(\sum_{i=1}^{n_t} w_i(t)\sigma_i^2(t) - \sum_{i=1}^{n_t}\sum_{j=1}^{n_t} w_i(t)w_j(t)\sigma_i(t)\sigma_j(t)\right) \\
&\geq \frac{1}{2}\left(\sum_{i=1}^{n_t} w_i(t)\sigma_i^2(t) - \left(\sum_{i=1}^{n_t} w_i(t)\sigma_i(t)\right)^2\right)
\end{aligned}
$$

Let $f : X \to \mathbb{R}$ be a differentiable convex real-valued function of a real variable such that its domain X is a subset of $\mathbb{R}$.

More generally, given a vector of weights $(w_1(t), \ldots, w_{n_t}(t)) \in (\mathbb{R}_+)^n$ such that $\sum_{i=1}^{n_t} w_i(t) = 1$ and $(\sigma_1(t), ..., \sigma_{n_t}(t)) \in X^{n_t}$ and given a real convex function f the Jensen's inequality can be

stated as:

$$f(w_1(t)\sigma_1(t) + \ldots + w_{n_t}(t)\sigma_{n_t}(t)) \leq w_1(t)f(\sigma_1(t)) + \ldots + w_{n_t}(t)f(\sigma_{n_t}(t)) \tag{1.26}$$

We apply the equation (1.26) to the convex function $f(x) = x^2$ and obtain:

$$\left(\sum_{i=1}^{n_t} w_i(t)\sigma_i^2(t)\right) \geq \left(\sum_{i=1}^{n_t} w_i(t)\sigma_i(t)\right)^2 \tag{1.27}$$

Hence the positivity of the excess growth rate $g_t^\star$.

1.B Relation Between the Excess Growth Rate Optimization and the MVR Portfolio

We wrote and solved the optimization problem of the excess growth rate maximization of a portfolio in a general setting with n_t risky assets. If we introduce a risk-free asset in the general setting with n_t risky assets as in Mantilla-Garcia (2016) then we can show that the portfolio that maximizes the excess growth rate is a mixed allocation between the risk-free asset and a risky asset portfolio. We define this risky asset portfolio as the Maximum Volatility Return MVR portfolio. The weights of the MVR portfolio verify:

$$w_t^{MVR} = \frac{\Sigma_t^{-1}(V_t \odot V_t)}{\mathbb{1}_{n_t}^\top \Sigma_t^{-1}(V_t \odot V_t)} \tag{1.28}$$

To demonstrate this result, we complete our general framework involving n_t risky assets with a risk free asset, an augmented correlation matrix $\tilde{C}_t = (\rho_{i,j}(t))_{\substack{1\leq i\leq n_t+1\\1\leq j\leq n_t+1}}$, an augmented volatility vector $\tilde{V}_t = (\sigma_i(t))_{1\leq i\leq n_t+1}$, an augmented covariance matrix $\tilde{\Sigma}_t = (\sigma_{i,j}(t))_{\substack{1\leq i\leq n_t+1\\1\leq j\leq n_t+1}} = (\rho_{i,j}(t)\sigma_i(t)\sigma_j(t))_{\substack{1\leq i\leq n_t+1\\1\leq j\leq n_t+1}}$ and an augmented weights vector $\tilde{w}_t = (w_i(t))_{1\leq i\leq n_t+1}$ with $w_i(t) = w_{r_t}$ the weight of the risk-free asset. Then the optimization problem of the excess growth rate maximization of the portfolio $\tilde{w_t^\star}$ can be formulated as:

$$\begin{aligned}\tilde{w}_t^\star &= \arg\max_{\tilde{w}_t} \left(\tilde{w_t}^\top \left(\tilde{V}_t \odot \tilde{V}_t\right) - \tilde{w_t}^\top \tilde{\Sigma}_t \tilde{w}_t\right)\\ &u.c.\ \tilde{w_t}^\top \mathbb{1}_{n_t} = 1\end{aligned}$$

and rewritten as:

$$\begin{aligned}&\arg\max_{w_t, w_{r_f}} \left(w_t^\top (V_t \odot V_t) - w_t^\top \Sigma_t w_t\right)\\ &u.c.\ w_t^\top \mathbb{1}_{n_t} + w_{r_t} = 1\end{aligned}$$

The Lagrange function of the optimization problem above is:

$$\mathcal{L}(w_t; w_{r_t}; \lambda) = \frac{1}{2}\left(w_t^\top (V_t \odot V_t) - w_t^\top \Sigma_t w_t\right) - \lambda(w_t^\top \mathbb{1}_n + w_{r_t} - 1) \tag{1.29}$$

The solution verifies the following first-order conditions:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial w_t}(w_t^\star; w_{r_t}^\star; \lambda) = \frac{1}{2}(V_t \odot V_t) - \Sigma_t w_t^\star - \lambda \mathbb{1}_{n_t} = \mathbb{0}_{n_t} \\ \frac{\partial \mathcal{L}}{\partial w_{r_t}}(w_t^\star; w_{r_t}^\star; \lambda) = \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda}(w_t^\star; w_{r_t}^\star; \lambda) = w_t^{\star\top} \mathbb{1}_{n_t} + w_{r_t} - 1 = 0 \end{cases} \tag{1.30}$$

From the first and second equations of (1.30), we have:

$$w_t^\star = \frac{1}{2}\Sigma_t^{-1}\left(V_t \odot V_t\right) \tag{1.31}$$

From the third equation of (1.30), we have:

$$w_{r_t}^\star = 1 - \frac{\mathbb{1}_{n_t}^\top \Sigma_t^{-1}\left(V_t \odot V_t\right)}{2} \tag{1.32}$$

Finally, the solution, defined by Mantilla-Garcia (2016) as the maximum volatility return (or MVR) portfolio, is a linear combination of the risk-free asset and a risky portfolio. In this chapter, we define the MVR portfolio as the risky portfolio, whose weights are given by: $\frac{\Sigma_t^{-1}(V_t \odot V_t)}{\mathbb{1}_{n_t}^\top \Sigma_t^{-1}(V_t \odot V_t)}$.

Chapter

2 Measuring Portfolio Rebalancing Benefits in Equity Markets

Abstract

The potential source of additional performance because of the simple act of resetting portfolio weights back to the original weights is referred as the rebalancing premium. It is also sometimes known as the volatility pumping effect or diversification bonus because volatility and diversification turn out to be key components of the rebalancing premium. The purpose of this chapter is to provide a thorough numerical and empirical analysis of the volatility pumping effect in equity markets and to examine the conditions under which it can be maximised. Our main contribution to the understanding of the rebalancing premium is an effort to disentangle and separately measure the isolated impact of various components of the total effect. Using Fama-French-Carhart four-factor model, we find that after controlling for factor exposures, the outperformance of the rebalanced strategy with respect to the corresponding buy-and-hold strategy remains substantial at an annualised level above 100 basis points over a 5-year time horizon for stocks in the S&P 500 universe. We also find that size, value, momentum and volatility are sorting characteristics which 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. We also show that the initial weighting scheme has a significant impact on the size of the rebalancing premium. Taken together, these results suggest that a substantial rebalancing premium can potentially be harvested in equity markets over reasonably long horizons for suitably selected types of stocks.

This chapter is the fruit of a collaboration with Lionel Martellini [1]. A part of this chapter led to the publication of an article published in *The Journal of Portfolio Management* in March 2020.

Contents

2.1 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

[1] Professor of finance at EDHEC Business School and director of EDHEC-Risk Institute

performance is known as the *rebalancing premium*, also sometimes referred as the *volatility pumping* effect (see Luenberger (1997)) or *diversification bonus* (see Booth and Fama (1992), Erb and Harvey (2006), Willenbrock (2011) and Qian (2012)) 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 (see Bernstein and Wilkinson (1997), Hallerbach (2014)), where the portfolio *growth rate* is the compounded geometric mean return of the portfolio (see Kelly (1956) and Latané (1959)), a meaningful measure of performance in a multi-period setting.[2].

The first analysis of the benefits of rebalancing in a continuous-time setting can be traced back to the seminal contribution by Merton (1971), who establishes that a continuously rebalanced fixed-weight portfolio strategy is optimal for a risk-averse investor in a multi-period setting with a constant opportunity set. This result suggests that if market conditions never change and if relative risk aversion is independent of wealth, there is no reason to change portfolio weights, so a fixed-mix is optimal. Following this first analysis, a large number of papers, including notably Fernholz and Shay (1982), Wise (1996), Fernholz (2002) and Gabay and Herlemont (2007) compare a fixed-weight portfolio with a buy-and-hold portfolio and analyse the probability that the continuously rebalanced portfolio outperforms the buy-and-hold portfolio. In practice, portfolios cannot be rebalanced continuously, risk and return parameters are time-varying, time horizon is finite and transaction costs can have a strong impact on the performance of a rebalanced strategy. Several authors have subsequently pointed out the limitations of an analysis of the rebalancing premium in such a stylised analytical setting (see Chambers and Zdanowicz (2014), Cuthbertson et al. (2015)) and others have analysed whether the alleged benefits of rebalancing subsist empirically when taking into account the impact of transaction costs, initial weighting schemes, investment universe, rebalancing frequency, time-varying correlations and volatilities and the performance metrics used to compare the performance and/or risk-adjusted performance of the portfolios. The conclusions of these empirical analysis of the rebalancing premium in realistic settings have been mostly positive. In particular, Arnott and Lovell (1993), Tsai (2001) and Harjoto and Jones (2006) find that a rebalanced strategy outperforms the corresponding buy-and-hold strategy considering a two-asset (equity and bond) universe.[4] A common limitation of these papers, however, is that their empirical protocol is based on a fixed historic period, which raises questions about the robustness of the reported findings. In an attempt to enhance the robustness of the empirical analysis of the rebalancing premium, Dichtl et al. (2014) consider three different asset classes (equity, bond and cash) in three countries (USA, UK and Germany) using a stationary bootstrap methodology (Politis and Romano (1994)) for a period ranging from January 1982 to December 2011. They find that rebalancing adds value in the three countries considered for different asset allocation strategies but they also specify that the best portfolio allocation strategy and the best rebalancing frequency strongly "depend on both the country and the period under investigation".

Another key limitation in the related literature is that it does not provide a clear understanding of the actual origins of the rebalancing premium. In particular, no attempt is made to control for differences in exposure of the rebalanced portfolio and the corresponding buy-and-hold portfolio with respect to rewarded risk factors. These differences are expected to have an impact on the

[2]The definition of the rebalancing premium in the academic literature actually varies amongst the authors: the rebalancing premium was first defined in the case of a fixed-weight portfolio as the difference between the expected growth rate of the portfolio and the weighted average of its underlying securities' expected growth rates (Bouchey et al. (2012)). However the weighted average of the underlying securities' expected growth rates of a portfolio does not represent the expected growth rate of any investable risky strategy (see Bernstein and Wilkinson (1997) and Chambers and Zdanowicz (2014)) and cannot be considered as a relevant benchmark to interpret the benefits of portfolio rebalancing. In fact it can be shown that this quantity can be regarded as one of the ingredients of the rebalancing premium defined as the difference between the expected growth rate of the rebalanced portfolio and the expected growth rate of the corresponding buy-and-hold portfolio[3] (see Equation (2.27)).

[4]Jaconetti et al. (2010) assert that there is no optimal frequency when considering a rebalancing strategy.

relative performance of the portfolios, an impact which is not directly related to the volatility pumping effect. For example, in an equity context and using the Fama-French-Carhart factor model to measure risk exposures, the intuition suggests that the buy-and-hold strategy would lead to growth, large cap and momentum biases relative to the corresponding fixed-mix strategy while the fixed-mix strategy would lead to value, small cap and contrarian biases relative to the corresponding buy-and-hold strategy. The main contribution of this chapter is precisely to measure and analyse the sources of the rebalancing premium effect by disentangling the isolated impact of the various aforementioned components. More generally, we aim at determining the main drivers of the rebalancing premium in a theoretical and numerical framework and then empirically analysing the difference of performance between a rebalanced portfolio and a buy-and-hold portfolio from a risk-adjusted performance perspective. The average outperformance of the initially equally-weighted rebalanced strategy with respect to the corresponding buy-and-hold strategy over a 5-year time horizon for stocks in the S&P 500 universe remains substantial at an annualised level above 100 basis points even after controlling for factor exposures.

In the first part of the chapter, we provide an mathematical and numerical analysis of the rebalancing premium. More specifically, we compute the probability of outperformance, the difference of the Sharpe ratios and the excess expected growth rate of a rebalanced portfolio compared with the corresponding buy-and-hold portfolio using analytical expressions and/or Monte Carlo simulations based on standard assumptions about underlying asset returns. We also conduct a number of robustness checks with respect to assumptions regarding the level of expected return and volatility of the risky assets, the rebalancing frequency, the initial portfolio weights, the introduction of serial correlation in asset returns and the time horizon. This mathematical and quantitative analysis allows us to provide useful insights regarding the magnitude of the rebalancing premium and the conditions under which it can be maximised, with a particular emphasis on the impact of the covariance matrix, portfolio weights and time horizon. We find out, in a stylised framework with n risky assets without serial correlation in asset returns, that the amplitude of the rebalancing premium is low (i.e., lower than 50 basis points) when considering realistic parameters for the number of risky assets, risky assets volatility, risky assets expected return, pairwise correlations, time horizon and rebalancing frequency. The introduction of negative serial correlation in asset returns via a fractional Brownian motion substantially enhances the amplitude of the rebalancing premium. On the other hand, this controlled quantitative analysis cannot shed insights regarding the actual size of the rebalancing premium that can be harvested in practice. Moreover, it does not allow us to measure the contribution to the rebalancing premium emanating from differences in risk exposures between the rebalanced and the buy-and-hold portfolio. In an attempt to address these important questions, a second part of the chapter is dedicated to an empirical analysis of the rebalancing premium in realistic settings. To control for differences in factor exposures, we regress the excess monthly returns of the rebalanced portfolio over the buy-and-hold portfolio on the market, size, value and momentum factors, and interpret the alpha of this regression as the marginal value added by a pure volatility pumping effect, defined as the fraction of the rebalancing premium that is not related to differences in factor exposures. As such, this analysis is related to Plyakha et al. (2012), who consider stocks in the major US equity indices over forty years, compare the performance of equally-weighted portfolios with price-weighted and value-weighted portfolios and analyse the difference of performance using inter alia Fama-French-Carhart four-factor linear regression. A key difference is that we perform an empirical comparison of the rebalanced versus buy-and-hold portfolios for the *same* initial weighting scheme, while Plyakha et al. (2012) compare the performance of a rebalanced *equally-weighted* portfolio to the performance of a buy-and-hold *cap-weighted* portfolio. In a second step, we analyse how this pure volatility pumping effect depend upon the covariance matrix of the underlying assets, the serial correlation in asset returns, the rebalancing frequency, the stocks' characteristics and the portfolio weights. We observe, for a 5-year time horizon, a rebalancing premium higher than 80 basis points (bps in short) in our base universe extracted from the S&P 500. We also show that a substantial idiosyncratic component earned by the rebalancing process

is not explained by the 4-factor model. Using Fama-French-Carhart four-factor model, we find that after controlling for factor exposures, the outperformance of the rebalanced strategy with respect to the corresponding buy-and-hold strategy remains susbtantial at an annualised level above 100 basis points over a 5-year time horizon for our base case analysis. This result suggests that the outperformance of the rebalanced portfolio with respect to the buy-and-hold portfolio is not purely explained by differences in factor exposures. In fact the negative exposure to the positively rewarded momentum factor of the excess return of the rebalanced portfolio with respect to the buy-and-hold portfolio implies that the idiosyncratic component of the volatility premium effect is higher than what its raw value is. Our empirical analysis on individual stocks' characteristics also 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 randomly selected portfolios for time horizons from 1 year to 5 years. We also show that the weighting scheme has a susbtantial impact on size of the rebalancing premium, which is larger for a portfolio construction aiming at maximising the portfolio excess growth rate compared to the equally-weighted base case portfolio construction scheme. 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 at least in segments where market frictions are relatively low.

The rest of the chapter is organised as follows. In Section 2, we provide a detailed analysis of the rebalancing premium in a stylised continuous time setting, using analytical expressions when available and turning to numerical Monte-Carlo simulations in situations where analytical expressions are not available. In Section 3, we turn to an empirical analysis of the rebalancing premium to assess whether its benefits survive in real world situations. Conclusions and suggestions for further research are found in Section 4. Technical details are delegated to a dedicated appendix.

2.2 Mathematical and Quantitative Analysis of the Rebalancing Premium

It can be shown that the buy-and-hold portfolio always displays an expected terminal wealth higher than or equal to that of the rebalanced portfolio (see Qian (2014) in a discrete time framework or Spinu (2015) in a continuous time framework). The expected terminal wealth of the two portfolios are equal if and only if all the expected returns of the assets are the same, which is of course unrealistic in practice. So if an investor's only concern was to maximize the expected wealth of a portfolio then rebalancing is not desirable. On the other hand, if we assume again that every asset has the same expected return then it can be shown that the variance of the rebalanced portfolio is lower than or equal to that of the buy-and-hold portfolio. As a result, the impact of rebalancing is not straightforward from a risk-adjusted perspective even in a stylized setting.

In this section, we wish to extend these results and provide a thorough theoretical analysis of these effects in the context of a general stochastic multi-period framework with time horizon T, where P_0 denotes the portfolio initial value and P_T its final value. The growth rate of a portfolio $G_P(0,T)$ on the period $[0;T]$ is defined as:

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

Note that this quantity is stochastic as seen from the initial date since the final portfolio value P_T is unknown as seen from the initial date. If we now consider a fixed-weight portfolio P^{reb} and a buy-and-hold portfolio $P^{b\&h}$ based on holding n assets with the same initial weights $(w_1,\ldots,w_n)$ (this weights are held constant in the rebalanced portfolio, while they drift away in the buy-and-hold portfolio), then the *rebalancing premium* $RP(0,t)$ over the period $[0;t]$ is simply defined

by:

$$RP(0,t) = \mathbb{E}[G_{P^{reb}}(0,t)] - \mathbb{E}[G_{P^{b\&h}}(0,t)] = \mathbb{E}\left[\frac{1}{t}\ln\left(\frac{P_t^{reb}}{P_t^{b\&h}}\right)\right] \tag{2.1}$$

Note that in what follows we will restrict our analysis to long-only rebalanced strategies:

$$\sum_{i=1}^{n} w_i = 1 \text{ and } \forall i \in [1;n], w_i \geq 0$$

The difference between the expected mean return and the expected growth rate of a portfolio is a critical component in our multi-period framework analysis of the rebalancing premium. Booth and Fama (1992) put forward the link between the expected arithmetic mean return and the expected growth rate of a portfolio with the following second-order Taylor series approximation:

$$\mathbb{E}[G_P] \approx \ln(1+\mu_P) - \frac{\sigma_P^2}{2(1+\mu_P)^2} \tag{2.2}$$

where μ_P is the portfolio expected arithmetic return and σ_P the portfolio volatility. This approximation has been extensively used in numerous academic and practitioner papers (see Bodie et al. (1999), Jordan and Miller (2008)) and is acceptable as long as daily to monthly returns are considered. It shows that a portfolio growth rate is negatively related to portfolio volatility. In fact, we shall demonstrate below that the rebalancing premium benefits from a situation where we start with highly volatile assets and combine them to form a low volatility portfolio (see Equation (2.2)). If we consider daily to monthly returns we can make the approximation $\ln(1+\mu_P) \approx \mu_P$ and $(1+\mu_P)^{-2} \approx 1 - 2\mu_P$ so that we obtain the approximation by neglecting the second order term $\sigma_P^2 \times \mu_P$:

$$\mathbb{E}[G_P] \approx \mu_P - \frac{\sigma_P^2}{2} \tag{2.3}$$

2.2.1 Simulation Framework

All the stochastic processes below are defined under the physical measure $\mathbb{P}$ and under a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, F_t)$ where Ω is a set of outcomes, $\mathcal{F}$ is a sigma-algebra that represents the set of all measurable events, $\mathbb{P}$ is a probability measure and F_t is a filtration. To set the stage for the mathematical and quantitative analysis of the rebalancing premium, we consider a universe with n assets assumed to follow the system of n stochastic differential equations

$$\frac{dS_t^i}{S_t^i} = \mu_i dt + \sigma_i dW^i(t), \quad i \in [\![1,n]\!] \tag{2.4}$$

Where each $(W^i(t))_{t\geq 0}$ is a one-dimensional standard Brownian motion and the correlations $\rho_{i,j}$ between asset returns are by:

$$Corr(W_t^i, W_t^j) = \rho_{i,j}$$

For each asset i, we solve its corresponding stochastic differential equation in (2.4) and obtain

$$S^i(T) = S^i(0)\exp\left[\left(\mu_i - \frac{\sigma_i^2}{2}\right)T + \sigma_i W^i(T)\right] \tag{2.5}$$

where T is the time horizon in years and each $W^i(T)$ follows a normal distribution ($W^i(T) \sim \mathcal{N}(0,T)$).

We denote $\boldsymbol{W}(T) = \left(W^i(T)\right)_{1\leq i\leq n}$ the random vector made of the n correlated random variables $W^i(T)$. $\boldsymbol{W}(T)$ follows a multivariate normal distribution:

$$\boldsymbol{W}(T) \sim \mathcal{N}(\mathbf{0}_n, T\boldsymbol{C})$$

where $\boldsymbol{C} = \left(\rho_{i,j}\right)_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}$ corresponds to the correlation matrix and $\mathbf{0}_n$ is a vector of size n full of zeros.

To generate the n correlated components of $\boldsymbol{W}(T)$, we first generate a vector $\boldsymbol{Z}$ of n random variables such that:

$$\boldsymbol{Z} \sim \mathcal{N}\left(\mathbf{0}_n, \mathbb{I}_n\right)$$

where $\mathbb{I}_n$ denotes the identity matrix of size $(n \times n)$. All the elements of $\boldsymbol{Z}$ are independent and distributed as standard normal. We then decompose the hermitian, positive-definite correlation matrix $\boldsymbol{C}$ as a product of a lower triangular matrix and its conjugate transpose via the Cholesky factorization:

$$\boldsymbol{C} = \boldsymbol{L}^t \boldsymbol{L}$$

We have $\sqrt{T}\boldsymbol{L}\boldsymbol{Z} \sim \mathcal{N}\left(\mathbf{0}_n, T\boldsymbol{L}^t\boldsymbol{L}\right)$ and we are finally able to simulate n correlated normal random variables from n independent normal random variables.

By replacing T by respectively $t + \Delta t$ and t in Equation (2.A), we have:

$$\begin{aligned} S^i(t+\Delta t) &= S^i(0)\exp\left[\left(\mu_i - \frac{\sigma_i^2}{2}\right) \times (t+\Delta t) + \sigma_i W^i(t+\Delta t)\right] \\ S^i(t) &= S^i(0)\exp\left[\left(\mu_i - \frac{\sigma_i^2}{2}\right) t + \sigma_i W^i(t)\right] \end{aligned} \tag{2.6}$$

We deduce from Equation(2.6) that:

$$S^i(t+\Delta t) = S^i(t)\exp\left[\left(\mu_i - \frac{\sigma_i^2}{2}\right)\Delta t + \sigma_i\left(W^i(t+\Delta t) - W^i(t)\right)\right] \tag{2.7}$$

In our general case with an investment universe with n risky assets, we discretize Equation (2.7) and generate 1,000 simulations for each of the asset price S_t^i over a time horizon T for a monthly time step $\Delta t = \frac{1}{12}$ and an initial price $S_0^i = 100$ as:

$$S^i_{t+\Delta t} = S^i_t \exp\left[\left(\mu_i - \frac{\sigma_i^2}{2}\right)\Delta t + \sigma_i\left(W^i_{t+\Delta t} - W^i_t\right)\right] \tag{2.8}$$

Where t is a monthly observation date such that $\frac{1}{\Delta t} \times t \in [\![0, \frac{T}{\Delta t}]\!]$ and where $\boldsymbol{W_t} = (W_t^i)_{1\leq i\leq n}$ represents the vector of n correlated brownian motions that verify:

$$\boldsymbol{W}_{t+\Delta t} - \boldsymbol{W}_t \sim \mathcal{N}\left(\mathbf{0}_n, \Delta t\boldsymbol{C}\right)$$

To simulate $\boldsymbol{W}_{t+\Delta t} - \boldsymbol{W}_t$, we need, as described above, to generate a vector $\boldsymbol{Z} \sim \mathcal{N}\left(\mathbf{0}_n, \mathbb{I}_n\right)$ of n random variables.

Assuming a time step of one month between two observations, a maximum time horizon of 100 years (i.e. 1,200 months), a number of Monte Carlo simulations made for each risky asset equal to 1,000 and a maximum number of risky assets equal to 100, we use in every setting the same $1,200 \times 1,000 \times 100$ core matrix of random numbers from a standard Gaussian distribution. This 3-dimensional matrix allows us to generate, for each of the 1,000 Monte Carlo simulations, the monthly stock prices of 100 risky assets over a time horizon equal to $1,200$ months. Once the time horizon T(in years) and the number of risky assets n (≤ 100) are fixed, we then extract from this three-dimensional matrix a $\frac{T}{\Delta t} \times 1,000 \times n$ matrix for the Monte Carlo sampling computation and will use this same matrix to generate the risky asset prices whatever the assumptions on volatility, correlation or expected return are. This choice allows to do consistent robustness checks which would have been difficult and time consuming if we used Monte Carlo sampling for each set of parameters considered. Doing robustness checks is an usual practice made to assess how the results of an empirical study evolve when changing some core parameter values of the study.

In what follows, we provide a number of additional mathematical and numerical results regarding the rebalancing premium based on the framework described above.

2.2.2 Comparison of a Rebalanced portfolio with its Corresponding Buy-and-hold Portfolio

We first consider a simple setting with one risky and one risk-free asset, then move to a setting with two risky assets, and finally consider a more general situation with n risky assets with no serial correlation for asset returns.

In addition to the naively diversified equally-weighted portfolio we will examine, two different weighting schemes[5], namely:

- Equal Risk Contribution (ERC), [6]
- the weighting scheme which, given the covariance matrix, maximises the excess growth rate term: $\max_w g^\star(w)$ (see Section 2.2.5.1 for the general definition of the excess growth rate).

We analyse the robustness checks by considering the following three indicators:

- The rebalancing premium, which is the rebalanced portfolio expected growth rate minus the buy-and-hold portfolio expected growth rate (see Equation 2.1);
- The outperformance probability, which is the probability that the rebalanced portfolio expected terminal wealth is greater than that of the buy-and-hold portfolio;
- The difference in Sharpe ratios, which is the rebalanced portfolio Sharpe ratio minus the buy-and-hold portfolio Sharpe ratio.

We also report the 95% confidence interval for the rebalancing premium $RP(0,t)$ defined as:

$$CI_{95\%}(RP(0,t)) = \left[\overline{x} - t_{n_{\text{sims}}-1,\frac{0.05}{2}} \frac{s}{\sqrt{n_{\text{sims}}}}; \overline{x} + t_{n_{\text{sims}}-1,\frac{0.05}{2}} \frac{s}{\sqrt{n_{\text{sims}}}}\right]$$

where:

- n_{sims} is the number of Monte Carlo simulations,
- $t_{n_{\text{sims}}-1,\frac{0.05}{2}}$ is the critical value of the Student distribution with $n_{\text{sims}}-1$ degrees of freedom $\mathcal{T}$(n_sims-1) such that $\mathbb{P}\left(Z > t_{n_{\text{sims}}-1,\frac{0.05}{2}}\right) = \frac{0.05}{2}$ for $Z \sim \mathcal{T}$(n_sims-1),
- $\overline{x}$ is the rebalancing premium estimate.
- s is the standard deviation estimate.

We refer the reader to Appendix 2.A for the mathematical details.

2.2.3 One risky asset and one risk-free asset

So as to obtain some simple and easy intuitions, we start with the case of a single risky asset.

[5] One has to note that these two weighting schemes are fixed-weighted because we assume a constant covariance matrix.

[6] The underlying idea of the ERC portfolio is to equalize the contribution to the overall risk (i.e., the portfolio volatility). The ERC portfolio is the tangency portfolio if all constituents have the same Sharpe ratio and if all pairs of correlation are identical. The first formal analysis of the ERC portfolio was subsequently given by Maillard et al. (2010), who establish its existence and uniqueness in a long-only framework and propose numerical algorithms to compute the portfolio.

2.2.3.1 Mathematical analysis

Let us consider a portfolio with a risk-free asset and a risky asset (as in Fernholz and Shay (1982) and Gabay and Herlemont (2007)). For simplicity, we assume that the risk-free asset has a constant expected return r. We also assume that the risky asset price $S(t)$ follows a geometric Brownian motion with an expected return μ and a volatility σ:

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t)$$

A straightforward application of Ito's lemma shows that:

$$S(t) = S(0)\exp(gt + \sigma W(t))$$

where $g = \mu - \frac{\sigma^2}{2}$. We have that:

$$\mathbb{E}\left[G_S(0,T)\right] = \mathbb{E}\left[\frac{1}{T}\ln\left(\frac{S_T}{S_0}\right)\right] = g$$

and g can therefore be regarded as the expected growth rate of the risky asset. The risk-free asset price $S_0(t)$ verifies: $S_0(t) = S_0(0)\exp(rt)$. [7]

If we now consider buy-and-hold and continuously rebalanced portfolios with values P^{reb} and $P^{b\&h}$, both of them with initial weight w invested in the risky asset and $1-w$ in the risk-free asset, we have:

$$\begin{aligned} P^{reb}(t) &= (S_0(t))^{1-w}(S(t))^w \exp(g^\star t) \\ P^{b\&h}(t) &= (1-w)S_0(t) + wS(t) \end{aligned} \tag{2.9}$$

where $g^\star = \frac{w(1-w)\sigma^2}{2}$, is a quantity known as *excess growth rate*, which is maximum for $w = \frac{1}{2}$.

Then we have:

$$\ln\frac{P^{reb}(t)}{P^{b\&h}(t)} = \frac{1}{2}\sigma^2 w(1-w)t + \ln\frac{S_0(t)^{1-w}S(t)^w}{(1-w)S_0(t) + wS(t)}$$

which implies that:

$$RP(0,t) = \mathbb{E}\left[\frac{1}{t}\ln\left(\frac{P_t^{reb}}{P_t^{b\&h}}\right)\right] = \underbrace{\frac{1}{2}\sigma^2 w(1-w)}_{\text{excess growth rate term}} + \underbrace{\frac{1}{t}\mathbb{E}\left[\ln\frac{S_0(t)^{1-w}S(t)^w}{(1-w)S_0(t)+wS(t)}\right]}_{\text{dispersion term}} \tag{2.10}$$

We can note that the rebalancing premium is always smaller or equal to $g^\star$ since $\ln\frac{S_0(t)^{1-w}S(t)^w}{(1-w)S_0(t)+wS(t)} \leq 0$ and is zero if and only if $S_0(t) = S(t)$ (from Jensen's inequality, see Appendix 2.D).

In the same setting, and with the assumption that $r = 0$, Wise (1996) shows that the outperformance probability of the rebalanced portfolio is:

$$\mathbb{P}(P^{reb}(t) \geq P^{b\&h}(t)) = \mathcal{N}\left(\frac{\ln y(t) - gt}{\sigma\sqrt{t}}\right) - \mathcal{N}\left(\frac{\ln x(t) - gt}{\sigma\sqrt{t}}\right) \tag{2.11}$$

where $\mathcal{N}()$ is the cumulative probability function for the standard normal distribution, $x(t)$ and $y(t)$ are respectively the lower and higher roots of the equation: $wS(t)+(1-w)-S(t)^w\exp\left(\frac{w(1-w)\sigma^2}{2}t\right) = 0$. This result suggests that the outperformance probability mainly depends on the expected growth rate g and volatility σ of the risky asset and on the allocation to the risky asset w. On the other hand Wise (1996) demonstrates that if the expected return of the risky asset follows the relationship:

$$-\frac{(1-w)\sigma^2}{2} < g < \frac{w\sigma^2}{2} \Leftrightarrow \frac{w\sigma^2}{2} < \mu < \frac{(1+w)\sigma^2}{2} \tag{2.12}$$

[7] We assume that $S_0(0) = S(0)$.

then the rebalanced portfolio outperforms the buy-and-hold portfolio with a probability which converges to 1 as the time horizon becomes long enough. Conversely, if the previous inequality is not verified then the buy-and-hold portfolio outperforms the rebalanced portfolio with a probability which converges to 1 as the time horizon goes to infinity.

We can generalise Equation (2.12) to the case where the risk-free asset expected return r is supposed constant but not necessarily non null[8]:

$$\frac{w\sigma^2}{2} < \mu - r < \frac{(1+w)\sigma^2}{2} \tag{2.13}$$

According to Equation(2.13), if $\frac{\sigma^2}{4} < \mu - r < \frac{3\sigma^2}{4}$ an equally-weighted rebalanced portfolio outperforms the buy-and-hold portfolio with a probability which converges to 1 as the time horizon goes to infinity. For instance for a volatility of the risky asset $\sigma = 20\%$, a risky asset weight $w = 50\%$ and risk-free rate $r = 2\%$ the aforementioned condition becomes: $3\% < \mu < 5\%$.

Finally, in the case where $w = \frac{1}{2}$, $g = 0$ and $r = 0$, Gabay and Herlemont (2007) establish the following expression:

$$\ln\left(\frac{P^{reb}(t)}{P^{b\&h}(t)}\right) = \frac{\sigma^2 t}{8} - \ln\cosh\left(\frac{1}{2}\sigma\sqrt{t}z\right)$$

where z follows a standard normal distribution. We have:

$$\mathbb{P}(P^{reb}(t) > P^{b\&h}(t)) = \mathbb{P}(-z_c(t) < z < z_c(t)) = 2\mathcal{N}(z_c(t)) - 1$$

where $z_c(t) = \frac{2}{\sigma\sqrt{t}} \ln\left(e^{\frac{\sigma^2 t}{8}} + \sqrt{e^{\frac{\sigma^2 t}{4}} - 1}\right)$ and $\mathcal{N}()$ is the cumulative probability function for the standard normal distribution. When t is small, $z_c \approx 1$ and we have $\mathbb{P}(P^{reb}(t) > P^{b\&h}(t)) \approx 68\%$ and $\mathbb{E}\left(\ln\frac{P^{reb}(t)}{P^{b\&h}(t)}\right) \approx 0$ (see Wise (1996) for further details).

Starting from Equation (2.9) we can show that the rebalanced portfolio outperforms the buy-and-hold portfolio in scenarios where the constituents have cumulative performances not too far apart. We have:

$$\begin{aligned} P^{reb}(t) - P^{b\&h}(t) &= (S_0(t))^{1-w} S(t)^w \exp\left(g^\star t\right) - (1-w) S_0(t) - wS(t) \\ &= S_0(t)\left[\left(\frac{S(t)}{S_0(t)}\right)^w \exp\left(g^\star t\right) - (1-w) - w\left(\frac{S(t)}{S_0(t)}\right)\right] \\ &= S_0(t) \times \phi\left(\frac{S(t)}{S_0(t)}\right) \end{aligned}$$

Where $\phi(x) = x^w \exp(g^\star t) - (1-w) - wx$.

Then:

$$\phi'(x) = w\left[x^{w-1}\exp(g^\star t) - 1\right]$$

We assess that:

1.
$$\forall x \in \mathbb{R}^+,\ \phi'(x) = 0 \Longleftrightarrow x = \exp\left(\frac{-g^\star t}{w-1}\right) = x_0, \tag{2.14}$$

2.
$$\forall x \in \mathbb{R}^+,\ \phi'(x) > 0 \Longleftrightarrow x < x_0, \tag{2.15}$$

[8]This result can be obtained by considering the framework with two risky assets (see Section (2.2.4.1) and Equation (2.24)), fixing the pairwise correlation $\rho_{1,2} = 0$, the second asset volatility $\sigma_2 = 0$ and the second asset expected return $\mu = r$.

3.
$$\forall x \in \mathbb{R}^+,\ \phi^{'}(x) < 0 \Longleftrightarrow x > x_0, \tag{2.16}$$

4.
$$\phi(0) = w - 1 < 0 \tag{2.17}$$

and
$$\lim_{+\infty} \phi(x) = +\infty, \tag{2.18}$$

5.
$$\phi(x) = \left[\exp\left(\frac{g^{\star}t}{1-w}\right) - 1\right][1-w] > 0 \tag{2.19}$$

From Equations (2.14), (2.15), (2.16), (2.17) and (2.19) we deduce that there exist x_1 and x_2 that verify $0 < x_1 < x_0 < x_2$ such as:

$$\phi(x) = 0 \Longleftrightarrow x_1 < x < x_2 \tag{2.20}$$

We finally obtain from Equation (2.20):

$$P^{reb}(t) \geq P^{b\&h}(t) \Longleftrightarrow x_1 \leq \frac{S(t)}{S_0(t)} \leq x_2$$

2.2.3.2 Quantitative analysis

We propose to compute the outperformance probability of the rebalanced portfolio relative to the corresponding buy-and-hold portfolio, and the distribution probability of the terminal wealth of the rebalanced portfolio in excess of the terminal wealth of the buy-and-hold portfolio as a function of the parameter values. We also show the distribution of the growth rate difference of the rebalanced portfolio minus the growth rate of the corresponding buy-and-hold portfolio and its main characteristics, assuming the following base case parameter values in Monte Carlo simulations:

- The risk-free rate r is constant and taken to be 2%;
- The expected return of the risky asset is 10% and the volatility of the risky asset is 20%;
- An initial weight $w = \frac{1}{2}$ invested in the risky asset;
- The rebalancing frequency is 1 month;
- The time horizon is 5 years.

The rebalancing premium, which is the expected value of the growth rate of the rebalanced portfolio minus that of the corresponding buy-and-hold portfolio, is decomposed into two components, the excess growth rate term and the dispersion term (see Equation (2.10)). We also compute the following statistics for each strategy: annualised performance, annualised volatility, Sharpe ratio, skewness, kurtosis, maximum drawdown, 1% Value at Risk (1-month horizon), certainty equivalent return for an investor who has a constant relative risk aversion (CRRA) utility function with risk aversion parameter γ, and finally annualised one-way turnover. We give further details of these statistics in Appendix 2.B. Robustness checks are performed relatively to assumptions regarding the level of expected return and volatility of the risky asset, the rebalancing frequency, the portfolio weights, the introduction of serial correlation in the risky asset returns through the use of a fractional Brownian motion process and the time horizon.

Base case scenario

Table 2.1 displays statistics relative to the rebalanced and the buy-and-hold portfolios in the base case for one risk-free asset and one risky asset. From a risk-adjusted performance perspective, the two portfolios are very similar (same Sharpe ratios of 0.39%, annualised volatility of 10.0% for the rebalanced portfolio and 10.7% for the buy-and-hold portfolio, expected return of 5.9% for the rebalanced portfolio and 6.1% for the buy-and-hold portfolio). The return distribution of the rebalanced portfolio is more skewed to the right than that of the buy-and-hold portfolio with a skewness of 0.1745 versus 0.1089 for the latter portfolio. From an extreme risk perspective the buy-and-hold portfolio is slightly riskier than the rebalanced portfolio (maximum drawdown of 14.7 % and Value at Risk of -5.7% for the rebalanced portfolio versus 15.9% and -6.4% for the buy-and-hold portfolio). The annualised one-way turnover of the rebalanced portfolio is 13.8%. The outperformance probability is 60% and the rebalancing premium is -17 bps (with a confidence interval at the 95% level equal to [-22, -11] bps) showing that the amplitude of the difference between the two portfolios expected growth rate is negative in the base case.

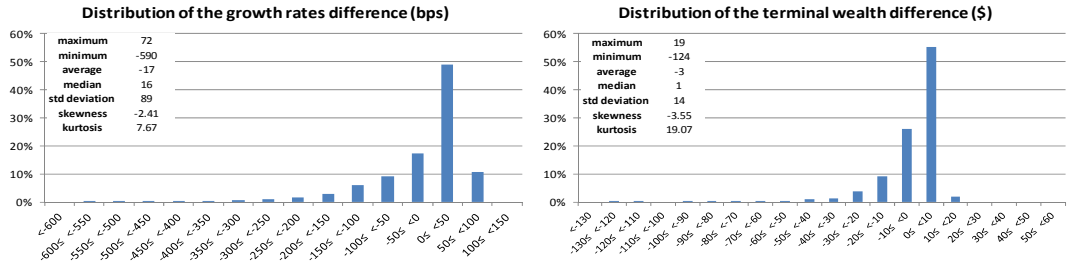
Table 2.1: Statistics for a 5-year time horizon

	Rebalanced Portfolio	**Buy-and-Hold Portfolio**
Expected Return	5.9%	6.1%
Annualised Volatility	10.0%	10.7%
Sharpe Ratio	0.39	0.39
Skewness	0.1745	0.1089
Kurtosis	2.9518	2.9894
Maximum Drawdown	14.7%	15.9%
Monthly VaR 1%	-5.7%	-6.4%
Turnover (%)	13.8%	0.0%
CEQ ($\gamma = 2$)	4.9%	5.0%
CEQ ($\gamma = 5$)	3.4%	3.3%
Outperformance probability	60%	
Rebalancing premium (bps) (1) + (2)	-17	
Rebalancing premium $CI_{95\%}$ (bps)	$[-22, -11]$	
Excess growth rate $g^\star$(bps) (1)	50	
Dispersion term (bps) (2)	-67	
Conditional (>0) Rebalancing premium (bps)	35	
Conditional (<0) Rebalancing premium (bps)	-93	

This table reports the main performance, risk and rebalancing statistics for the rebalanced portfolio and the corresponding buy-and-hold portfolio in the base case scenario with an investment universe composed of one risky asset and one risk-free asset. The time horizon is 5 years.

Figure 2.1 displays the distribution of the rebalanced portfolio growth rate minus the buy-and-hold portfolio growth rate. The distribution is negatively skewed (-2.41) the minimum difference in the 1000 scenarios is -590 bps whereas the maximum difference is 72 bps. In nearly 50% of the scenarios, the growth rates difference is between 0 bps and 50 bps. Figure 2.1 also displays the distribution of the rebalanced portfolio terminal value minus the buy-and-hold portfolio terminal value. Each portfolio has an initial value of $100. The distribution is negatively skewed (-3.55) the minimum difference in the 1,000 scenarios is -$124 whereas the maximum difference is $19. In 55% of the scenarios, the terminal wealth difference is between $0 and $10.

Figure 2.1: Distributions of the growth rates difference (bps) and of the terminal wealth difference($) for a 5-year time horizon

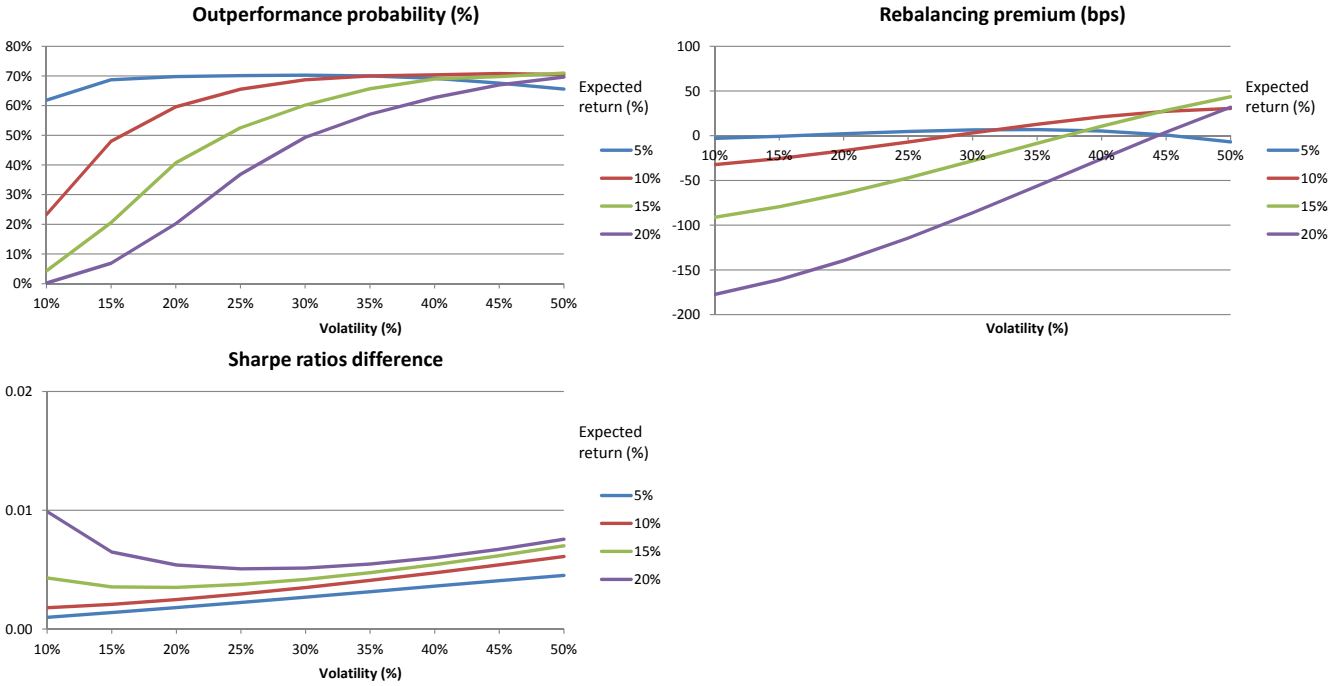


This figure displays the probability distribution of the difference between the growth rate of the rebalanced portfolio and that of the corresponding buy-and-hold portfolio in the base case for an investment universe made of one risky asset and one risk-free asset.It also displays the probability distribution of the difference between the terminal wealth of the rebalanced portfolio and that of the corresponding buy-and-hold portfolio in the base case for an investment universe made of one risky asset and one risk-free asset. The initial wealth of each portfolio is 100$. The time horizon is 5 years.

Two-dimensional robustness check with respect to volatility and expected return

Figure 2.2 displays the outperformance probability, the rebalancing premium and the difference in Sharpe ratios for different values of the risky asset expected return (from 5% to 20%) and volatility (from 10% to 50%). If the volatility of the risky asset is greater than or equal to 35% then the outperformance probability is greater than 50%. On the other hand if we set the volatility of the risky asset at 20% then the outperformance probability is greater than or equal to 50% if the expected return of the risky asset is lower than or equal to 10%. The range of volatility for which the outperformance probability is greater than 50% is wider if the expected return of the risky asset is low. Turning now to the magnitude of the rebalancing premium, we find that it is positive for a 30% volatility and a 10% expected return on the risky asset but it is very low (i.e., 3 bps). Even in the unrealistic case where the volatility is 50% and the expected return is 15% the rebalancing premium is merely 44 bps. If the volatility of the risky asset is lower than or equal to 15%, then the rebalancing premium is negative. This first robustness check underlines that with realistic assumptions on the expected return and volatility of the risky asset the amplitude of the rebalancing premium without taking into account the transaction costs is low. The Sharpe ratios difference between the rebalanced and the buy-an-hold portfolio is lower than 0.01 but positive in all the cases considered.

Figure 2.2: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of one risky asset and one risk-free asset



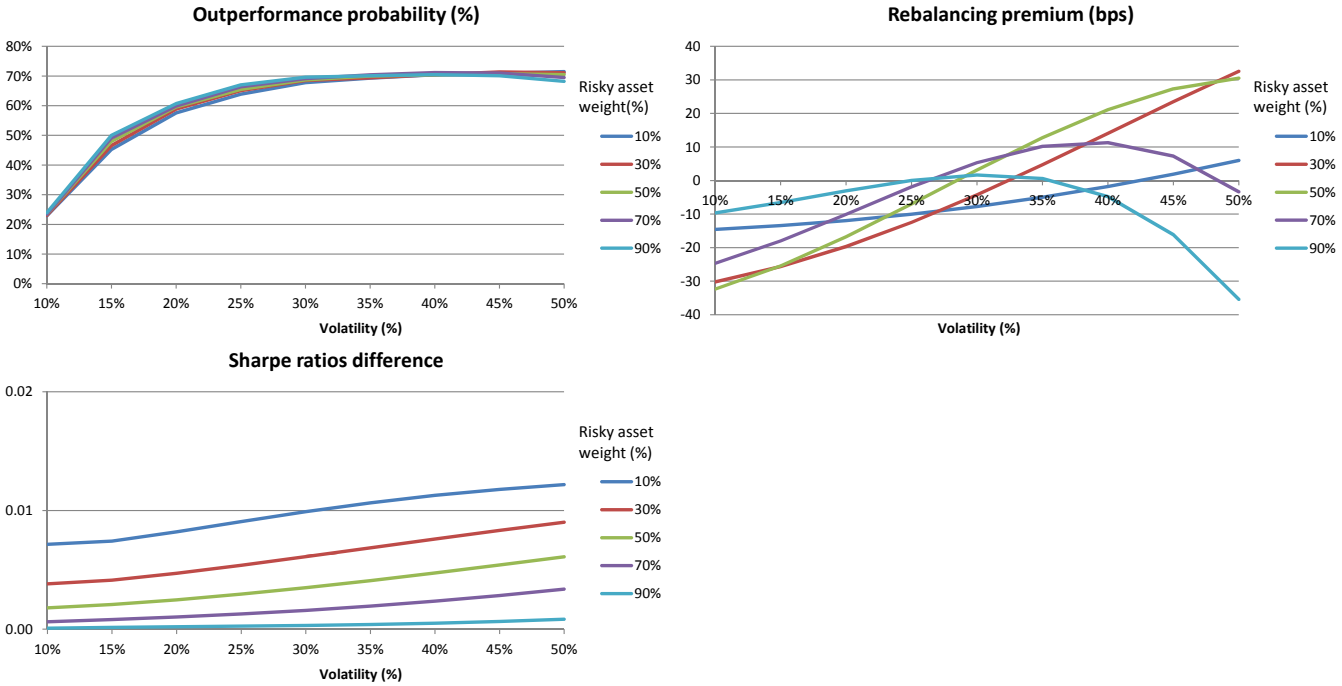
This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky asset for different values of the risky asset expected return. The time horizon is 5 years, the rebalancing frequency is 1 month and the pairwise correlation is 0%.

Two-dimensional robustness check with respect to volatility and risky asset weight

Figure 2.3 displays the outperformance probability, the rebalancing premium and the Sharpe ratios difference for different values of the risky asset weight (from 10% to 90%) and volatility (from 10% to 50%). If we consider a risky asset volatility greater than or equal to 20%, then the outperformance probability is greater than 57%. For a given volatility, the risky asset weight has little impact on the outperformance probability: for instance, for a risky asset volatility of 20%, the outperformance probability varies between 58% and 61%. In all cases, the rebalancing premium is less than or equal to 33 bps (reached for a risky asset volatility of 50% and a risky asset weight of 30%). For a 20% volatility, the rebalancing premium is negative regardless of the risky asset weight considered. In all cases, the difference in Sharpe ratios is less than 0.02 but positive.

In the case of one risky asset and one risk-free asset, the equal risk contribution portfolio is not relevant because the risk-free asset has zero volatility. The weighting scheme that maximises the excess growth rate $g^\star$ in this particular setting is the equally-weighted portfolio.[9]

Figure 2.3: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of one risky asset and one risk-free asset



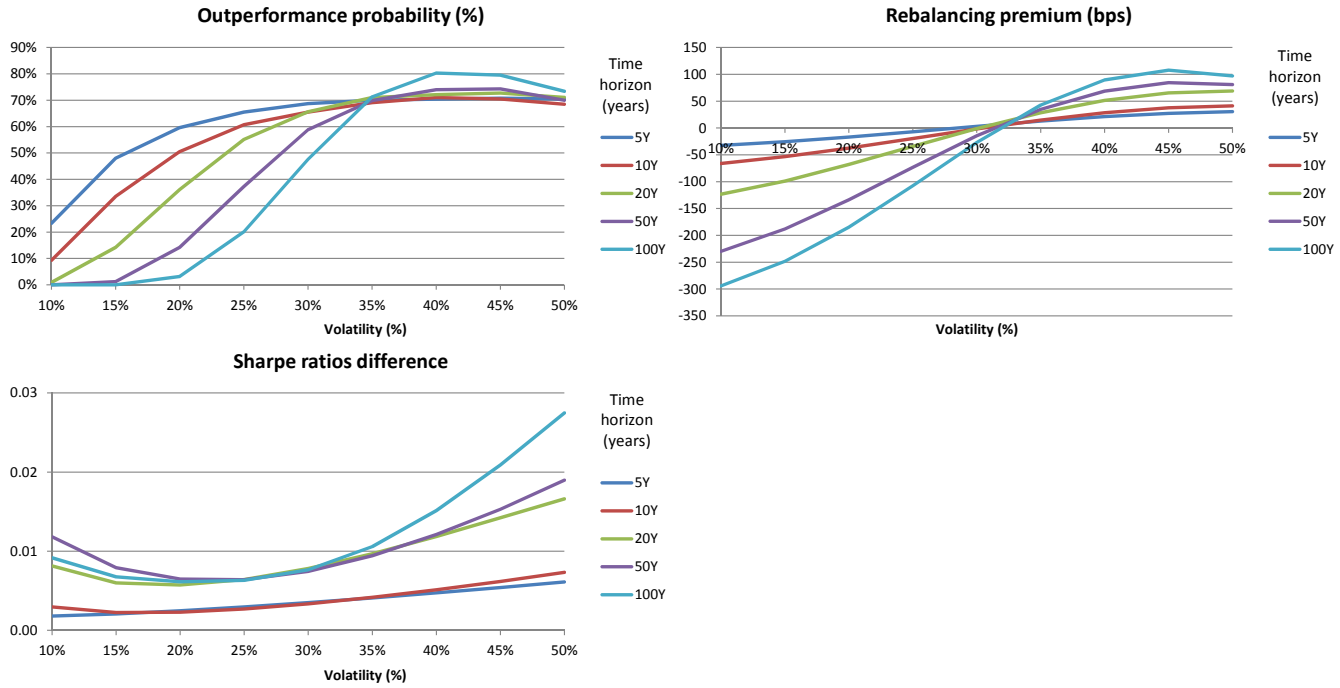
This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky asset for different values of the risky asset weight. The time horizon is 5 years and the rebalancing frequency is 1 month.

[9]The function $f(x) = x(1-x)$ reaches its maximum on the $[0;1]$ interval for $x = \frac{1}{2}$.

Two-dimensional robustness check with respect to volatility and time horizon

Figure 2.4 displays the outperformance probability, the rebalancing premium and the Sharpe ratios difference for different values of the time horizon (from 5 years to 100 years) and the risky asset volatility (from 10% to 50%). The outperformance probability is greater than or equal to 69% if we assume a risky asset volatility greater than or equal to 35% . If the risky asset volatility is less than 35% the outperformance probability decreases as the time horizon increases. If we consider a time horizon shorter than or equal to 20 years, then the rebalancing premium is less than 70 bps. If the risky asset volatility is less than 35%, then the rebalancing premium is negative and decreases as the time horizon increases. If the risky asset volatility is greater than or equal to 35% the rebalancing premium is positive (51 bps for a 40% volatility and a 20-year horizon). The Sharpe ratios difference between the rebalanced and the buy-and-hold portfolio is less than 0.03, and for a given level of volatility it increases as the time horizon increases.

Figure 2.4: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of one risky asset and one risk-free asset

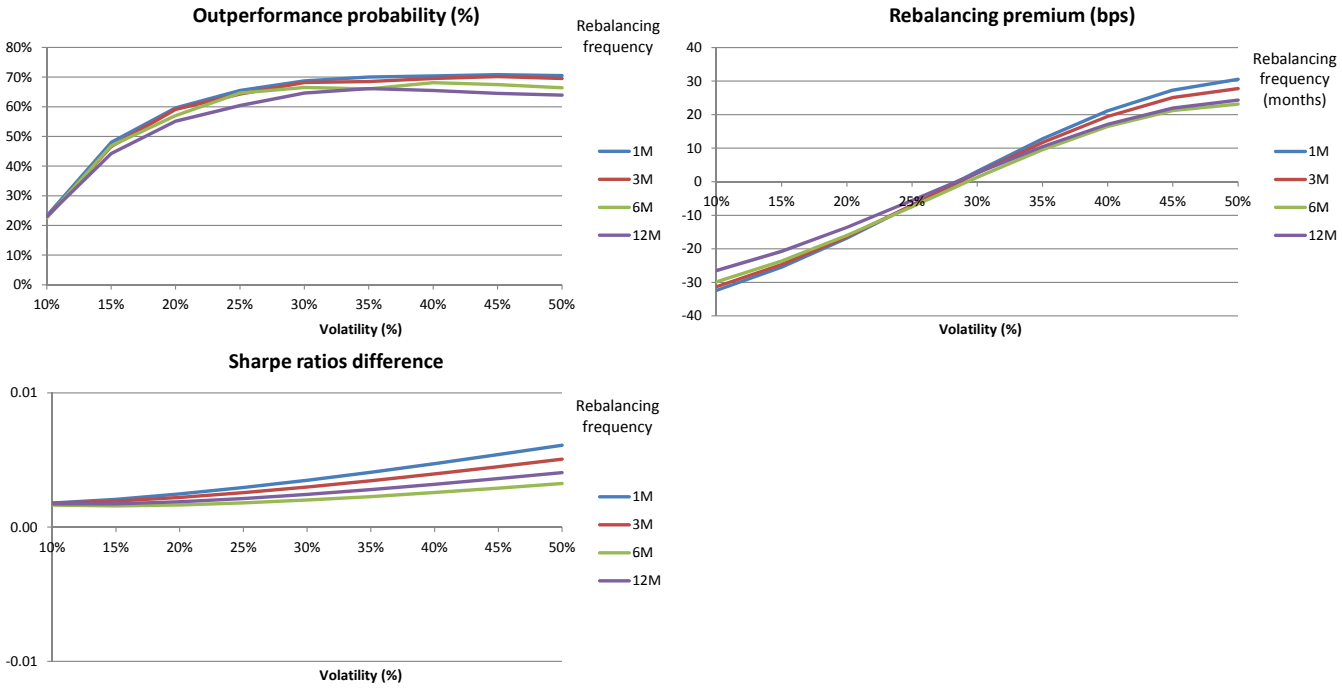


This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky asset for different values of the time horizon. The rebalancing frequency is 1 month, the volatility of the risky asset is 20% and the expected return of the risky asset is 10%.

Two-dimensional robustness check with respect to volatility and rebalancing frequency

Figure 2.5 displays the outperformance probability, the rebalancing premium and the difference in Sharpe ratios for different values of the rebalancing frequency (from 1 month to 12 months) and the risky asset volatility (from 10% to 50%). For a given level of volatility the rebalancing frequency seems to have little impact on the rebalancing metrics: if we fix the risky asset volatility to 20%, the outperformance probability is 60% for a 1-month rebalancing frequency and 57% for a six-month rebalancing frequency. For volatility greater than or equal to 30%, the rebalancing premium is always maximal for a monthly rebalancing frequency: if we fix the risky asset volatility to 40%, the rebalancing premium is 21 bps for a one-month rebalancing frequency and 16 bps for a six-month rebalancing frequency. The Sharpe ratios difference, for a given level of volatility, is not a monotonic function of the rebalancing frequency.

Figure 2.5: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of one risky asset and one risk-free asset



This figure displays the outperformance probability, rebalancing premium and Sharpe ratios difference as functions of the volatility of the risky asset for different values of the rebalancing frequency. The time horizon is 5 year, the volatility of the risky asset is 20% and the expected return of the risky asset is 10%.

2.2.4 Two risky assets

We now consider a setting with two risky assets following geometric Brownian motions with respective expected returns μ_1 and μ_2, respective volatilities σ_1, σ_2, covariance $\sigma_{1,2} = \sigma_1\sigma_2\rho_{1,2}$ (where $\rho_{1,2}$ is the pairwise correlation), and respective weights w_1 and $w_2 = 1 - w_1$.

2.2.4.1 Mathematical analysis

Using the same notation as previously, we have:

$$\ln\left(\frac{P^{reb}(t)}{P^{b\&h}(t)}\right) = g^\star t + \ln\frac{S_1(t)^{w_1}S_2(t)^{w_2}}{w_1S_1(t) + w_2S_2(t)} \tag{2.21}$$

where:

$$g^\star = \frac{1}{2}\left(w_1\sigma_1^2 + w_2\sigma_2^2 - w_1^2\sigma_1^2 - w_2^2\sigma_2^2 - 2w_1w_2\sigma_{1,2}\right) = \frac{1}{2}w_1(1-w_1)(\sigma_1^2 + \sigma_2^2 - 2\sigma_{1,2})$$

In Chapter 1, we show in a general setting with n assets that $g^\star$ is also the difference between the portfolio expected growth rate and the weighted-average expected growth rate of the assets in the portfolio. For this reason, $g^\star$ is called *excess growth rate* (of the rebalanced portfolio with respect to the weighted average of the individual expected growth rates). For fixed risky asset volatilities σ_1 and σ_2, $g^\star$ is maximum for $w_1 = w_2 = \frac{1}{2}$ and $\sigma_{1,2} = -1$.

In the end, Equation (2.21) can be rewritten as follows:

$$\ln\left(\frac{P^{reb}(t)}{P^{b\&h}(t)}\right) = g^\star t + \ln\frac{\left(\frac{S_1(t)}{S_2(t)}\right)^{w_1}}{w_1\frac{S_1(t)}{S_2(t)} + w_2} \tag{2.22}$$

so that

$$RP(0,t) = \mathbb{E}\left[\frac{1}{t}\ln\left(\frac{P_t^{reb}}{P_t^{b\&h}}\right)\right] = \underbrace{g^\star}_{\text{excess growth rate}} + \underbrace{\frac{1}{t}\mathbb{E}\left[\ln\frac{\left(\frac{S_1(t)}{S_2(t)}\right)^{w_1}}{w_1\frac{S_1(t)}{S_2(t)} + w_2}\right]}_{\text{dispersion term}} \tag{2.23}$$

Wise (1996) considers a notional derivative security defined by $S(t) = \frac{S_1(t)}{S_2(t)}$ mixed with a risk-free asset and calculates the outperformance probability based on the previous case with one risk-free asset and one risky asset. Denoting $w = w_1$, $\sigma^2 = \sigma_1^2 - 2\sigma_{1,2} + \sigma_2^2$, and defining $g = g_1 - g_2$ the difference between the expected growth rate of asset 1, $g_1 = \mu_1 - \frac{\sigma_1^2}{2}$, and that of asset 2, $g_2 = \mu_2 - \frac{\sigma_2^2}{2}$, we then have the following equation similar to Equation 2.11:

$$\mathbb{P}(P^{reb}(t) \geq P^{b\&h}(t)) = N\left(\frac{\ln y(t) - gt}{\sigma\sqrt{t}}\right) - N\left(\frac{\ln x(t) - gt}{\sigma\sqrt{t}}\right)$$

where $x(t)$ and $y(t)$ are respectively the lower and higher roots of the equation: $wS(t) + (1-w) - S(t)^w \exp\left(\frac{w(1-w)\sigma^2}{2}t\right)$.

Similarly to Equation (2.12), if the difference of the expected growth rates $g = g_1 - g_2$ satisfies:

$$-\frac{(1-w)\sigma^2}{2} < g_1 - g_2 < \frac{w\sigma^2}{2} \tag{2.24}$$

then the rebalancing portfolio outperforms the buy-and-hold portfolio with a probability which converges to 1 as the time horizon goes to infinity. Conversely, if the previous inequality is not verified (i.e., the gap between the two expected growth rates is too large), then the buy-and-hold

portfolio outperforms the rebalanced portfolio with a probability which converges to 1 as the time horizon goes to infinity. If we assume $w = \frac{1}{2}$, then Equation(2.24) becomes:

$$\frac{\sigma_1^2 - 3\sigma_2^2 + 2\sigma_{1,2}}{4} < \mu_1 - \mu_2 < \frac{3\sigma_1^2 - \sigma_2^2 - 2\sigma_{1,2}}{4} \tag{2.25}$$

If we consider the case where asset 2 is risk-free (i.e., $\mu_2 = r = 2\%$, $\sigma_2 = 0\%$ and $\sigma_{1,2} = 0\%$) and $w_1 = w_2 = \frac{1}{2}$, then Equation (2.25) can be simplified as $\frac{\sigma_1^2}{4} < \mu_1 - r < \frac{3\sigma_1^2}{4}$ which is the case that has been discussed before. If we now consider the case of two risky assets with respective volatilities of 30% and 20 % and with $w_1 = w_2 = \frac{1}{2}$, then Equation (2.25) becomes $-2.0\% \leq \mu_1 - \mu_2 \leq 2.0\%$ if we assume null pairwise correlation; and $-1.2\% \leq \mu_1 - \mu_2 \leq 1.2\%$ if we assume a 40% pairwise correlation. For two risky assets with a 30% volatility, null pairwise correlation and $w_1 = w_2 = \frac{1}{2}$, then Equation (2.25) becomes $-4.5\% \leq \mu_1 - \mu_2 \leq 4.5\%$.

These results suggest that as long as the difference in expected returns is not too large, the rebalanced portfolio is expected to outperform the buy-and-hold portfolio over long periods of time. In the particular case where $g_1 = g_2$, the condition (2.24) is always verified and then the rebalancing portfolio outperforms the buy-and-hold portfolio with a probability which converges to 1 as the time horizon goes to infinity. The amplitude of the rebalancing premium remains to be determined numerically.

2.2.4.2 Quantitative analysis

The quantitative analysis will allow us to obtain a quantitative assessment of the sign and magnitude of the rebalancing premium for different parameter values. In terms of outputs, as in the previous case, we propose to compute the outperformance probability and the distribution probability of the terminal value of the rebalanced portfolio minus the terminal value of the buy-and-hold portfolio with its main characteristics. We also show the distribution probability for the growth rate difference of the rebalanced portfolio minus that of the corresponding buy-and-hold portfolio and its main characteristics (not only its expected value but also its volatility, conditional expected value, etc.), assuming the following base case parameter values in Monte Carlo simulations:

- The assets are uncorrelated;
- The expected return on both assets is 10%;
- The volatility for both assets is 20%;
- An initial allocation $w_i = \frac{1}{2}$ invested in each asset;
- The rebalancing frequency is 1 month;
- The time horizon is 5 years.

The rebalancing premium, which is the expected value of the growth rate difference of the rebalanced portfolio over the corresponding buy-and-hold portfolio, is decomposed in the excess growth rate term and the dispersion term (see Equation (2.23)).

We will also compute the following statistics for each strategy: annualised performance, annualised volatility, Sharpe ratio, skewness, kurtosis, maximum drawdown, 1% Value at Risk (one month horizon), certainty-equivalent return for an investor who has a constant relative risk aversion (CRRA) utility function with risk aversion parameter γ, and finally annualised one-way turnover. Robustness checks will then be performed with respect to assumptions regarding the level of expected return and volatility of the risky assets, the rebalancing frequency, the portfolio weights and the time horizon.

Base case scenario

Table 2.2 displays statistics relative to the rebalanced and the buy-and-hold portfolios in the base case for two risky assets. From a risk-adjusted performance perspective, the two portfolios are similar (expected return of 4.6%, annualised volatility of 14.2% and certainty equivalent return (for $\gamma = 2$) of 2.6% for the rebalanced portfolio versus 4.6%, 14.5% and 2.5% for the buy-and-hold portfolio). From an extreme risk perspective the statistics of the two portfolios are close (maximum drawdown of 24.2% and monthly 1% Value at Risk of -8.7% for the rebalanced portfolio versus 24.7% and -8.9% for the buy-and-hold portfolio). The annualised one-way turnover of the rebalanced portfolio is 19.4%.

For a 5-year time horizon, the outperformance probability is 68% and the rebalancing premium is 2 bps (with a confidence interval at the 95% level equal to [-6, 10] bps) showing that the amplitude of the difference between the expected growth rates of the two portfolios is low in our base case scenario. As a reminder, the rebalancing premium when considering a risky asset and a risk-free asset was -17 bps in the base case scenario.

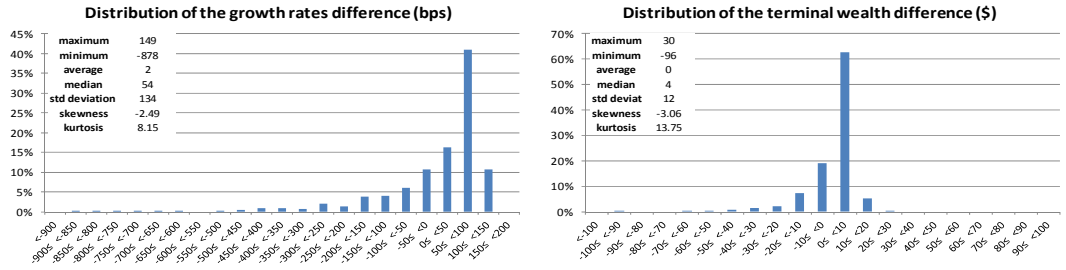
Table 2.2: Statistics for a 5-year time horizon

	Rebalanced Portfolio	**Buy-and-Hold Portfolio**
Expected return	9.6%	9.7%
Annualised volatility	14.2%	14.5%
Sharpe Ratio	0.54	0.53
Skewness	0.11	0.11
Kurtosis	2.95	2.96
Maximum Drawdown	19.4%	19.9%
Monthly VaR 1%	-8.3%	-8.5%
Turnover	19.4%	0.0%
CEQ (γ=2)	7.6%	7.5%
CEQ (γ=5)	4.6%	4.4%
Outperformance probability	68%	
Rebalancing premium (bps)	2	
Rebalancing premium $CI_{95\%}$ (bps)	$[-6, 10]$	
Excess growth rate g*(bps)	100	
Dispersion term (bps)	-98	
Conditional (>0) Rebalancing premium (bps)	70	
Conditional (<0) Rebalancing premium (bps)	-144	

This table reports the main performance, risk and rebalancing statistics for the rebalanced portfolio and the corresponding buy-and-hold portfolio in the base case scenario with an investment universe composed of 2 risky assets. The time horizon is 5 years.

Figure 2.6 displays the distribution of the rebalanced portfolio growth rate minus the buy-and-hold portfolio growth rate. The distribution is negatively skewed (-2.49), the minimum difference in the 1000 scenarios is -878 bps whereas the maximum difference is 149 bps. In 41% of the scenarios, the growth rates difference is between 50 bps and 100 bps. Figure 2.6 also displays the distribution of the rebalanced portfolio terminal value minus the buy-and-hold portfolio terminal value. Each portfolio has an initial wealth of $100. The distribution is negatively skewed (-3.06), the minimum difference in the 1,000 scenarios is -$96 whereas the maximum difference is $30. In 63% of the scenarios, the terminal wealth difference is between $0 and $10.

Figure 2.6: Distributions of the growth rates difference (bps) and of the terminal wealth difference($) for a 5-year time horizon



This figure displays the probability distribution of the difference between the growth rate of the rebalanced portfolio and that of the corresponding buy-and-hold portfolio in the base case for an investment universe made of 2 risky assets. It also displays the probability distribution of the difference between the terminal wealth of the rebalanced portfolio and that of the corresponding buy-and-hold portfolio in the base case for an investment universe made of two risky assets. The initial wealth of each portfolio is $100. The time horizon is 5 years.

We study the sensitivity of the rebalancing premium with respect to the expected return of the risky assets for a given level of volatility, all things being equal. We find that if we assume that both assets have the same expected return (and not necessarily the same volatility), then the level of expected return has no influence on the amplitude of the rebalancing premium. As this result can be demonstrated with the calculation below, we do not even test the effect of changing the expected return.

Let us replace $S_1(t) = S(0)\exp((\mu - \frac{\sigma_1^2}{2})t + \sigma_1 W_1(t))$ and $S_2(t) = S(0)\exp((\mu - \frac{\sigma_2^2}{2})t + \sigma_2 W_2(t))$ in Equation (2.22). Then we have:

$$\begin{aligned}\ln\left(\frac{P^{reb}(t)}{P^{b\&h}(t)}\right) &= g^\star t + \ln\frac{S_1(t)^{w_1}S_2(t)^{w_2}}{w_1 S_1(t) + w_2 S_2(t)} \\ &= g^\star t + \left(\ln\frac{\exp\left(-\left(w_1\frac{\sigma_1^2}{2} + w_2\frac{\sigma_2^2}{2}\right)t\right)\exp\left(\sigma_1 W_1(t) + \sigma_2 W_2(t)\right)}{w_1\exp\left(-\frac{\sigma_1^2}{2}t\right)\exp\left(\sigma_1 W_1(t)\right) + w_2\exp\left(-\frac{\sigma_2^2}{2}t\right)\exp\left(\sigma_2 W_2(t)\right)}\right)\end{aligned}$$

So in that particular case, the rebalancing premium does not depend on the expected return. This result is not valid anymore if the two risky assets have different expected returns.

Robustness check with respect to the weighting scheme considered

In the case of 2 risky assets with the assumption of a constant covariance matrix, the weighting scheme that maximises the excess growth rate $g^\star$ is the equally-weighted one.[10] In addition to the naively diversified equally-weighted portfolio considered in the base case we examine another weighting scheme: the Equal Risk Contribution (ERC).

The marginal risk contributions verify (see Roncalli (2013) for further details):

$$\begin{cases}\frac{\partial\sigma(w_1,w_2)}{\partial w_1} = \frac{w_1\sigma_1^2 + (1-w_1)\rho_{1,2}\sigma_1\sigma_2}{\sigma} \\ \frac{\partial\sigma(w_1,w_2)}{\partial w_2} = \frac{(1-w_1)\sigma_2^2 + w_1\rho_{1,2}\sigma_1\sigma_2}{\sigma}\end{cases}$$

where $\rho_{1,2} = \frac{\sigma_{1,2}}{\sigma_1\sigma_2}$ is the pairwise correlation of assets 1 and 2, and σ is the rebalanced portfolio volatility.

The ERC rebalanced portfolio verifies the equality of the two risk contributions [11]:

$$w_1^2\sigma_1^2 = (1-w_1)^2\sigma_2^2$$

.

The unique [12] solution is given by:

$$w_1 = \frac{\frac{1}{\sigma_1}}{\frac{1}{\sigma_1} + \frac{1}{\sigma_2}} = \frac{\sigma_2}{\sigma_1 + \sigma_2}$$

We assess that the asset weights on the portfolio are inversely proportional to their volatilities and do not depend on the pairwise correlation $\rho_{1,2}$. In the case where the two risky assets have different volatilities (and potentially different expected returns) the EW portfolio and the ERC portfolio have not the same weights anymore.

We consider firstly in this robustness check an investment universe composed by a first risky asset with expected return $\mu_1 = 10\%$ and volatility $\sigma_1 = 20\%$, and a second risky asset with

[10] $g^\star = \frac{1}{2}w_1(1-w_1)(\sigma_1^2 + \sigma_2^2 - 2\sigma_{1,2})$ so maximising $g^\star$ is equivalent to maximise the function $f(x) = x(1-x)$ on the [0;1] interval.

[11] The risk contribution of asset 1 to volatility is equal to its marginal contribution (i.e., $\frac{\partial\sigma(w_1,w_2)}{\partial w_1}$) times the weight of asset 1 in the portfolio (i.e., w_1).

[12] Because our framework is long-only.

expected return $\mu_2 = 14\%$ and volatility $\sigma_2 = 30\%$, so the assets have the same Sharpe ratio of 0.40. All other parameters (pairwise correlation, rebalancing frequency and time horizon) are unchanged compared with the base case. Table 2.3 shows that the Sharpe ratio of the ERC portfolio is slighty higher than that of the EW portfolio (0.54 versus 0.52). The outperformance probability is the same for both portfolios (68%), the Sharpe ratios difference is nearly the same (0.014 for the EW portfolio and 0.015 for the ERC portfolio) and the rebalancing premium is slightly higher for the EW portfolio (6 bps with a confidence interval at the 95% level equal to [2, 10] bps) than for the ERC portfolio (3 bps with a confidence interval at the 95% level equal to [0, 6] bps).

Table 2.3: Statistics of an equally-weighted portfolio and an equal risk contribution portfolio for a 5-year time horizon

	Equally Weighted	**Equal Risk**
Expected return	11.5%	11.2%
Volatility	18.2%	17.1%
Sharpe ratio	0.52	0.54
Outperformance prob	68%	68%
Sharpe ratios difference	0.014	0.015
Rebalancing premium (bps) (1)+(2)	6	3
Rebalancing premium $CI_{95\%}$ (bps)	$[2, 10]$	$[0, 6]$
Excess growth rate g*(bps) (1)	163	156
Dispersion term (bps) (2)	-157	-153

This table reports on the first three raws the expected return, the volatility and the Sharpe ratio of an equally-weighted portfolio and an equal risk contribution portfolio when the investment universe is made of 2 risky assets with respective expected return of 10% and 14% and respective volatilities of 20% and 30%. The five other raws display the rebalancing statistics of each portfolio. The time horizon is 5 years.

We now consider another investment universe composed by a first risky asset with expected return $\mu_1 = 10\%$ and volatility $\sigma_1 = 20\%$, and a second risky asset with expected return $\mu_2 = 22\%$ and volatility $\sigma_2 = 50\%$, with all other parameters unchanged. Table 2.4 shows that the volatility of the EW portfolio is sensibly higher (27.3%) than that of the ERC portfolio (23.3%) and the ERC portfolio displays a better Sharpe ratio (0.52) than the EW portfolio (0.49). The outperformance probability is the same for both portfolios (68%), the Sharpe ratios difference is sensibly higher for the ERC portfolio (0.028) than for the EW portfolio (0.017) and the rebalancing premium is higher for the EW portfolio (37 bps with a confidence interval at the 95% level equal to [12, 62] bps) than for the ERC portfolio (29 bps with a confidence interval at the 95% level equal to [5, 54] bps).

Table 2.4: Statistics of an equally-weighted portfolio and an equal risk contribution portfolio for a 5-year time horizon

	Equally Weighted	**Equal Risk**
Expected return	15.3%	14.0%
Volatility	27.3%	23.3%
Sharpe ratio	0.49	0.52
Outperformance prob	68%	68%
Sharpe ratios difference	0.017	0.028
Rebalancing premium (bps) (1)+(2)	37	29
Rebalancing premium $\text{CI}_{95\%}$ (bps)	$[12, 62]$	$[5, 54]$
Excess growth rate $g^\star$(bps) (1)	363	344
Dispersion term (bps) (2)	-326	-315

This table reports on the first three raws the expected return, the volatility and the Sharpe ratio of an equally-weighted portfolio and an equal risk contribution portfolio when the investment universe is made of 2 risky assets with respective expected return of 10% and 22% and respective volatilities of 20% and 50%. The five other raws display the rebalancing statistics of each portfolio. The time horizon is 5 years.

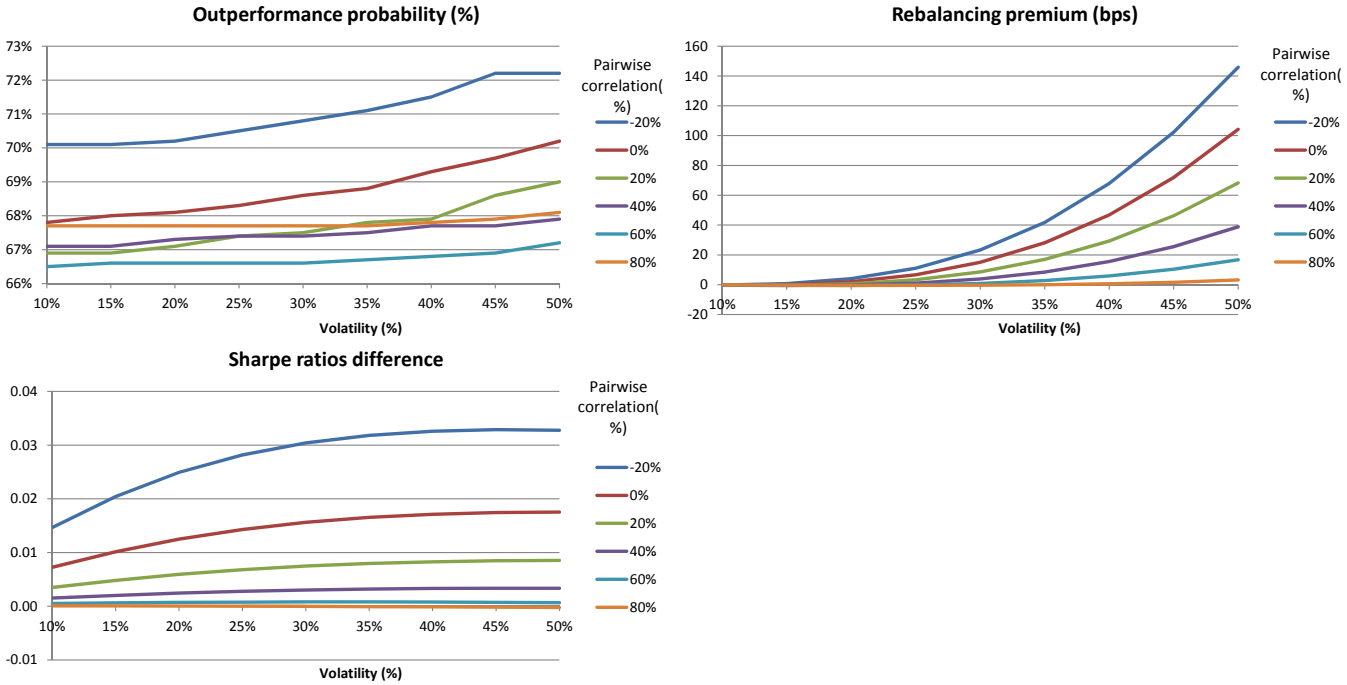
Overall, the previous comparisons between an EW portfolio and an ERC portfolio considering two different investment universes seems to indicate that the ERC weighting scheme does not enhance the rebalancing premium compared with a naive EW weighting scheme.

In the following robustness checks all the portfolios considered are initially equally-weighted.

Two-dimensional robustness check with respect to volatility and pairwise correlation

Figure 2.7 displays the outperformance probability, the rebalancing premium and the Sharpe ratios difference relatively to different values of risky assets pairwise correlation (from -20% to 80%) and volatility (from 0% to 50%). The outperformance probability is between 66% and 73%. For a given level of volatility, the rebalancing premium increases as the pairwise correlation ρ decreases: for σ=45%, the rebalancing premium is 26 bps if ρ=40% and 102 bps if ρ=-20%. The Sharpe ratios difference, for a given level of volatility, decreases as the pairwise correlation increases.

Figure 2.7: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of 2 risky assets

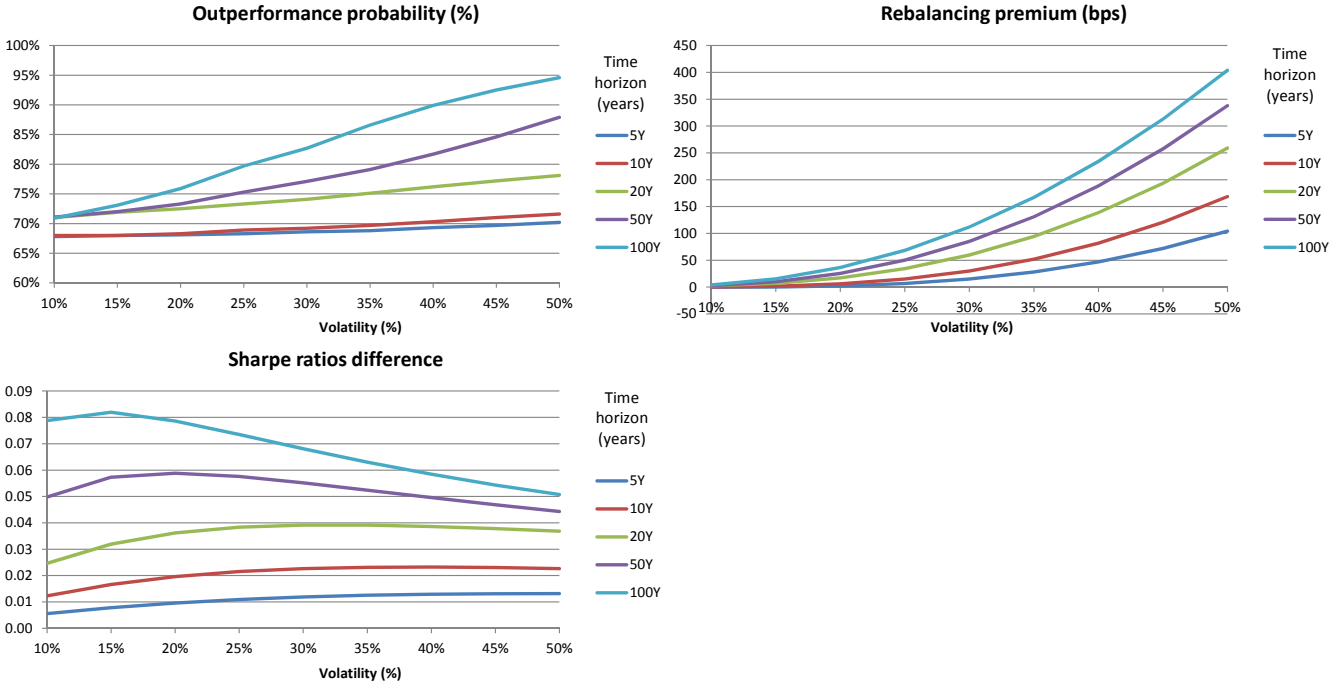


This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky assets for different values of the pairwise correlation. The rebalancing frequency is 1 month and the time horizon is 5 years.

Two-dimensional robustness check with respect to volatility and time horizon

Figure 2.8 displays the outperformance probability, the rebalancing premium and the Sharpe ratios difference relatively to different values of the time horizon (from 5 years to 100 years) and risky assets volatility (from 10% to 50%). For a given level of volatility, the outperformance probability increases as the time horizon increases: for σ=30%, the outperformance probability is 69% for a 5-year horizon and 83% for a 100-year horizon. For a given level of volatility, the rebalancing premium increases as the time horizon increases: for σ=40%, the rebalancing premium is 47 bps for a 5-year horizon and 188 bps for a 50-year horizon. The Sharpe ratios difference between the rebalanced and the buy-and-hold portfolios increases as the time horizon increases.

Figure 2.8: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of 2 risky assets

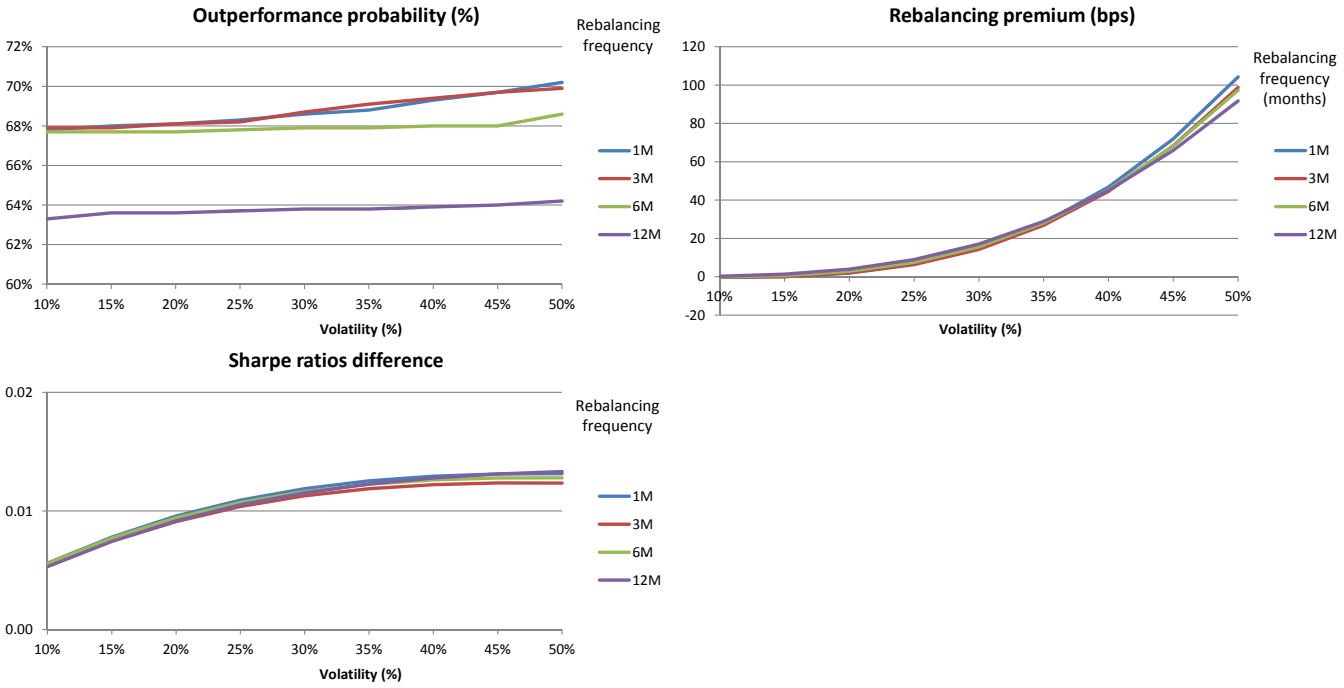


This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky assets for different values of the time horizon. The rebalancing frequency is 1 month and the pairwise correlation is null.

Two-dimensional robustness check with respect to volatility and rebalancing frequency

Figure 2.9 displays the outperformance probability, the rebalancing premium and the Sharpe ratios difference relatively to different values of the rebalancing frequency (from 1 month to 12 months) and risky assets volatility (from 10% to 50%). The outperformance probability is between 63% and 71%. For a given level of volatility, the impact of the rebalancing frequency on the rebalancing metrics is low.

Figure 2.9: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of 2 risky assets



This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky assets for different values of the rebalancing frequency. The time horizon is 5 years and the pairwise correlation is null.

2.2.5 n risky assets (n > 2)

We now turn to the general case involving n risky assets with expected returns $(\mu_i)_{1\leq i\leq n}$, covariance matrix $(\sigma_{i,j})_{1\leq i,j\leq n}$ and weights $(w_i)_{1\leq i\leq n}$. We denote by $g_i = \mu_i - \frac{\sigma_i^2}{2}$ the expected growth rate on asset i[13]. This is the setting considered by Fernholz (2002) in its continuous time analysis of stock market structure and portfolio behavior using stochastic calculus (see Chapter 1 for more details). Following Fernholz and Shay (1982), Fernholz (2002) and Gabay and Herlemont (2007), we use this setup in our analytical study to compare a rebalanced strategy with the corresponding static buy-and-hold strategy in a continuous-time framework.

2.2.5.1 Mathematical analysis

In this setting, it can be shown (see Appendix 2.C) that the value of the rebalanced portfolio $P^{reb}(t)$ and of the corresponding passive buy-and-hold portfolio $P^{b\&h}(t)$ are:

$$P^{reb}(t) = \exp(g^\star t)\prod_{i=1}^{n} S_i(t)^{w_i}$$

$$P^{b\&h}(t) = \sum_{i=1}^{n} w_i S_i(t)$$

where $g^\star = \frac{1}{2}\left(\sum_{i=1}^{n} w_i\sigma_i^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{i,j}\right)$ is half of the difference between the weighted-average variance of the assets in the rebalanced portfolio and the rebalanced portfolio variance.

If we assume null pairwise correlation between all assets, initial equal weights and same volatility σ for all assets then:

$$g^\star = \frac{\sigma^2}{2}\left(1-\frac{1}{n}\right) \underset{n\to+\infty}{\longrightarrow} \frac{\sigma^2}{2}$$

It can also be shown that under the approximation made in Equation (2.3), we have:

$$g^\star = g_p - \sum_{i=1}^{n} w_i g_i \text{ (see Chapter 1 for more details)}$$

where g_p is the expected growth rate of the rebalanced portfolio, which is the reason why $g^\star$ is known as the excess growth rate (of the rebalanced portfolio with respect to the weighted average expected growth rate of individual assets). It turns out that $g^\star$ is a key component of the rebalancing premium. To see this, consider the following equation:

$$d\left(\ln\frac{P^{reb}(t)}{P^{b\&h}(t)}\right) = g^\star dt + d\left(\ln\frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right) \tag{2.26}$$

which implies:

$$RP(0,t) = \mathbb{E}\left[\frac{1}{t}\ln\left(\frac{P_t^{reb}}{P_t^{b\&h}}\right)\right] = g^\star + \frac{1}{t}\mathbb{E}\left[\ln\left(\frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right)\right] \tag{2.27}$$

[13] We define $\sigma_i = \sqrt{\sigma_{i,i}}$ as the volatility of asset i.

The first (positive) term $g^\star$ is the excess growth rate of the rebalanced portfolio (with respect to the weighted average expected growth rate of individual assets), which is a measure of portfolio diversification that is higher in situations involving highly volatile assets that are combined to form a low volatility portfolio through an efficient use of the pairwise correlations. On the other hand the second term is negative by concavity of the ln function (see Appendix 2.D for more details). We note that if at any time t every asset has the same price (i.e., $\forall i \in [\![1;n]\!]$, $S_i(t) = S(t)$) then this second term is null independently of the weighting scheme considered and could be as such interpreted as a dispersion term. Moreover, we can show that the dispersion term is linked to the weighted variance of the growth rates across the assets and goes to zero in case the terminal value is identical for all assets in the portfolio (see Appendix 2.E for further details).

A quantitative analysis is required to measure the net outcome resulting from these competing effects. In particular, it will allow us to spell out conditions under which the rebalancing premium, which results from the sum of a positive and a negative component, turns out to be positive.

2.2.5.2 Quantitative analysis

We propose to compute the outperformance probability of a rebalanced portfolio compared with the corresponding buy-and-hold portfolio. We also show the distribution probability for the growth rate difference of the rebalanced portfolio minus that of the corresponding buy-and-hold portfolio and its main characteristics (not only its expected value but also its volatility, conditional expected value, etc.), assuming the following base case setting in Monte Carlo simulations:

- The investment universe is composed by 30 risky assets;
- The assets are uncorrelated;
- The expected return of each asset is 10%;
- the volatility of each asset is 20%;
- Initial weights $w_i = \frac{1}{n}$ invested in each asset;
- The rebalancing frequency is 1 month;
- The time horizon is 5 years.

The rebalancing premium, which is the expected value of the excess growth rate of the rebalanced portfolio over that of the corresponding buy-and-hold portfolio, is decomposed in the excess growth rate term and the dispersion term (see Equation (2.27)).

We will also compute the following statistics for each strategy: annualised performance, annualised volatility, Sharpe ratio, skewness, kurtosis, maximum drawdown, 1% Value at Risk (1 month horizon), certainty-equivalent return for an investor who has a constant relative risk aversion (CRRA) utility function with risk aversion parameter γ, and finally annualised 1-way turnover. Robustness checks are performed relatively to assumptions regarding the level of expected return and volatility of the risky assets, the rebalancing frequency, the portfolio weights and the time horizon.

Base case scenario

Table 2.5 displays statistics relative to the rebalanced and the buy-and-hold portfolios made of 30 risky assets. The two portfolios are similar from a risk-adjusted performance perspective: the rebalanced portfolio has an expected return of 5.0%, an annualised volatility of 3.7% and a certainty equivalent return (for $\gamma = 2$) of 9.9% while the buy-and-hold portfolio displays corresponding statistics of 5.0%, 3.9% and 9.9%). From an extreme risk perspective the statistics of the two portfolios are close (maximum drawdown of 3.6 % and monthly 1% Value at Risk of -2.0% for the rebalanced portfolio versus 3.9% and -2.1% for the buy-and-hold portfolio. The annualised 1-way turnover of the rebalanced portfolio is 27.2%.

For a 5-year time horizon, the outperformance probability is 56% and the rebalancing premium is 2 bps (with a confidence interval at the 95% level equal to [-1, 5] bps). The amplitude of the rebalancing premium in the base case scenario is as low as that of the previous base case scenario with 2 risky assets.

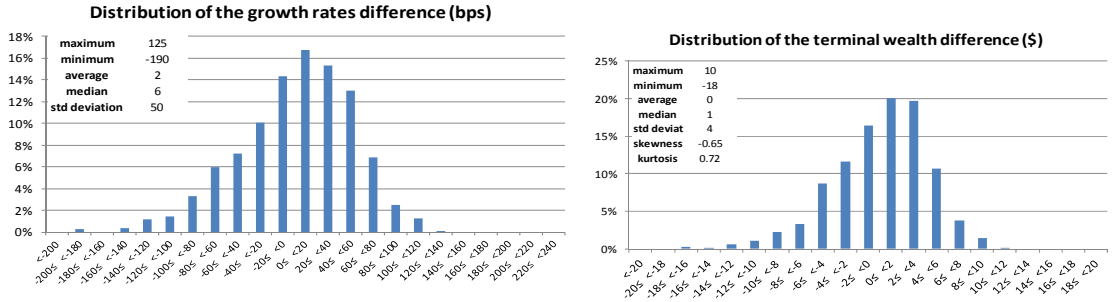
Table 2.5: Statistics for a 5-year time horizon

	Rebalanced Portfolio	**Buy-and-Hold Portfolio**
Expected return	10.1%	10.0%
Annualised volatility	3.7%	3.9%
Sharpe Ratio	2.18	2.08
Skewness	0.03	0.03
Kurtosis	2.91	2.91
Maximum Drawdown	3.6%	3.9%
Monthly VaR 1%	-2.0%	-2.1%
Turnover	27.2%	0.0%
CEQ (γ=2)	9.9%	9.9%
CEQ (γ=5)	9.7%	9.7%
Outperformance probability	56%	
Rebalancing premium (bps)	2	
Rebalancing premium $CI_{95\%}$ (bps)	$[-1, 5]$	
Excess growth rate $g^\star$ (bps)	193	
Dispersion term (bps)	-191	
Conditional (>0) Rebalancing premium (bps)	38	
Conditional (<0) Rebalancing premium (%)	-43	

This table reports the main performance, risk and rebalancing statistics in the base case scenario for the rebalanced portfolio and the corresponding buy-and-hold portfolio in the case where the investment universe is composed of 30 risky assets. The time horizon is 5 years.

Figure 2.10 displays the distribution of the rebalanced portfolio growth rate minus the buy-and-hold portfolio growth rate. The distribution is negatively skewed (-0.53) the minimum difference in the 1000 scenarios is -190 bps whereas the maximum difference is 125 bps. In 45% of the scenarios, the growth rates difference is between 0 bps and 60 bps. Figure 2.10 also displays the distribution of the rebalanced portfolio terminal wealth minus the buy-and-hold portfolio terminal wealth. Each portfolio has an initial wealth of 100$. The distribution is negatively skewed (-0.65), the minimum difference in the 1000 scenarios is -18$ whereas the maximum difference is 10$. In 56% of the scenarios, the terminal wealth difference is between 0$ and 10$.

Figure 2.10: Distributions of the growth rates difference (bps) and of the terminal wealth difference($) for a 5-year time horizon



This figure displays the probability distribution of the difference between the growth rate of the rebalanced portfolio and that of the corresponding buy-and-hold portfolio in the base case for an investment universe made of 30 risky assets. It also displays the probability distribution of the difference between the terminal wealth of the rebalanced portfolio and that of the corresponding buy-and-hold portfolio in the base case for an investment universe made of 30 risky assets. The initial wealth of each portfolio is 100$. The time horizon is 5 years.

Similarly to the previous case with 2 risky assets, if we consider $n (n > 2)$ risky assets with the same expected return μ and possibly different volatilities then the level of expected return has no influence on the amplitude of the rebalancing premium. If we replace for each asset i $(i \in [\![1;n]\!])$ $S_i(t) = S(0)\exp\left(\left(\mu - \frac{\sigma_i^2}{2}\right)t + \sigma_i W_i(t)\right)$ on Equation (2.26), we obtain:

$$\ln\left(\frac{P^{reb}(t)}{P^{b\&h}(t)}\right) = g^\star t + \left(\ln \frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right)$$

$$= g^\star t + \ln\left(\frac{\exp\left(-\left(\sum_{i=1}^{n} w_i \frac{\sigma_i^2}{2}\right)t\right)\exp\left(\sum_{i=1}^{n}\sigma_i W_i(t)\right)}{\sum_{i=1}^{n} w_i \exp\left(-\frac{\sigma_i^2}{2}t\right)\exp\left(\sigma_i W_i(t)\right)}\right)$$

Robustness check with respect to the weighting scheme considered

In addition to the naively diversified equally-weighted portfolio considered in the base case we examine 2 different weighting schemes: the Equal Risk Contribution (ERC) portfolio and the Max $g^\star$ weighting scheme which, given the covariance matrix maximises the excess growth term: $\frac{1}{2}\left(\sum_{i=1}^{n} w_i(t)\sigma_{i,i}(t) - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i(t)\sigma_{i,j}(t)w_j(t)\right)$.

Under the assumption that every pair of risky assets has the same constant pairwise correlation (i.e., $\forall (i,j) \in [\![1;n]\!]^2,\ i \neq j\ :\ \rho_{i,j} = \frac{\sigma_{i,j}}{\sigma_i \sigma_j} = \rho$), the weights of the ERC portfolio verify the following relationship:

$$w_i\sigma_i\left((1-\rho)\, w_i\sigma_i + \rho\sum_{k=1}^{n} w_k\sigma_k\right) = w_j\sigma_j\left((1-\rho)\, w_j\sigma_j + \rho\sum_{k=1}^{n} w_k\sigma_k\right) \tag{2.28}$$

Equation (2.28) implies that $\forall (i,j) \in [\![1;n]\!]^2,\ i \neq j\ :\ w_i\sigma_i = w_j\sigma_j$ and we have:

$$w_i = \frac{\sigma_i^{-1}}{\sum_{j=1}^{n}\sigma_j^{-1}}$$

The $(n \times 1)$ weight vector $w^* = (w_1, ..., w_n)$ which maximises the excess growth rate $g^\star$ is:

$$w^* = \frac{1}{2}\Sigma^{-1}V - \frac{1}{2}\frac{{}^tI\Sigma^{-1}V - 2}{{}^tI\Sigma^{-1}I}\Sigma^{-1}I$$

where Σ is the $(n \times n)$ covariance matrix, V the $(n \times 1)$ stocks' variance vector and I a $(n \times 1)$ vector of ones.

This analytical solution can lead to possibly negative weights, so in our long-only framework we have to compute the solution numerically. If we consider, like in the base case, assets with the same volatility and pairwise correlation then the 3 weighting schemes mentioned above are the same. We will assume the following alternative case setting:

- 30 risky assets: 10 low-volatility risky assets with 15% volatility, 10 medium-volatility risky assets with 30% volatility and 10 high-volatility risky assets with 45% volatility
- Each asset expected return μ_i is defined such as every asset has the same Sharpe ratio of 0.40 (i.e., $\mu_i = 2\% + 0.40 \times \sigma_i$);

- All pairwise correlations are equal to 20%;
- Rebalancing frequency is 1 month;
- Time horizon is 5 years;

We constrain the weights to be positive and higher than or equal to $\frac{1}{3n}$ for the Max $g^\star$ portfolio. Table 2.6 displays that the EW and ERC portfolios have the highest Sharpe ratios (0.82 and 0.83). The Sharpe ratio of the Max $g^\star$ portfolio is slightly below (0.78). The Max $g^\star$ portfolio has the highest volatility (20.8%). The rebalancing premium is higher if we initially consider the Max $g^\star$ weighting scheme: 15 bps. So are the outperformance probability (60%) and the Sharpe ratios difference (0.041). This simple example illustrates the potential impact of considering alternative initial weighting schemes to enhance the rebalancing premium: the rebalancing premium is 40 bps higher for the Max $g^\star$ weighting scheme than for the EW weighting scheme. However the ERC and Max $g^\star$ weighting schemes become time-varying in the empirical analysis where the covariance matrix is time-varying: in the empirical analysis, we will then have to estimate the covariance matrix on the whole sample for computing the initial weights relative to this two weighting schemes.

Table 2.6: Statistics of an equally-weighted portfolio, an equal risk contribution portfolio and a Max $g^\star$ portfolio for a 5-year time horizon

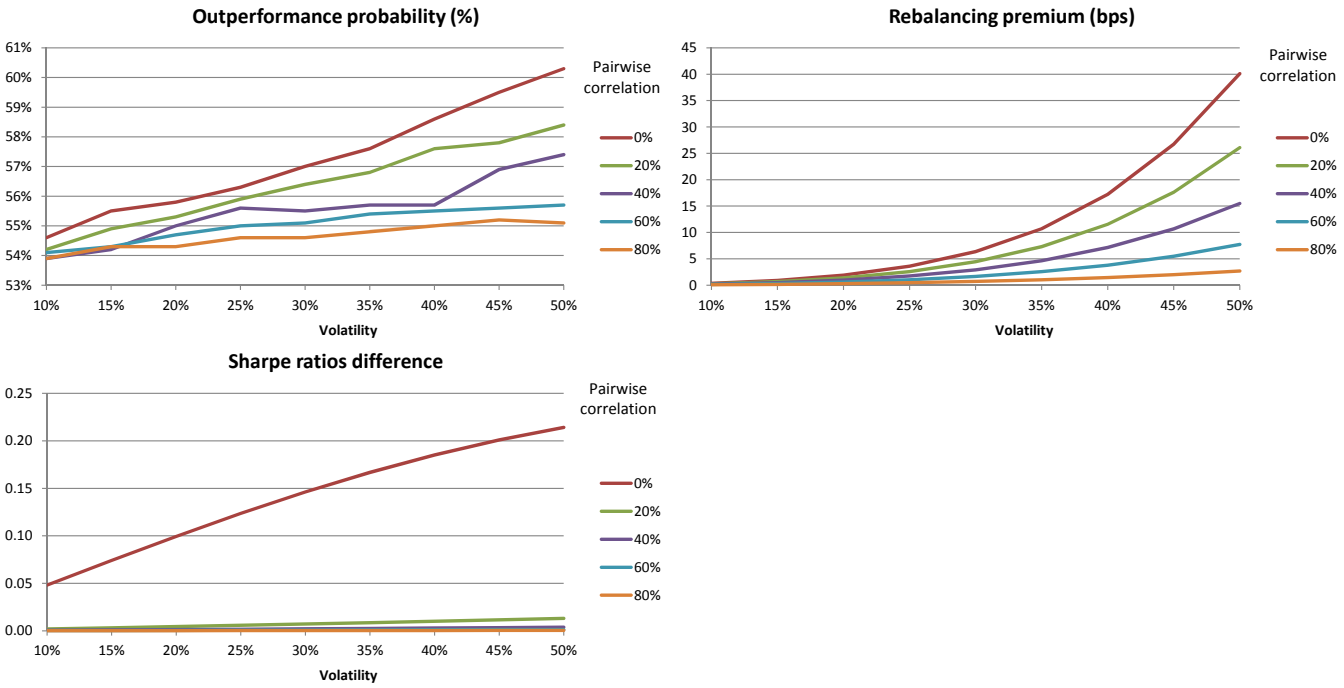
	Equally Weighted	**Equal Risk**	**Max g***
Expected return	14.0%	11.8%	18.2%
Volatility	14.6%	11.8%	20.8%
Sharpe ratio	0.821	0.827	0.779
Outperformance prob	54%	48%	60%
Sharpe ratios difference	0.028	0.018	0.041
Rebalancing premium (bps) (1)+(2)	-25	-33	15
Rebalancing premium $\text{CI}_{95\%}$ (bps)	$[-36, -14]$	$[-41, 25]$	$[-2, 33]$
Excess growth rate $g^\star$(bps) (1)	421	300	640
Dispersion term (bps) (2)	446	333	625

This table reports on the first three raws the expected return, the volatility and the Sharpe ratio of an equally-weighted portfolio, an equal risk contribution portfolio and a Max $g^\star$ portfolio when the investment universe is made of 30 risky assets (10 low volatility risky assets with 15% volatility, 10 medium volatility risky assets with 30% volatility and 10 high volatility risky assets with 45% volatility) with the same pairwise correlation of 20%. Each asset expected return is defined such as every asset has the same Sharpe ratio of 0.40. The five other raws display the rebalancing statistics of each portfolio. The time horizon is 5 years.

Two-dimensional robustness check with regard to volatility and pairwise correlation

Figure 2.11 displays the outperformance probability, the rebalancing premium and the Sharpe ratios difference relatively to different values of the risky assets pairwise correlation (from 0% to 80%) and volatility (from 0% to 50%). The outperformance probability is between 53% and 61% when the volatility varies between 10% and 50%, and the pairwise correlation varies between 0% and 80%. For a given level of volatility, the rebalancing premium increases as the pairwise correlation ρ decreases (for $\sigma = 40\%$, the rebalancing premium is 7 bps if $\rho = 40\%$ and 17 bps if $\rho = 0\%$). The Sharpe ratios difference between the rebalanced and the buy-and-hold portfolio increases as the pairwise correlation decreases (for a given level of volatility). This difference strongly increases when ρ goes from 20% to 0%.

Figure 2.11: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of 30 risky assets

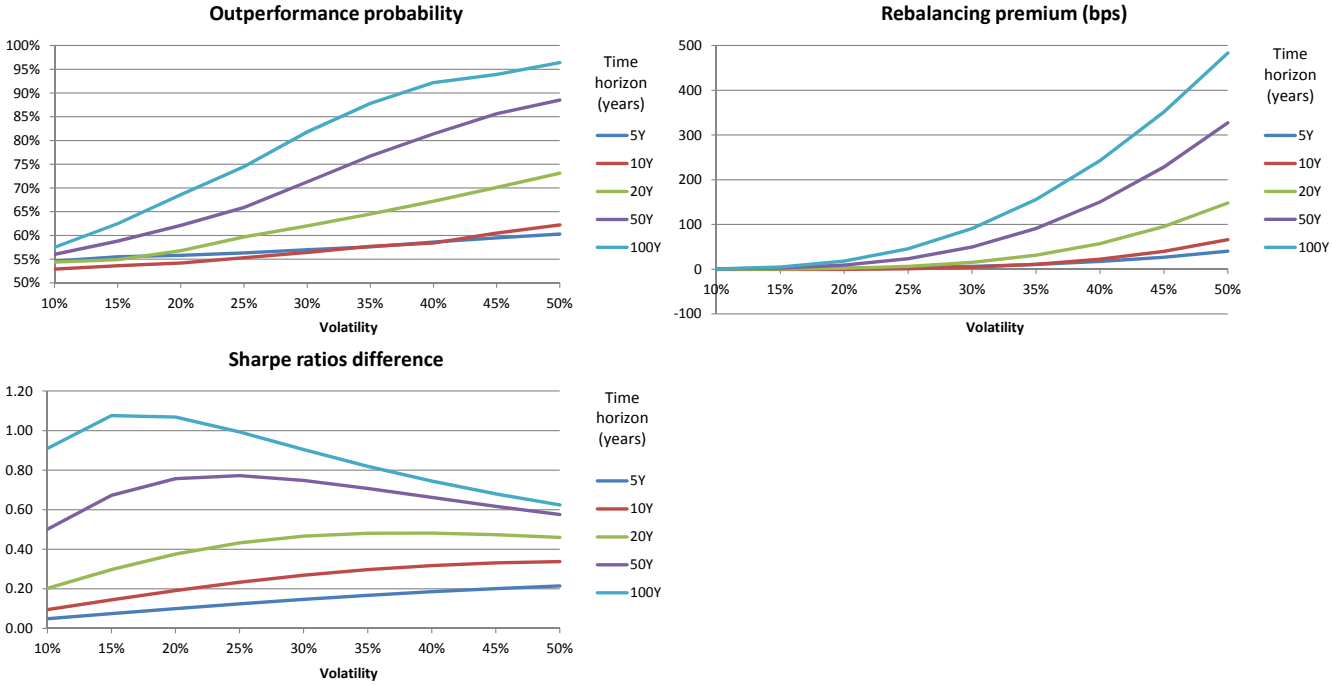


This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky assets for different values of the pairwise correlation. The rebalancing frequency is 1 month and the time horizon is 5 years.

Two-dimensional robustness check with respect to volatility and time horizon

Figure 2.12 displays the outperformance probability, the rebalancing premium and the Sharpe ratios difference relatively to different values of the time horizon (from 5 years to 100 years) and risky assets volatility (from 10% to 50%). For a given level of volatility, the outperformance probability increases as the time horizon increases: for σ=30%, the outperformance probability is 69% for a 5-year horizon and 83% for a 100-year horizon. For a given level of volatility, the rebalancing premium increases as the time horizon increases: for σ=40%, the rebalancing premium is 17 bps for a 5-year horizon and 150 bps for a 50-year horizon. The Sharpe ratios difference between the rebalanced and the buy-and-hold portfolios increases as the time horizon considered increases.

Figure 2.12: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of 30 risky assets

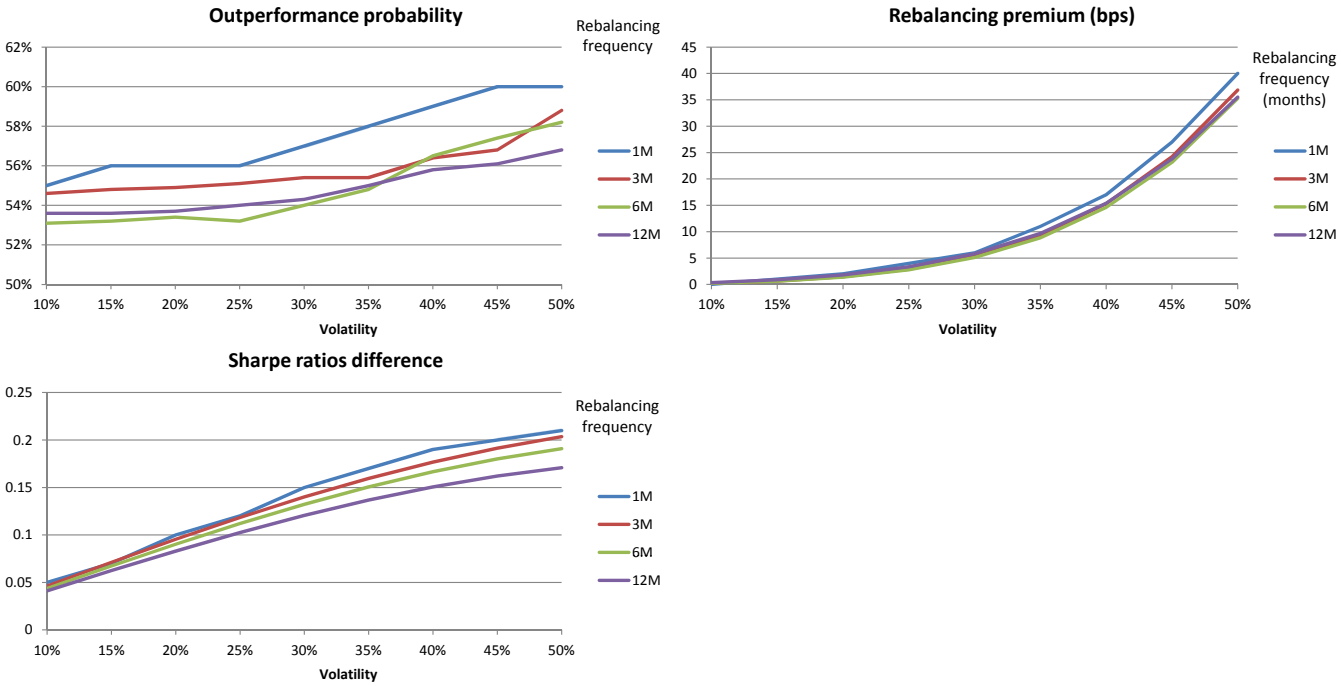


This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky assets for different values of the time horizon. The rebalancing frequency is 1 month and the pairwise correlation is null.

Two-dimensional robustness check with respect to volatility and rebalancing frequency

Figure 2.13 displays the outperformance probability, the rebalancing premium and the Sharpe ratios difference relatively to different values of the rebalancing frequency (from 1 month to 12 months) and risky assets volatility (from 10% to 50%). The outperformance probability is between 53% and 61%. For a given level of volatility, the 1-month rebalancing frequency optimises the outperformance probability, the rebalancing premium and the Sharpe ratios difference, but there are no meaningful differences on the rebalancing metrics when considering the different rebalancing frequencies.

Figure 2.13: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of 30 risky assets

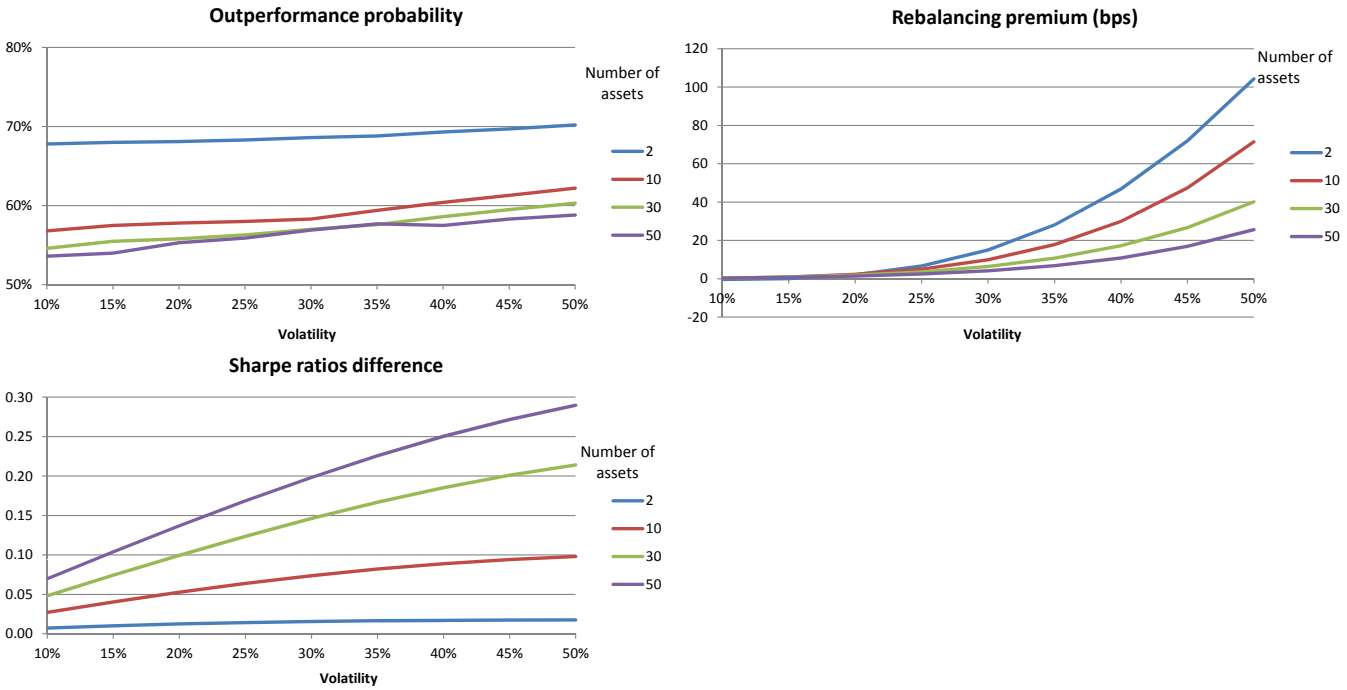


This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky assets for different values of the rebalancing frequency. The time horizon is 5 years and the pairwise correlation is null.

Two-dimensional robustness check with respect to volatility and number of risky assets

Figure 2.14 reports the outperformance probability, the rebalancing premium and the Sharpe ratios difference relatively to different values of assets number (from 2 to 50) and risky assets volatility (from 10% to 50%). We assess that for a given level of risky asset volatility, the outperformance probability decreases with the number or assets in the portfolio. The outperformance probability is between 53% and 70%. For a given level of volatility, the outperformance probability decreases as the number of assets in the portfolio increases; and so does the rebalancing premium: for a 40% volatility, the rebalancing premium is 47 bps for a portfolio of 2 assets and 11 bps for a portfolio of 50 assets. For a given volatility, the Sharpe ratios difference increases as the number of assets in the portfolio increases.

Figure 2.14: Outperformance probability, rebalancing premium and Sharpe ratios difference for investment universes of different size



This figure displays the outperformance probability, the rebalancing premium (in bps) and the Sharpe ratios difference as functions of the volatility of the risky assets for different values of the number of assets. The time horizon is 5 years, the rebalancing frequency is 1 month and the pairwise correlation is null.

2.2.6 Framework Extension with Serial Correlation Modeling

The framework defined in Section 2.2.1 implicitly assumes that there is no serial correlation in asset returns. In this section, we extend this framework for a universe made of 1 risky asset and 1 risk-free asset by modeling serial correlation in the risky asset returns through the use of a fractional Brownian motion process. This Gaussian process was introduced by Kolmogorov (1940) and Mandelbrot and Van Ness (1968). We do not consider a universe with several risky assets in this framework because it would have required the use of a multivariate fractional Brownian motion, for which the cross-covariance matrix is more complex to exploit for a relevant quantitative analysis (Lavancier (2009), Amblard et al. (2010), Amblard and Coeurjolly (2011)).

Let H be a real number in the range $]0;1[$. The fractional Brownian motion with index of self-similarity H (also named Hurst exponent) is a stochastic process $\{W^H(t)\}_{t\geq 0}$ with continuous paths and Gaussian distribution with mean 0 and covariance function:

$$\mathbb{E}[W^H(t)W^H(s)] = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H})$$

This general definition can be seen as an extension of the ordinary Brownian motion which corresponds to the particular case where $H = \frac{1}{2}$. This framework allows us to introduce negative serial correlation in asset returns for $0 < H < \frac{1}{2}$ and positive serial correlation in asset returns for $\frac{1}{2} < H < 1$ by taking into account long-range dependence. Indeed we can prove that for $t \geq s$:

$$\mathbb{C}ov(W^H(t) - W^H(s), W^H(s)) = \frac{1}{2}(t^{2H} - s^{2H} - (t-s)^{2H})$$

and find that $\mathbb{C}ov(W^H(t) - W^H(s), W^H(s))$ is strictly positive if $H > \frac{1}{2}$, null if $H = \frac{1}{2}$ and negative if $H < \frac{1}{2}$. We report below the main properties of the fractional Brownian motion (see Biagini et al. (2008) for the theory and application of the fractional Brownian motion and Amblard et al. (2010) for more details on the multivariate fractional Brownian motion):

- $W^H(0) = 0$;
- $W^H(t+s) - W^H(s)$ has the same distribution as $W^H(t) \forall s, t \geq 0$ (stationary increment);
- $W^H(Tt)$ has the same distribution as $T^H W^H(t)$ (self-similarity);
- $Var[W^H(t)] = t^{2H}$, $\forall t \geq 0$.

The fractional Brownian motion process allows us to model serial correlation in the risky asset returns by generalising the Brownian motion framework. We model the stock price as a geometric fractional Brownian motion

$$\frac{dS_t}{S_t} = \mu dt + \sigma d W_t^H \tag{2.29}$$

where $W^H(t)$ is a fractional Brownian motion with Hurst exponent H.

We solve the stochastic differential equation (2.29) with respect to the Wick Ito Skorohod integral introduced by Duncan et al. (2000) and Hu and Øksendal (2010) and also defined in Biagini et al. (2008):

$$S(T) = S(0) \exp\left[\left(\mu T - \frac{\sigma^2}{2} T^{2H}\right) + \sigma_i W_H(T)\right] \tag{2.30}$$

Cheridito (2003) put forward that in the case where $H \neq \frac{1}{2}$, the fractional Brownian motion is not a semimartingale and allows arbitrage opportunities which could be problematic in a pricing context. We define the autocorrelation at lag k of the first difference of fractional Brownian motion, which corresponds to the autocorrelation at lag k of the asset returns, as:

$$\gamma(k) = Corr(W^H(t) - W^H(t-1), W^H(t+k) - W^H(t+k-1)) = \frac{1}{2}((k+1)^{2H} - 2k^{2H} + (k-1)^{2H}) \tag{2.31}$$

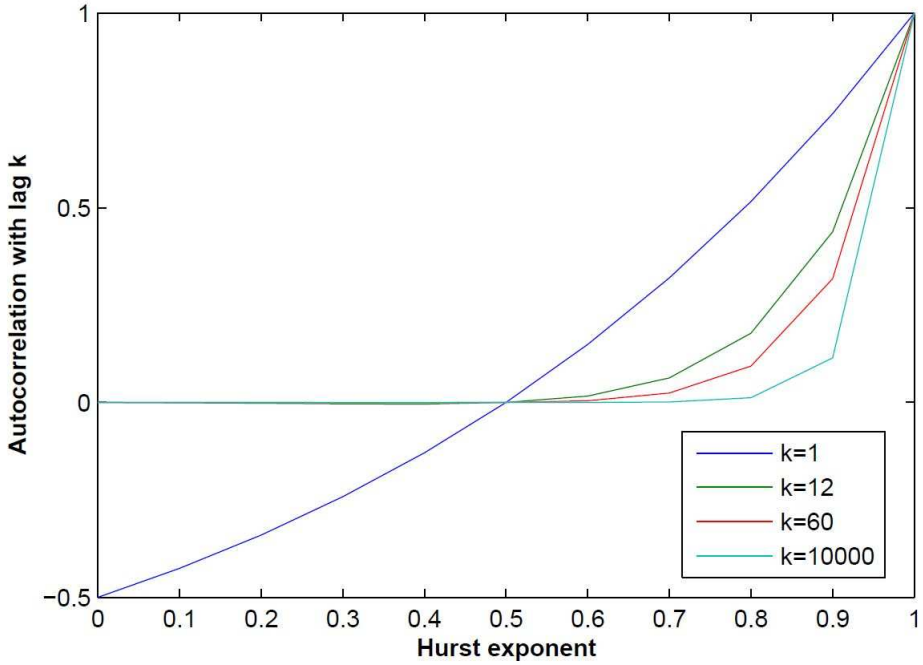
For $H = \frac{1}{2}$ we find that $\gamma(k) = 0$ which is the classical result of independent increments for the Brownian motion.

An order 2 Taylor expansion of Equation (2.31) when k tends to $+\infty$ implies :

$$\gamma(k) = Corr(W^H(t) - W^H(t-1), W^H(t+k) - W^H(t+k-1)) \underset{+\infty}{\sim} H(2H-1)k^{2H-2}$$

Figure 2.15 displays the autocorrelation at lag k for $k = 1, 12, 60$ and 10000 of the first difference of fractional Brownian motion as a function of the Hurst exponent. The values of Hurst exponent we consider in the quantitative analysis (0.3, 0.4, 0.5, 0.6 and 0.7) correspond to respective levels of autocorrelation at lag 1 of -24%, -13%, 0%, 15% and 32%. For instance, the estimated autocorrelation at lag 1 for the monthly returns S&P 500 Total Return Index on the February 1988-April 2016 period is -4.2%.

Figure 2.15: Autocorrelation at lag k of the first difference of fractional Brownian motion as a function of the Hurst exponent H



This figure displays the autocorrelation at lag k of the first difference of fractional Brownian motion as a function of the Hurst exponent H. The lags considered are 1, 12, 60 and 10000 months.

By replacing T by respectively $t + \Delta t$ and t in Equation (2.30), we have:

$$\begin{aligned} S^H(t+\Delta t) &= S^H(0)\exp\left[\mu \times (t+\Delta t) - \frac{\sigma^2}{2} \times (t+\Delta t)^{2H} + \sigma W^H(t+\Delta t)\right] \\ S^H(t) &= S^H(0)\exp\left[\mu t - \frac{\sigma^2}{2} t^{2H} + \sigma W^H(t)\right] \end{aligned} \tag{2.32}$$

We deduce from Equation(2.32) that:

$$S^H(t+\Delta t) = S^H(t)\exp\left[\mu \Delta t - \frac{\sigma^2}{2}\left((t+\Delta t)^{2H} - t^{2H}\right) + \sigma\left(W^H(t+\Delta t) - W^H(t)\right)\right] \tag{2.33}$$

Finally, we discretize Equation (2.33) and for values of $H \in \{0.3, 0.4, 0.6, 0.7\}$ we generate 1,000 simulations for the price of the risky asset S_t over a time horizon T for a monthly time step $\Delta t = \frac{1}{12}$ and an initial price $S_0 = 100$ as:

$$S^H_{t+\Delta t} = S^H_t \exp\left[\mu \Delta t - \frac{\sigma^2}{2}\left((t+\Delta t)^{2H} - t^{2H}\right) + \sigma\left(W^H_{t+\Delta t} - W^H_t\right)\right] \tag{2.34}$$

Where t is a monthly observation date such that $\frac{1}{\Delta t} \times t \in [\![0, \frac{T}{\Delta t}]\!]$ and where the increment process $\left(W^H_{t+\Delta t} - W^H_t\right)$ is defined as fractional Gaussian noise in the literature (see Elliott and Van Der Hoek (2003) for more details about the theory and Craigmile (2003) for more details about the simulation). We used the following first-order Taylor series approximation to obtain Equation (2.34) from Equation (2.33):

$$(t+\Delta t)^{2H} - t^{2H} \approx \Delta t \times 2H$$

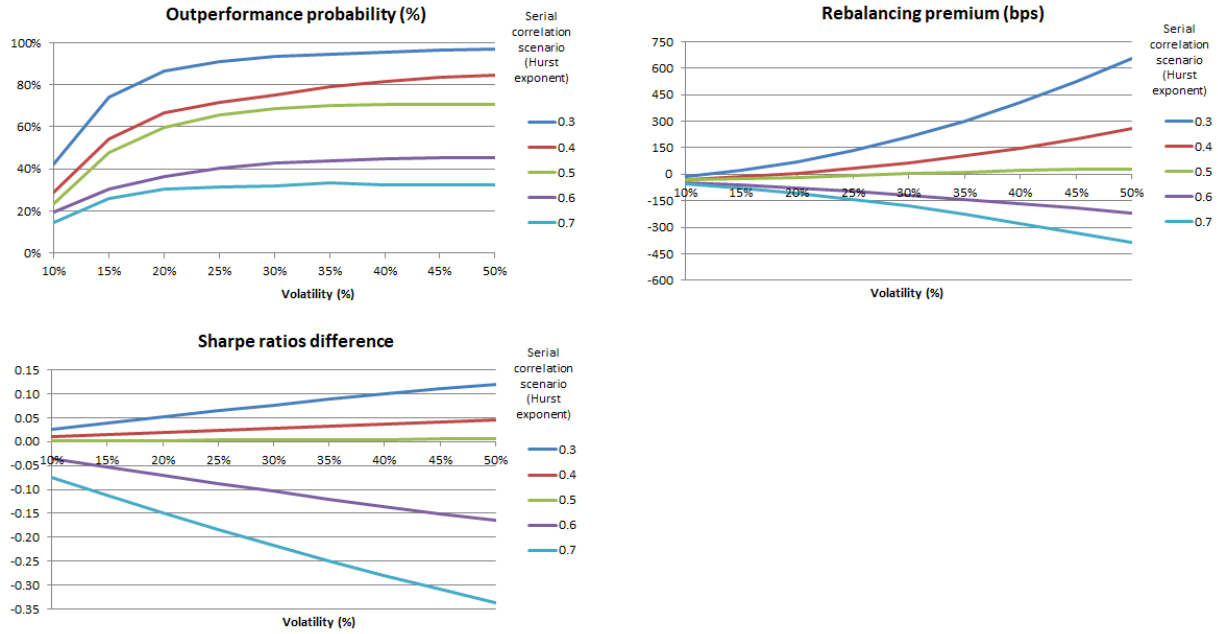
We simulate each of the fractional Gaussian noise $\left(W^H_{t+\Delta t} - W^H_t\right)_t$ via the Davies and Harte (1987) method using the Python package *fbm*. This method is known to be faster than other methods detailed in Hosking (1984) and Asmussen (1998).

Two-dimensional robustness check with respect to volatility and serial correlation

Figure 2.16 displays the outperformance probability, the rebalancing premium and the difference in Sharpe ratios for different values of the volatility σ (from 10% to 50%) and different serial correlation levels (Hurst exponent H varying from 0.3 to 0.7). The serial correlation on asset returns impacts significantly the outperformance probability: if we consider a volatility of 20%, the outperformance probability is 60% if we assume no serial correlation on asset returns (i.e., an Hurst exponent of 0.5), 87% if we assume a Hurst exponent of 0.3 (i.e., a negative correlation of -24% between the returns on two consecutive periods [14]) and 36% if we assume a Hurst exponent of 0.6 (i.e., a positive serial correlation with lag 1 of +15% on the risky asset returns). The serial correlation impacts significantly the rebalancing premium too: if we consider a volatility of 20%, the rebalancing premium is -17 bps if we assume no serial correlation on asset returns (i.e., an Hurst exponent of 0.5), 71 bps if we assume a Hurst exponent of 0.3 (i.e., a negative serial correlation) and -76 bps if we assume a Hurst exponent of 0.6 (i.e., a positive serial correlation).

[14] A period corresponds to one month.

Figure 2.16: Outperformance probability, rebalancing premium and Sharpe ratios difference for an investment universe made of one risky asset and one risk-free asset



This figure displays the outperformance probability, rebalancing premium and Sharpe ratios difference as functions of the volatility of the risky asset for different serial correlation scenarios on asset returns. The time horizon is 5 year, the rebalancing frequency is 1 month and the expected return of the risky asset is 10%.

2.2.7 Summary

In the setting with one risky asset and one risk-free asset, the robustness checks with respect to the risky asset expected return, volatility, risky asset weight, time horizon and rebalancing frequency display in most of the (realistic) cases disappointing results: even if the outperformance probability is 60% in our base case, the rebalancing premium is negative (-17 bps). Overall, these results do not support the possibility to harvest a significant rebalancing premium when considering a portfolio of one risky asset and one risk-free asset without assuming serial correlation in the risky asset returns.

In a more general framework with n risky assets, if we assume that all the assets have the same expected return then the rebalancing premium does not depend on the expected return. The robustness checks relatively to the number of risky assets, risky assets volatility, pairwise correlations, time horizon and rebalancing frequency display in most of the realistic cases an outperformance probability higher than 50% with a low amplitude rebalancing premium (i.e., lower than 50 bps). The Sharpe ratios difference, if a 30 risky assets setting, is substantial when we consider uncorrelated assets (0.10 in the base case) but decreases quickly if we consider a higher level of pairwise correlation (Sharpe ratios difference of 0.00 for a pairwise correlation of 20% anything else equal). The most influential parameters relative to the amplitude of the rebalancing premium seem to be volatility, time horizon, pairwise correlation and number of risky assets.

In the framework extension where we considered a universe with one risk-free and one risky asset and introduced serial correlation in the risky asset returns via a fractional Brownian motion, we assess that the introduction of negative serial correlation in the risky asset return lead to a sensibly higher rebalancing premium: for a risky asset volatility equal to 20%, the rebalancing premium is -17 bps if we assume no serial correlation on asset returns (i.e., an Hurst exponent of 0.5) and increases to 71 bps if we assume a Hurst exponent of 0.3 (i.e., a negative serial correlation). This finding suggests that serial correlation is an influential parameter with respect to the amplitude of the rebalancing premium amplitude.

The analytical and quantitative analysis made in that section provided us some insights regarding the magnitude of the rebalancing premium and the conditions under which it can be maximised, with a particular emphasis on the impact of the covariance matrix, portfolio weights, time horizon and serial correlation. We acknowledge that this controlled quantitative analysis cannot shed insights regarding the actual size of the rebalancing premium that can be harvested in practice. Moreover, it does not allow us to measure the contribution to the rebalancing premium emanating from differences in risk exposures between the rebalanced and the buy-and-hold portfolio. In an attempt to address these important questions, we conduct in the next section an empirical analysis of the rebalancing premium in realistic settings.

2.3 Empirical Analysis of the Rebalancing Premium

This section is organised as follows. We first analyse the rebalancing premium in realistical settings for stocks in the S&P 500 universe. Then we realise two robustness checks on the rebalancing frequency and on the portfolios' initial weights. Next we measure the sensitivity of the excess monthly arithmetic return of the rebalanced portfolio over the corresponding buy-and-hold portfolio to the Fama-French-Carnhart four factors and finally we test for the empirical relationship between the (out-of-sample) historical rebalancing premium ans standard characteristics such as market capitalisation, book-to-market ratio, past performance, volatility and serial correlation. The last section provides conclusions and suggestions for further research.

2.3.1 Analysis of Risk, Performance and Rebalancing Statistics

In this section we provide an empirical analysis of the rebalancing premium in realistic settings. The base universe of our 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 [15]. We choose a constant base universe in order to have a totally passive buy-and-hold portfolio and so taking into account only the impact of portfolio rebalancing. Note that this obviously implies the presence of a survivorship bias in equity portfolio performance. We use a resampling procedure (as defined in Plyakha et al. (2012)) and build one set of 30 randomly selected equally-weighted portfolios and another set with the 30 randomly selected corresponding buy-and-hold portfolios; each portfolio consisting of N stocks selected randomly in the aforementioned base case universe. Then we average the performance, risk and rebalancing metrics for the two sets of 30 random portfolios. More precisely, the metrics (such as volatility or outperformance probability) that we use to compare the rebalanced portfolios with the buy-and-hold portfolios are reported in this setting as the average of the metrics over the 30 random portfolios. This particular procedure mitigates the impact of stock selection biases. We consider the following set of values for N: 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, then we obtain a single portfolio, as opposed to 30 different portfolios. We also study the impact of the rebalancing frequency while introducing transaction costs and the impact of the initial weights of the portfolios on the rebalancing premium.

Differences in exposures of the rebalanced portfolios and the corresponding buy-and-hold portfolios with respect to rewarded risk factors are expected to have an impact on the rebalancing premium, an impact which is not directly related to the volatility pumping effect. In an attempt to control for differences in factor exposures and disentangle the two effects, we regress the excess monthly returns of the rebalanced portfolios over the buy-and-hold portfolios on the market, size, value and momentum factors, and interpret the alpha of this regression as the marginal value added by a pure volatility pumping effect, defined as the fraction of the rebalancing premium that is purely related to a volatility pumping effect. We then seek to analyse the contribution of diversification (portfolio weights) and the risk structure of assets (the covariance matrix) on this pure volatility pumping effect. We also seek to control for the presence of serial correlation in asset returns, which can have a positive impact (case of a negative serial correlation) or a negative impact (case of a positive serial correlation) on the rebalancing premium. Beyond serial correlation, we also study the impact of other characteristics (market capitalisation, book-to-market-ratio, past performance and volatility) on the rebalancing premium by performing standard portfolio sorts. We use monthly returns in our analysis. Equity stocks returns are extracted from the CRSP database and firm characteristics (such as value or size), when needed, are extracted from CRSP and COMPUSTAT databases.

We make the following assumptions:

- We consider the base case universe as our investment universe;
- The initial weight invested in each asset is $w_i = \frac{1}{N}$;
- The rebalancing frequency is 1 month.

As already mentioned, we mitigate selection biases by using the following resampling procedure (see Plyakha et al. (2012)):

1. We fix a time horizon T (in years);
2. We build, for each (end-of-month) date t_0 on the period November 1985-December (2015-T), 30 random monthly rebalanced equally-weighted portfolios and 30 corresponding buy-and-hold portfolios of N stocks ($N \in [\![2; 132]\!]$);

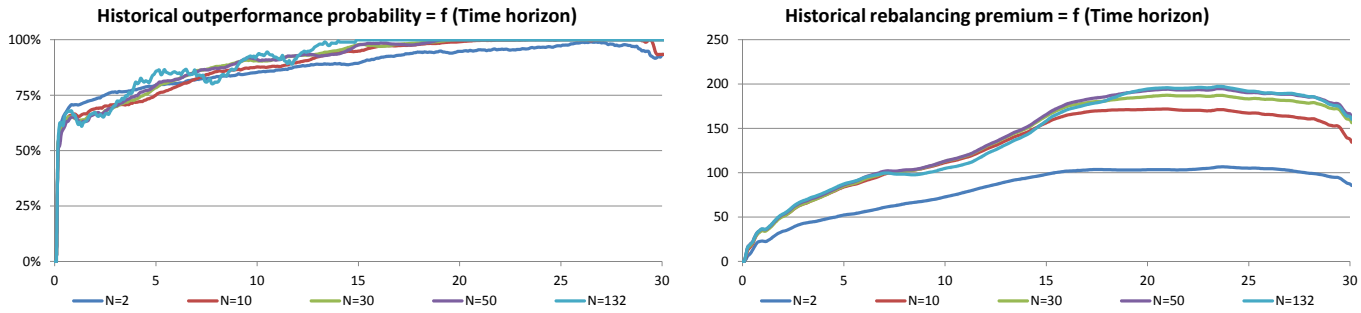
[15] The returns time-series start in December 1985 and end in December 2015.

3. We report the global metrics (such as the rebalancing premium) by first averaging each metric over the 30 pair of random portfolios (sample average) and then averaging each metric over all dates $t = t_0$ on the period November 1985-December (2015-T) (historical average).

We discuss for different values of N (2, 10, 30, 50 and 132) the main reporting measures as functions of time horizon. In particular, we compute the historical outperformance probability and the historical rebalancing premium [16]. As mentioned above, we consider rolling time periods for a given time horizon over the period November 1985 to December 2015. For instance if we fix the time horizon to 5 years we have 302 subperiods of 5 years to take into account.[17] We then take the average of each statistic on the set of the subperiods and apply the same process for each time horizon considered. We limit ourselves to a maximum time horizon of 20 years for which we have 122 subperiods. We examine the rebalancing metrics from three different perspectives.

Firstly, by examining Figure 2.17, we focus on the historical outperformance probability and the historical rebalancing premium as functions of time horizon (ranging from 1 month to 20 years) for different number of stocks in the portfolios. For a 5-year time horizon and $N = 30$, the outperformance probability is 78% and 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 and the outperformance probability is 91%. The outperformance probability increases with time horizon, while the number of stocks considered has almost no influence as long as we exceed a minimum number 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 time horizon of at least 5 years. If we take the 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.

Figure 2.17: Historical outperformance probability and historical rebalancing premium for different number of stocks in the portfolios



This figure plots the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of the time horizon (in years) considered for different number of stocks in the portfolios. The stock universe is made of of the 132 stocks extracted from CRSP which have continuously been in the S&P 500 index from November 1985 to December 2015. The portfolios are built according to the resampling procedure defined in Plyakha et al. (2012).

[16] We call *historical rebalancing premium* the difference between the realised growth rate of the rebalanced portfolio and that of the buy-and-hold portfolio on the time horizon considered.

[17] The first period is November 1985-November 1990 and the last period is December 2010-December 2015.

In a second time, we now fix the time horizon to 5 years to compute statistics related to the rebalanced portfolios and the corresponding buy-and-hold portfolios for a number of stocks $N = 50$. We also study the distribution of the difference in growth rates, the distribution of the terminal wealth difference and the decomposition of the 5-year historical rebalancing premium.

We examine in Figure 2.7 the statistics relative to the rebalanced and the corresponding buy-and-hold portfolios for a number of stocks $N = 50$. From a risk-adjusted performance perspective the rebalanced portfolios have on average a higher Sharpe ratio (0.83 versus 0.77 for the buy-and-hold portfolios), a higher volatility (14.6% versus 14.0% for the buy-and-hold portfolios) and a higher level of extreme risk (respective maximum drawdown and monthly 1% Value-at-Risk of 23.2% and -11.3% versus 22.4% and 11.1% for the buy-and-hold portfolios). The average turnover of the rebalanced portfolios is 30.2%. We compute the excess growth rate g^* on each 5-year subperiod by estimating the covariance matrix with the monthly returns on this subperiod: the average value is 323 bps.

Table 2.7: Empirical analysis statistics for portfolios with equally-weighted initial weights and a 5-year time horizon

	Rebalanced Portfolio	Buy-and-Hold Portfolio
Expected return	14.3%	13.4%
Annualised volatility	14.6%	14.0%
Sharpe Ratio	0.83	0.77
Maximum Drawdown	23.2%	22.4%
Monthly VaR 1%	-11.3%	-11.1%
Turnover	30.2%	0.0%
CEQ ($\gamma = 2$)	12.2%	11.4%
CEQ ($\gamma = 5$)	9.0%	8.4%
Outperformance probability	80%	
Rebalancing premium (bps)	86	
Excess growth rate $g^\star$(bps)	323	
Dispersion term (bps)	-237	
Conditional (>0) Rebalancing premium (bps)	116	
Conditional (<0) Rebalancing premium (bps)	-72	
Annualised standard deviation (bps)	112	

This table reports the main average performance, risk and rebalancing statistics for the rebalanced and corresponding buy-and-hold portfolios in the empirical base case scenario. The time horizon is 5 years. The stock universe is made of of the 132 stocks extracted from CRSP which have continuously been in the S&P 500 index from November 1985 to December 2015. The portfolios are built according to the resampling procedure defined in Plyakha et al. (2012).

Figure 2.18 shows the distribution of the difference in growth rates for a 5-year time horizon as well as the distribution of the terminal wealth differences for a 5-year time horizon and an initial investment of 100$ in both portfolios. The average growth rates difference is 86 bps. The growth rates difference is higher than 50 bps in 61% of the historical scenarios. The average realised terminal wealth difference is 7$. The realised terminal wealth difference is higher than 5$ in 61% of the historical scenarios. The analysis of the 5-year realised growth rates difference and 5-year terminal wealth distribution difference enables to have a more precise view on all the 5-year historical scenarios and not only their average.

Lastly we consider the analytical decomposition of the growth rates difference for each 5-year period derived from Gabay and Herlemont (2007):

$$\underbrace{\frac{1}{T}\ln\left(\frac{P_T^{reb}}{P_t^{b\&h}}\right)}_{\text{growth rates difference}} = \underbrace{g^*}_{\text{excess growth rate}} + \underbrace{\frac{1}{T}\left(\ln\frac{\prod_{i=1}^{n} S_i(T)^{w_i}}{\sum_{i=1}^{n} w_i S_i(T)}\right)}_{\text{dispersion term}} \tag{2.35}$$

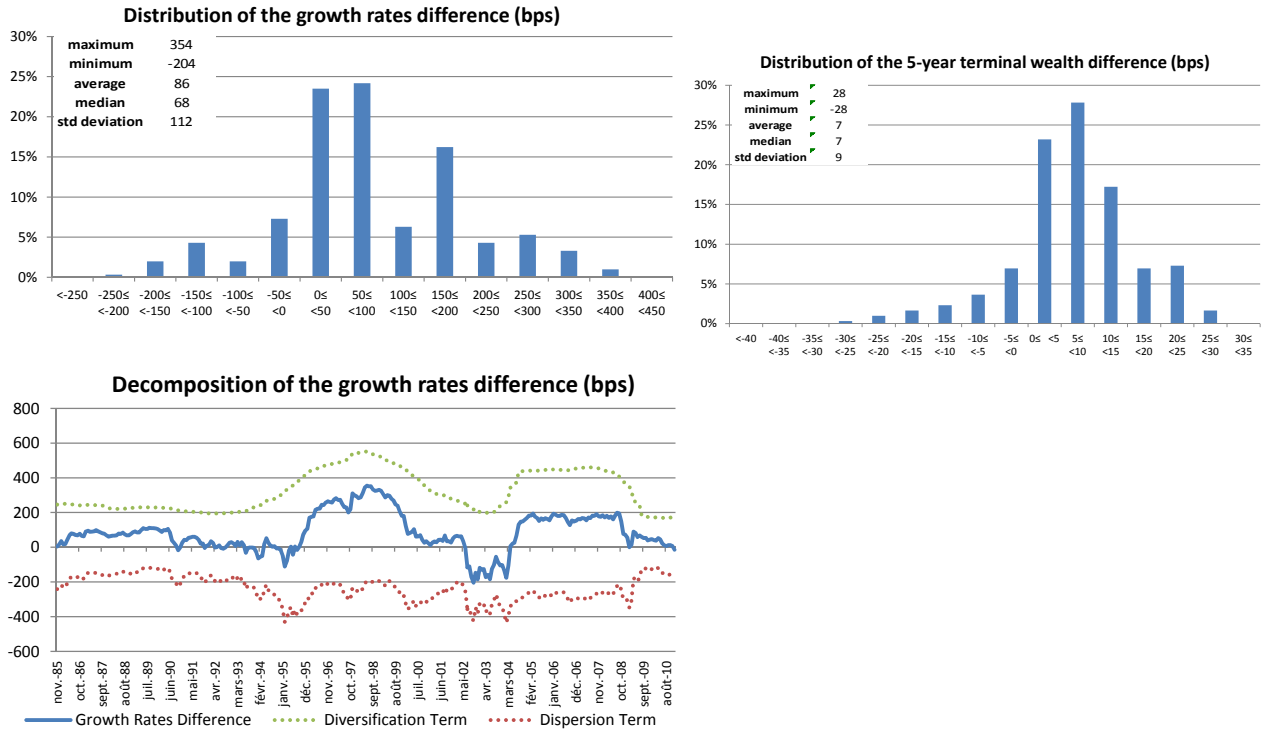
Note that this equality is valid only if we assume that (i) the stock prices follow geometric Brownian motions and (ii) that the rebalanced portfolio is continuously rebalanced. In an empirical analysis, none of these two stylised conditions are verified. We will thus decompose the growth rates difference as follows:

$$\underbrace{\frac{1}{T}\ln\left(\frac{P_T^{reb}}{P_T^{b\&h}}\right)}_{\text{growth rates difference}} = \underbrace{g^*}_{\text{excess growth rate}} + \underbrace{D(T)}_{\text{dispersion term}}$$

where the dispersion term $D(T)$ is simply defined as the empirical difference between the the growth rate difference and the excess growth rate, as opposed to being defined as the empirical realisation of the second term on the right-hand side of Equation (2.35).

Figure 2.18 also contains, for $N = 50$, the decomposition and the evolution over time of the 5-year growth rates difference as the sum of a diversification term and a dispersion term, as well as the evolution of each term over time on the period November 1985-December 2010. Each date on the chart correspond to a 5-year period starting date. The growth rates difference achieve the highest value (higher than 100 bps) for the starting dates in the period January 1996-January 2000 and 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% display a growth rates difference higher than 100 bps, 61% a growth rates difference higher than 50 bps and 16% a negative growth rates difference.

Figure 2.18: Distribution of the 5-year growth rates difference (bps), distribution of the 5-year terminal wealth difference($) and decomposition of the 5-year growth rates difference



This figure plots 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. It also displays the distribution of the average difference between the terminal wealth of the rebalanced portfolios and that of the corresponding buy-and-hold portfolios. The time horizon considered is 5 years, the number of stocks in the portfolios is $N = 50$ and the initial amount invested in each portfolio is 100$. It also displays the decomposition of the 5-year average growth rates difference as the sum a diversification term and a dispersion term and the evolution of each term over time. The stock universe is made of of the 132 stocks extracted from CRSP which have continuously been in the S&P 500 index from November 1985 to December 2015. The portfolios are built according to the resampling procedure defined in Plyakha et al. (2012).

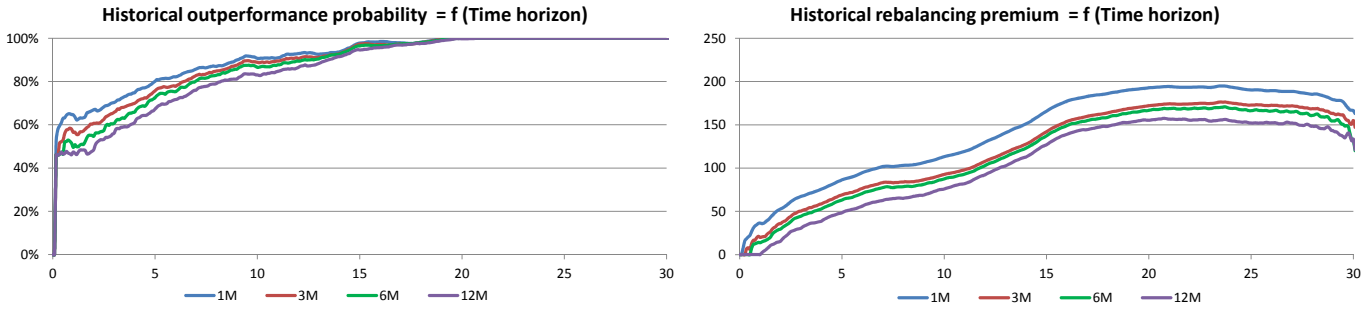
2.3.2 Robustness Check on the Rebalancing Frequency and on the Initial Weights

We perform for $N = 50$ a robustness check on the rebalancing frequency of the rebalanced portfolios.

The historical outperformance probability and the historical rebalancing premium are represented in Figure 2.19 as functions of time horizon (from 1 month to 20 years) for $N = 50$ stocks in every portfolio. We consider four possible rebalancing frequencies: 1,3,6 and 12 months and find that the rebalancing premium increases as expected with the rebalancing frequency: for a 5-year time horizon, the historical rebalancing premium are respectively 86, 69, 63 and 48 bps for 1,3,6 and 12 months rebalancing frequencies. Interestingly, the rebalancing premium remains sizable, at about half a percent, even for the lowest rebalancing frequency that we test (12 months). The drawback of increasing the rebalancing frequency is the corresponding increase in transaction costs. DeMiguel et al. (2009) and Plyakha et al. (2012) propose a conservative estimation of the transaction costs of a portfolio strategy as the 2-way annualised turnover of the portfolio multiplied by 50 bps. We use this basic model to estimate the transaction costs for different rebalancing frequencies and for a 5-year time horizon.

The historical 5-year rebalancing premium, when transaction costs are taken into account, are reported in Figure 2.20. We find that the presence of a trade-off between a higher gross value of the rebalancing premium and a higher level of transaction costs in case of increases in rebalancing frequency results in a relatively tight range for the net values (after transaction costs) of the rebalancing premium, which are 51 bps minimum and 56 bps maximum over the tested rebalancing frequencies. We will assume a monthly rebalancing frequency in the rest of the study.

Figure 2.19: Historical outperformance probability and historical rebalancing premium for different rebalancing frequencies



This figure gives the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of the time horizon (in years) considered for $N = 50$ stocks in every portfolio. We consider four rebalancing frequencies: 1, 3, 6 and 12 months. The stock universe is made of of the 132 stocks extracted from CRSP which have continuously been in the S&P 500 index from November 1985 to December 2015. The portfolios are built according to the resampling procedure defined in Plyakha et al. (2012).

Figure 2.20: Historical 5-year rebalancing premium with transaction costs

	1M	3M	6M	12M
Historical Rebalancing Premium (bps)	86	69	63	48
Annualised 1-Way Turnover (%)	30.2%	17.2%	12.0%	8.3%
Transaction Costs (bps)	30.2	17.2	12.0	8.3
Historical Rebalancing Premium with transaction costs (bps)	**56**	**52**	**51**	**40**

This figure contains the historical 5-year rebalancing premium when transaction costs are taken into account. The transaction costs are estimated as the 2-way annualised turnover of the portfolio multiplied by 50 bps. The stock universe is made of of the 132 stocks extracted from CRSP which have continuously been in the S&P 500 index from November 1985 to December 2015. The portfolios are built according to the resampling procedure defined in Plyakha et al. (2012).

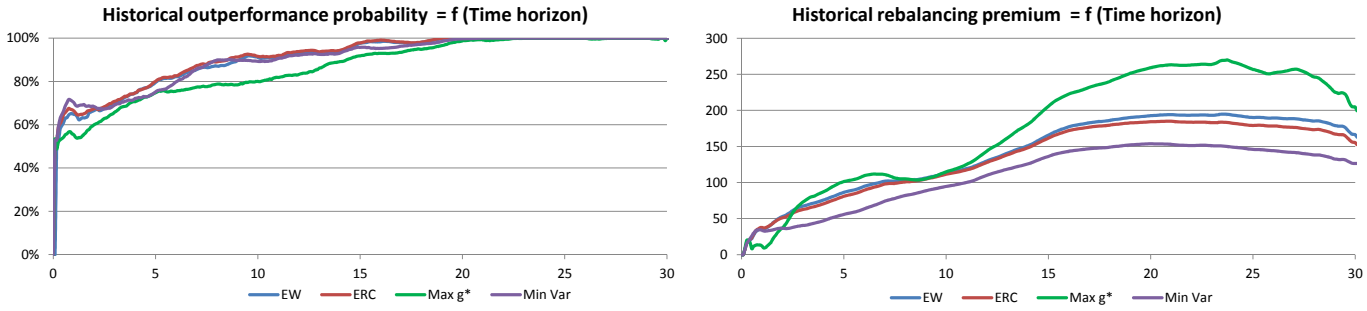
We carry out for $N = 50$ a robustness check on the initial weights of the monthly rebalanced and buy-and-hold portfolios. In addition to the heuristic equally-weighted initial weights, we consider three alternatives initial weights corresponding to portfolio optimisation procedures that are performed over the whole 30 years sample period and as such suffer from a look-ahead bias:

- Equal Risk Contribution with positive weights constraints;
- Max g^*, i.e., the weighting scheme that maximises the excess growth rate $g^\star$ with positive weights constraints and an additional constraint that the portfolio's effective number of constituents [18] (ENC) must be higher than or equal to one third of the number of constituents so as to avoid excessively concentrated corner solutions;
- Minimum Variance with positive weights constraints and the same constraint on the portfolio effective number of constituents.

We examine in Figure 2.21, for each initial weighting scheme, the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of time horizon (in years) for $N = 50$ stocks in every portfolio. We find that the portfolios computed with the Max g^* initial weights display the highest historical rebalancing premium: for a 5-year time horizon the historical rebalancing premium is 101 bps (versus 86 bps for the equally-weighted, 81 bps for the equal risk contribution and 55 bps for the minimum variance). This robustness check emphasises that the choice of the initial weights of the portfolios may have a substantial impact on the amplitude of the rebalancing premium. It also suggests that thorough out-of-sample empirical testing would be needed to assess the benefits of maximising a portfolio excess growth rate as an optimisation objective.

[18] The effective number of constituents (ENC), defined as the inverse of the sum of squared weights, is a measure of the deconcentration of a portfolio in terms of dollars. The ENC constraint aims to impose a minimum level of heuristic diversification in terms of dollars across assets.

Figure 2.21: Historical outperformance probability and historical rebalancing premium for different initial weighting schemes



This figure reports the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of the time horizon (in years) considered for $N = 50$ stocks in every portfolio. We consider four initial weighting schemes: Equally-Weighted (EW), Equal Risk Contribution (ERC), Max $g^\star$ and Minimum Variance (Min Var). The stock universe is made of of the 132 stocks extracted from CRSP which have continuously been in the S&P 500 index from November 1985 to December 2015. The portfolios are built according to the resampling procedure defined in Plyakha et al. (2012).

2.3.3 Analysis of Factor Exposures

We still consider the 132 stocks which were in the S&P 500 index over the period November 1985-December 2015 as our base case universe, and we measure by means of time series regressions the sensitivity of the excess monthly arithmetic return of the rebalanced portfolio over the corresponding buy-and-hold portfolio with respect to the standard Fama-French-Carhart four factors: market, size, value and momentum. We consider 5-year rolling sub-periods on the whole November 1985 to December 2015 period for the regressions.

We proceed in three steps:

1. For each (end-of-month) date t_0 on the period November 1985-December (2015-5), we consider the corresponding set of 30 random (monthly rebalanced) equally-weighted and buy-and-hold portfolios. For each randomly selected group of stocks, we compute the 5-year monthly excess returns of the rebalanced equally-weighted portfolio versus the corresponding buy-and-hold portfolio;

2. For each randomly selected group of stocks, we apply the Fama-French-Carhart regression to the monthly excess return time-series;

3. We average the coefficients of the regression across all randomly selected groups of stocks (sample average) and across all the times $t = t_0$ on the period November 1985-December (2015-5) (historical average).

We use the yield on Secondary Market US Treasury Bills (3M) as a proxy for the risk-free rate. The marker factor corresponds to the ERI Scientific Beta Long-Term US capitalisation-weighted excess return time series. The size, value and momentum factors are taken from the long-term ERI Scientific Beta database. For each factor, stocks are sorted on a given criterion and the universe is divided into two equal groups, the top and the bottom 30 percent groups. Each long/short factor has a long leg, which is a top or the bottom group depending on the factor, and a short leg, which is the other group. The choice of the criterion and the definition of the long leg are the following:

- Size: stocks are sorted on their market capitalisation and the long leg is the bottom size group.

- Value: stocks are sorted on their book-to-market (BE/ME) ratio and the long leg is the top BE/ME group.

- Momentum: stocks are sorted on their past 52-week return, skipping the last month, and the long leg is the past winner group.

Figure 2.22 shows, for different number N of stocks in the portfolios, the arithmetic average of the regression coefficients on all 5-year periods as well as the percentage of 5-year periods in which the coefficients are statistically significant at 95% confidence level. The annualised monthly alpha of the regression (i.e., the difference of performance not explained by the factor exposures) is between 103 to 107 bps for N =10, 30, 50 or 132 and this alpha is statistically significant in 41% of the historical scenarios for N =132. The momentum factor is the most significant factor and has the largest average loading in absolute value for all N. The exposure to the momentum factor is negative, as can be explained by the contrarian nature of the rebalancing strategy. The value and size factors displays relative average loadings of 0.0650 and 0.0116 for N =50 with respective percentages of significant historical simulations of 52% and 31%. The market factor loading varies from -0.0043 (N =132) to -0.0005 (N =50). Its percentage of significant historical simulations varies from 19% (N =2) to 37% (N =132). The average adjusted R-squared varies from 0.0886 (N =2) to 0.5514 (N =132). Overall our four-factor model shows that after controlling for factor exposures, the outperformance of the rebalanced strategy with respect to the corresponding buy-and-hold strategy remains susbtantial. This result suggests that the outperformance of the rebalanced portfolio with respect to the buy-and-hold portfolio is not purely explained by differences in factor exposures. In fact the negative exposure to the positively rewarded momentum factor (294 bps, see last row of Figure 2.22) implies the presence of a drag on performance.

Figure 2.22: Factor exposures of the arithmetic returns difference for a 5-year time horizon

	Excess Return	Alpha	Market (ER)	Size	Momentum	Value	Adj-R2
N=2							
Coefficients		0.0005	-0.0008	0.0060	-0.0283	0.0264	0.0886
% of significant simulations at 95%		10%	19%	18%	19%	17%	
Corresponding annualised returns (bps)	56	64	-1	2	-8	10	
N=10							
Coefficients		0.0009	-0.0026	0.0197	-0.0635	0.0424	0.2207
% of significant simulations at 95%		15%	23%	24%	51%	22%	
Corresponding annualised returns	90	103	-2	5	-19	16	
N=30							
Coefficients		0.0009	-0.0038	0.0111	-0.0706	0.0579	0.3572
% of significant simulations at 95%		21%	25%	31%	79%	39%	
Corresponding annualised returns	93	105	-3	3	-21	22	
N=50							
Coefficients		0.0009	-0.0005	0.0116	-0.0737	0.0650	0.4382
% of significant simulations at 95%		28%	27%	31%	87%	52%	
Corresponding annualised returns	97	107	0	3	-22	25	
N=132							
Coefficients		0.0009	-0.0043	0.0148	-0.0731	0.0679	0.5514
% of significant simulations at 95%		41%	37%	41%	96%	67%	
Corresponding annualised returns	98	104	-3	4	-21	26	
Factor average returns (bps)			696	276	294	388	

The investment universe is the base universe defined in the empirical analysis: we consider different number N of stocks in the portfolios (2, 10, 30, 50 and 132). For each of the 5-year period considered we perform the regression of the difference of the monthly returns on the market, size, momentum, value and low volatility factors. Then we take the average of each coefficient on all the 5-year periods. The first column represents the average excess arithmetic return of the 30 pairs of random portfolios on all the 5-year periods considered. The four following columns report the average coefficients of the linear regressions. The last column displays the average adjusted R-squared of the regressions. We also compute for each factor the percentage of significant simulations. The first 5-year period is November 1985-November 1990 and the last period is December 2010-December 2015. In the last row, we also display for each factor its annualised average returns in basis points across all the 5-year periods.

2.3.4 Analysis of 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 t_0 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 time t_0;
2. the second set of portfolios (low) is made of the 30 worst performing stocks of the base universe according to the characteristic at time t_0.

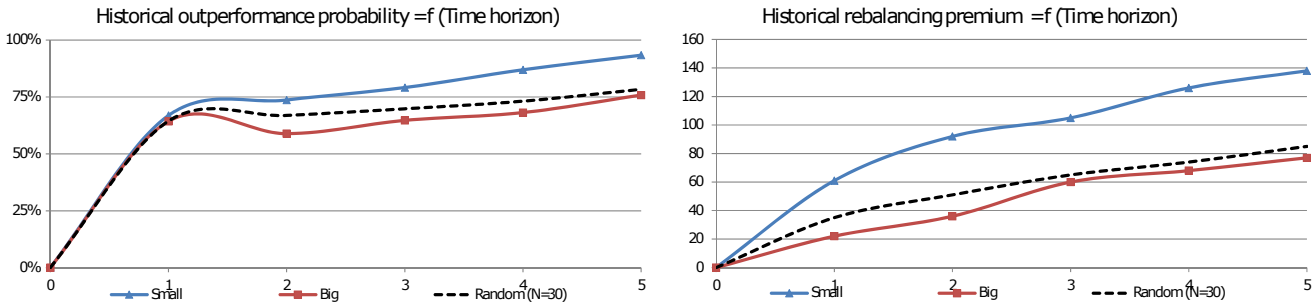
The investment universe for each portfolio is held constant over the corresponding time horizon, which allows us to analyse the influence of the characteristic on the volatility pumping effect. We then compute for each characteristic and each time horizon the average rebalancing statistics of the high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

2.3.4.1 Market Capitalisation

We consider for each stock its market capitalisation at time t for building the portfolios at time t. We see in Figure 2.23 that for a 5-year time horizon, the (historical average) outperformance

probability is 93% for the set of portfolios built with small cap stocks versus 76% for the set of portfolios built with large cap stocks and the (historical average) rebalancing premium is 138 bps versus 77 bps. By way of comparison, in the previous section framework with 30 sets of random portfolios built with $N = 30$ stocks randomly selected from the S&P500 base universe, the 5-year rebalancing premium is 85 bps and the outperformance probability is 78%. Sorting stocks by their market capitalisation increases the 5-year rebalancing premium of 53 bps compared with a random selection. These results show that the market capitalisation (size) as a sorting characteristic has a significant impact on the rebalancing premium, and that the rebalancing premium is higher for smaller stocks. Note that transaction costs would also be higher for smaller stocks so it is not clear what the next effect would be in implementation.

Figure 2.23: Historical outperformance probability and historical rebalancing premium with market capitalisation as a sorting characteristic

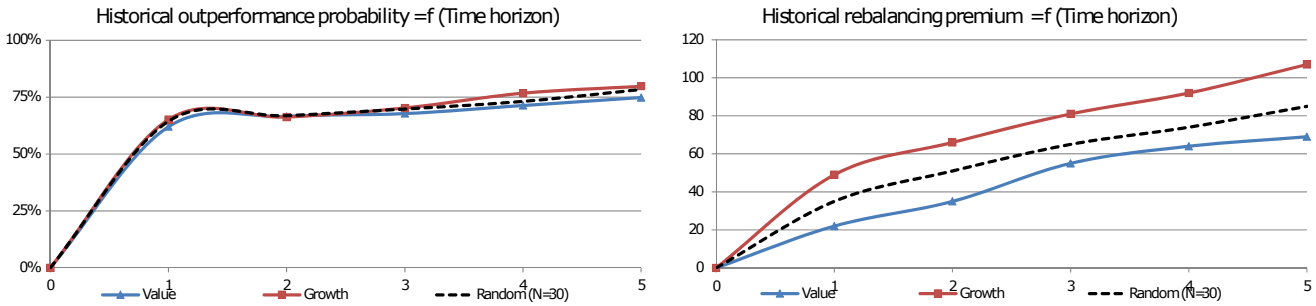


This figure displays the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of the time horizon (in years) when portfolios are sorted by the market capitalisation of the stocks. We also display the functions (dashed lines) when the $N = 30$ stocks are randomly selected from the base S&P 500 universe. The time horizons considered are 1, 2, 3, 4 and 5 years.

2.3.4.2 Book-to-Market Ratio

For each stock we look at its ratio of book value over the market value at time t to build the portfolios at time t. We find in Figure 2.24 that for a 5-year time horizon, the outperformance probability is 80% for the deepest growth stocks versus 75% for the deepest value stocks in our universe, and the rebalancing premium is 107 bps versus 69 bps. We note that sorting stocks by their book-to-market-ratio leads to an increase of 23 bps of the rebalancing premium for a 5-year time horizon, compared with the random setting. These results show that the book-to-market ratio (Value) as a sorting characteristic has a significant impact on the rebalancing premium, and that the rebalancing premium tends to be higher for growth versus value stocks.

Figure 2.24: Historical outperformance probability and historical rebalancing premium with book-to-market ratio as a sorting characteristic

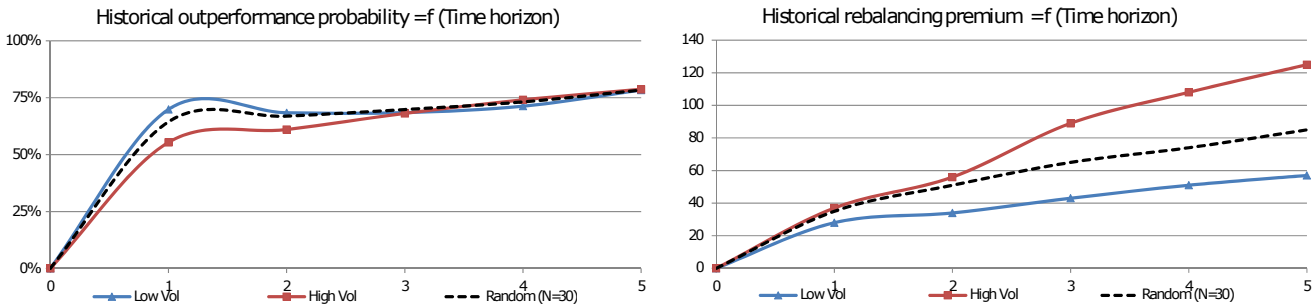


This figure displays the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of the time horizon (in years) when portfolios are sorted by the book-to-market of the stocks. We also display the functions (dashed lines) when the $N = 30$ stocks are randomly selected from the base S&P 500 universe. The time horizons considered are 1, 2, 3, 4 and 5 years.

2.3.4.3 Volatility

We regard for each stock its historical volatility computed with the 24 months past returns to build the portfolios at time t. We find in Figure 2.25 that for a 5-year time horizon, the outperformance probability is 78% for the set of portfolios built with low volatility stocks versus 79% for the set of portfolios built with high volatility stocks and the rebalancing premium is 57 bps versus 125 bps. Selecting high volatility stocks rather than random stocks allows to enhance the rebalancing premium of 40 bps for a 5-year time horizon. This figure confirms that high volatility stocks are as expected better candidates for harvesting the benefits of the rebalancing effect compared to low volatility stocks.

Figure 2.25: Historical outperformance probability and historical rebalancing premium with volatility as a sorting characteristic

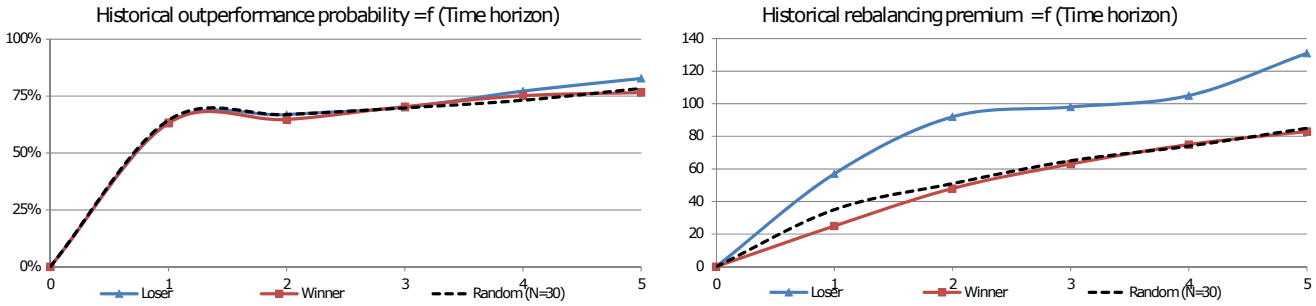


This figure displays the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of the time horizon (in years) when portfolios are sorted by the volatility of the stocks. We also display the functions (dashed lines) when the $N = 30$ stocks are randomly selected from the base S&P 500 universe. The time horizons considered are 1, 2, 3, 4 and 5 years.

2.3.4.4 Past Performance

We consider for each stock its performance between t-12 months and t-1 month to build the portfolios at time t. We see in Figure 2.26 that for a 5-year time horizon, the outperformance probability is 77% for the set of portfolios built with winner stocks versus 83% for the set of portfolios built with loser stocks and the historical rebalancing premium is 83 bps versus 131 bps. These numbers show that past performance (momentum) as a sorting characteristic has a significant impact on the rebalancing premium, which is highest for past losers compared to past winners.

Figure 2.26: Historical outperformance probability and historical rebalancing premium with past performance as a sorting characteristic



This figure displays the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of the time horizon (in years) when portfolios are sorted by the past performance of the stocks. We also display the functions (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.

2.3.4.5 Serial Correlation in Asset Returns

The intuition suggests that the rebalancing premium will be higher for stocks that display negative serial correlation, and that sorting stocks as a function of their serial correlation can be used to maximise the benefits of volatility harvesting as long as this indicator is persistent over time. Before analysing the impact of serial correlation on the rebalancing premium, we thus test the persistence of serial correlation in asset returns. To do so, we consider for three 10-year sub-periods (December 1985-December 1995, December 1995-December 2005 and December 2005-December 2015) monthly total returns of the 132 stocks of the base universe. For a given 10-year sub-period we split the corresponding stock returns time series into two 5-year time series of 132

stock returns.[19] We estimate, for each stock, the autocorrelation at lag 1 and the Hurst exponent via the detrended fluctuation analysis methodology (see Appendix 2.F for further details) on each one of the six 5-year periods.

We then define four classes:

- Class 1 (C1): stocks which have negative serial correlation on both 5-year sub periods;
- Class 2 (C2): stocks which have negative serial correlation on the first 5-year sub period and positive serial correlation on the second 5-year sub period ;
- Class 3 (C3): stocks which have positive serial correlation on the first 5-year sub period and negative serial correlation on the second 5-year sub period ;
- Class 4 (C4): stocks which have positive serial correlation on both 5-year sub periods.

We finally compute for each 10-year sub-period the following quantities:

- The probability that a stock with a negative serial correlation on the first sub-period has a negative serial correlation in the second sub-period: $P_1 = \frac{Card(C_1)}{Card(C_1)+Card(C_2)}$;
- The probability that a stock with a negative serial correlation on the first sub-period has a positive serial correlation in the second sub-period: $1 - P_1$;
- The probability that a stock with a positive serial correlation on the first sub-period has a positive serial correlation in the second sub-period: $P_2 = \frac{Card(C_4)}{Card(C_3)+Card(C_4)}$;
- The probability that a stock with a positive serial correlation on the first sub-period has a negative serial correlation in the second sub-period: $1 - P_2$;
- The corresponding transition matrix: $\begin{pmatrix} P_1 & 1-P_1 \\ 1-P_2 & P_2 \end{pmatrix}$.

We find in Table 2.8 that for each sub-period, at least at a global level, some evidence of persistence in negative serial correlation (P_1 higher than or equal to 65%) and a tendency for stocks returns with initial positive serial correlation to become negatively serially correlated ($1-P_2$ higher than or equal to 73%).

Table 2.8: Transition matrix for autocorrelation with lag 1

First period	72%	28%
December 1985 - December 1995	81%	19%
Second period	65%	35%
December 1995 - December 2005	73%	27%
Third period	74%	26%
December 2005 - December 2015	79%	21%

This table shows for three distinct periods the matrix transition according to the level of autocorrelation with lag 1 for the 132 stocks of the base universe. For instance, if we consider the first period, the probability that a stock with a negative estimated autocorrelation with lag 1 on the period December 1985-December 1990 has still a negative autocorrelation with lag 1 on the period December 1990-December 1995 is 72%. Likewise the probability that a stock with a positive estimated autocorrelation with lag 1 on the period December 1985-December 1990 has still a positive autocorrelation with lag 1 on the period December 1990-December 1995 is 19%.

[19] The 5-year periods considered are the following December 1985-December 1990, December 1990-December 1995, December 1995-December 2000, December 2000-December 2005, December 2005-December 2010 and December 2010 -December 2015.

We now consider for each stock its autocorrelation with lag 1 computed with the 60 months past returns to build the portfolios at time t. We see in Figure 2.27 that , for a 5-year time horizon, the outperformance probability is 71% for the set of portfolios built with stocks with negative serial correlation versus 84% for the set of portfolios built with stocks with positive serial correlation and the rebalancing premium is 67 bps versus 135 bps. This figure shows that the serial correlation as a sorting characteristic does not allow to maximise the rebalancing premium for time horizons longer than 1 year.

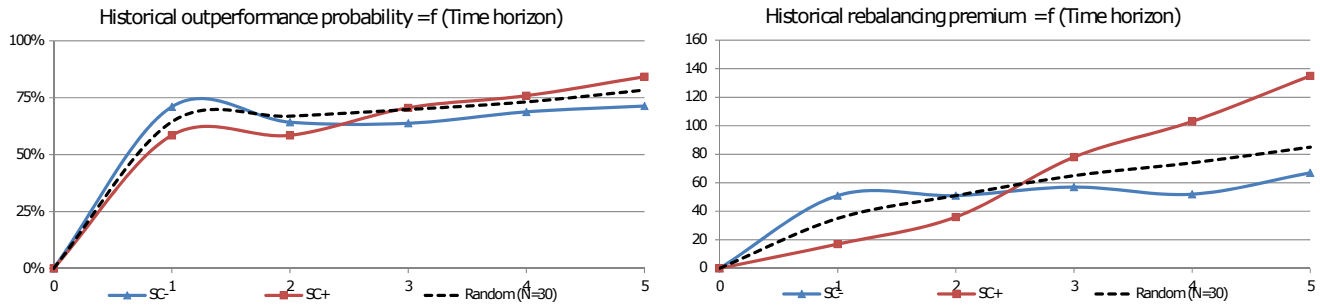
We use the same methodology by considering the Hurst exponent H estimated via a detrended fluctuation analysis and considering stocks with $H < 0.5$ (negative serial correlation) and $H > 0.5$ (positive serial correlation). We confirm in Table 2.9 the presence of persistence in negative serial correlation (P_1 higher than or equal to 54%) over each 10-year period and a tendency for stocks returns with initial positive serial correlation to become negatively serially correlated ($1 - P_2$ higher than or equal to 67%). While this mild evidence of persistence over a 5-year horizon is encouraging, it does not imply the presence of persistence over shorter periods of time. In fact, we find below that selecting stocks with negative serial correlation does not allow for a robust increase in the rebalancing premium, which suggests a lack of persistence over the tested horizons.

Table 2.9: Transition matrices for Hurst exponent

First period	54%	46%
December 1985 - December 1995	71%	29%
Second period	73%	27%
December 1995 - December 2005	90%	10%
Third period	64%	36%
December 2005 - December 2015	67%	33%

This table shows for three distinct periods the matrix transition according to the value of the Hurst exponent for the 132 stocks of the base universe. For instance, if we consider the first period, the probability that a stock with an estimated Hurst exponent lower than 0.50 on the period December 1985-December 1990 an Hurst exponent lower than 0.50 on the period December 1990-December 1995 is 54%. Likewise the probability that a stock with an estimated Hurst exponent higher than 0.50 on the period December 1985-December 1990 has still Hurst exponent higher than 0.50 on the period December 1990-December 1995 is 29%.

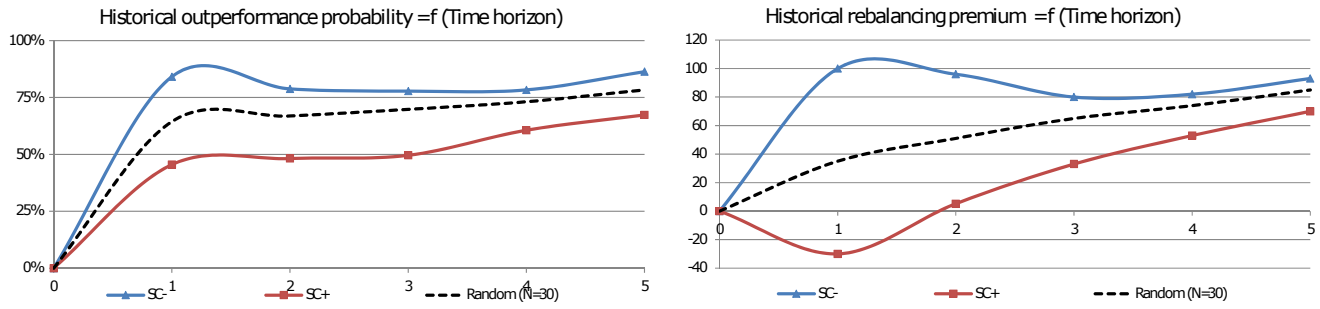
Figure 2.27: Historical outperformance probability and historical rebalancing premium with serial correlation as a sorting characteristic



This figure displays the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of the time horizon (in years) when portfolios are sorted by the serial correlation of the stocks. We also display the functions (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.

If we now consider for each stock its autocorrelation with lag 1 computed with the 60 months next returns (i.e., with a look-ahead bias) to build the portfolios at time t, we find that (see Figure 2.28) for a 5-year time horizon, the average outperformance probability is 86% for the set of portfolios built with stocks with negative serial correlation versus 67% for the set of portfolios built with stocks with positive serial correlation and the average historical rebalancing premium is 93 bps versus 70 bps. Figure 2.28 shows that the serial correlation computed ex-post has an explanatory power on the historical rebalancing premium. The difference of results between ex-post serial correlation and ex-ante serial correlation emphasises the possible lack of persistence of serial correlation at the stock individual level for time horizons longer than 1 year.

Figure 2.28: Historical outperformance probability and historical rebalancing premium with serial correlation as an explanatory characteristic



This figure displays the historical outperformance probability (in %) and the historical rebalancing premium (in bps) as functions of the time horizon (in years) when portfolios are sorted by the 5-year ex-post serial correlation of the stocks. We also display the functions (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.

2.3.4.6 Summary

The main results of the empirical analysis applied to characteristics are:

- Market capitalisation, book-to-market ratio, past performance and volatility are sorting characteristics which have a significant impact on the out-of-sample maximisation of the rebalancing premium. Small cap, growth, past loser and high volatility portfolios display a stronger out-of-sample rebalancing premium effect compared to random portfolios for time horizons from 1 year to 5 years;
- Serial correlation as a sorting characteristic does not allow to maximise the rebalancing premium out-of-sample for time horizons higher than 1 year;

- Serial correlation, when computed ex-post, has an explanatory power on the rebalancing premium. Taken together the last two results suggest a relative lack of persistence in serial correlation for time horizons higher than 1 year.

2.4 Conclusions and Extensions

This chapter provides a thorough numerical and empirical analysis of the volatility pumping effect and the conditions under which it can be maximised in equity markets. 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. We find that a significant part of the difference in arithmetic returns of the equally-weighted and buy-and-hold portfolios is explained by exposures to risk factors (mainly momentum and secondly size and value). Using the standard Fama-French-Carhart factor model, we find however an annualised alpha higher than 100 bps, implying that there is a substantial part of the outperformance of the rebalanced strategy with respect to the corresponding buy-and-hold strategy that is not explained by standard exposures with respect to rewarded risk factors. Our analysis on individual stocks' characteristics highlights that size, value, momentum and volatility are sorting characteristics which 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. We also show that the initial weighting scheme has a susbtantial impact on the size of the rebalancing premium. 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.

Our work can be extended in several dimensions. 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 can 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. We leave these, and other related, questions for further research.

2.A Confidence interval of the rebalancing premium

The rebalancing premium in a general case with n risky assets is defined in the simulation framework as:

$$RP(0,t) = \mathbb{E}\left[\frac{1}{t}\ln\left(\frac{P_t^{reb}}{P_t^{b\&h}}\right)\right] = g^\star + \frac{1}{t}\mathbb{E}\left[\ln\left(\frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right)\right]$$

where $g^\star = \frac{1}{2}\left(\sum_{i=1}^{n} w_i\sigma_i^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{i,j}\right)$ is half of the difference between the weighted-average variance of the assets in the rebalanced portfolio and the rebalanced portfolio variance.

Each asset i price verifies:

$$S^i(t) = S^i(t)\exp\left[\left(\mu_i - \frac{\sigma_i^2}{2}\right)t + \sigma_i W^i(t)\right]$$

where t is the time horizon in years and each $W^i(t) \sim \mathcal{N}(0,t)$ and each $(W^i(t))_{t\geq 0}$ is a one-dimensional standard Brownian motion where the correlations $\rho_{i,j}$ between asset returns are defined by $Corr(W_t^i, W_t^j) = \rho_{i,j}$. We denote $\boldsymbol{W}(t) = \left(W^i(t)\right)_{1\leq i\leq n}$ the random vector made of the n correlated random variables $W^i(t)$. We have:

$$\boldsymbol{W}(t) \sim \mathcal{N}\left(\boldsymbol{0}_n, t\boldsymbol{C}\right)$$

We define the random variable X as:

$$X = g^\star + \frac{1}{t}\ln\left(\frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right)$$

We have by construction:

$$\mathbb{E}[X] = RP(0,t)$$

We define the following quantities:

- A sample of size n_{sims} $X_1, \ldots, X_{n_{\text{sims}}}$ of independent and identically distributed random variables with the same distribution as X.
- A realization of the sample $x_1, \ldots, x_{n_{\text{sims}}}$, which are the observed values taken by $X_1, \ldots, X_{n_{\text{sims}}}$ on that given sample.

The sample mean[20] estimator obtained from this sample is defined as:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Finally, for a random variable X of mean $\mu = RP(0,t)$ and standard deviation σ, we can assume from the central limit theorem that the sampling distribution of $\overline{X}$ is a normal distribution with mean $\mu_{\overline{X}} = \mu$ and variance $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n_{\text{sims}}}$ as long as n_{sims} is large enough (typically $n_{\text{sims}} \geq 30$).

$$\bar{X} \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

Since the variance σ^2 of X is unknown we need to estimate it via the sample variance estimator. The sample variance of the sample $X_1, \ldots, X_n$ is a random variable defined as:

$$S^2 = \frac{1}{n_{\text{sims}} - 1}\sum_{i=1}^{n_{\text{sims}}} (X_i - \bar{X})^2$$

The core result to build a confidence interval for $RP(0,t)$ at the $1-\alpha$ level, labelled as $CI_{1-\alpha}(RP(0,t))$ is:

$$Z = \frac{\bar{X} - RP(0,t)}{\frac{S}{\sqrt{n_{\text{sims}}}}} \sim \mathcal{T}(n_{\text{sims}} - 1)$$

Where $\mathcal{T}(n_{\text{sims}} - 1)$ denotes the Student distribution with n_{sims} degrees of freedom.

Finally we obtain a confidence interval at the $1-\alpha$ level for $RP(0,t)$ with respect to a sample realization $x_1, \ldots x_{n_{\text{sims}}}$:

$$CI_{1-\alpha}(RP(0,t)) = \left[\overline{x} - t_{n_{\text{sims}}-1,\frac{\alpha}{2}}\frac{s}{\sqrt{n_{\text{sims}}}}; \overline{x} + t_{n_{\text{sims}}-1,\frac{\alpha}{2}}\frac{s}{\sqrt{n_{\text{sims}}}}\right]$$

Where:

[20] The sample mean is a random variable that is sample dependent.

- $t_{n_{\text{sims}}-1,\frac{\alpha}{2}}$ is the critical value of the Student distribution with $n_{\text{sims}}-1$ degrees of freedom $\mathcal{T}$(n_sims-1) such that $\mathbb{P}(Z > t_{n_{\text{sims}}-1,\frac{\alpha}{2}} = \frac{\alpha}{2})$,
- s is the standard deviation estimate.

For $\alpha = 5\%$ and $n_{\text{sims}} = 1,000$, we have $t_{n_{\text{sims}}-1,\frac{\alpha}{2}} \approx 1.9623$.

2.B Performance, risk and turnover statistics definition

We define in this appendix the main statistics computed on the numerical and empirical analysis.

- Sharpe ratio: the Sharpe ratio is the annualised return in excess of the risk-free rate divided by the annualised volatility of monthly excess returns. The risk-free asset considered is the US 3-month Treasury Bill Index.
- Skewness: skewness measures the asymmetry of the data around the sample mean. A negative skewness indicates that the data are spread out more to the left of the mean than to the right whereas a positive skewness indicates that the data are spread out more to the right. The skewness of the normal distribution (or any perfectly symmetric distribution) is zero. The skewness of a distribution is defined as: $S = \frac{\mathbb{E}[(r-\mu)^3]}{\sigma^3}$ where μ is the mean of r and σ the standard deviation of r. An unbiased estimator of the distribution skewness is:

$$\hat{S} = \frac{\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^2}\right)^3} \times \frac{\sqrt{n(n-1)}}{n-2}$$

- Kurtosis: kurtosis quantifies how outlier-prone a distribution is. The kurtosis of the normal distribution is 3. Distributions that are more outlier-prone than the normal distribution have kurtosis greater than 3; distributions that are less outlier-prone have kurtosis less than 3. The kurtosis of a distribution is defined as: $K = \frac{\mathbb{E}[(r-\mu)^4]}{\sigma^4}$ where μ is the mean of r and σ the standard deviation of r. An unbiased estimator of the distribution kurtosis is:

$$\hat{k} = 3 - \frac{3(n-1)^2}{(n-2)(n-3)} + \frac{\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^2\right)^2} \times \frac{n^2-1}{(n-2)(n-3)}$$

- γ-certainty-equivalent return: like DeMiguel et al. (2009) we define the γ-certainty-equivalent return of a portfolio P as:

$$\hat{\mu}_P - \gamma\frac{\hat{\sigma}_P^2}{2}$$

It corresponds to the risk-free rate than a mean-variance investor is willing to accept rather than investing a specific risky portfolio strategy. γ can be interpreted as the risk aversion of the investor, $\hat{\mu}_P$ is the annualised total return of the portfolio strategy and $\hat{\sigma}_P$ the annualised volatility of the portfolio strategy.

- Annualised 1-way turnover: at each rebalancing date t, the one-way turnover of the portfolio is computed as:

$$\theta_t = \frac{1}{2}\sum_{i=1}^{K}|w_{i,t} - w_{i,t_-}|$$

where $w_{i,t}$ is the weight of constituent i imposed on date t and $w_{i,t-}$ the effective weight just before the rebalancing. The annualised 1-way turnover is the average of the θ_t multiplied by 12 to convert it to an annual quantity.

2.C Wealth process for portfolios continuously rebalanced to fixed weights

We consider a set of n risky assets where each risky asset price $S_i(t)$ follows a geometric Brownian motion with an expected return μ_i and a volatility σ_i:

$$\frac{dS_i(t)}{S_i(t)} = \mu_i dt + \sigma_i dW_i(t)$$

We apply Ito's lemma to $\ln S_i(t)$:

$$d(\ln S_i(t)) = (\mu_i - \frac{\sigma_i^2}{2})dt + \sigma_i dW_i(t) \tag{2.36}$$

The wealth $P^{reb}(t)$ of a continuously rebalanced portfolio to constant weights $(w_1, \ldots, w_n)$ verifies:

$$\begin{aligned}
\frac{dP^{reb}(t)}{P^{reb}(t)} &= \sum_{i=1}^{n} w_i \frac{dS_i(t)}{S_i(t)} \\
&= \sum_{i=1}^{n} w_i \mu_i dt + \sum_{i=1}^{n} w_i \sigma_i dW_i(t) \\
&= \sum_{i=1}^{n} w_i \mu_i dt + d\bar{W}(t) \left(\sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{i,j} \right)^{1/2}
\end{aligned} \tag{2.37}$$

where $\bar{W}(t) = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{i,j} \right)^{-1/2} \sum_{k=1}^{n} w_k \sigma_k dW_k(t)$ is a normalised Wiener process.

Ito's lemma applied to Equation (2.37) implies:

$$d(\ln(P^{reb}(t))) = \left(\sum_{i=1}^{n} w_i \mu_i - \frac{1}{2} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{i,j} \right) \right) dt + \sum_{i=1}^{n} w_i \sigma_i dW_i(t) \tag{2.38}$$

Then Equations (2.36) and (2.38) imply:

$$d(\ln(P^{reb}(t))) - \sum_{i=1}^{n} w_i d(\ln(S_i(t))) = \frac{1}{2} \left(\sum_{i=1}^{n} w_i \sigma_i^2 - \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{i,j} \right) dt$$

and:

$$d(\ln(P^{reb}(t))) = d \left(\sum_{i=1}^{n} \ln(S_i(t)^{w_i}) \right) + \frac{1}{2} \left(\sum_{i=1}^{n} w_i \sigma_i^2 - \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{i,j} \right) dt$$

$$\ln(P^{reb}(t)) = \sum_{i=1}^{n} \ln(S_i(t)^{w_i}) + \frac{1}{2}\left(\sum_{i=1}^{n} w_i\sigma_i^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{i,j}\right) t$$
$$= \ln\left(\prod_{i=1}^{n}(S_i(t)^{w_i})\right) + \frac{1}{2}\left(\sum_{i=1}^{n} w_i\sigma_i^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{i,j}\right) t$$

Finally:

$$P^{reb}(t) = \prod_{i=1}^{n} S_i(t)^{w_i} \times \exp\left[\frac{1}{2}\left(\sum_{i=1}^{n} w_i\sigma_i^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{i,j}\right) t\right] = \exp\left(g^* t\right) \prod_{i=1}^{n} S_i(t)^{w_i}$$

2.D Jensen's inequality for concave functions

Let $f : X \to \mathbb{R}$ be a differentiable real-valued function of a real variable such that its domain X is a subset of $\mathbb{R}$.

f is concave $\Leftrightarrow \forall (a,b) \in X^2, \forall \alpha \in [0,1] :\ f((1-\alpha)x + \alpha y)) \geq (1-\alpha) f(x) + \alpha f(y)$.

More generally, given a vector of weights $(w_1, \ldots, w_n) \in (\mathbb{R}_+)^n$ such that $\sum_{i=1}^{n} w_i = 1$ and $(S_1, ..., S_n) \in X^n$ then we have:

$$f(w_1 S_1 + \ldots + w_n S_n) \geq w_1 f(S_1) + \ldots + w_n f(S_n)$$

If we consider $f = \ln$ which is a concave function, then the Jensen's inequality applied to f can be written as:

$$\ln(\sum_{i=1}^{n} w_i S_i) \geq \sum_{i=1}^{n} w_i \ln(S_i)$$

which is equivalent to:

$$\ln(\sum_{i=1}^{n} w_i S_i) \geq \ln\left(\prod_{i=1}^{n}(S_i)^{w_i}\right)$$

and finally by substracting the two terms of the previous inequality we have:

$$\ln\left(\frac{\prod_{i=1}^{n}(S_i)^{w_i}}{\sum_{i=1}^{n} w_i S_i}\right) \leq 0$$

2.E Dispersion term

Formally, we can write the rebalancing premium as:

$$RP(0,t) = \mathbb{E}\left[\frac{1}{t}\ln\left(\frac{P_t^{reb}}{P_t^{b\&h}}\right)\right] = g^* + \frac{1}{t}\mathbb{E}\left[\ln\left(\frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right)\right]$$

Where $\frac{1}{t}\mathbb{E}\left[\ln\left(\frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right)\right]$ is defined as the dispersion term.

Let us consider a given trajectory for the asset prices. We note the asset prices at time t: $(S_1(t), \ldots, S_n(t))$. By definition of the (realised) growth rate, we have for each asset i: $g_i^{rea}(0,t) = \frac{1}{t}\ln\left(\frac{S_i(t)}{S_i(0)}\right)$ and $S_i(t) = S_i(0)\exp(g_i^{rea}(0,t)\times t)$.[21] We have:

$$\begin{aligned}\ln\left(\prod_{i=1}^{n} S_i(t)^{w_i}\right) &= \sum_{i=1}^{n} w_i \ln(S_i(t)) \\ &= \ln(S(0)) + \sum_{i=1}^{n} w_i g_i^{rea}(0,t)t\end{aligned}$$

and:

$$\begin{aligned}\ln\left(\sum_{i=1}^{n} w_i S_i(t)\right) &= \ln\left(\sum_{i=1}^{n} w_i \exp\left(\ln(S_i(0)) + g_i^{rea}(0,t)t\right)\right) \\ &= \ln\left(S(0)\sum_{i=1}^{n} w_i \exp\left(g_i^{rea}(0,t)t\right)\right) \\ &= \ln(S(0)) + \ln\left(\sum_{i=1}^{n} w_i \exp\left(g_i^{rea}(0,t)t\right)\right)\end{aligned}$$

We can then rewrite the building block of the dispersion term (i.e., without taking its the expected value) as:

$$\begin{aligned}\frac{1}{t}\ln\left(\frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right) &= \frac{1}{t}\ln\left(\prod_{i=1}^{n} S_i(t)^{w_i}\right) - \frac{1}{t}\ln\left(\sum_{i=1}^{n} w_i S_i(t)\right) \\ &= \sum_{i=1}^{n} w_i g_i^{rea}(0,t) - \frac{1}{t}\ln\left(\sum_{i=1}^{n} w_i \exp\left(g_i^{rea}(0,t)t\right)\right)\end{aligned} \tag{2.39}$$

We set $\bar{g}(0,t) = \sum_{i=1}^{n} w_i g_i^{rea}(0,t)$ the weighted average realised growth rate of the individual assets.

[21] We assume that $\forall i \in [\![1;n]\!]$, $S_i(0) = S(0)$.

Equation (2.39) becomes:

$$\begin{aligned}\frac{1}{t}\ln\left(\frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right) &= \sum_{i=1}^{n} w_i g_i^{rea}(0,t) - \frac{1}{t}\ln\left(\sum_{i=1}^{n} w_i \exp\left(\left(g_i^{rea}(0,t) - \bar{g}(0,t)\right)t\right)\exp\left(\bar{g}(0,t)t\right)\right)\\ &= \sum_{i=1}^{n} w_i g_i^{rea}(0,t) - \bar{g}(0,t) - \frac{1}{t}\ln\left(\sum_{i=1}^{n} w_i \exp\left(\left(g_i^{rea}(0,t) - \bar{g}(0,t)\right)t\right)\right)\\ &= -\frac{1}{t}\ln\left(\sum_{i=1}^{n} w_i \exp\left(\left(g_i^{rea}(0,t) - \bar{g}(0,t)\right)t\right)\right)\end{aligned}$$

Finally, a second order Taylor expansion valid for $(g_i^{rea}(0,t) - \bar{g}(0,t))t << 1$ gives:

$$\begin{aligned}\frac{1}{t}\ln\left(\frac{\prod_{i=1}^{n} S_i(t)^{w_i}}{\sum_{i=1}^{n} w_i S_i(t)}\right) &\approx -\frac{1}{t}\ln\left(\sum_{i=1}^{n} w_i\left(1 + (g_i(0,t) - \bar{g}(0,t))t + (g_i(0,t) - \bar{g}(0,t))^2\frac{t^2}{2}\right)\right)\\ &\approx -\frac{1}{t}\ln\left(1 + \underbrace{\sum_{i=1}^{n} w_i (g_i(0,t) - \bar{g}(0,t))t}_{=0} + \sum_{i=1}^{n} w_i (g_i(0,t) - \bar{g}(0,t))^2\frac{t^2}{2}\right)\\ &\approx -\frac{1}{t}\sum_{i=1}^{n} w_i (g_i(0,t) - \bar{g}(0,t))^2\frac{t^2}{2}\\ &\approx -\frac{t}{2}\ Var(g)\end{aligned}$$

Where $Var(g) = \sum_{i=1}^{n} w_i (g_i(0,t) - \bar{g}(0,t))^2$ is the weighted variance of the growth rates across the assets.

2.F Hurst exponent estimation

The Hurst exponent is not only defined as the self-similarity index in a fractional Brownian framework but more generally as a measure of long memory in time series. This indicator was first proposed by the English hydrologist Hurst (1951) who developed the empirical rescaled range methodology for measuring long-range dependence of the Nile.

If the estimated Hurst exponent is less than 0.5 the process displays "anti-persistence", meaning that positive excess return is more likely to be reversed and therefore the next period's performance is likely to be below average. If the estimated Hurst exponent is greater than 0.5, the process displays "persistence". This means that positive excess return is more likely to be continued and therefore the next period's performance is likely to remain above average. Finally if the Hurst index is equal to 0.5, the returns do not display any memory and thus positive excess return is equally likely to be followed by above or below average performance.

Weron (2002) considers three methods for estimating the Hurst exponent of an asset return time series: the rescaled range (R/S) analysis, the detrended fluctuation analysis (see Peng et al.

(1994) for further details) and the periodogram regression method (see Geweke and Porter-Hudak (1983) for further details).

The R/S analysis, one of the best known method, was first proposed by Mandelbrot and Wallis (1968). Its rationale is:

1. Consider a time series $(Z_t)_{1 \leq t \leq T}$ of length T,
2. We split $(Z_t)_{1 \leq t \leq T}$ into m subseries $(Z_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ of length n $(T = m \times n)$[22],
3. For each subseries $j \in [\![1, m]\!]$ of length n, calculate its mean E_j and its standard deviation S_j,
4. Normalize each subseries j as:

$$X_{i,j} = Z_{i,j} - E_j \ (i \in [\![1; n]\!])$$

,

5. Calculate for each subseries j its cumulative time series:

$$Y_{i,j} = \sum_{k=1}^{i} X_{k,j} \ (i \in [\![1; n]\!])$$

,

6. Calculate for each subseries j the range:

$$R_j = \max(Y_{1,j}, \ldots, Y_{n,j}) - \min(Y_{1,j}, \ldots, Y_{n,j})$$

,

7. For all the m subseries of length n the rescaled range R/S statistics is defined as:

$$(R/S)_n = \frac{1}{m} \sum_{k=1}^{m} \frac{R_k}{S_k}$$

.

The asymptotic relation $(R/S)_n \sim cn^H$ (see Weron (2002)) allows to estimate H with the following linear regression:

$$\ln(R/S)_n = \ln c + H \ln n$$

over different subseries length n.

The detrended fluctuation analysis main steps are:

1. Consider a time series $(Z_t)_{1 \leq t \leq T}$ of length T (like for the R/S analysis),
2. Split $(Z_t)_{1 \leq t \leq T}$ into m subseries $(Z_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ of length n $(T = m \times n)$ (like for the R/S analysis),
3. Calculate for each subseries j its cumulative time series:

$$Y_{i,j} = \sum_{k=1}^{i} X_{k,j} \ (i \in [\![1; n]\!])$$

,

[22] We assume that T is a multiple of n.

4. For each subseries j ($j \in [\![1; m]\!]$), determine the parameters (a_j, b_j) that minimise:

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_{i,j} - a_j i - b_j)^2}$$

,

5. Calculate the root mean square standard deviation of each of the detrended subseries j:

$$F(j) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_{i,j} - a_j i - b_j)^2}$$

,

6. Finally calculate the mean value of the root mean square fluctuation for all subseries of length n:

$$\bar{F}(n) = \frac{1}{m}\sum_{k=1}^{m} F(k)$$

The asymptotic relation $\bar{F}(n) \sim cn^H$ allows to estimate H with the following linear regression:

$$\ln \bar{F}(n) = \ln c + H \ln n$$

.

The periodogram regression method is based on the periodogram (or squared absolute value of the Fourier transform) of a time series $(x_t)_{t=1,\dots,L}$ defined as:

$$I_L\left(\frac{k}{L}\right) = \frac{1}{L}\left|\sum_{t=1}^{L} x_t \exp\left(-2\Pi i(t-1)\frac{k}{L}\right)\right|^2, \text{ with } k = 1, \dots, \left[\frac{L}{2}\right]$$

Then the linear regression

$$\ln\left(I_L\left(\frac{k}{L}\right)\right) = a - b\ln\left(4\sin^2\left(\frac{k}{2L}\right)\right) + \varepsilon_k$$

is computed and finally the estimation of the Hurst exponent $H = b + 0.5$ is obtained.

Chapter

3 Holistic Goals-Based Investing Framework for Analyzing Efficient Retirement Investment Decisions in the Presence of Long-Term Care Risk

Abstract This chapter presents a formal analysis of efficient investment strategies for individuals and households in the decumulation phase of their life-cycle. Our main contribution is to develop a comprehensive and flexible framework for providing personalized advice to retirement investment decisions in the presence of life event risk, with a relatively rich menu of investment opportunities that includes balanced funds, target date funds, but also various types of annuity products, for which we use realistic market quotes. The main finding of the chapter is that the presence of long-term care risk strongly reduces the optimal demand for annuities for most individuals. This result suggests that the costly reversibility of annuitization decisions contributes to explaining the annuity puzzle for individuals facing life event uncertainty. We also find that the introduction of annuity products with upside potential (e.g., variable annuities) has a strongly positive impact on both investor welfare and the demand for annuitization.

This chapter is the fruit of a joint work with Lionel Martellini [1] and Vincent Milhau [2].

An article will be derived from this chapter and submitted to *Insurance: Mathematics and Economics*.

Contents

[1] Professor of finance at EDHEC Business School and director of EDHEC-Risk Institute
[2] Research director at EDHEC-Risk Institute

3.1 Introduction

A major crisis is threatening the sustainability of pension systems across the globe. The first pillar of pension systems, which is made of public social security benefits and aims at providing a universal core of pension coverage to address basic consumption needs in retirement, is strongly impacted by rising demographic imbalances. Life expectancy at age 65 in OECD countries is expected to grow by 4.2 years for women and 4.6 years for men between 2020 and 2065. As a result, the number of individuals aged 65 and over per 100 individuals aged between 20 and 64 rose from 13.9 in 1950 to 27.9 in 2015, and is expected to grow to 58.6 by 2075.[3] In parallel, the second pillar of pension systems, which is expected to provide additional replacement income for retirees via public or private occupational pensions, is also weakening. In particular, private pension funds have been strongly impacted by the shift in accounting standards towards the valuation of pension liabilities at market rates instead of fixed discount rates, which have resulted in increased volatility for the value of liabilities. The impact of this new constraint has been reinforced by stricter solvency requirements following the 2000-2003 pension fund crisis. As a result of these changes in accounting and prudential regulations, a large number of corporations have closed their defined-benefit pension schemes to new members and increasingly to further accrual of benefits so as to reduce the impact of pension liability risk on their balance sheets and income statements. Overall, a massive shift from defined-benefit pension schemes to defined-contribution pension schemes is taking place across the world, implying a transfer of retirement risks from corporations to individuals.

As an almost universal rule, pillar I and pillar II pension arrangements deliver replacement income that is inferior to the needs of individuals in retirement, and the inadequacy risk is sometimes severe. According to the aforementioned OECD report, an individual earning the average income in the United States can expect to enjoy a mere 49.1% replacement rate upon retirement, a number that falls to 29.0% in the United Kingdom. With the need to supplement public and private retirement benefits via voluntary contributions, the so-called third pillar of pension systems, individuals are becoming increasingly responsible for their own retirement savings and investment decisions. This global trend poses substantial challenges to individuals, who often lack the expertise required to make such complex financial decisions in the face of multiple sources of uncertainty that include longevity risk, market and inflation risks as well as risk of unforeseen expenses due to life events such as long-term care needs.

In principle, annuity products are attractive since they can be used to generate a target level of replacement income throughout retirement and can therefore solve the problem of longevity, market and inflation risks. This intuition has been confirmed by many papers, including Pfau (2013), Huang et al. (2012), Milevsky (2013), Milevsky and Macqueen (2015), Vernon (2014, 2015) and Pang and Warshawsky (2012) inter alia, which have demonstrated the value of annuity contracts in a retirement investment context. In practice, however, the demand for such products is extremely low despite their risk-free nature in a retirement investment context. Using the RAND Health and Retirement Study dataset for the cohort aged 65-75 in 1998, Pashchenko (2013) for example reports that only 5% of individuals receive income from annuities, with a peak at 12.2% among the highest income quintile and a low at 0.4% for the lowest quintile. Common explanations of the "annuity puzzle" are related to the fact that annuities involve counterparty risk and high levels of fees, and also that they do not contribute to bequest objectives. One additional key drawback of annuity products is their severe lack of flexibility. Indeed, annuitization is an

[3] Figures cited here are from the OECD report *Pensions at a Glance 2017*.

almost irreversible decision, unless one is willing to bear the costs of high surrender charges, which can amount to several percentage points of the invested capital. Brown and Warshawsky (2001) argue that this lack of flexibility is a major shortcoming in the presence of life event uncertainties such as marriage and children, changing jobs, changing locations to lower or higher cost cities or countries, decisions about retirement dates, and also, and perhaps most notably, long-term care needs implied by health-related issues in the later stage of retirement. These uncertainties require modification to retirement plans on a regular basis, annually or when life events occur, which is simply not possible with annuities.

One of the main goals of this chapter is to assess whether the opportunity cost of such substantial rigidity may partly explain the annuity puzzle, and to provide a quantitative measure of the associated impact on the optimal demand for annuities. To perform this analysis, we introduce a simulation framework that incorporates a market simulation engine, involving Monte Carlo simulations coupled with realistic long-term capital market assumptions (CMAs), a product simulation engine incorporating scenarios for stocks, bonds as well as annuity products, and an individual investor simulation engine incorporating mortality risk scenarios, as well as target levels of replacement income cash-flows including random shocks to replacement income needs due to life events such as long-term care needs.

This chapter is not the first attempt to measure the impact of long-term care risk on the demand for annuities. In a related effort, Pashchenko (2013) develops an individual optimization model to disentangle the impacts of various potential explanations for the annuity puzzle on participation rates in the annuity market, and finds that the presence of uncertain medical expenses does reduce predicted annuitization rates, except among the individuals with the 20% largest earnings. On the one hand, this factor appears to be secondary compared to the bequest motives, the presence of pre-annuitized wealth, housing ownership and the imposition of a minimum amount on annuity purchase. Reichling and Smetters (2015) consider the effect of negative health shocks, which reduce an individual's self-estimated survival probabilities, hence the expected present value of future annuity benefits, and also incur medical costs and possible losses of wages. If agents are concerned with hedging against the events that adversely affect their welfare, as in the inter-temporal portfolio choice model of Merton (1973a), this implies that annuities are less desirable. On the other hand, the model of Pang and Warshawsky (2010) predicts that uncertainty over health expenditure can increase the optimal holding in annuities, even though they are illiquid. A possible explanation is that annuities may still be attractive to older individuals, for they act as a hedge against future medical expenses and they are expected to outperform bonds thanks to their embedded mortality premium. This work is also closely related to Peijnenburg et al. (2017), who use a life-cycle investment and consumption model with health care risk, and find that the presence of health cost risk can explain a large proportion of empirically observed annuity choices, and in particular that it lowers the optimal annuity demand at retirement.

We complement this existing body of research by considering a more flexible framework that allows for a richer set of assets including various forms of annuities. More specifically, we consider in this chapter Single Premium Immediate Annuities with a Cost-of-Living Adjustment (SPIA-COLA in short) as well as also Variable Annuities (VA in short) (see Section 3.3.1.1 for more details). A SPIA-COLA is purchased with a lump sum of money and offers a guaranteed source of income for retirement over the life of the policyholder. The income received by the policyholder increases each year based on a cost-of-living adjustment. Unlike a SPIA-COLA which does not offer upside potential, a VA offers the opportunity to the policyholder, if certain favorable conditions are met, to withdraw more than the minimum guaranteed amount at the expense of minoring the future minimum guaranteed amount.

Our main findings are as follows. We first confirm that the presence of long-term care risk strongly reduces the optimal demand for immediate annuities for most individuals, at least those who are sufficiently risk-averse to show appetite for annuities in the first place. For example in the case of an individual targeting a 4% withdrawal rate, we find that the demand for annuities from

a moderately risk-averse individual decreases from 100% to 23% when long-term care needs are accounted for (see Section 4 for more details). In the case of a 5% initial target withdrawal rate, the demand for annuities for the same moderate individual decreases from 100% to 47% when long-term care needs are accounted for. Overall these results suggest that the costly reversibility of annuitization decisions can contribute to explaining the annuity puzzle for individuals facing life event uncertainty. We also find that the optimal demand for annuities is substantially higher when variable annuities are considered instead of SPIA-COLAs, suggesting that the upside potential of VAs make them a strong competitor for stock and bond products for individuals mostly interested in performance. We also find that the optimal allocation strategies are relatively close in the presence and in the absence of life events for relatively high risk aversion levels. This last result can be explained by the increased flexibility offered by the VA compared to the SPIA-COLA in case of liquidity needs due to the occurrence of a life event and/or unfavorable market conditions.

The rest of the chapter is organized as follows. Section 2 presents a detailed review of the related literature. Section 3 provides a stylized description of a variety of existing annuity products, including single premium immediate annuities (SPIA), deferred income annuities (DIA) and variable annuities (VA). It also contains a detailed discussion of annuity pricing methodologies. Section 4 presents the main numerical results. Section 5 contains the main numerical results. More detailed specifications and examples of annuity products can be found in a dedicated Appendix.

3.2 Related Literature

In this section we review the related literature, which we sort into two main categories, namely existing research on the demand for annuities and the annuity puzzle in the absence of long-term care risk, and existing research on life events and long-term care risks and their impact on annuitization decisions by individuals and households.

3.2.1 The Demand for Annuities in the Absence of Long-Term Care Risk

Yaari (1965) argues that if an individual faces uncertain lifetime, has no utility from leaving a positive bequest but is subject to the constraint of leaving non-negative assets upon death, then they should hold their assets entirely in the form of "actuarial notes", defined as perpetual notes that pay a higher interest rate than regular notes in exchange for stopping their payments at death time. This argument provides conceptual justification for full annuitization of wealth through the purchase of lifetime annuities. Yaari (1965) assumes that annuities are "fairly" priced, that is the promised rate of return on actuarial notes should equal the rate on standard notes plus the survival intensity [4], but Davidoff et al. (2005) show that this condition is not required and that full annuitization remains optimal as long as there is no bequest motive, annuities deliver a premia over alternative financial assets and the annuity market is "complete" in the sense that any consumption path can be matched with selected annuities. To have a quantitative measure of the benefits of having access to annuities, Mitchell et al. (1999) and Brown and Warshawsky (2001) calculate the additional amount of wealth that makes an individual with uncertain lifetime indifferent between purchasing an actuarially fair annuity or getting no income from an annuity. Mitchell et al. (1999) find that he would be willing to sacrifice some 25% to 40% of savings at age 65 to include annuities in his portfolio.

In practice, individuals appear to be reluctant to purchase the sort of lifetime annuities considered in these papers, a stylized fact that is known as the "annuity puzzle" since it is at odds with the theoretical evidence for the benefits of annuitization. Using figures from the Health and Retirement Study in 1998, Pashchenko (2013) finds that only 5% of individuals aged 70 or more report receiving income from annuities, a percentage that is higher among those with the

[4]The survival intensity is the probability of death per unit of time.

highest income and lower among the group with the lowest income. Brown and Warshawsky (2001) survey a range of possible explanations. First, annuities could be marketed at prices above the actuarially fair level, for they include fees charged by providers and an adverse selection provision to anticipate the fact that annuity buyers tend to live longer than the general population. In support for this explanation, Mitchell et al. (1999) calculate the "money's worth" of representative annuities by discounting future benefits and weighting them by survival probabilities. A value of 1 would mean that annuities are sold at actuarially fair prices, but the actual value is found to range approximately from 0.75 to 0.90. Second, annuities do not contribute to bequest. Third, individuals face the possibility of "life events", which are analyzed in greater detail in Section 3.2.2 and are defined as events that generate potentially large expenses, such as the need for long-term care. The quasi-impossibility of recovering the annuitized capital in such situations discourages the purchase of annuities and encourages self insurance via precautionary savings that are therefore unavailable for annuitization. Fourth, annuities may not be properly understood by consumers whose behavior is not well described by rational decision models. Fifth, there is a general trend away from defined-benefit towards defined-contribution (DC in short) pension plans, and participants in DC plans are less prone to choose an annuity, if only because this option is not always available to them. Several authors examined different frameworks with annuities to determine the optimal strategy for retirees, based on trade-offs supporting minimum spending needs and preserving financial assets. Pfau (2013) considered an asset allocation among stocks, bonds, inflation-adjusted single premium immediate annuities (SPIAs), fixed SPIAs and variable annuities with guaranteed living benefit riders (VA/GLWBs). He finds that optimal retirement income portfolio generally consists of stocks and fixed SPIAs. In the same vein, Huang et al. (2012) introduce the Efficient Income Frontier (EIF) concept based on the trade-off between income risk and legacy potential with asset allocation among traditional investments, single premium immediate annuities (SPIAs), and variable annuities with guaranteed living benefit riders (VA/GLWBs). They assess that the inclusion of annuities in a retiree's portfolio reduces income risk, increases legacy potential or increases the amount of income that can likely be sustained throughout retirement. Milevsky (1998) complements this body of research by analyzing the optimal age at which to annuitize and concluding that the decision to anuitize should come relatively late in retirement.

Milevsky (2013) and Milevsky and Macqueen (2015) develop the concept of product allocation, a technique that seeks to hedge against key retirement risks by allocating savings across a variety of investment and insurance-based products. They design their optimal asset allocation with traditional investments, income annuities (immediate or deferred income) and deferred annuities with a guaranteed withdrawal benefit rider and obtain that the trade-off of income sustainability versus financial legacy is at the heart of the product allocation strategy. Vernon (2014, 2015) develops an analytical framework to determine the optimal retirement income generators (RIGs) for retirees that includes: drawing from Social Security, a Systematic Withdrawal Plan, buying a guaranteed lifetime annuity, a reverse mortgage, interest and dividend income and real estate rental income. He observes that (1) no single retirement income generator delivers on all potential retirement goals, (2) clients will need to prioritize and make trade-offs among various retirement goals and (3) annuities can be used strategically to improve retirement security. Pang and Warshawsky (2012) compared the following six strategies: (1) systematic withdrawal from mutual funds, (2) fixed payout immediate life annuity, (3) immediate variable annuity for life, (4) variable annuity (VA in short) plus guaranteed minimum withdrawal benefit (GMWB in short), (5) systematic withdrawal from mutual funds plus fixed payout immediate life annuity and (6) systematic withdrawal from mutual funds plus fixed payout life annuity, gradual annuitization over time. They find that none of the six strategies considered dominates the others and that the best advice may be to segment wealth to establish minimum necessary income and hedge longevity risk, then focus on growth. Brown et al. (2001), Feldstein and Ranguelova (2001), Charupat and Milevsky (2002), Horneff et al. (2009), Horneff et al. (2010), Koijen et al. (2011) and Maurer et al. (2013) also studied optimal asset allocation problems in a retirement context

with variable annuities in the investment universe.

Overall these authors put forward the added value of the annuity in a retirement portfolio context. The main aspects we want to test in this chapter are (1) the potential benefits of annuities in the decumulation phase documented in the literature and (2) the robustness of these benefits when life events are taken into account.

3.2.2 Life Events and Annuitization

One reason for the lack of popularity of lifetime annuities is that the decision to buy them is not easily reversed, except at the cost of substantial surrender charges. So, an annuity buyer cannot easily recover the invested capital, even if they need a big amount of money to bear the large expenses associated with certain life events. In our framework, a life event is defined as an event that does not occur for sure during the decumulation phase but generates expenses, either in the form of upfront costs, or in the form of multiple payments that can be temporary or perpetual. A severe illness with a costly medical treatment fits in the former category, and a permanent disabled condition belongs to the latter. Implicit in the definition of a life event is the assumption that the individual has no option to give up on or to delay the payments. Moreover, a life event must be either uninsurable or not usually insured against, so as to justify dedicated precautionary savings that constitute self insurance. Such events are a potential obstacle to annuitization because individuals concerned with life events can prefer to save money in liquid form.

The typical example of a life event, and one that is of concern to many individuals, is that of a negative shock on health condition that calls for costly medical care and/or long-term care, either in a nursing home or at the person's home. As noted by French and Jones (2004), a medical event can be of concern either because it generates huge expenses in a single year or because it gives rise to small but permanent expenses. Quoting figures from the Congressional Budget Office, Brown and Finkelstein (2011) report that about 60% of expenditures on long-term care are paid by public insurance (Medicaid and Medicare) and only 4% by private insurance, so that about one third are paid out of pocket. Indeed, neither Medicare nor Medigap insurance plans cover the part of long-term care that does not correspond to medical services, so most of stays in nursing home or assistance with daily living is borne by individuals. Using figures of the 2008 Health and Retirement Study, Brown and Finkelstein (2011) find that only 13.8% of respondents aged 60 and above own long-term care insurance, a rate that varies across wealth groups, with wealthier individuals being more prone to purchase insurance.

A first model of health care costs has been developed by French and Jones (2004), and has been used in the subsequent studies of De Nardi et al. (2010), Pashchenko (2013) and Peijnenburg et al. (2017), with a few variants. It covers the time series dynamics and the cross section dispersion of health care costs, by decomposing them (in fact, their logarithm) into an expected part and a shock. In French and Jones (2004), the expected term is a function of sex, married status, age, income, the presence of employer-provided or privately purchased insurance and the presence of Medicaid coverage. De Nardi et al. (2010) model it as a function of health status, classified as "good" or "bad", sex, age and permanent income. The shock is the sum of a fixed person effect, a transitory component that follows a moving average process (French and Jones, 2004) or a Gaussian white noise (De Nardi et al., 2010) and a persistent component that follows an AR(1) process. In De Nardi et al. (2010), the variance of the shock is a function of the same four individual characteristics as the expected costs, while Pashchenko (2013) assumes a uniform value across individuals and periods. Health care costs thus modeled enter the budget constraint of an individual as an expense, but also as a driver of the Government transfers that they receive, since these transfers make up for the gap, if any, between a minimum consumption level plus medical expenses and the income from financial assets and annuities. De Nardi et al. (2010) and Pashchenko (2013) use this budget equation as a constraint in a consumption-investment choice problem.

Turning to the impact of long-term care risk and the demand for annuities, Pashchenko (2013) develops an individual optimization model to disentangle the impacts of various potential explanations for the annuity puzzle on participation rates in the annuity market, and she finds that the presence of uncertain medical expenses does reduce predicted annuitization rates, except among the individuals with the 20% largest earnings, but this factor appears to be secondary compared to the bequest motives, the presence of pre-annuitized wealth, housing ownership and the imposition of a minimum amount on annuity purchase. Reichling and Smetters (2015) considers the effect of negative health shocks, which reduce an individual's self-estimated survival probabilities, hence the expected present value of future annuity benefits, and also incur medical costs and possible losses of wages. If agents are concerned with hedging against the events that adversely affect their welfare, as in the inter-temporal portfolio choice model of Merton (1973a), this implies that annuities are less desirable. On the other hand, the model of Pang and Warshawsky (2010) predicts that uncertainty over health expenditure can increase the optimal holding in annuities, even though they are illiquid. A possible explanation is that annuities may still be attractive to older individuals, for they act as a hedge against future medical expenses and they are expected to outperform bonds thanks to their embedded mortality premium. Turra and Mitchell (2008) studied the annuity demand in the presence of health expense risk. Ameriks et al. (2011) studied the impact of long-term care risk and bequest motives on the demand for annuities: they found that long-term care risk diminishes the optimal annuity demand. Koijen et al. (2016) focused on the impact of long-term care risk on the demand for annuities but also on the demand for life insurance and for long-term care insurance: they acknowledged a low level of annuitization with respect to a life-cycle model with no bequest motive and an unexplained variation in the annuity demand across households. Our work is also closely related to Peijnenburg et al. (2017), who use a life-cycle investment and consumption model with health care risk, and find that the presence of health cost risk can explain a large proportion of empirically observed annuity choices, and in particular that it lowers the optimal annuity demand at retirement. We complement this existing body of research by considering a more flexible framework that allows for a rich set of assets, including various forms of annuities.

3.3 A Review of Existing Annuity Products and Associated Pricing Methods

In this section, we first provide a brief description of standard annuity products at the level of generality and formalization that is required for the modeling of these products in the context of the stochastic simulations that we intend to perform. We then turn to a brief discussion of competing annuity pricing methodologies, and we refer the interested reader to Milevsky (2013) for a general discussion of research on life annuities, including institutional aspects and valuation issues.

3.3.1 Overview of Annuity Products

Annuities are contracts written for a specific individual, called the annuitant and usually identical to the beneficiary, so they are attached to a physical person. They can neither be sold nor liquidated, so there is no secondary market that is comparable in size and liquidity with the financial markets. However, some contracts allow their owner to *surrender* them, often at significant costs. Annuities have benefits determined at the issuance date, either in the form of a fixed stream of income or in the form of a rule that specifies how they depend on the value of some underlying investment. A key distinction exists between *live benefits*, which are paid if the annuitant is alive, and *death benefits*, which are paid to heirs upon the annuitant's death. Finally, most annuities provide a hedge against longevity risk at the individual level by guaranteeing that live benefits are paid as long as the annuitant is alive. This characteristic is critically important in the context

of retirement investing since no financial asset offers similar protection.

Annuities come in large variety. The simplest contract is a lifetime annuity, which pays fixed income for an individual's lifetime, but as noted in the previous section, it is not as popular as microeconomic theory of consumer choices predicts it should be. In practice, most of the annuity market is represented by more complex products that aim to address some of the shortcomings of plain vanilla lifetime annuities by providing a range of additional options. "Fixed-index annuities" and "variable annuities" are examples of such contracts in which the benefits depend on the performance of some underlying investment, and they form a large part of annuity sales.

This section presents a set of formal expressions for the benefits of annuities in terms of the premium paid. It is important to point that these expressions does not rely on any specific model for the dynamics of interest rates or financial markets. Such a model is needed to perform numeric implementation and run simulations, but it is irrelevant to specify contract rules.

3.3.1.1 Characteristics Common to All Annuity Contracts

Beyond their specific features, all annuities contracts have a number of characteristics in common, which we review now, and for which we introduce dedicated mathematical notations.

State Vector Among the characteristics are functions that describe the state of an annuity at a given point in time, e.g. the fees to be paid in a given period or the value of the underlying fund for variable annuities. These are functions of time, but also of "state variables", of economic and financial nature, that are relevant for the contract considered. The values of these variables at a date t are collected in a vector $\mathbf{X}_t$, the content of which is left unspecified in the general presentation of annuities but will be made explicit for each product below. Formally, $\mathbf{X}_t$ is a stochastic vector that contains any information useful to describe the state of the annuity at a given point in time. It must be measurable for the sigma-algebra $\mathcal{F}_t$, which is the information set available to the user of the framework at time t. The sequence of sigma-algebras is increasing, so it forms a filtration $\mathcal{F}$, to which the stochastic process $\mathbf{X}$ must be adapted. In what follows, the value of a stochastic process at time t is denoted with a subscript t, and it is always implicitly assumed to be measurable for $\mathcal{F}_t$.

Premium Each contract is purchased by paying a premium P at the issuance date. It can be thought as a "price" for the contract, but does not represent a no-arbitrage price in the sense of financial theory because insurance contracts are not usually priced by arbitrage arguments like options are in financial markets. A first reason is that annuities do not have a tradable underlying security, so they do not lend themselves to dynamic replication. Second, arbitrage opportunities with annuities would not be easy to exploit given that they cannot be liquidated and that no entity can be both a purchaser and a seller of annuities, while taking advantage of an arbitrage opportunity means purchasing underpriced securities and selling overpriced ones to make money with a zero investment. Last, annuities involve exposure to the risk of longevity of a single individual, and this risk cannot be hedged with financial instruments. Section 3.3.2 reviews pricing approaches for annuities.

Payment Dates for Benefits Live benefits to the beneficiary are paid as long as the annuitant, who is usually the same person as the beneficiary and to whom we refer in what follows as the "individual", is alive, but they take place on predefined dates. We stack these dates in a vector $\boldsymbol{\Theta}$, so for each element t of this vector the individual receives income from the annuity at date t if she is alive. Variable contracts offer flexibility with respect to the amount of money withdrawn from the contract, since this amount does not necessarily coincide with the benefits specified in the contract. We assume in the framework that withdrawals, whether equal to or different from the contractual benefits, take place only at dates t.

In addition to live benefits, some annuities also provide death benefits, which are paid to a designated beneficiary, usually the individual's heirs. These payments occur at the individual's death time, which we denote with τ and is by nature a random date.

Payout Rate Function All annuities have "contractual benefits", defined as the maximum amount that the individual can withdraw from the contract without penalty. Depending on the annuity type and the rules that govern the contract, she can withdraw less or more than the contractual amount, which depends on the date but is also possibly a function of state variables like the performance of an underlying fund, or the realization of a price index. So, the contractual benefits at time t are a random variable e_t.

They are calculated as the product of a *payout rate* δ_t and a *benefit base* B_t:

$$e_t = \delta_t \times B_t.$$

The description of each annuity should include a function that calculates the payout rate from the current date and state vector. With a slight abuse of notation, we denote it with δ, so we have

$$\delta_t = \delta(t, \mathbf{X}_t).$$

An important special case is when δ is a constant function, that is when the payout rate is fixed once for all at the issuance of the contract and depends neither on time nor on the realized scenario. The contractual benefits can still be time- and scenario-dependent if the benefit base is. Alternatively, some lifetime annuities have their payout rate indexed on realized inflation, in which case the vector $\mathbf{X}_t$ contains the growth in the price index between a reference date and date t.

Benefit Base Function The second function that enters the description of an annuity explains how the benefit base is calculated from the date and the state. Again committing a minor abuse of notation, we denote this function with B, so that

$$B_t = B(t, \mathbf{X}_t)$$

For vanilla lifetime annuities, the benefit base is constant, but for variable annuities, the benefit base depends on the performance of an underlying fund and the withdrawals that were made on previous dates, so it becomes a genuine random variable.

Contract Value Function We introduce the contract value as a third function that characterize an annuity. It corresponds to the value the annuity is worth if the individual decides to cancel her contract. It is calculated from the date and the state. Again committing a minor abuse of notation, we denote this function with C, so that

$$C_t = C(t, \mathbf{X}_t)$$

For vanilla lifetime annuities, which are totally illiquid and no additional amount can be withdrawn if need be, the contract value is null, but for variable annuities, the contract value depends on the performance of an underlying fund and the withdrawals that were made on previous dates, so it becomes a genuine random variable.

Fee and Charge Function In some annuity contracts, fees are already "factored" in the payout rate, so they are not explicitly charged to the annuitant during the contract's life. This applies for instance to immediate annuities with fixed benefits, which are set to a lower value than what would be implied by the fair actuarial valuation formula (see Section 3.3.2), so as to compensate for the insurer's costs. On the other hand, for contracts with an underlying investment, fees are

billed periodically and they are subtracted from the contract value, so they indirectly impact benefits paid on subsequent dates. Like the other functions, they are calculated from the time and the state vector, so the fees to be paid at time t can be written as

$$\Phi_t = \Phi(t, \mathbf{X}_t)$$

The function Φ will be taken to be the constant function that returns a value of zero for annuities with implicit fees.

Death Benefit Function Some annuities pay a last benefit when the annuitant dies. Such covenants are intended to mitigate an often cited concern raised by annuities, which is that individuals are reluctant to annuitize if they are afraid of dying sooner than expected, thereby leaving most of the capital they had invested to the insurer. With a death benefit, heirs receive money when the annuitant dies, so that the concern over the possibility of making a bad deal is alleviated. In the most general specification, death benefits can be written as a function of the date and the state vector, so the benefit paid at date t in the event of the annuitant's death has the form

$$D_t = D(t, \mathbf{X}_t)$$

3.3.1.2 Presentation of Specific Annuity Contracts

We first review single premium immediate annuities, before moving on to discuss more complex annuity products. While our simulation framework is flexible enough to handle a variety of annuity products, we will mostly focus on single premium immediate annuities and variable annuities in this chapter.

Single Premium Immediate Annuities (SPIA) and Deferred Income Annuities (DIA) Single premium immediate annuities (SPIA) and deferred income annuities (DIA) have similar characteristics, the main difference is that single premium immediate annuities start to pay benefits in the first year, while deferred income annuities will not begin to pay benefits until some time after the first year. We first present plain vanilla contract before discussing contracts with a cost-of-living adjustment.

Plain Vanilla Contracts The simplest form of annuities pays constant live benefits that are determined at the time the contract is written and begin within one year of purchase. So, if the issuance date is denoted as date 0 by convention, the first payment date, Θ_0, is less than or equal to 1 year.

Following the standard rule, benefits are expressed as the product of a payout rate and the benefit base, both of which are constant here. The benefit base is equal to the premium paid at the issuance date, so the contractual benefits are

$$e_t = \begin{cases} \delta_t \times P & \text{for } t \text{ in } \boldsymbol{\Theta} \text{ and } t < \tau \\ 0 & \text{for } t \geq \tau \end{cases}.$$

The beneficiary has no option to delay the payments or to withdraw more than e_t, so the actual withdrawals c_t will equal the contractual benefits.

By default, the payout rate is constant, so no state vector is needed to characterize the state of the annuity. Then, $\mathbf{X}_t$ is an empty vector, and in the notations of Section 3.3.1.1, we have

$$\delta(t, \mathbf{X}_t) = \delta, \quad B(t, \mathbf{X}_t) = P, \quad C(t, \mathbf{X}_t) = 0 \text{and} \quad \Phi(t, \mathbf{X}_t) = 0$$

If present in the contract rules, the death benefit has the form of a cash refund, meaning that it is equal to any fraction of the premium that has not been paid in the form of benefits, to ensure

that the sum of the benefits paid by the insurer is at least equal to the premium that it received. Not all annuities have this feature, so we introduce a 0-1 variable DB that is 1 only if the annuity has a death benefit. So, D is a function of time alone given by

$$\begin{aligned} D_t &= \mathrm{DB}\left[P - \sum_{\substack{t \in \boldsymbol{\Theta} \\ t<\tau}} e_t \right]^+ \\ &= \mathrm{DB} \times P \left[1 - \sum_{\substack{t \in \boldsymbol{\Theta} \\ t<\tau}} \delta_t \right]^+ \end{aligned}$$

where $[x]^+$ is the positive part of x. Note that individuals can also have access to Deferred Income Annuities (DIA). DIAs behave similarly to SPIAs, except that the start date of payments is deferred by at least 13 months after purchase. The deferral period can be as long as 30 or 40 years.

SPIA and DIA with a Cost of Living Adjustment or Inflation Indexation The life expectancy at age 65 of a generic US individual is close to 20 years, a duration that is long enough for inflation to erode the purchasing power of savings fixed in nominal terms. To compensate for this effect, some annuities have a payout rate indexed on some cost of living index or on realized inflation. In the former case, the cost of living adjustment is done by applying a fixed growth rate π to the payout rate every year, starting from an initial payout rate. Taking date 0 (the issuance date) to be the reference date for indexation, we have

$$\delta(t) = \delta_0[1+\pi]^t.$$

Alternatively, the payout rate can be proportional to realized inflation, a rule that introduces state-dependency in the annuity and requires the introduction of a state variable I, which is the price index. Then, the state vector is actually a scalar, equal to realized inflation:

$$X_t = \frac{I_t}{I_0}.$$

The payout rate function in this case is given by

$$\delta(t, X_t) = \delta_0 X_t.$$

As for plain vanilla contracts, SPIA and DIA with a cost of living adjustment or inflation indexation verify

$$B(t, \mathbf{X}_t) = P, \quad C(t, \mathbf{X}_t) = 0 \text{and} \quad \Phi(t, \mathbf{X}_t) = 0$$

and have a death benefit

$$D_t = \mathrm{DB} \times P \left[1 - \sum_{\substack{t \in \boldsymbol{\Theta} \\ t<\tau}} \delta_t \right]^+$$

In what follows we will denote a SPIA with a cost of living adjustment as SPIA-COLA and a DIA with a cost of living adjustment as DIA-COLA.

Variable Annuities (VA) One of the key limits of SPIAs and DIAs with or without a cost of living adjustment is that they do not offer any upside potential. As such, they are best suited for individuals who can afford to make contributions that are sufficiently large to meet their replacement income targets. So as to provide upside potential, which may be of interest to both the wealthiest and those who cannot afford to buy as many SPIAs as necessary, a number of annuity products with upside potential have been introduced, including most notably variable annuities (VAs). Since these products provide access to income growth, it is possible that the amount c_t withdrawn by the policyholder at time t exceeds contractual benefits e_t for a VA. In details, we have

$$e_t = \delta_0 \times B_{t-}$$

where B_{t-} the benefit base at time t of the variable annuity before penalties.[5] The payout rate $\delta_t = \delta_0$ is a constant function.

The main variables that drive the dynamic of the variable annuity are the benefit base before and after penalties respectively denotes as B_{t-} and B_t, the contract value before and after withdrawal, fees and surrender charges respectively denoted as C_{t-} and C_t, the penalty rate ν_t, the underlying fund value U_t, the underlying fund return r_t, the fees fee_t, the surrender charges surr_t and the actual withdrawal c_t and the contractual withdrawal e_t.

The initial contract value before withdrawal and fees C_{0-}, the initial benefit base before penalty B_{0-} and the initial underlying fund value U_0 respectively verify:

$$C_{0-} = P, \quad B_{0-} = P, \quad U_0 = 1$$

Where P is the premium paid by the policyholder at date $t = 0$ to buy the variable annuity.

Then we detail below the equations that characterize the evolution of the main variables driving the dynamic of the variable annuity.

- The underlying fund value U_t for a VA will typically correspond to a portfolio invested in stocks, bonds and cash with a maximum allowed stock allocation of 80% for a VA. The underlying fund value U_t at time t follows:

$$U_t = U_{t-1} \times (1 + r_t) \quad \forall t \geq 1$$

- The contract value C_{t-} at time t before withdrawal, fees and surrender charges follows:

$$C_{t-} = C_{t-1} \times \frac{U_t}{U_{t-1}} \quad \forall t \geq 1$$

- The benefit base B_{t-} at time t before penalty verifies:

$$B_{t-} = \max\left(B_{t-1}, C_{t-}\right) \quad \forall t \geq 1$$

- The contract value C_t at time t after withdrawal, fees and surrender charges respectively verifies:

$$C_t = \left[C_{t-} - c_t - \text{fee}_t - \text{surr}_t\right]^+ \quad \forall t \geq 0$$

- The maximum amount e_t that can be withdrawn at time t without penalties verifies:

$$e_t = \delta_0 \times B_{t-} \quad \forall t \geq 0$$

[5] A penalty is different from surrender charges and occurs if the policyholder withdraws more than the maximum amount stated.

- The fees fee_t at time t are given by:

$$\text{fee}_t = f_1 \times C_{t-} + f_2 \times B_{t-} = \Phi(t, X) \quad \forall t \geq 0$$

  The expense rates on C_{t-} and B_{t-} used in the empirical simulations are respectively $f_1 = 0.65\%$ and f_2=1.25%: they are taken from a data provider (see Section 3.3.2).

- The surrender charges surr_t can occur during the first seven years on withdrawals that exceed the annual $\alpha = 10\%$ free withdrawal amount with respect to the contract value C_{t-} with the time-dependent yearly rate s_t following the declining yearly schedule $(8\%, 8\%, 7\%, 6\%, 5\%, 4\%, 3\%)$. The surrender charges surr_t at time t are:

$$\text{surr}_t = s_t \times [c_t - \alpha C_{t-}]^+ \quad \forall t \geq 0$$

- The penalty rate ν_t at time t, which is non null if and only if the policyholder has withdrawn at time t more than the contractual withdrawal e_t, follows:

$$\nu_t = \frac{[c_t - e_t]^+}{C_{t-}} \quad \forall t \geq 0$$

- The benefit base B_t at time t after penalty verifies:

$$B_t = B_{t-} \times [1 - \nu_t] \quad \forall t \geq 0$$

The policyholder cannot withdraw at time t more than $e_t + C_{t-}$. If she totally liquidates her annuity at time t she will receive $e_t + C_{t-}$ and will have to deduce from this amount the possible surrender charges.

The death benefit D_t verifies the equation:

$$D_t = C_{t-} \times DB$$

Where DB is a 0-1 variable that is 1 only if the annuity has a death benefit.

To make things more concrete, Table 3.1 displays how the different variables described above drive the dynamic of the variable annuity and details the income flows c_t paid by the VA. In that example, we assume that the individual withdraws at each date t exactly the maximum contractual withdrawal $e_t = \delta_0 \times B_{t-}$ and as such we have $B_t = B_{t-}$. At date $t = 0$, the individual buys a variable annuity for a premium $P = \$100,000$. The initial benefit base before penalty B_{0-} is equal to $P = \$100,000$. The income c_0 received by the individual at time $t = 0$ to support his needs between $t = 0$ and $t = 1$ is equal to $\delta_0 \times B_{0-} = \$4,720$. The contract value C_0 at time $t = 0$ after withdrawal, fees and surrender charges is equal to $C_{0-} - f_1 \times C_{0-} - f_2 \times B_{0-} - c_0 = \$93,380$. We assume that the underlying fund return between time $t = 0$ and time $t = 1$ is 20%. Then the contract value C_{1-} at time $t = 1$ after withdrawal, fees and surrender charges is equal to $C_0 \times (1 + r_1) = \$112,056$. Since C_{1-} is higher than B_0 we have $B_{1-} = C_{1-} = \$112,056$. The income c_1 received by the individual at time $t = 1$ to support his needs between $t = 1$ and $t = 2$ is equal to $\delta_0 \times B_{1-} = \$5,289$. The contract value C_1 at time $t = 1$ after withdrawal, fees and surrender charges is equal to $C_{1-} - f_1 \times C_{1-} - f_2 \times B_{1-} - c_1 = \$106,767$.

Table 3.1: VA Rationale Example on a Given Scenario

Premium ($)	100,000	(P)
Payout rate (%)	4.72	(δ_0)
Expenses on C_{t-} (%)	0.65	(f_1)
Expenses on B_{t-} (%)	1.25	(f_2)

	$r_t(\%)$	C_{t-}	C_t	$B_{t-} = B_t$	$c_t = \delta_0 \times B_{t-}$	D_t
Year 0	NA	100,000	93,380	100,000	4,720	
Year 1	20	112,056	106,767	112,056	5,289	
Year 2	9	116,376	110,883	116,376	5,493	
Year 3	-3	107,557	102,064	116,376	5,493	
Year 4	-30	71,445	65,952	116,376	5,493	
.						
.						
.						
Year τ					0	$C_{\tau-}$

This table details the income flows c_t paid by the VA on a given scenario. In the example above, we assume that the individual does not withdraw more than e_t and as such we have $B_t = B_{t-}$.

3.3.2 Pricing Methodologies for Annuity Products

All simulations conducted in this chapter are based on actual quotes for annuity products obtained from CANNEX, a data provider which compiles information and calculations about a variety of financial products, including annuity products, and makes that information available to financial service providers through a central exchange. More specifically, the data that we use in what follows is shown in Tables 3.2 and 3.3, which respectively contain for male and female individual aged between 50 and 80 the initial payout rate of the single premium immediate annuity with a 2% cost-of-living-adjustment indexation (SPIA-COLA in short) and the initial payout rate of the variable annuity, which are the two annuity products we focus on in the numerical results that follow. These payout rates will be used as inputs later in Section 3.5.

Table 3.2: Payout rate for the SPIA-COLA

Age	**Male**	**Female**
50	2.78	2.70
51	2.84	2.75
52	2.90	2.81
53	2.96	2.87
54	3.02	2.92
55	3.08	2.98
56	3.15	3.04
57	3.22	3.10
58	3.27	3.17
59	3.40	3.28
60	3.49	3.33
61	3.60	3.42
62	3.72	3.52
63	3.83	3.62
64	3.91	3.69
65	4.02	3.79
66	4.14	3.90
67	4.26	4.01
68	4.39	4.12
69	4.53	4.24
70	4.61	4.37
71	4.75	4.48
72	4.91	4.63
73	5.07	4.78
74	5.16	4.87
75	5.32	5.02
76	5.49	5.17
77	5.66	5.33
78	5.85	5.49
79	6.00	5.66
80	6.09	5.77

This table gives the initial payout rate (in percentage) of a single premium immediate annuity indexed by a 2% COLA (SPIA-COLA) as a function of the individual sex and age.

Table 3.3: Payout rate for the VA

Age	Male	Female
50	2.90	2.90
51	2.90	2.90
52	2.90	2.90
53	2.90	2.90
54	2.90	2.90
55	3.30	3.30
56	3.30	3.30
57	3.30	3.30
58	3.30	3.30
59	3.30	3.30
60	3.93	3.93
61	3.93	3.93
62	3.93	3.93
63	3.93	3.93
64	3.93	3.93
65	4.72	4.72
66	4.97	4.97
67	4.97	4.97
68	4.97	4.97
69	4.97	4.97
70	5.00	5.00
71	5.00	5.00
72	5.00	5.00
73	5.00	5.00
74	5.00	5.00
75	5.33	5.33
76	5.33	5.33
77	5.33	5.33
78	5.33	5.33
79	5.33	5.33
80	5.33	5.33

This table gives the initial payout rate (in percentage) of a variable annuity (VA) as a function of the individual sex and age.

Model prices can be used in place of actual market quotes for annuities. Their advantage is apparent when one needs to estimate the contract value after the contract signing date. In what follows, we only need initial contract values, which is the reason why we have favored market quotes, which are by construction the most realistic estimates of actual contract terms.[6] While we do not calculate model prices, we provide in this section a broad discussion of three competing approaches used to price annuity products. The first approach is based on the no-arbitrage pricing assumption that is used in finance to price derivative products. The second assumes that the insurance company calculates the premium of its contracts so as to hedge its liabilities: the premium depends on the chosen criterion. The third approach is the one that is most commonly used, and it consists in the calculation of the "actuarially fair" value for a contract.

3.3.2.1 No-Arbitrage Pricing

Financial theory uses the assumption of no arbitrage as the key argument to calculate the prices of new securities. Informally, prices must be sufficiently consistent with each other so that no trader can make money out of nothing by selling an "expensive" asset and buying a "cheap" one. As shown by Duffie (2001) (Proposition 6K), the absence of arbitrage (plus technical conditions in continuous-time markets) implies that there exists an equivalent martingale measure (EMM), or risk-neutral probability measure, under which discounted prices follow martingales. Thus, the price of a security is the expected discounted value of future cash flows from holding it and eventually selling it. When the market is complete, all conceivable payoffs can be synthesized with existing assets, so they have a uniquely defined price, and the EMM is unique. The pricing problem then reduces to a technical problem, which is to calculate the risk-neutral expectation of a discounted, possibly complex, payoff. The option pricing model of Black and Scholes (1973) and Merton (1973b) is a well known application of the no arbitrage principle: when the underlying asset price follows a random walk with time-invariant volatility of returns and interest rates are constant, the payoff of a European call or put written on this asset can be perfectly replicated by a dynamic trading strategy involving this asset and a money market account, so the option price is the capital to invest in the replicating strategy to arrive at an exactly zero net wealth at expiration.

No-arbitrage pricing raises additional difficulties when the market is incomplete, that is when some sources of risk are not hedgeable with available securities. The price of each of these risks, i.e. the additional expected return that is gained for being exposed to this risk, is then undefined, and in fact, an infinite number of values for this price are compatible with the absence of arbitrage opportunities. In other words, in an incomplete market, the assumption of no arbitrage is insufficient to fix a price for unhedgeable risks. This situation arises when one wants to price a lifetime annuity for a given individual in the presence of longevity risk, because no securities are traded with a payoff contingent to an individual's life or death. To do no-arbitrage pricing in this case, one has to arbitrarily fix a price for the individual's longevity risk, or equivalently choose an EMM among the infinite set of possible EMMs, keeping in mind that other choices would lead to a different price.

Formally, consider the problem of pricing at time 0 a lifetime annuity that pays a fixed benefit ℓ at dates 1, 2, 3, ... as long as the individual is alive, and denote with τ the uncertain death time. It is assumed that there exists a maximum longevity. Let $\mathbb{1}_{\{\tau>t\}}$ denote the indicator function that is 1 if the individual is alive at time t, and 0 otherwise. Thus, the actual cash flow of the annuity at time $t = 1, 2, 3, \ldots$ is $\ell\mathbb{1}_{\{\tau>t\}}$. Let also B_t denote the value at time t of a cash account in which \$1 invested at time 0 is continuously rolled over at the short-term rate of interest. For

[6] Of course one can reconcile model prices with observed quotes by introducing a "loading factor" defined as the relative difference between the market price and the actuarial price for identical benefits. This definition is the same as those of Horneff et al. (2008) and Pang and Warshawsky (2010), who respectively assume values of 0% and 15% for this factor. Similarly, Pashchenko (2013) takes the "administrative load" to be 10%. In practice, one can adjust the loading factor so that the model price coincides with the market quote.

a given EMM $\mathbb{Q}$, the no-arbitrage price of the annuity at time 0 is the risk-neutral expectation of future discounted cash flows, i.e.,

$$P_0^{\text{no arb.}}(\mathbb{Q}) = \mathbb{E}^{\mathbb{Q}}\left[\sum_{t=1}^{\infty} \frac{1}{B_t} \ell \mathbb{1}_{\{\tau > t\}}\right]. \tag{3.1}$$

The standard assumption that the individual's longevity is independent (in the probabilistic sense) from all financial sources of risk, hence from the cash account value, allows the price to be rewritten as

$$P_0^{\text{no arb.}}(\mathbb{Q}) = \sum_{t=1}^{\infty} \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{B_t}\right] \ell \mathbb{Q}(\tau > t). \tag{3.2}$$

The expectation of $1/B_t$ is the price of a pure discount bond that pays \$1 at time t regardless of the individual's condition. Assuming that such bonds are either traded or replicable by stripping coupon-paying bonds, they have a uniquely defined no-arbitrage price, so their risk-neutral expectation is independent from the chosen measure $\mathbb{Q}$, and the only source of dependency of the annuity price with respect to $\mathbb{Q}$ is the vector of the survival probabilities, $\mathbb{Q}(\tau > 1)$, $\mathbb{Q}(\tau > 2)$, ... Denote with $b_{0,t}$ the price at time 0 of the pure discount bond that matures at time t.

In the absence of any information on how the market would price exposure to longevity risk, it is tempting to replace risk-neutral survival probabilities with "physical" probabilities representing the anticipations of survival chances in the real world. Then, the price becomes

$$P_0^{\text{no arb.}}(\mathbb{Q}) = \sum_{t=1}^{\infty} b_{0,t} \ell \mathbb{P}(\tau > t). \tag{3.3}$$

Note that the survival probabilities are for a specific individual, so they should in principle reflect all of the information available on the individual's characteristics that is relevant to assess chances of survival. This includes the age, the gender (women tend to live longer than men) and the health status, but also assets and income (wealthier individuals tend to live longer) as well as lifestyle (smoking, risky activities, etc.) Put differently, Equation (3.3) calls for the use of individual-specific survival probabilities as opposed to generic probabilities derived from life tables, except when we have no other information about the individual than the characteristics on which tables are based, which are typically the age and the gender.

The presence of time-varying and possibly stochastic benefits, like those of variable annuities, does not substantially alter the framework, but it brings in technicalities. The main difference with respect to constant benefits is that benefits cannot be taken out of the risk-neutral expectations, so Equation (3.2) is replaced by

$$P_0(\mathbb{Q}) = \sum_{t=1}^{\infty} \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{B_t} \ell_t\right] \mathbb{Q}(\tau > t).$$

The separation between the expectation of ℓ_t/B_t and the survival probability is made possible by the assumption of independence between financial risks and individual longevity risk.[7] The problem of estimating risk-neutral survival probabilities is bypassed by replacing them with physical probabilities, so we are back to a purely financial problem, which is to price the payoff ℓ_t. Assuming that ℓ_t is a replicable payoff, its price is uniquely defined, so the choice of the EMM is irrelevant. For instance, if the annuity offers inflation-linked payouts, ℓ_t is proportional to realized inflation, so its no-arbitrage price is the price of an inflation-linked pure discount bond.

As a conclusion, the no-arbitrage framework that is traditionally used in finance applies to annuity pricing provided benefits are independent from individual longevity and one fixes the values of risk-neutral survival probabilities, the default choice being the physical probabilities.

[7]This assumption excludes products in which the benefits depend on the mortality experience of the insurer.

3.3.2.2 Pricing by Hedging Arguments

No-arbitrage pricing is rooted in the idea that securities can be purchased and sold, so that price inconsistencies resulting in arbitrage opportunities cannot subsist. One may argue that this framework does not apply well to the "annuity market", which is mostly a primary market where insurance companies write contracts for individuals who purchase them and cannot exit them, at least as far as lifetime annuities are concerned. With this one-way organization, no entity can be both a seller and a buyer of annuities, so arbitrage opportunities would be impossible to exploit, unlike in the stock, bond and option markets. There does exist a secondary market for annuities, but it is of very modest size. Moreover, longevity risk for a given individual is not hedgeable with financial securities, so any attempt to price a contract by arbitrage arguments involves the choice of an arbitrary pricing measure to fix risk-neutral survival probabilities.

If no substitute for annuities is available in financial markets, one may question the premise that contracts are priced with reference to financial securities so as to eliminate arbitrage opportunities. The existence of multiple quotes offered by different providers for the same profile defined by an age and a gender supports the view that pricing assumptions differ across insurers. Differences in ratings can partially account for these discrepancies, since in an equilibrium pricing model with rational agents, companies with higher default risk must promise higher benefits per dollar of premium, but it seems difficult to dismiss the idea that pricing practices can vary across companies. Annuities are sold in a competitive market, so prices cannot be disconnected from the economic and demographic backdrop against which all providers operate, but insurers have often multiple business lines, sometimes with the possibility of compensating gains and losses across them, and they may follow different marketing strategies, leading to more or less aggressive pricing policies in their annuity business. Moreover, insurers are required by the law to maintain actuarial reserves to meet their liabilities, and the current level of reserves may have an impact on the pricing policy too.

A quantitative pricing model that incorporates such features is well beyond the scope of this chapter, but we now give two simple examples to show how different but economically motivated pricing assumptions lead to different prices for a given contract. The starting point is not the absence of arbitrage opportunities, but a hedging criterion for the insurer's liabilities. We consider an insurer who does risk pooling by selling many contracts and is committed to make a series of payments at dates 1, 2, ..., T, with values ℓ_1, ℓ_2, ..., ℓ_T. These payments represent the aggregate amounts of benefits to be paid to beneficiaries, so for life-contingent benefits, ℓ_t is the sum of benefits paid to those alive at date t. Technically, the payments must be nonnegative almost surely and positive with non-zero probability.

The insurer's assets initially consist of the total premia paid by contract buyers, and they subsequently decrease as payments are made. The value of net assets at time t just after a payment is denoted with A_t, and the value just before the payment is denoted with A_{t-}, so that on a payment date, the net asset value jumps downwards as

$$A_t = A_{t-} - \ell_t.$$

We are interested in finding the minimum capital to start with at date 0, A_0, in order to satisfy one of the following two hedging criteria:

- Hedging on average: $\mathbb{E}[A_T] \geq 0$;
- Hedging almost surely: $A_T \geq 0$ with 100% probability.

The second criterion is clearly stronger than the first, which allows for negative net asset values, so it is expected to lead to a higher capital requirement.

It is shown in Appendix 3.A that under the assumptions that the payments ℓ_t are independent from the returns on the funds in which assets are invested, the minimum capital to invest to hedge

on average, or the aggregate premium to be charged to contract buyers, is

$$A_0^{\text{hed. av.}} = \sum_{t=1}^{T} \mathbb{E}[\ell_t] \frac{[1+\rho_{t,T}]^{T-t}}{[1+\rho_{0,T}]^T}, \tag{3.4}$$

where $\rho_{t,T}$ is the annualized expected return between dates t and T. This formula provides some justification for discounting future liability payments with the expected return on assets.

To derive the minimum capital requirement for almost sure hedging, we assume that each payment is the payoff of a pure discount bond that matures at the payment date, and that there are no arbitrage opportunities. We do not claim that these assumptions are realistic: the former is violated in many cases, e.g. when payments depend on the number of living individuals in a population of annuitants, and as discussed previously, the latter may not hold in insurance contract markets. However, our objective here is not to derive an operational pricing formula based on a realistic model, but to illustrate how sensitive the calculation of a minimum capital requirement is to the choice of a hedging criterion. We show in Appendix 3.A that the minimum capital to invest here is

$$A_0^{\text{hed. a.s.}} = \sum_{t=1}^{T} c_{0,t},$$

where $c_{0,t}$ is the price at time 0 of the pure discount bond that replicates the payment of date t. When payments are non-stochastic, the pure discount bonds are nominal zero-coupon bonds, so we can rewrite the minimum capital in the more familiar form

$$A_0^{\text{hed. a.s.}} = \sum_{t=1}^{T} l_t b_{0,t}, \tag{3.5}$$

that is the sum of future payments discounted with nominal zero-coupon rates.

An important difference between Equations (3.4) and (3.5) is that the latter makes no reference to the characteristics of the portfolio in which assets are invested. In a low interest rate environment, it is costly to hedge almost surely, and hedging on average leads to a lower capital requirement because the expected return on the asset portfolio contains a risk premium. An insurer who would adopt the on-average hedging criterion could offer lower prices for the same guarantees than another who requires almost sure hedging, but it would have to make up for losses, i.e. negative net asset values, with reserves or profits from other business lines.

3.3.2.3 Actuarial Pricing

The actuarially fair price, or actuarial price in shorter form, of an annuity is defined as the sum of expected future discounted benefits, weighted by the probability of receiving these benefits. The contract is said to be actuarially fair if the premium paid equals this price. Mathematically, assume that the benefits are paid at dates 1, 2, 3, ..., let ℓ_t denote the benefit paid at date t, which can be life-contingent, and let $D_{0,t}$ be the discount factor applied at time 0 to cash flows occurring in t periods. The actuarial price is defined as

$$P_0^{\text{act}} = \sum_{t=1}^{T} \mathbb{E}[\ell_t] D_{0,t} \tag{3.6}$$

Equation (3.6) should be regarded as a definition rather than as a result derived from a pricing model, but it can be noted that it bears some resemblance with the previously introduced prices. Consider for instance a lifetime annuity that pays fixed benefits to an individual as long as she is alive. Then, the expectation of ℓ_t equals ℓ, the promised benefit, times the probability of being alive at time t. Also, take the discount factors to be the prices of pure discount bonds. Then,

the actuarial price coincides with the no-arbitrage price given by Equation (3.3). More generally, for time-varying benefits, the actuarial price equals the no-arbitrage price when the lifetime is assumed to be independent from discounted benefits, and the risk-neutral survival probabilities in the latter price are equated with physical probabilities.

When the insurer promises a fixed benefit, say ℓ, to each survivor, the expected value of ℓ_t equals the promised benefit times the expected number of survivors, so the premium to be charged to each annuitant at time 0 to hedge liabilities on average under the forward contract strategy is

$$P_0^{\text{hed. av.}} = \sum_{t=1}^{T} \ell p_{0,t} b_{0,t}$$

where $p_{0,t}$ is the survival probability at horizon t estimated at time 0.

Discounting and Mortality Assumptions Equation (3.6) solves the conceptual issues raised by the choice of a pricing measure in incomplete markets and the choice of a hedging criterion in pricing by hedging, but it still leaves room for the specification of some parameters. The first is the probability measure under which expected benefits are calculated. To make the discussion more concrete, consider the case of life-contingent benefits $\ell_t = \ell \mathbb{1}_{\{\tau > t\}}$, where ℓ is fixed and τ is the death time of an individual. Then, the expectation of ℓ_t is ℓ times the probability of living until date t. Survival probabilities are generally inferred from mortality tables, but there are many of such tables, relating to different populations (e.g., the general population versus the population of annuitants, who tend to live longer), and they are periodically revised. Different insurers may choose different tables, which lead to different prices for the same annuity. Moreover, as we explain in Section 3.4.3 below, survival probabilities may be adjusted for expected improvements in longevity, but not all insurers will choose to perform this adjustment.

The second set of parameters that is not unambiguously specified is the discount factors, or equivalently the discount rates defined as

$$y_{0,t} = -\frac{1}{T-t} D_{0,t}$$

Given the long maturity of annuity payouts, some of which are to be paid in more than ten years, even small changes in discount rates can have a sizable impact on the actuarial price. The choice most consistent with financial theory and risk-neutral valuation principles is to use the default-free zero-coupon rates, but since the actuarial value formula is not derived from this theory, it does not formally exclude other choices. As an illustration of this relative indeterminacy, Mitchell et al. (1999) calculate the present value of annuity payouts under various alternative assumptions on discount rates: they use the Treasury bond yield curve or the BAA corporate bond yield curve. As their results show (see their Table 3), the choice of Treasury versus corporate bond rates has a material impact, with the latter rates implying a decrease in the present value of the annuity payouts per dollar of annuity premium. The biggest decrease is obtained for the youngest individuals, since they have the longest payout maturities. Later in this section, we explain how discount rates can be adjusted to capture the difference between actuarial prices and observed prices for annuities.

Equation (3.6) is the basis of the pricing exercises that we conduct in this chapter. Specifically, we calculate benefits for a given premium by equating the actuarial value with the premium, a procedure that is standard in the literature.

Actuarial Prices versus Observed Prices Even under reasonable assumptions for discount rates and survival probabilities, actuarial annuity prices for a given individual profile may differ from quotes offered by insurers for a variety of reasons:

- Market prices include fees and administrative costs that are not factored in the actuarial price;

- Insurers use different assumptions for discount rates and mortality;
- They deliberately deviate from actuarial prices for marketing reasons;
- They use a different pricing method than actuarial valuation.

There are at least two methods to assess the difference between the actuarial price and a market quote.

Loading Factor The first method consists in calculating a "loading factor". We propose two approaches to define the loading factor λ.

1. In the first approach, we consider an annuity that pays benefits and for which P_0^{mkt} and P_0^{act} denote the market and the actuarial prices of these benefits.

$$\lambda = \frac{P_0^{\text{mkt}}}{P_0^{\text{act}}} - 1,$$

2. In the second approach, we assume constant benefits. Then the actuarial price is clearly proportional to the magnitude of these benefits, so as long as the market price satisfies the same condition, the loading factor can also be obtained by comparing the benefits implied by the actuarial valuation and the benefits offered by the insurer for the same premium:

$$\lambda = \frac{\ell^{\text{act}}}{\ell^{\text{mkt}}} - 1.$$

The loading factor is positive if real-world annuities are less generous than a contract priced at fair value (i.e. if real-world annuities pay a constant benefit ℓ^{mkt} lower than the constant benefit of the contract priced at faire value ℓ^{act}), and negative otherwise. We expect to be in the former situation if differences between actuarial and market prices only reflect costs and fees that are not captured in the former prices, but if differences also arise from the use of alternative assumptions, the loading factor can have either sign.

Spread The second method is to calculate the (positive or negative) spread that needs to be added to each discount rate in the actuarial value to match the market price. Formally, the spread s is defined by the condition that

$$P_0^{\text{mkt}} = \sum_{t=1}^{T} \mathbb{E}[l_t] D_{0,t} e^{-ts}.$$

Hence, the relation between the loading factor and the spread can be written as

$$\lambda = \frac{\sum_{t=1}^{T} \mathbb{E}[\ell_t] D_{0,t} e^{-ts}}{\sum_{t=1}^{T} \mathbb{E}[\ell_t] D_{0,t}} - 1.$$

This equation shows that λ and s have opposite signs.

Actuarial Pricing for Complex Products Pricing annuities essentially means to calculate the benefits of the contract for a given premium, or the premium to be charged for a given set of benefits. This is relatively straightforward for lifetime annuity, when one can calculate an "actuarial value" by discounting future benefits and weighting them by the probability that they are paid. Thus, a life benefit is weighted by the probability of being alive, while a death benefit is multiplied by the death probability. The situation is more complex for variable annuities since the benefits are uncertain. Gaillardetz and Lakhmiri (2011) develops a fair valuation approach to evaluate at date t an equity-linked product of maturity n that pays a payoff $D(s)$ at date $s \in [t, n]$ if the individual is alive. He writes the fair value of this equity-linked product as a weighted average of the financial contingent claim prices of various maturities, the weights being determined by the probability that the policyholder dies instantaneously or survives to n:

$$FV(x,t,n) = \int_0^{n-t} \Pi(t,t+s) f(t+s|T(x) \geq t) ds + \Pi(t,n)_{n-t} p_{x+t}.$$

Here, $T(x)$ is a random variable representing the future lifetime of an individual x at time $t = 0$ with f being probability density function of $T(x)$, t is the time of evaluation and n is the maturity date. $\Pi(t,n) = \mathbb{E}^{\mathbb{Q}}[D(n)\exp(-r(n-t))|F(t)]$ represents the no-arbitrage price at time t of a contingent claim with payoff $D(n)$ payable at time $n \geq t$.

From this fair value, we can price more complex annuity products such as fixed indexed annuities and variable annuities. Bacinello et al. (2009, 2010) and Bacinello et al. (2011) also propose a pricing setup for variable annuities. Bacinello et al. (2009) define a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ where the residual lifetime τ is a $\mathcal{F}$-stopping time and use a least-square Monte-Carlo (LSMC in short) approach to value an equity-linked endowment contract with a surrender option. The contract price with the surrender option is given by an optimal stopping problem and its resolution via the LSMC requires discretization in the time dimension and simulation of all random processes: the term structure of interest rates, the market dynamic and the mortality intensity. Bacinello et al. (2011) adopt a unifying approach to value variable annuities with different GMxB guarantees[8]. In the same vein, Bauer et al. (2008) and Bauer et al. (2010) introduce a framework to price variable annuity contracts with guaranteed minimum benefits and finite maturity, assuming that the policyholder follows a given strategy with respect to surrender and withdrawals.

3.4 Simulation Framework for Retirement Investing Decisions

In this section, we present the simulation framework that will be used to analyze retirement investing decisions for individuals in decumulation.

3.4.1 Modeling Individuals and Individual Decisions

The decision making framework has inputs and outputs, and explains how the former are processed to obtain the latter. A key category of inputs consists of the individual's characteristics, namely the age and the gender, which have an impact on the expected time in retirement and also the current wealth. The next key individual input is the targeted consumption level, which can be either constant (up possibly to a cost-of-living adjustment) or depend on the occurrence of a "life event"and as such is stochastic, as discussed in section 3.4.3.3.

The current date is referred to as date 0. The individual retires at date 0 and cannot live past date $T_{\max}$.

An individual is described by the following characteristics:

[8]The capital letters G, M and B respectively stand for Guaranteed, Minimum and Benefit. The x letter can be replaced by A for accumulation, D, for death, I for income or W for withdrawal

- The age at date 0, denoted with x and expressed in years. In the proposed framework, x can be any integer value between 50 and 80;
- The sex, male or female. This parameter will implicitly determine the mortality probability distribution used in the Monte Carlo framework;
- A stochastic process $(L_t)_{0\leq t\leq T_{\max}}$, which describes the target withdrawals to be made at time t. These withdrawals include an expected component, which is typically modeled as a constant or cost-of-living-adjusted fraction of initial wealth. They also have an unexpected component, which is contingent upon the occurrence of a life event. These events are described in section 3.4.3.3;
- An allocation to financial products, which include balanced funds and target date funds (see Section 3.4.2). This allocation is summarized in a column vector with m elements, where m is the number of products in which the individual invests, and each element is the dollar allocation to a product. m is zero if the individual owns no financial product. The vector is denoted with $\boldsymbol{\theta}_{0-}$ in what follows, and the subscript $0-$ emphasizes the fact that the dollar amounts are measured before any withdrawal or contribution that would be made at time 0;
- An allocation to various annuity products, namely SPIA, SPIA-COLA, DIA, DIA-COLA and VA. It is summarized in a column vector of size n denoted with $\mathbf{P}$, in which each element is the premium paid for a contract at time 0. If the individual purchases no annuity, then n is zero and the vector $\mathbf{P}$ is empty.

3.4.1.1 Investment Strategies

It is important to note that the allocation to financial products and the allocation to annuities – the vectors $\boldsymbol{\theta}_{0-}$ and $\mathbf{P}$ with the above notation – do not necessarily represent the *actual* allocations of an *actual* client. Instead, they represent *hypothetical* initial conditions for an *hypothetical* individual who has the same age and gender as an actual client, and simulations of future replacement income and bequest are performed conditional on this starting point. The client's current allocation is a natural default choice for these conditions, but precisely because they are hypothetical, the framework allows them to be modified, so that the user (e.g., a financial adviser) can see their impact on the outcome. If a "good" or "optimal" (in a sense to be specified in Section 3.5) allocation is found, then it can be taken as a recommendation to the client, who ultimately decides whether or not to switch from their current allocation to the suggested one.

Implicit in the definition of an individual's profile is that she follows buy-and-hold strategies both in financial products and in annuities. All these products and contracts are purchased at time 0, and the individual can subsequently withdraw money from them following the rules described in Section 3.4.1.3 below, but she does not rebalance between them. This assumption does not mean that the financial products would themselves be of the buy-and-hold type, and they can be dynamically rebalanced funds, as explained in Section 3.4.2.

3.4.1.2 Target Withdrawal Strategy

Like the allocation to financial products and annuities, the withdrawals as represented by the stochastic process (L_t) are an assumption taken as an input to the simulations, but even if the client's wishes represent a natural reference point, they are meant to be varied. Indeed, the client may be interested in seeing the impact of the magnitude of withdrawals on the probability of being able to sustain them for lifetime, and on the distribution of the bequest. So, the hypothetical individual that represents the client in the framework may have different withdrawal objectives.

The target withdrawals relate to two of the objectives of an individual in retirement, namely to generate replacement income and to face life events that generate new expenses. The expected

and the unexpected parts of withdrawals correspond to these two motives. The third objective of an individual is to leave some bequest to heirs, but we do not use it as an input to the framework because the bequest is any unspent money or any death benefit of annuities at the time of death: as such, it is an output, rather than an input, of the simulations.

We decompose the target withdrawal for a given date t as

$$L_t = L_t^{\text{exp}} + L_t^{\text{unexp}},$$

where L_t^{exp} is the withdrawal that the individual thinks she needs to make to meet foreseeable expenses and L_t^{unexp} is the withdrawal that would be needed in case a life event occurs.

The expected target withdrawal L_t^{exp} can be:

- Constant if the target withdrawal level is a fixed nominal amount for all dates and states of the world: $L_t^{\text{exp}} = L_0 =$ constant for all $t \geq 0$;
- Deterministic if the target withdrawal level contains a cost of living adjustment: $L_t^{\text{exp}} = L_0 \times \frac{I_t}{I_0}$ where I_t is a non-stochastic quantity that grows by a fixed rate every year;
- Stochastic if the target withdrawal level is indexed with respect to realized inflation: $L_t^{\text{exp}} = L_0 \times \frac{I_t}{I_0}$ where I_t is a random quantity that grows by a stochastic rate every year.

Given the number of outstanding questions related to the simulation (and implementation) of a state-dependent withdrawal strategy, we shall focus in our base case simulations on state-independent withdrawal strategies with a COLA adjustment component. We do not want to simulate L_t as a function of wealth or allocation. The target withdrawal represents a recurrent or exceptional expenditure objective (life event), and not, as in Merton (1971), an optimal level of consumption at each moment. Moreover in our framework, the actual withdrawal may differ from the target withdrawal. While the analysis of whether the joint identification of improved rule-based investment strategies and rule-based withdrawal strategies could dominate (in some sense to be defined) simpler state-independent strategies is a possibly relevant topic, this question is somewhat beyond the scope of the core focus of this chapter, and is left for further research.

Note that in all cases the unexpected target withdrawal L_t^{unexp} is stochastic since it is based on the occurrence of a life event with uncertainty with respect to the timing and severity of the event in terms of replacement income - see Section 3.4.3.3. Note also that even in the absence of a life event and in the simple case of a constant target withdrawal strategy, the actual withdrawal strategy involves a random component related to the possibility that the individual is not able to make the target withdrawal at any given point in time and for any scenario because the outstanding wealth level may be insufficient after liquidation of all of the assets in portfolio. The specification of such liquidation rules are precisely the focus of the next section.

3.4.1.3 Actual Withdrawal Strategy

We denote:

- C_{t-}^i: the contract value at time t before withdrawal, fees and penalties of annuity i (if the annuity is totally illiquid, as a SPIA-COLA for example, then $C_{t-}^i = 0$),
- e_t^i: the contractual withdrawal at time t of annuity i,
- W_{t-}: the liquid wealth of the individual at time t before withdrawal,
- surr_t^i: the surrender charges at time t of annuity i (if the annuity is totally illiquid, as a SPIA-COLA for example, then $\text{surr}_t^i = 0$),
- L_t: target withdrawal for a given date t.

To achieve a target withdrawal, the individual uses in priority the benefits that annuities pay without penalties or surrender charges (step 1), but if the sum of these benefits across all contracts is less than the target, she must liquidate part of the financial assets and/or make additional withdrawals from the contracts. We use a mixture of both options, by applying the following liquidation rules:

1. First liquidate the financial assets bucket (step 2);
2. If the financial bucket plus the income provided by the annuities with lifetime income is not sufficient to meet the target expense, then annuities with lifetime income are partially or fully liquidated (step 3).

The aforementioned three steps can be written as follows:

1. Step 1: if $\sum_{i=1}^{n} e_t^i \geq L_t$ then the replacement income need is satisfied;
2. Step 2: if the replacement income need is not satisfied after step 1, the policyholder withdraws $c_t^{\text{liq}} = \min\left(\left[L_t - \sum_{i=1}^{n} e_t^i\right]^+, W_{t-}\right)$ from the financial asset bucket. If $L_t - \sum_{i=1}^{n} e_t^i \leq W_{t-}$ then the replacement income need is satisfied;
3. Step 3: if the replacement income need is still not satisfied after step 2, then the policyholder has to totally use up her financial asset bucket ($c_t^{\text{liq}} = W_{t-}$) and has to partially/totally liquidate the variable annuity (if she has one in her portfolio among the n annuities) and recovers an amount from it equal to $L_t - \sum_{i=1}^{n} e_t^i - W_{t-}$. If we assume that annuity n is a variable annuity and that the other annuities are fully illiquid then $C_{t-}^n - \text{surr}_t$ represents the maximum amount the policyholder will receive at time t in addition to the contractual withdrawal e_t^n if she fully liquidates her variable annuity. If $C_{t-}^n - \text{surr}_t^n > L_t - \sum_{i=1}^{n} e_t^i - W_{t-}$ then the replacement income need is satisfied.
4. If the replacement income need is not satisfied after the three steps then the target expense at date t is not met.

We choose these withdrawal rules because liquidating financial assets generates few (or no, in our stylized framework) fees, while liquidating annuities gives rise to various fees and penalties, hence the desire to minimize the amount taken from the contracts. To summarize, the proposed algorithm minimizes, at each date, the amount of fees, assuming that these are an increasing function of the amount liquidated.

3.4.2 Modeling Financial Products

In addition to the annuity-type products mentioned in Section 3.3, we assume that the individual has also access to a set of standard investment products, including most importantly balanced funds and target date funds.

Generally speaking, these investment opportunities are defined as possibly time-varying mixtures of equity and fixed-income components, which are typically accessible in a mutual fund format. These mutual funds are treated as liquid investment vehicles, which can be purchased and sold at any time, and have a well-defined market price, or net asset value for mutual funds. This makes it possible to evaluate an individual's positions on a mark-to-market basis at any point in time.

The main underlying ingredients and related notation for these investment products are listed below.

- $(S_t)_{0\leq t\leq T_{\max}}$ is the random discrete time series representing the total return price series of a stock index. We set $S_0 = 1$;
- $(B_t)_{0\leq t\leq T_{\max}}$ is the random discrete time series representing the total return price series of a bond index. We set $B_0 = 1$;
- $(R_t)_{0\leq t\leq T_{\max}}$ is the random discrete time series representing the total return price series of cash. We set $R_0 = 1$;
- $(w^{\text{BF}}) = (w^S, w^B, w^R)$ is the (constant) weight vector that corresponds to the proportion of stock, bond and cash invested in the balanced fund (in the numerical analysis, we consider a balanced fund invested 50% in stocks and 50% in bonds - see Section 3.5 for more details);
- $(w_t^{\text{TDF}})_{0\leq t\leq T_{\max}} = (w_t^S, w_t^B, w_t^R)$ is the (time-varying) weight vector that corresponds to the proportion of stock, bond and cash invested in the target date fund (see Section 3.5 for more details regarding the glide path).
- Θ^{liq} represents the vector of the initial dollar allocation between the m financial products.
- $\boldsymbol{\theta}_{t-}$ is the dollar amount $(m \times 1)$ vector invested in each financial product at time t before withdrawal;
- $\boldsymbol{\theta}_t$ is the dollar amount $(m \times 1)$ vector invested in each financial product at time t after withdrawal;
- W_{t-} is the liquid wealth of the individual at time t, before withdrawal;
- W_t is the liquid wealth of the individual at time t, after withdrawal;
- $\boldsymbol{\Upsilon}_t = (\Upsilon_t^i)_{1\leq i\leq m}$ is a $(m \times 1)$ vector that represents the total return price series of each financial product at time t.

Using this notation, the amount withdrawn by the individual from the financial assets $(c_t^{\text{liq}})_{0\leq t\leq T_{\max}}$ at time t is

$$c_t^{\text{liq}} = \min\left(\left[L_t - \sum_{i=0}^{n} e_t^i\right]^+, W_{t-}\right)$$

The liquid wealth of the individual at time t, after withdrawal is

$$W_t = W_{t-} - c_t^{\text{liq}}$$

The dollar amount vector that corresponds to the dollar amount invested in each financial product at time t after withdrawal[9] is

$$\begin{aligned}\boldsymbol{\theta}_t &= \left[1 - \frac{c_t^{\text{liq}}}{(\boldsymbol{\theta}_{t-})^\top \mathbf{1}_m}\right]\boldsymbol{\theta}_{t-} \\ &= \left[1 - \frac{c_t^{\text{liq}}}{W_{t-}}\right]\boldsymbol{\theta}_{t-}\end{aligned}$$

The dollar amount vector that corresponds to the dollar amount invested in each financial product at time $t+1$ before withdrawal is

[9]Weights remain unchanged after withdrawal.

$$\begin{aligned}\boldsymbol{\theta}_{[t+1]-} &= \boldsymbol{\theta}_t \odot (\boldsymbol{\Upsilon}_{t+1}./\boldsymbol{\Upsilon}_t) \\ &= \left[1 - \frac{c^{\text{liq}}_{t+1}}{W^{\text{liq}}_{[t+1]-}}\right] \boldsymbol{\theta}_{t-} \odot (\boldsymbol{\Upsilon}_{t+1}./\boldsymbol{\Upsilon}_t)\end{aligned} \tag{3.7}$$

The liquid wealth at time $t+1$ before withdrawal is

$$\begin{aligned}W_{[t+1]-} &= (\boldsymbol{\theta}_{[t+1]-})^\top \mathbf{1}_m \\ &= [\boldsymbol{\theta}_t \odot [\boldsymbol{\Upsilon}_{t+1}./\boldsymbol{\Upsilon}_t]]^\top \mathbf{1}_m \\ &= [\boldsymbol{\theta}_t]^\top [\boldsymbol{\Upsilon}_{t+1}./\boldsymbol{\Upsilon}_t]\end{aligned} \tag{3.8}$$

The last equality in Equation (3.8) comes from $[\boldsymbol{u} \odot \boldsymbol{v}]^\top \mathbf{1}_m = \boldsymbol{u}^\top \boldsymbol{v}$ where u and v are two $(m \times 1)$ vectors. Here $\odot$ refers to the Hadamard product, which is the term-by-term product of two matrices. $\mathbf{1}_m$ is a $(m \times 1)$ vector filled with ones. ./ denotes the term-by-term division.

We substitute Equation (3.8) into Equation (3.7):

$$\boldsymbol{\theta}_{t+1} = \left[1 - \frac{c^{\text{liq}}_{t+1}}{(\boldsymbol{\theta}_t)^\top (\boldsymbol{\Upsilon}_{t+1}./\boldsymbol{\Upsilon}_t)}\right] \boldsymbol{\theta}_t \odot [\boldsymbol{\Upsilon}_{t+1}./\boldsymbol{\Upsilon}_t]$$

Balanced funds are represented by a strategy involving the equity index, the bond index and cash, which is assumed to be rebalanced at an annual frequency towards constant weights w^S, w^B and w^R. The total return price series of the balanced fund is calculated as follows:

$$\forall t \in [\![0; T_{\max} - 1]\!], \frac{\Upsilon^i_{t+1}}{\Upsilon^i_t} = \left(w^S \times \frac{S_{t+1}}{S_t} + w^B \times \frac{B_{t+1}}{B_t} + w^R \times \frac{R_{t+1}}{R_t}\right)$$

Target date funds are represented by a strategy involving the equity index, the bond index and cash, which is assumed to be rebalanced at an annual frequency towards time-varying weights w^S_t, w^B_t and w^R_t. Thus, the target mix of stocks, bonds and cash evolves in time until a date called the target date of the fund, with a deterministic decrease in equity allocation according to a predetermined glide path, which we specify in Section 3.5.

The total return price series of the target date fund is calculated as follows:

$$\forall t \in [\![0; T_{\max - 1}]\!], \frac{\Upsilon^i_{t+1}}{\Upsilon^i_t} = \left(w^S_t \times \frac{S_{t+1}}{S_t} + w^B_t \times \frac{B_{t+1}}{B_t} + w^R_t \times \frac{R_{t+1}}{R_t}\right)$$

3.4.3 Modeling Risk Factors

This section describes the dynamics of the various stochastic processes that enter the Monte-Carlo simulation model, as well as the base case parameter values. While asset management products are often marketed on the basis of their past track records, historical scenarios are not of direct relevance to compare retirement strategies, even though they can be used to perform stress test analysis. The proper evaluation criterion for a retirement investment solution is its ability to secure the needed level of replacement income with the highest confidence level while offering attractive probabilities of reaching higher goals. Monte-Carlo analysis is well suited for this purpose because it can be used to simulate a wide range of possible market environments. All the stochastic processes below are defined under the physical measure $\mathbb{P}$ and under a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, F_t)$ where Ω is a set of outcomes, $\mathcal{F}$ is a sigma-algebra that represents the set of all measurable events, $\mathbb{P}$ is a probability measure and F_t is a filtration.

3.4.3.1 Financial Market Uncertainty

We model the yield curve as a function of a single state variable which is the short term interest rate r_t and as such adopt a one-factor model for the term structure. We assume that the short term interest follows the model of Vasicek (1977), described by the mean-reverting process:

$$dr_t = a(b - r_t)dt + \sigma_r dz_{r_t} \tag{3.9}$$

z_{r_t} is a standard Brownian motion.

There are four parameters to estimate: the speed of mean reversion a, the long-term mean b, the short-term volatility σ_r and the initial value r_0.

It can be shown (see for instance Brigo and Mercurio (2001) page 109) that the short rate r_t verifies:

$$r_t = r_0 \exp(-at) + b(1 - \exp(-at)) + \sigma_r \int_0^t \exp(-a(t-u))\, dz_{r_u} \tag{3.10}$$

We note that r_t is normally distributed with mean

$$\mu_t = \mathbb{E}(r_t) = r_0 \exp(-at) + b(1 - \exp(-at))$$

and variance

$$\sigma_t^2 = \mathbb{V}(r_t) = \frac{\sigma_r^2}{2a}(1 - \exp(-2at))$$

.

We can also prove from Equation(3.10) that:

$$\mathbb{C}orr[r_t, r_{t+1}] = \exp(-a)$$

We define the long-term moments as the limits of the expected value, the volatility and the one-year autocorrelation as time goes to infinity (and therefore memory of initial conditions is lost):

$$\mu_\infty = \lim_{t\to\infty} [\mathbb{E}(r_t)]$$

$$\sigma_\infty = \lim_{t\to\infty} \sqrt{[\mathbb{V}(r_t)]}$$

$$\rho_\infty = \lim_{t\to\infty} \mathbb{C}orr[r_t, r_{t+1}]$$

We then deduce that the long-term moments implied by the Vasicek model are

$$\mu_\infty = b, \quad \sigma_\infty^2 = \frac{\sigma_r^2}{2a}, \text{ and } \quad \rho_\infty = \exp(-a)$$

so that the parameters can be recovered from the long-term moments as $a = -\ln \rho_\infty$, $b = \mu_\infty$ and $\sigma_r = \sigma_\infty \sqrt{2a}$. Moments ρ_∞ , μ_∞ and σ_∞ are estimated as the sample moments over the estimation period as in Chapman et al. (1999). We use historical estimation of the Vasicek model parameters as of June, 1st, 2020 from monthly observations of the secondary market rate on three-month Treasury bills over the 20-year period June, 1st, 2000 - June, 1st, 2020. Data are available from the Fed website at https://www.federalreserve.gov/releases/h15/. All series are sampled at the monthly frequency . Specifically, b is estimated as the mean of the nominal three-month rate over 20 years. a and σ_r are chosen in such a way that the model-implied long-term volatility and one-year autocorrelation match the sample moments of the short-term rate estimated over the same time frame. r_0 is set as the secondary market rate on three-month Treasury bills as of June, 1st, 2020. We obtain: $a = 0.3833$, $b = 0.0153$, $\sigma_r = 0.0147$ and $r_0 = 0.0014$.

Table 3.4: Capital market assumptions and correlation matrix

(A) Capital Market Assumptions

Asset Classes	Arithmetic return (%)	Volatility (%)	Fees (%)	Geometric return after fees (%)
US Equity	9.90	18.18	0.50	8.25
US Fixed Income	3.89	5.17	0.45	3.76
Cash	2.83	1.69	0.18	2.82

(B) Correlation matrix

Correlation	US Equity	US Fixed Income	Cash
US Equity	1	−0.09	0.26
US Fixed Income	−0.09	1	0.19
Cash	0.26	0.19	1

This table reports the capital market assumptions and the correlation matrix used for the Monte Carlo market simulations.

The stock and bond indices that are used as building blocks for both balanced funds and target dates funds are modeled as diffusion processes with a constant volatility and a constant expected excess return over the nominal short-term rate. The stock and the bond indices evolve as:

$$\begin{aligned} \frac{dS_t}{S_t} &= (r_t + \sigma_S \lambda_S)dt + \sigma_s dz_{S_t} \\ \frac{dB_t}{B_t} &= (r_t + \sigma_B \lambda_B)dt + \sigma_B dz_{B_t} \end{aligned} \tag{3.11}$$

where z_S and z_B are Brownian motions correlated to the Brownian motion of the nominal short-term rate z_r.

We set $Z_t = \begin{bmatrix} \log S_t \\ \log B_t \\ r_t \end{bmatrix}$ and write the continuous-time vector autoregression (VAR in short) model[10] as

$$dZ_t = (A + BZ_t)dt + \sigma' dW_t$$

with $A = \begin{bmatrix} \sigma_S \lambda_S - \frac{\sigma_S^2}{2} \\ \sigma_B \lambda_B - \frac{\sigma_B^2}{2} \\ ab \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -a \end{bmatrix}$ and σ such that $\sigma'\sigma = \Sigma = \begin{bmatrix} \sigma_S^2 & \sigma_{SB} & \sigma_{Sr} \\ \sigma_{SB} & \sigma_B^2 & \sigma_{Br} \\ \sigma_{Sr} & \sigma_{Br} & \sigma_r^2 \end{bmatrix}$

where Σ is the covariance matrix of the random vector Z_t. σ is determined by Cholesky decomposition.

We finally obtain the discretized equation of the continuous-time VAR model (see Appendix 3.B for the proof):

$$Z_{t+h} = \Psi_0 + \Psi_1 Z_t + \varepsilon_{t+h}$$

where $\Psi_0 = \left(\int_0^h \exp(sB)ds\right) A$, $\Psi_1 = \exp(hB)$ and $\varepsilon \sim \mathcal{N}\left(\mathbf{0}, \int_0^h \exp(sB)\,\sigma'\sigma\,\exp(sB')\,ds\right)$.

We finally use the following long-term capital market assumptions provided by the Merrill Lynch Chief Investment Office at Bank of America (see Table 3.4). In our notation, they imply that $\sigma_S = 18.2\%$, $\lambda_S = 0.361$, $\sigma_B = 5.2\%$, $\lambda_B = 0.118$, $\rho_{SB} = -0.09$, $\rho_{Sr} = 0.26$ and $\rho_{Br} = 0.19$.

[10] We apply Ito's lemma in Equation (3.11) to obtain the instantaneous log return processes.

3.4.3.2 Longevity Uncertainty

An important source of risk in retirement investing decisions relates to the uncertainty with respect to the date of death. In this chapter, we use death probabilities taken from the 2012 Individual Annuitant Mortality Table of the Society of Actuaries. We select individual annuity tables based on annuity buyers: the 2012 IAM Basic Tables (2585/2586) [11]. We also select mortality factor reduction tables from the 2012 IAM Basic Tables (2583/2584)[12] to take into account mortality improvement over time.

We define t as the current date (which corresponds to the policyholder retirement date) and assume that the policyholder is aged x at date t. Assuming that future mortality probabilities are equal to the ones that prevail at date t, the probability $h(t,x,s)$ that an individual aged x at current date t will still be alive in s years can be computed from the reference mortality table as:

$$h(t,x,s) = \prod_{i=1}^{s-t}[1 - q_{x+i-1}] \tag{3.12}$$

where q_{x+i-1} is the probability for an individual aged $x+i-1$ to die within one year evaluated at date t.

A limitation of the assumption of stationary longevity distributions is that they are not consistent with the possibility of improvements in mortality. An individual aged x at date t who reaches age $x+i$ will do so in year $t+i$, and the probability of her dying within this year is likely to be lower than the probability of a $(x+i)$-old dying within year t. To take into account the decreasing trend in mortality, we apply a projected reduction factor in mortality, in which case Equation (3.12) becomes (see Dickson et al. (2013) p. 68):

$$h(t,x,s) = \prod_{i=1}^{s-t}[1 - q_{x+i-1}(t+i-1)] \tag{3.13}$$

where $q_{x+i-1}(t+i-1)$ is the probability for an individual aged $x+i-1$ at date $t+i-1$ to die within the year. The adjusted probability is given by

$$q_{x+i-1}(t+i-1) = q_{x+i-1}(t) \times [1 - \varphi(x+i-1)]^{i-1}$$

where $\varphi(x+i-1)$ is the mortality reduction factor at age $x+i-1$.

These combinations of a mortality table with a projection scale factor are internally consistent because the documentation on the SoA website indicates that the Projection Scale G2 was developed "in conjunction with" the 2012 IAM tables. The Projection Scale G was released at the same time as the 1983 Individual Annuity Mortality Tables, that is in 1981, and was also developed "in conjunction with" these tables.

We defined the probability $h(t,x,s)$ as a function of the current date t because the death probabilities $q_x(t_0)$ we use from the 2012 Individual Annuitant Mortality Table of the Society of Actuaries have been calculated at date t_0 which is anterior to the current date t. To adjust the death probabilities $q_x(t)$ for the mortality improvement when t is greater than t_0 and calculate the probability $h(t,x,s)$, we need to know the number of years between t_0 and t. For instance if we assume that the current date t corresponds to year 2020 then $t_0 = t-8$ since t_0 corresponds to year 2012.

We assume that the possible ages of death are integers between x and 120. As such the probability that an individual aged x at current date t dies in s years is $h(t,x,s-1) - h(t,x,s)$

[11] The male and female versions are available from the Society of Actuaries website at: http://mort.soa.org/ViewTable.aspx?&TableIdentity=2585 and http://mort.soa.org/ViewTable.aspx?&TableIdentity=2586

[12] The mortality reduction factor is the Projection Scale G2 factor. The male and female versions are available at http://mort.soa.org/ViewTable.aspx?&TableIdentity=2583 and http://mort.soa.org/ViewTable.aspx?&TableIdentity=2584

and we finally obtain the death probability distribution by calculating, for each integer s varying from 0 to $120 - x$ inclusive, the quantity $h(t, x, s-1) - h(t, x, s)$.

If we denote τ the random variable that represents the remaining time-to-live of an individual aged x at current date t, then then the unconditional probability distribution of τ is given by:

$$\mathbb{P}(\tau = s) = h(t, x, s-1) - h(t, x, s) \quad \forall s \in [\![0, 120 - x]\!]$$

3.4.3.3 Life Event Uncertainty

We focus in this chapter on a specific type of life event related to the emergence of long-term care need, possibly triggered by health-related problems. (A possible extension of the framework could be to also take into other life events such as divorce, for example.) For this reason, the denominations "life event" and "long-term care" or "long-term care needs" are used interchangeably in what follows. More precisely, we propose to model the unexpected component of expenses with a set of cash-flows that can start at a random age and last until the uncertain date of death of the individual (see Section 3.4.3.2).

We derive a long-term care probability distribution conditional on age and gender from the incidence and continuance information based on Appendices D and G of the Society of Actuaries 1994-2004 Intercompany Experience Study.[13] For each age ranging from 50 until 120, Table 3.5 reports the incidence rates by attained age and the probability of the event requiring nursing home coverage. The probability of entering a nursing home was derived from Appendix D8 of the Society of Actuaries 2004 LTC study. The incidence rates were presented in Appendix D2A of the same study. The data is sorted by attained age and gender and an exponential curve is fitted with a cap of 12.20% applied to both male and female rates in order to prevent unrealistic rates. The caps are derived from the maximum incidence rate observed at senior ages in the study. The following two equations represent the fitted exponential incidence rates:

$$\begin{aligned}
&\text{LTC Incidence Rate}_{\text{Male}}(x) = \min\left(12.2\%, 0.0005 \times \exp\left(0.1153 \times (x - 49)\right)\right) \\
&\text{LTC Incidence Rate}_{\text{Female}}(x) = \min\left(12.2\%, 0.0006 \times \exp\left(0.1146 \times (x - 49)\right)\right)
\end{aligned}$$

where x represents the age of the individual.

Assuming incidence rates and the probability of entering a nursing home given the occurrence of a health event to be constant over time, we can then calculate from Table 3.5 the unconditional probability $h^{\text{LE}}(t, x, s)$ that an individual aged x at current date t will have an incident leading her to enter a nursing home at date $t + s$ (she will be aged $x + s$ at date $t + s$, $s \geq 0$) as

$$\begin{aligned}
h^{\text{LE}}(t, x, s) = \left(\prod_{i=1}^{s} \left[1 - \text{LTC Incidence Rate}(x + i - 1) \times \text{NH}(x + i - 1)\right]\right) \times \\
\text{LTC Incidence Rate}(x + s) \times \text{NH}(x + s) \times h(t, x, s)
\end{aligned}$$

where $\text{NH}(x + i - 1)$ is the probability that an individual aged $x + i - 1$ has to enter a nursing home given the occurrence of a life event. $h(t, x, s)$ is the probability that an individual aged x at current date t will still be alive in s years as defined in Equation (3.13).

If we denote τ^{LE} the random variable that represents the remaining time before a life event occurs to an individual aged x at current date t, then the unconditional probability distribution of τ^{LE} is given by:

[13] see https://www.soa.org/resources/experience-studies/2005-2009/research-ltc-study-1984/ for more details.

Table 3.5: Long-term care incidence rates

Attained Age	Male Incidence Rate %	Female Incidence Rate %	Nursing Home %
50	0.06	0.07	15
51	0.06	0.08	15
52	0.07	0.08	15
53	0.08	0.09	15
54	0.09	0.11	15
55	0.10	0.12	15
80	1.78	2.09	86
81	2.00	2.35	86
82	2.25	2.63	86
83	2.52	2.95	86
84	2.83	3.31	86
115	12.20	12.20	93
116	12.20	12.20	93
117	12.20	12.20	93
118	12.20	12.20	93
119	12.20	12.20	93

This table details the incidence rates by attained age and the probability of the event requiring nursing home coverage. We filled in values only for ages ranging from 50 years to 55 years inclusive, 80 years to 84 years inclusive, and 115 years to 119 years inclusive. The data is based on Appendices D8 and D2A of the Society of Actuaries 2004 LTC study.

$$\mathbb{P}(\tau^{\mathrm{LE}} = s) = h^{\mathrm{LE}}(t, x, s), \quad \forall s \in [\![0, 120 - x]\!]$$
$$\mathbb{P}(\tau^{\mathrm{LE}} = 1,000 - x) = 1 - \sum_{s=0}^{120-x} \mathbb{P}(\tau^{\mathrm{LE}} = s)$$

where $(\tau^{\mathrm{LE}} = 1,000 - x)$ corresponds to a scenario where no life event occurs to the individual over her lifetime.

Using the data from the Society of Actuaries we have been able to obtain, for an individual aged x at current date t, the unconditional probability distribution of her remaining time-to-live τ and the unconditional probability distribution of the remaining time before a life event occurs τ^{LE}. However τ and τ^{LE} are clearly not independent and we need to determine the joint probability distribution of τ and τ^{LE} to perform numerical simulations. Assuming that an individual enters a nursing home once and only once during her life and that she dies at the end of her stay, we can deduce from the empirical distributions of the length of stay in nursing home by age group given in Table 3.6 the conditional distribution of the random variable τ given τ^{LE}. For instance, if we consider an individual aged $x = 65$ years today to whom a life event occurs when she is 76 years old (i.e. $\tau^{\mathrm{LE}} = 11$). Then we obtain from the fourth column and fifteenth row of Table 3.6 that $\mathbb{P}(\tau = 13 | \tau^{\mathrm{LE}} = 11) = 11.59\%$.

Table 3.6: Nursing home stay duration distribution

Duration (years)	**55-64 (%)**	**65-74 (%)**	**75-84 (%)**	**85-89 (%)**	**90+ (%)**
0.0000	0.79	0.70	0.29	0.17	0.10
0.0027	0.32	0.49	0.31	0.26	0.30
0.0055	0.31	0.51	0.33	0.19	0.21
0.0082	0.39	0.55	0.39	0.28	0.20
0.0110	0.38	2.82	0.38	0.25	0.29
0.0137	2.31	0.00	2.17	1.57	1.24
0.0274	3.82	4.41	3.87	3.09	3.32
0.0548	3.59	3.97	3.57	2.90	2.91
0.0822	8.54	8.83	7.40	6.54	7.21
0.1644	6.72	5.87	5.14	4.75	5.63
0.2466	5.10	4.59	4.39	4.29	4.90
0.3288	8.00	6.72	6.32	6.58	7.56
0.4931	13.91	12.28	13.19	14.95	17.24
1	11.70	13.71	15.97	19.20	20.30
2	7.35	9.32	11.59	13.28	14.07
3	6.96	7.76	8.97	9.55	7.61
4	4.76	5.25	5.42	4.73	3.11
5	4.26	3.46	3.00	2.87	2.00
6	1.79	2.22	1.92	1.73	0.37
7	1.75	1.41	1.46	1.05	0.36
8	1.52	1.24	1.15	0.59	1.07
9	2.16	0.90	1.03	0.74	0.00
10	1.24	0.75	0.42	0.44	0.00
11	0.81	0.59	0.12	0.00	0.00
12	0.30	0.46	0.30	0.00	0.00
13	0.00	0.60	0.90	0.00	0.00
14	0.00	0.59	0.00	0.00	0.00
15	1.22	0.00	0.00	0.00	0.00

This table details the duration distribution of a nursing home stay for different age classes: 55-64, 65-74, 75-84, 85-89 and more than 90 years old. The data is based on Appendices D8 and D2A of the Society of Actuaries 2004 LTC study. See the report on page 36 downloadable with the link https://www.soa.org/resources/experience-studies/2005-2009/research-ltc-study-1984/.

Table 3.7: Nursing home stay duration distribution modified

Duration (years)	55-64 (%)	65-74 (%)	75-84 (%)	85-89 (%)	90-100 (%)	>100 (%)
1	65.88	65.45	63.72	65.02	71.41	100.00
2	7.35	9.32	11.59	13.28	14.07	0.00
3	6.96	7.76	8.97	9.55	7.61	0.00
4	4.76	5.25	5.42	4.73	3.11	0.00
5	4.26	3.46	3.00	2.87	2.00	0.00
6	1.79	2.22	1.92	1.73	0.37	0.00
7	1.75	1.41	1.46	1.05	0.36	0.00
8	1.52	1.24	1.15	0.59	1.07	0.00
9	2.16	0.90	1.03	0.74	0.00	0.00
10	1.24	0.75	0.42	0.44	0.00	0.00
11	0.81	0.59	0.12	0.00	0.00	0.00
12	0.30	0.46	0.30	0.00	0.00	0.00
13	0.00	0.60	0.90	0.00	0.00	0.00
14	0.00	0.59	0.00	0.00	0.00	0.00
15	1.22	0.00	0.00	0.00	0.00	0.00

This table is derived from Table 3.6. It groups all the durations lower than or equal to one year into a single one-year duration category. It also contains a column with the duration probability distribution conditional upon the occurrence of a long-term care event when the individual is older than 100 years.

We then make the following two assumptions: (1) since the time step we have for mortality is one year, we will group the durations less than one year with the one equal to one year and (2) we restrict ourselves to the 90-100 age range for the last column and we make the assumption that if a life event occurs when the individual is over 100 years old, then she dies within a year with probability equal to 100%. We finally use Table 3.7 to derive, for an individual aged x, the conditional probability that the remaining time-to-live τ is equal to s given that the remaining time before the occurrence of a life event τ^{LE} is equal to s^{LE} is:

$$\begin{aligned}
&\mathbb{P}(\tau = s | \tau^{\text{LE}} = s^{\text{LE}}) = \text{dur}(x + s^{\text{LE}}, s - s^{\text{LE}}), \quad \forall (s^{\text{LE}}, s) \in [\![0, 120 - x]\!]^2 \quad \text{such that } s > s^{\text{LE}} \\
&\mathbb{P}(\tau = s | \tau^{\text{LE}} = s^{\text{LE}}) = 0, \quad \forall (s^{\text{LE}}, s) \in [\![0, 120 - x]\!]^2 \quad \text{such that } s \leq s^{\text{LE}}
\end{aligned} \tag{3.14}$$

where $x + s^{\text{LE}}$ is the age of the individual when the life event occurs, $s - s^{\text{LE}}$ corresponds to the duration between the the occurence of the life event and the individual death and $\text{dur}(x + s^{\text{LE}}, s - s^{\text{LE}})$ is the probability that an individual to whom a life event occurs when she is $x + s^{\text{LE}}$ years old will die $s - s^{\text{LE}}$ years after the life event occurrence.

From Equation (3.14), we obtain the joint distribution of τ et τ^{LE}:

$$\mathbb{P}(\tau^{\text{LE}} = s^{\text{LE}}, \tau = s) = \mathbb{P}(\tau = s | \tau^{\text{LE}} = s^{\text{LE}}) \times \mathbb{P}(\tau^{\text{LE}} = s^{\text{LE}}), \quad \forall (s^{\text{LE}}, s) \in [\![0, 120 - x]\!]^2$$

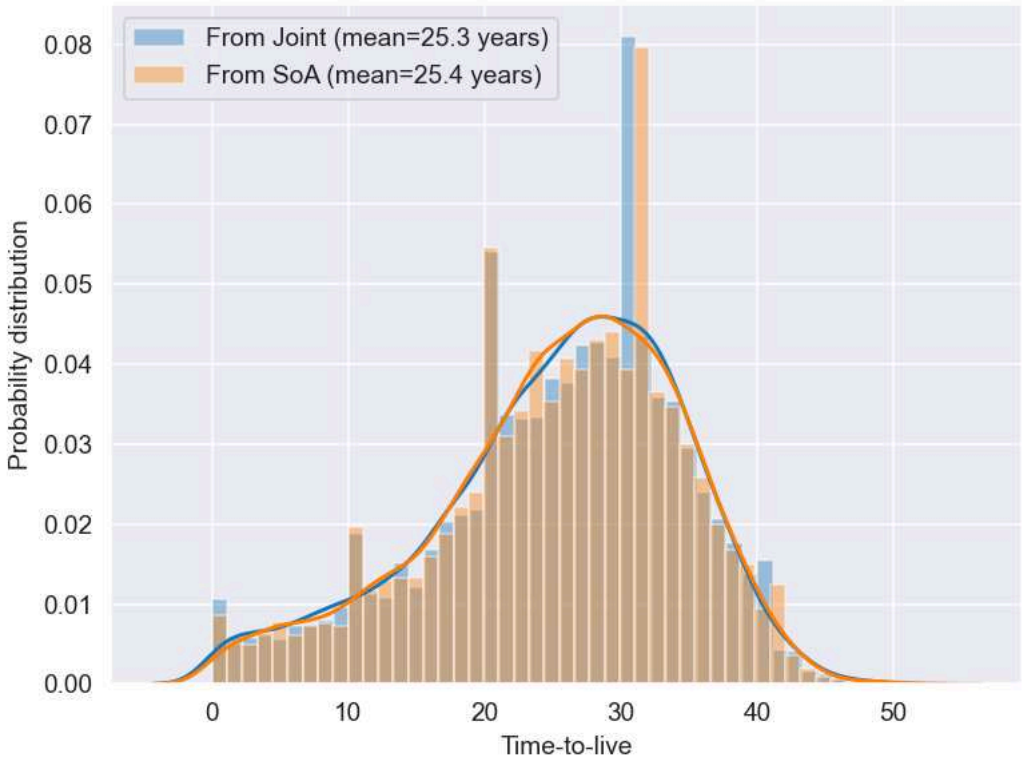
We also have:

$$\mathbb{P}(\tau^{\text{LE}} = 1,000 - x, \tau = s) = \mathbb{P}(\tau = s) - \sum_{s^{\text{LE}}=0}^{120-x} \mathbb{P}(\tau^{\text{LE}} = s^{\text{LE}}, \tau = s), \quad \forall s \in [\![0, 120 - x]\!]$$

and finally obtain the joint distribution of τ^{LE} and τ.

Figure 3.1 shows the empirical probability distribution of the time-to-live of the individual τ under SoA tables and the empirical probability distribution of the time-to-live of the individual derived from the joint density distribution of τ^{LE} and τ. We assess that the two distributions are close to each other.

Figure 3.1: Probability distributions of the individual remaining time-to-live



This figure shows two empirical probability distributions of the remaining time-to-live of the 65-year-old female individual. The orange one is directly built from the SoA mortality tables and the blue one is derived from the joint density distribution of τ^{LE} and τ (see Section 3.4.3.3 for more details). 10,000 draws were made for each probability distribution.

Table 3.8: Nursing home stay duration distribution for durations lower than or equal to one year

Duration (years)	55-64 (%)	65-74 (%)	75-84 (%)	85-89 (%)	90+ (%)
0.0000	1.20	1.07	0.46	0.26	0.14
0.0027	0.49	0.75	0.49	0.40	0.42
0.0055	0.47	0.78	0.52	0.29	0.29
0.0082	0.59	0.84	0.61	0.43	0.28
0.0110	0.58	4.31	0.60	0.38	0.41
0.0137	3.51	0.00	3.41	2.41	1.74
0.0274	5.80	6.74	6.07	4.75	4.65
0.0548	5.45	6.07	5.60	4.46	4.08
0.0822	12.96	13.49	11.61	10.06	10.10
0.1644	10.20	8.97	8.07	7.31	7.88
0.2466	7.74	7.01	6.89	6.60	6.86
0.3288	12.14	10.27	9.92	10.12	10.59
0.4931	21.11	18.76	20.70	22.99	24.14
1	17.76	20.95	25.06	29.53	28.43

This table is derived from Table 3.6 and reports the duration distribution of a nursing home stay for different age classes: 55-64, 65-74, 75-84, 85-89 and more than 90 years old conditional on the duration being less than or equal to one year.

In the scenarios where long-term care needs materialize (i.e. where $\tau^{\text{LE}} \neq 1000$), we add at each date an unexpected target withdrawal corresponding to the cost of such long-term care needs for a time period equal to $\tau - \tau^{\text{LE}}$. In the particular case where the time period is equal to one year which represents the most probable outcome, we will draw a random duration in years (1/365, 2/365, 5/365, 30/365, ..., 1) from the nursing home stay duration distributions when the duration is lower than or equal to one year in Table 3.8 to generate the most realistic nursing home costs possible.

For the cost of the life event, estimates depend upon the option chosen to deal with long-term care needs. In particular, the individual can use in-home care services, community and assisted living services, or nursing home facilities with various types of arrangements. We have obtained estimates for the USA national median annual costs of six options and their expected annual growth over the next five years from the Genworth Cost of Care Survey 2019:

1. \$51,480 and 3.44% for homemaker services;
2. \$52,624 and 3.09% for home health aide;
3. \$19,500 and 2.90% for adult daily health care;
4. \$48,612 and 2.97% for private one bedroom;
5. \$90,155 and 3.10% for semi-private room;
6. \$102,200 and 3.14% for private room.

While the framework can accommodate any cost estimate, we have chosen for the analysis that follows to use the fifth option, namely the semi-private room accommodation in a nursing home.

3.4.4 Monte Carlo Simulations

We first estimate the VAR model for the short rate and the stock and bond prices and also calculate the joint distribution of the death time and the age at which a life event occurs. Next, we specify the target additional replacement income L_t^{unexp} which is the additional withdrawal that would be needed in case a life event occurs. Then, we perform the following simulations:

1. Simulate 10,000 paths for the short rate, the stock price and the bond price under the historical probability measure;

2. For each path, we sample a random pair $(\tau^{\mathrm{LE}}, \tau)$ of remaining time before the life event occurrence and remaining time-to-live;

3. For each path where a life event occurs, if $\tau - \tau^{\mathrm{LE}} = 1$, then we sample a random duration from the nursing home stay duration when duration is lower than or equal to one year.

3.5 Numerical Analysis of Optimal Investment Strategies in Decumulation

In this section, we apply the framework to a 65-year-old woman who is already retired (and presumably just retired). Her initial and unique contribution is $[\boldsymbol{\theta}_{0-}]^{\top}\mathbf{1}_2 + [\boldsymbol{P}]^{\top}\mathbf{1}_2$: this equation states that she allocates her initial wealth to (1) a maximum of two financial assets via an investment in a balanced fund (BF in short), an equity index or a target date fund (TDF in short) and (2) annuities (in the analysis that follows, we will consider both a SPIA with a 2% COLA and a VA). We implement and test different sets of buy-and-hold strategies with withdrawal rates of 3%, 4% and 5% of the initial wealth, and a 2% annual cost-of-living adjustment.[14] We systematically report the results obtained in the absence and in the presence of life events to check for the impact of long-term care needs on the demand for annuities. As explained in Section 3.4.3.3, we assume that the cost of long-term care in year t is equal to $\$90,155 \times (1 + 3.10\%)^t$.

The performance and risks associated with any given strategy will be measured with a number of key indicators, which can be broadly sorted into figures of merit (to be maximized) and figures of risk (to be minimized). In terms of figures of merit, we first report in particular the median discounted income deficit. The income deficit is defined for a given scenario as the discounted sum of the differences between actual withdrawals and target withdrawals (which are defined as 3%, 4% or 5% of initial wealth subject to a 2% COLA, plus cost of long-term care needs, if and when they are incurred) and the actual withdrawal.[15] By definition, this quantity is equal to zero at best, when the individual has enough wealth to finance all target withdrawals [16]. As a related indicator we also report the median discounted percentage of lifetime income (abbreviated as PLI) achieved, which is defined as the median value across all scenarios of the ratio between the sum of the individual's discounted actual withdrawals and the sum of the discounted target withdrawals until death. We finally report the median discounted bequest value (MB in short), which is unbounded. In terms of figures of risk, we first report a short-term risk indicator defined as the median (over the scenarios) maximum (over time) annual loss on the liquid portion of the portfolio (invested in BFs or TDFs), denoted by MAL. We also report several long-term risk indicators, including (1) the shortfall probability, defined as the percentage of scenarios where the individual outlives her assets, (2) the median and extreme shortfall durations, defined respectively as the median and the ninety fifth percentile of the number of years when the actual withdrawal is strictly lower than the target withdrawal, (3) the extreme discounted income deficit defined as the fifth percentile of the discounted differences between actual and target withdrawals.

An optimization exercise requires the identification of a proper optimization objective (and possibly some constraints). This raises two main questions which are (1) the integration of potentially conflicting goals and (2) aggregation of risk and return dimensions for a given goal.

[14] The initial withdrawal rate is expressed as a percentage of the initial wealth.

[15] For each scenario, we use the sovereign US zero-coupon yield curve as of 1, June, 2020 for discounting purposes.

[16] We remind that in the particular case where the benefits paid by annuities is higher than the target withdrawal then the difference is reinvested into the liquid bucket of the portfolio. If the liquid bucket of the portfolio is completely depleted or if there is no liquid bucket then the difference is invested in a 50% stock 50% bond balanced fund.

Meaningful goals include expected lifetime income needs, unexpected lifetime income needs (long-term care) and bequest. The first challenge is to aggregate these three goals which can conflict with each other in the framework developed above. Expected and unexpected lifetime income needs can naturally be aggregated so that when the probability of a shortfall or expected shortfall is reported, it includes both the expected and unexpected components as part of the target. There still remains to aggregate the bequest objective with the total (expected plus unexpected) income objective.

Formally we take the performance indicator to be the median discounted bequest (note that this quantity is always positive and not bounded) over the 10,000 Monte Carlo simulations, which corresponds to MB. We take the risk indicator to be the fifth percentile of the discounted income deficit over the 10,000 Monte Carlo simulations, which corresponds to EID. By definition, EID is equal to zero at best, when the individual has enough wealth to finance all target withdrawals. Defining λ as a risk-aversion parameter that characterizes the risk appetite of the individual, we can write the objective function as

$$\arg \max_{w_1,\ldots,w_n} [\text{MB}(w_1,\ldots,w_n) + \lambda \text{EID}(w_1,\ldots,w_n)] \tag{3.15}$$

where $w_1,\ldots,w_n$ represent the percentage weights of each asset, encompassing financial liquid assets and annuities[17]. The formulas for the discounted bequest B and the discounted income deficit ID on a given Monte-Carlo scenario are respectively $\text{B} = W(\tau)\exp(-\tau R_{0,\tau})$ and $\text{ID} = \sum_{t=0}^{\tau-1}(AW(t) - TW(t))\exp(-tR_{0,t})$. Here $TW(t)$ is the target withdrawal level at time t (given by 3%, 4% or 5% of the initial wealth plus a possible cost of living adjustment), $AW(t)$ is the actual withdrawal level at time t (which is equal to $TW(t)$ when possible given the available wealth, and less than $TW(t)$ otherwise), $R_{0,t}$ is the annualized continuously compounded discount rate at time 0 for the maturity t, $W(t)$ is the wealth available at time t, and τ is the uncertain date of death. MB is computed as the median of B across all the Monte-Carlo scenarios and EID as the fifth percentile of ID across all the Monte-Carlo scenarios. In the analysis that follows, we take five values for λ ($\lambda = 0.5, 1, 2, 4, 6$), which we interpret as defining the Aggressive, Moderately Aggressive, Moderate, Moderately Conservative and Conservative investor, respectively. We also report results for two limit cases, namely $\lambda = 0$ (pure focus on performance) and $\lambda = 1,000$ (pure focus on risk).

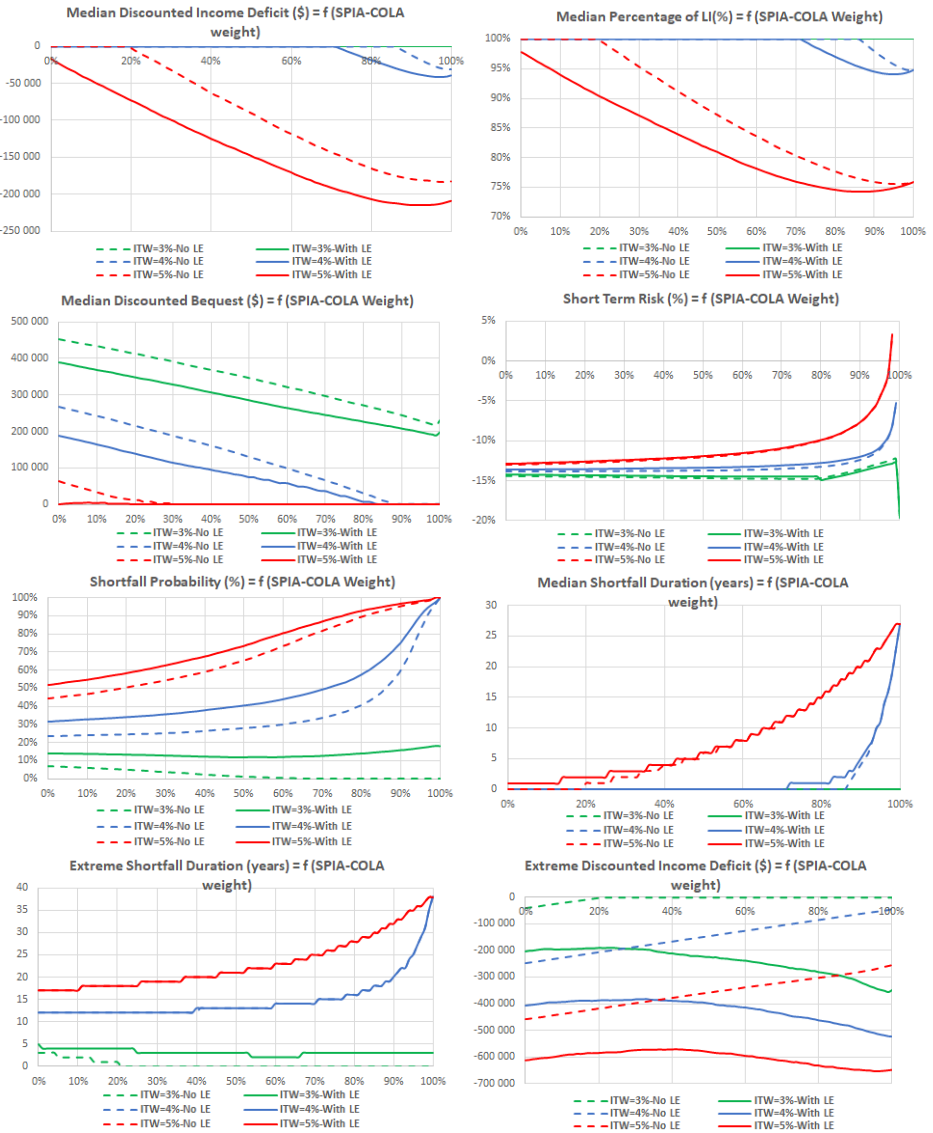
3.5.1 Base Case Analysis

We first present a base case analysis, where we assume that the individual (a 65-year-old female) is endowed with $500,000. We also assume that she has access to a simple investment universe that contains a balanced fund with a 50%/50% stock/bond annually rebalanced mix, and an immediate annuity with a 2% cost-of-living adjustment (SPIA-COLA). In this base case analysis, we systematically test three levels of initial target withdrawal rates, namely 3%, 4% and 5%, and we let target withdrawals grow by 2% per year to account for expected growth in for the cost of living. Reading in Table 3.2 that the initial payout rate of the SPIA-COLA annuity for a 65-year-old female is 3.79%, we can define the individual funding ratio as the percentage of her replacement income needs that she can secure with certainty (and in the absence of a life event) by investing her assets into a SPIA-COLA annuity. These funding ratios for initial target withdrawals of 3%, 4% and 5% are respectively $\frac{3.79}{3} = 126.3\%$, $\frac{3.79}{4} = 94.8\%$ and $\frac{3.79}{5} = 75.8\%$. These results suggest that the individual who has an aggressive target withdrawal rate (5%) is initially underfunded (by a bit less than 25%, with a funding ratio at 75.8%) while the individual who has a conservative target withdrawal rate (3%) is initially overfunded (by a bit more than 25%, with a funding ratio at 126.3%).

[17] We have $\sum_{i=1}^{n} w_i = 1$.

Figure 3.2 displays selected charts representing the various indicators introduced at the beginning of Section 3.5 as a function of the initial percentage allocation to the SPIA-COLA, with values ranging from 0% to 100%, with a grid step of 1%. We observe that for initial target withdrawal rates of 4% and 5%, the shortfall probability increases with the allocation to the SPIA-COLA. This is because investing a large fraction of the portfolio in annuities does not generate the amount of upside potential needed to finance a higher target level of consumption in retirement. On the other hand, in the case of a 3% withdrawal rate, the individual is sufficiently funded (with a funding ratio at 126.3%) to be able to meet target levels of withdrawals without substantial upside potential, that is without a significant investment in the balanced fund. Similarly, we find that the extreme shortfall durations increase with the SPIA-COLA weight, except in the case where the initial target withdrawal rate is equal to 3%. When accounting for life events and when the initial withdrawal rates are 4% or 5%, the median percentage of lifetime income first decreases as the allocation to annuities increases, then reaches its minimum for corresponding annuity weights of 95% and 87% and slightly increases beyond that point. We note that the median discounted bequest is a decreasing function of the allocation to annuities, as expected. On the other hand we observe that, when accounting for life events and for a 4% initial withdrawal rate, the median discounted income deficit decreases (i.e. is more negative) as the SPIA-COLA weight increases when the allocation to the SPIA-COLA is between 0% and 97%. We confirm in particular that the presence of long-term care needs has a strong impact on the distribution of the discounted income deficit with a median value that is significantly lower (more negative), especially in the case of a 5% withdrawal rate, which is the situation where the expected component is particularly sizable with respect to the initial wealth.

Figure 3.2: Reporting for the two-asset universe made of a SPIA-COLA and a BF



This figure reports the main statistics described at the beginning of Section 3.5 for the base case,i.e. for a 65-year-old female individual with a $500,000 initial wealth and a two-asset universe made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and a 50% stock / 50% bond balanced fund (BF). We consider three levels of initial target withdrawal rate: 3%, 4% and 5%. We display the results with and without taking into account life events. We use a grid weight step of 1% to simulate the 101 corresponding strategies. We report the results both in the absence and in the presence of life events.

Table 3.9: Optimal strategies for the two-asset universe made of a SPIA-COLA and a BF

Initial target withdrawal		**Optimal allocation = (% SPIA-COLA, %BF)**						
	RA	**0**	**0.5**	**1**	**2**	**4**	**6**	**1000**
3%	Without life events	(0, 100)	(0, 100)	(10, 90)	(21, 79)	(21, 79)	(21, 79)	(21, 79)
	With life events	(0, 100)	(0, 100)	(0, 100)	(1, 99)	(5, 95)	(6, 94)	(23, 77)
4%	Without life events	(0, 100)	(0, 100)	(0, 100)	(100, 0)	(100, 0)	(100, 0)	(100, 0)
	With life events	(0, 100)	(0, 100)	(0, 100)	(6, 94)	(14, 86)	(14, 86)	(33, 67)
5%	Without life events	(0, 100)	(100, 0)	(100, 0)	(100, 0)	(100, 0)	(100, 0)	(100, 0)
	With life events	(8, 92)	(41, 59)	(41, 59)	(41, 59)	(41, 59)	(41, 59)	(41, 59)

This table displays the optimal allocation to the single premium immediate annuity with a 2% COLA indexation (SPIA-COLA) and the balanced fund (BF) for different levels of risk aversion (RA) and initial target withdrawal rates. The target withdrawal rates grow by 2% every year. We report the results both in the absence and in the presence of life events.

Turning to the two key ingredients in the optimization problem, namely MBS (the median discounted bequest) and EID (the 5% value-at-risk of the discounted income deficit), we confirm in Figure 3.2 the presence of a typical risk-return trade-off. On the one hand, an increase in the allocation to annuities leads to a decrease in the median bequest (MB), which is not desirable since the median bequest is a measure of performance. We confirm that the median bequest is decreasing with respect to the target withdrawal rate, as expected since it is more difficult to leave a bequest starting from a given wealth level (here $500,000) when replacement income needs are higher. We also confirm, as expected, that the presence of life events (solid lines) leads to a smaller median bequest compared to a situation without life events (dotted lines), and this for all values of the initial withdrawal rate. On the other hand, when not accounting for life events, an increase in the allocation to annuities leads to a higher (less negative) value for EID, which is desirable since EID is a measure of risk. When life events are introduced, the monotonic relationship between the fifth percentile of the income deficit and the allocation to annuities no longer holds. In particular, the fifth percentile of the discounted income deficit distribution ceases to increase beyond a certain allocation to annuities. In details, an allocation to annuities beyond the 23%, 33% and 41% levels in the 3%, 4% and 5% withdrawal rate cases, not only implies a decrease in performance (measured by MB) but also an increase in risk. Intuitively, this is because a minimum amount of upside potential, which is generated by a non-zero allocation to the balanced fund, is needed in the 5% worst scenarios to ensure that replacement income needs (including both the expected and unexpected components) are met.

Table 3.9 shows the optimal strategies for the base case universe in the absence and in the presence of life events for various risk aversion levels. We assess that, for all initial withdrawal rates and both in the absence and in the presence of life events, the optimal demand for annuities increases when risk aversion increases. One important finding in this analysis for initial target withdrawal rates of 4% and 5%, is that the presence of long-term care risk strongly reduces the optimal demand for annuities for most individuals, at least those that are sufficiently risk-averse to show some appetite for annuities in the first place. Focusing for example on the case of a 4% withdrawal rate, we find that the demand for annuities from the moderate individual ($\lambda = 2$) decreases from 100% to 6% when long-term care needs are accounted for. In the case of a 5% initial target withdrawal rate, the demand for annuities for the same moderate individual ($\lambda = 2$) decreases from 100% to 41% when long-term care needs are accounted for. Overall, these results suggest that the costly reversibility of annuitization decisions contributes to explaining the annuity puzzle for individuals facing life event uncertainty.

3.5.2 Robustness Checks

In this section we conduct three robustness checks.

1. For a base case value of 4% initial target withdrawal rate, we test for the impact of changes in initial wealth on risk and return indicators and on the optimal demand for annuities. More specifically, we test two additional initial wealth values, namely $250,000 and $1,000,000.

2. For a base case characterized by a 4% initial target withdrawal rate and a $500,000 initial wealth, we analyze how the risk and return indicators and the optimal demand for annuities change if we replace the balanced fund by a target date fund (TDF) or by an equity index within the base case investment universe. We conduct a similar test for a 4% initial target withdrawal rate and a $500,000 initial wealth where we replace in the base case investment universe the SPIA-COLA by the VA. In both cases, our goal is to assess whether these new instruments (TDF/Equity Index or VA) contribute to increasing investor welfare.

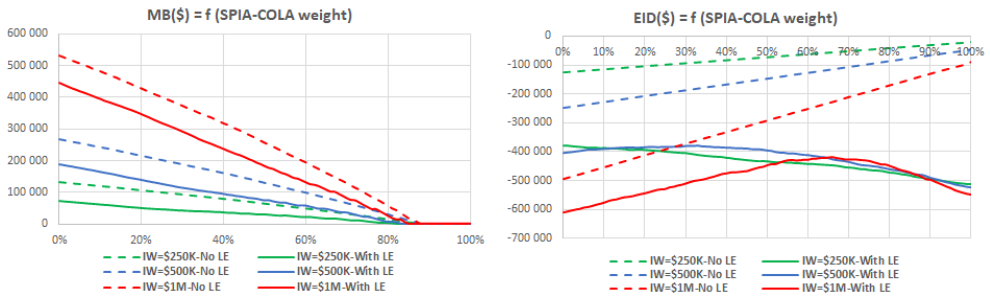
3. Starting from the same base case as before, we analyze the impact (on the risk and return indicators and on the optimal demand for annuities) of the introduction of a variable annuity in the investment universe in addition to (as opposed to in substitution to) the SPIA-COLA. In this case, we therefore have a three-asset universe containing the BF, the SPIA-COLA and the VA.

3.5.2.1 Robustness Check 1: Impact of Changes in the Initial Wealth Level

We examine how the initial wealth of the individual impacts the optimal demand for annuities. For given values of the initial target withdrawal rate and the risk aversion parameter, in the absence of life events, the optimal allocation will be independent of initial wealth since the replacement income needs of the individual will represent the same deterministic percentage of her initial wealth. However, when life events are taken into account, their relative costs will have a stronger impact on individuals with a lower initial wealth.

The charts in Figure 3.3 confirm our intuition. In particular, we find that the location and shape of EID indicator are strongly impacted by changes in the individual initial wealth. The maxima of the EID charts respectively correspond to SPIA-COLA allocations of 0%, 33% and 66% for initial wealth levels of $250,000, $500,000 and $1,000,000 when life events are accounted for.

Figure 3.3: Median bequest (MB) and extreme income deficit (EID) for the two-asset universe made of a SPIA-COLA and a BF with different values of the initial wealth (IW)



This figure reports the main statistics described at the beginning of Section 3.5 for the first robustness check,i.e. for a 65-year-old female individual with three different levels of initial wealth ($250,000, $500,000 and $1,000,000) and a two-asset universe made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and a 50% stock / 50% bond balanced fund (BF). We consider an initial target withdrawal rate of 4%. We apply a grid weight step of 1% to simulate the 101 corresponding strategies. We report the results both in the absence and in the presence of life events.

Table 3.10: Optimal strategies for the two-asset universe made of a SPIA-COLA and a BF with different values of the initial wealth

Initial wealth		**Optimal allocation = (% SPIA-COLA, %BF)**						
		Initial Target Withdrawal =4%						
	RA	**0**	**0.5**	**1**	**2**	**4**	**6**	**1000**
$250,000	Without life events	(0, 100)	(0, 100)	(0, 100)	(100, 0)	(100, 0)	(100, 0)	(100, 0)
	With life events	(0, 100)	(0, 100)	(0, 100)	(0, 100)	(0, 100)	(0, 100)	(0, 100)
$500,000	Without life events	(0, 100)	(0, 100)	(0, 100)	(100, 0)	(100, 0)	(100, 0)	(100, 0)
	With life events	(0, 100)	(0, 100)	(0, 100)	(6, 94)	(14, 86)	(14, 86)	(33, 67)
$1,000,000	Without life events	(0, 100)	(0, 100)	(0, 100)	(100, 0)	(100, 0)	(100, 0)	(100, 0)
	With life events	(0, 100)	(0, 100)	(0, 100)	(55, 45)	(55, 45)	(56, 44)	(66, 34)

This table displays the optimal allocation to the single premium immediate annuity with a 2% COLA indexation (SPIA-COLA) and the balanced fund (BF) for different levels of risk aversion (RA), an initial target withdrawal rate of 4% and three different levels of initial wealth: $250,000, $500,000 and $1,000,000. Withdrawals grow by 2% per year to provide an adjustment for the cost of living. We report the results both in the absence and in presence of life events.

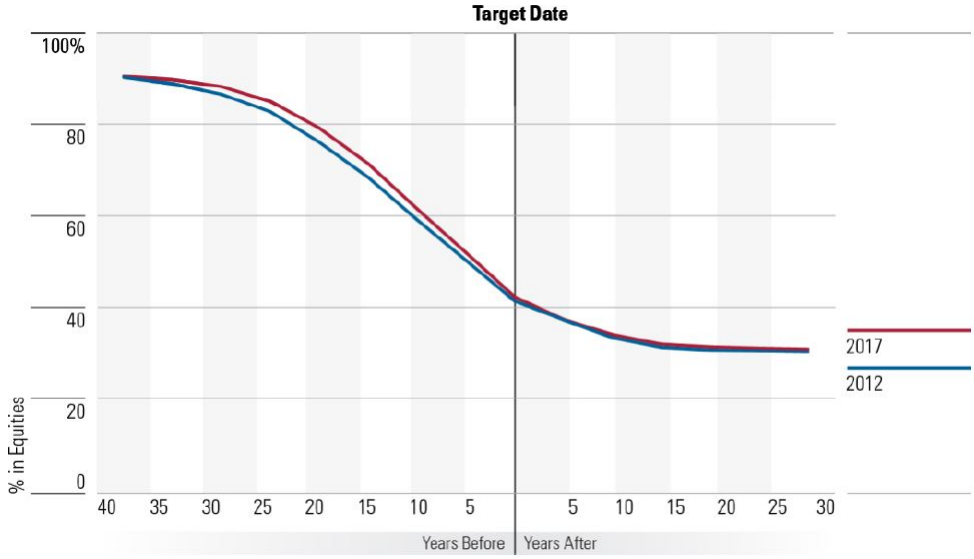
Table 3.10 shows the weight of the optimal strategies for the base case universe in the absence and presence of life events for these three different initial wealth levels. We again find that the introduction of long-term care needs have a strong impact on the demand for annuities for sufficiently risk-averse individuals ($\lambda \geq 2$), an impact that is decreasing in the initial wealth level. For example, when the initial wealth is $250,000, the optimal demand for annuities decreases from 100% to 0% when the life event is introduced for the moderately conservative investor ($\lambda = 4$), while it decreases from 100% to 14% for an initial wealth of $500,000 (our base case value) and only from 100% to 55% for an initial wealth of $1,000,000. Overall, these results suggest that the impact of life events is stronger for individuals with less initial endowment.

3.5.2.2 Robustness Check 2: Impact of Changes in the Investment Universe and Value Added by Target Date Funds, Equity Index and Variable Annuities

We study three alternative two-asset universes in addition to the base case universe made of a SPIA-COLA and a balanced fund, for an individual with a $500,000 initial wealth and a 4% initial target withdrawal rate.

The first alternative universe that we consider is obtained by replacing the balanced fund with a target date fund invested in stocks and bonds with a glide path corresponding to the red line of Figure 3.4, obtained from Morningstar. During the decumulation (i.e. post-retirement) phase of the individual, the allocation of the target date fund will be 42% stocks when the individual is between 65 and 70, 37% stocks when she is between 70 and 75, 34% stocks between 75 and 80, 32% stocks between 80 and 85, and 30% stocks above 85.

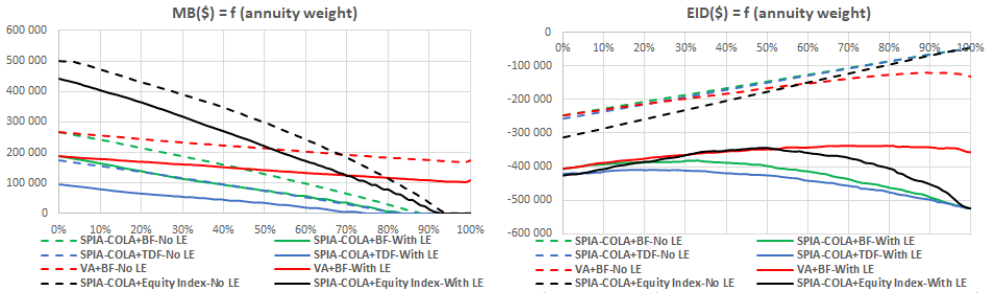
Figure 3.4: Target date fund glide path



This figure, reproduced from Morningstar, represents the average strategic equity glide path in 2017 and in 2012. We use the 2017 curve as the glide path of the target date fund in our case study (Source: Morningstar, Inc.Data as of 12/31/16).

Overall, this target date fund represents an only slightly different (less aggressive) investment opportunity compared to the balanced fund, and we do not expect to see particularly strong differences with respect to the base case investment universe as a result of this substitution. Looking at Figure 3.5, we find that the MB curves for the (SPIA-COLA + BF) and the (SPIA-COLA + TDF) universes both decrease as the SPIA-COLA weight increases. We also remark that the slope of the curve is slightly steeper for the (SPIA-COLA + BF) than for the (SPIA-COLA + TDF) universe, which again is consistent with the fact that the target date fund is a slightly more conservative investment opportunity (lower equity allocation) compared to the BF.

Figure 3.5: Median bequest (MB) and extreme income deficit (EID) for four two-asset universes made of a SPIA-COLA and a BF, a SPIA-COLA and TDF, a SPIA-COLA and an Equity Index, and a VA and BF



This figure reports the statistics MB and EID described at the beginning of section 3.5 for the second robustness check, i.e. for a 65-year-old female individual with an initial wealth of $500,000 and four different universes with two assets: the first universe is made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and a 50% stock / 50% bond balanced fund (BF), the second one is made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and a target date fund (TDF), the third one is made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and an equity index and the last one is made of a variable annuity (VA) and a 50% stock / 50% bond balanced fund (BF). We consider an initial target withdrawal rate of 4%. The target withdrawal rates grow by 2% per year. We apply a grid weight step of 1% to simulate the 101 corresponding strategies. We report the results both in the absence and in the presence of life events.

Table 3.11: Optimal strategies for four two-asset universes made of a SPIA-COLA and a BF, a SPIA-COLA and a TDF, a SPIA-COLA and an equity index, and a VA and a BF

		Optimal allocation - Initial Target Withdrawal =4%						
	RA	**0**	**0.5**	**1**	**2**	**4**	**6**	**1000**
(% SPIA, %BF)	Without LE	(0, 100)	(0, 100)	(0, 100)	(100, 0)	(100, 0)	(100, 0)	(100, 0)
	With LE	(0, 100)	(0, 100)	(0, 100)	(6, 94)	(14, 86)	(14, 86)	(33, 67)
(% SPIA, %TDF)	Without LE	(0, 100)	(0, 100)	(100, 0)	(100, 0)	(100, 0)	(100, 0)	(100, 0)
	With LE	(0, 100)	(0, 100)	(0, 100)	(0, 100)	(16, 84)	(16, 84)	(18, 82)
(% SPIA, %Eq. I)	Without LE	(0, 100)	(0, 100)	(3, 97)	(51, 49)	(100, 0)	(100, 0)	(100, 0)
	With LE	(0, 100)	(0, 100)	(0, 100)	(13, 87)	(34, 66)	(46, 54)	(48, 52)
(% VA, %BF)	Without LE	(0, 100)	(0, 100)	(83, 17)	(89, 11)	(89, 11)	(89, 11)	(89, 11)
	With LE	(0, 100)	(0, 100)	(35, 65)	(65, 35)	(65, 35)	(65, 35)	(79, 21)

This table displays the optimal allocation for different levels of risk aversion (RA), an initial target withdrawal rate of 4%, an initial wealth of $500,000 and four two-asset universes: the first universe is made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and a 50% stock / 50% bond balanced fund (BF), the second one is made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and a target date fund (TDF), the third one is made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and an equity index (Eq. I.) and the last one is made of a variable annuity (VA) and a 50% stock / 50% bond balanced fund (BF). The target withdrawal rates grow by 2% per year. We report the results both in the absence and in the presence of life events. RA stands for risk aversion.

As a result of such small differences, it is not surprising to find in Table 3.11 that the optimal allocation strategies for the (SPIA-COLA + BF) and the (SPIA-COLA + TDF) universes are relatively close, except for the moderately aggressive investor ($\lambda = 1$) in the absence of life events. In particular, we find that for risk aversion levels of 0 or 0.5, the optimal allocation is a portfolio fully invested in the liquid product (balanced fund or target date fund), whether we take into account the life event or not. For risk aversion parameters of 4 and 6, the optimal allocation without life events is a full investment in a SPIA-COLA for both universes while the optimal allocation with life events for a risk aversion parameter of 6 is 14% SPIA-COLA and 86% fund for the (SPIA-COLA + BF) universe versus 16% SPIA-COLA and 84% fund for the (SPIA-COLA + TDF) universe.

The second alternative universe that we consider is obtained by replacing the balanced fund with an equity index. We read in Table 3.11 that the optimal asset allocation with and without taking into account life events is a full investment in the equity index for risk aversion levels lower than or equal to 0.5. The optimal allocation in the presence of life events contains 46% of SPIA-COLA for a conservative investor ($\lambda = 6$), compared with 14% for the (SPIA-COLA, BF) universe and 16% for the (SPIA-COLA, TDF) universe other things being equal. We assess in Figure 3.5 that the MB curve for the (SPIA-COLA, Equity Index) universe is much steeper than for the (SPIA-COLA, BF) and the (SPIA-COLA, TDF) universes. We also assess that a 100% allocation in the equity index implies a value for MB that is sensibly higher than the one obtained with a 100% allocation to the balanced fund or a 100% allocation to the target date fund. We note that for the (SPIA-COLA, Equity Index) universe in the presence of life events, the EID curve reaches its maximum (less negative) for a SPIA-COLA weight of 48%.

The third alternative universe that we consider is obtained by replacing the SPIA-COLA with a variable annuity (VA). We read in Table 3.3 that the initial payout rate of the variable annuity for a 65-year-old female is 4.72%. It can be noted that this initial payout rate is lower than that of the SPIA-COLA (3.79%) but higher than that of a SPIA with no COLA (5.18%). Recall that the variable annuity payout rate is not indexed by a COLA but can increase if the benefit base of the variable annuity increases over time. The underlying fund of the annuity considered has 80% in stocks and 20% in bonds, and we refer to Sections 3.3.1.1 and to Table 3.1 for the detailed contract rules. We find that replacing the SPIA-COLA with the VA leads to much stronger differences compared to replacing the balanced fund with a target date fund or a equity index. In contrast with the (SPIA-COLA + BF) universe, the median bequest indicator is much less sensitive to the allocation to the annuity component (a variable annuity here). This can be explained by the upside potential offered by the variable annuity, which was absent with the SPIA-COLA, and justifies a high allocation to this asset even for individuals with low risk aversion. The EID curves for the (SPIA-COLA + BF) universe and the (VA + BF) universe are also quite different. In particular, the EID indicator keeps improving, even for an exceedingly large allocation to the VA, which is probably explained by the flexibility offered by the variable annuity for which a partial or total liquidation is possible. It is only when the allocation to the VA exceeds 79% that the EID indicator starts to deteriorate (when accounting for life events). Overall, we find that VAs can be regarded as competitors for BFs in terms of upside potential, while they can be regarded as competitors for SPIA-COLA in terms of downside risk. This attractive property of generating upside potential while offering downside protection relative to replacement income needs is the main motivation for introducing VAs in the first place.

From Table 3.11, we find that the optimal demand for annuities is overall much higher when variable annuities are considered instead of SPIA-COLAs, suggesting that the upside potential of VAs make them a strong competitor for the balanced fund for individuals mostly interested in performance. Without taking into account life events, the optimal portfolio is always (89%, 11%)[18] when the risk aversion level is higher than or equal to 2. When accounting for life events, the optimal portfolio is (65%, 35%) for risk aversion levels ranging from 2 and 6. This last result can be explained by the increased flexibility offered by the VA compared to the SPIA-COLA in case of liquidity needs due to the occurrence of a life event and/or unfavorable market conditions.

We compute in Table 3.12 the value of the optimization function for the optimal allocations when the two-asset universe and the investor risk aversion vary. For a given risk aversion level between 0.5 and 6, whether life events may occur or not, the two-asset universe in which welfare is highest is the (SPIA-COLA, Equity Index) universe. When excluding the (SPIA-COLA, Equity Index) universe, accounting for life events and for a given risk aversion level between 0.5 and 6, the (VA, BF) universe displays the highest optimization function value, followed by the (SPIA-COLA, BF) universe and then by the (SPIA-COLA, TDF) universe. This result suggests that in our retirement investing framework in decumulation, the equity index could be a relevant

[18]The first element of the vector is the portfolio VA weight and the second element is the BF weight.

Table 3.12: Optimization function value of the optimal strategies for four two-asset universes made of a SPIA-COLA and a BF, a SPIA-COLA and a TDF, SPIA-COLA and Equity Index, and a VA and a BF

Portfolio allocation		**Optimal allocation - Initial Target Withdrawal =4% - Optimization Function Value**						
	RA	**0**	**0.5**	**1**	**2**	**4**	**6**	**1000**
(SPIA, BF)	Without LE	266,317	142,733	19,148	−87,645	−175,291	−262,936	−43,822,683
	With LE	187,494	−15,471	−218,436	−614,262	−1,396,443	−2,171,571	−380,398,896
(SPIA, TDF)	Without LE	173,352	44,789	−43,823	−87,645	−175,291	−262,936	−43,822,683
	With LE	97,050	−113,127	−323,305	−743,660	−1,564,078	−2,382,021	−408,473,709
(SPIA, Eq. I.)	Without LE	501,401	348,304	196,608	−52,142	−175,291	−262,936	−43,822, 683
	With LE	443,302	230,027	16,752	−402,163	−1,125,335	−1,832,205	−344,267,640
(VA, BF)	Without LE	266,317	142,733	58,606	−61,093	−297,705	−534,316	−118,130,254
	With LE	187,494	−15,471	−200,371	−546,838	−1,223,209	−1,899,110	−337,400,186

This table reports the optimization function value of the optimal allocation for different levels of risk aversion (RA), an initial target withdrawal rate of 4%, an initial wealth of $500,000 and four two-asset universes: the first universe is made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and a 50% stock / 50% bond balanced fund (BF), the second one is made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and a target date fund (TDF), the third one is made of a single premium immediate annuity with a 2% COLA (SPIA-COLA) and an equity index (Eq. I.) and the last one is made of a variable annuity (VA) and a 50% stock / 50% bond balanced fund (BF). The target withdrawal rates grow by 2% per year. We report the results both in the absence and in the presence of life events.

alternative to a balanced fund with a 50% equity - 50% bond allocation or a target date fund.

Overall, we find that that the introduction of variable annuities, which are annuity products with upside potential, has a positive impact on welfare due to an improvement in the median bequest and the value-at-risk at the 5% threshold of income deficit. Since VAs can be regarded as competitors for both balanced funds (for their ability to offer upside potential in good market conditions) and SPIA-COLAs (for their ability to generate steady replacement income in less favorable market conditions), one might wonder what would be the optimal demand for this product in a three-asset universe containing a SPIA-COLA, a balanced fund and a VA. This is precisely what we turn to next.

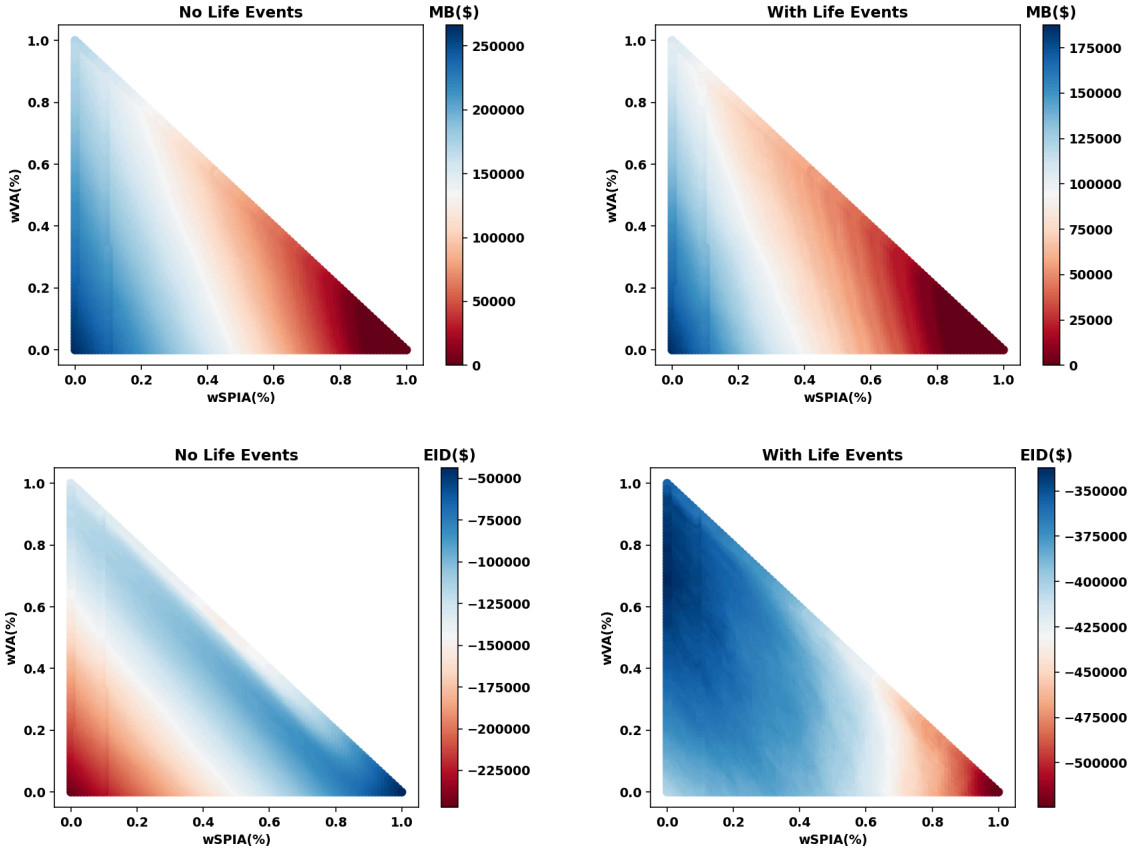
3.5.2.3 Robustness Check 3: Impact of the Introduction of VAs in Addition to SPIA-COLAs

This section is a natural extension of the previous robustness check, where we enlarge the base case universe by adding a variable annuity product. Figure 3.6 reports the MB and EID statistics for a 65-year-old female individual with a $500,000 initial wealth and a three-asset universe made of a SPIA with a 2% COLA, a variable annuity and a balanced fund. Again, we perform the simulations for an initial target withdrawal rate of 4%.

We display two charts for MB, one in the absence of life events and the other in the presence of life events. Similarly, we display two charts relative to EID so as to analyze the impact of life events. We note that the charts for MB are very similar whether or not life events are accounted for. The main difference relates to the median discounted bequest level, which is higher, as expected, in the absence of a life event. We also observe in the presence as well as in the absence of a life event that MB has a low sensitivity to changes in weight invested in the VA for a given allocation to SPIA-COLA. Since what is not invested in the SPIA-COLA or the VA is invested in a balanced fund, this suggests that the VA and the fund could be good substitutes for each other when it comes to upside potential. This result confirms some of the findings discussed in the previous robustness check. However, for a given allocation to VA, MB decreases significantly as the weight invested in SPIA-COLA increases. This is due to the fact that the SPIA-COLA offers no upside potential with respect to the need to generate replacement income, and therefore increasing (respectively, decreasing) the allocation to this asset, which implies decreasing (respectively, increasing) the allocation to the BF, leads to a lower median discounted bequest. Again, we emphasize that these results hold both in the absence and in the presence of life events, the main difference being that the median bequest is lower in the latter situation.

The situation is very different when we consider instead risk indicators. Indeed EID charts differ not only in levels, but also in shape, whether or not the presence of long-term care needs is accounted for. In the absence of life events, EID tends to increase as the allocation to SPIA-COLA increases and/or as the allocation to VA increases. When we account for life events, the EID indicator is a non-monotonic function of the allocation to SPIA-COLA. Indeed, a high allocation to SPIA-COLA (greater than 85%) leads to the lowest values for EID, a result which is consistent with the findings from Figure 3.6 . We also observe that without life events, the allocations with the highest value for EID tend to be the ones with a high allocation to SPIA-COLA while with life events, the allocations with the highest value for EID have a significant percentage invested in the VA.

Figure 3.6: Median bequest (MB) and extreme income deficit (EID) for the three-asset universe made of a SPIA-COLA, a BF and a VA



This figure reports graphically the MB and EID statistics described at the beginning of Section 3.5 for a 65-year-old female individual with a $500,000 initial wealth and a three-asset universe made of a single premium immediate annuity with a 2% COLA, a variable annuity and a balanced fund. We consider an initial target withdrawal rate of 4%. The target withdrawal rates grow by 2% per year. We apply a grid weight step of 1% to simulate the 5,151 corresponding strategies. We report the results both in the absence and presence of life events.

Table 3.13: Optimal strategies for the three-asset universe made of a SPIA-COLA, a BF and a VA

	Optimal allocation = (% SPIA-COLA, %VA, %BF) -Initial Wealth=$500,000						
RA level	**0**	**0.5**	**1**	**2**	**4**	**6**	**1000**
Without LE	(0, 0, 100)	(0, 6, 94)	(0, 79, 21)	(100, 0, 0)	(100, 0, 0)	(100, 0, 0)	(100, 0, 0)
With LE	(0, 0, 100)	(0, 0, 100)	(0, 35, 65)	(0, 65, 35)	(0, 65, 35)	(0, 69, 31)	(1, 69, 30)

This table reports the optimal allocation for different levels of risk aversion, an initial target withdrawal rate of 4%, an initial wealth of $500,000 and a three-asset universe made of a single premium immediate annuity with a 2% COLA, a variable annuity and a balanced fund. The target withdrawal rates grow by 2% per year. We report the results both in the absence and in the presence of life events.

The results in Table 3.13 show that the optimal allocations in the absence of life events are a full investment in the balanced fund for a risk aversion coefficient λ equal to 0, an allocation of 6% to the VA and 94% to the balanced fund for a risk aversion coefficient equal to 0.5 and an allocation of 79% to the VA and 21% to the balanced fund for a risk aversion coefficient equal to 1. For risk aversion coefficients λ greater than or equal to 2 and in the absence of life events, the optimal portfolio is a full investment into the SPIA-COLA. In all cases where life events are taken into account, the optimal SPIA-COLA weight is zero, and for risk aversion coefficients λ greater than or equal to 1, the weight of VA evolves between 35% for $\lambda = 1$ and 69% for $\lambda = 1,000$. We assess that for an investor with a large risk aversion level ($\lambda \geq 2$) the optimal allocation is strongly impacted by the presence of life events: when long-term care needs are accounted for, a mix that consists of 65% of VA and 35% of a balanced fund is the optimal portfolio for λ between 2 and 6, while a 100% SPIA-COLA allocation proves to be optimal in the absence of life events. Overall, these results confirm again that the VA is a good substitute for both the SPIA-COLA (because of a similar downside protection) and the balanced fund (because of a similar upside potential), and its presence forces these two assets out of the optimal portfolio.

Table 3.14: Computation Time of Numerical Simulations

Simulation	Number of Files	Computation Time (min)	Comp.Time per File (min)
BC (No LE)	303	8.31	0.027
BC (With LE)	303	58.00	0.191
RC1 (No LE)	0	NA	NA
RC1 (With LE)	202	41.14	0.204
RC2 (No LE)	303	8.52	0.028
RC2 (With LE)	303	55.18	0.182
RC3 (No LE)	5,151	142.99	0.028
RC3 (With LE)	5,151	950.17	0.184

This table reports the number of files generated, the computation time and the average computation time per file in minutes of the numerical simulations performed for different settings in section 3.5. RC stands for robustness check and LE for life events.

3.5.2.4 Computation Time of the Simulations

The optimization problem studied in the previous section in a universe with n assets for a given risk aversion level λ is:

$$\arg \max_{w_1,\ldots,w_n} [\text{MB}(w_1,\ldots,w_n) + \lambda \text{EID}(w_1,\ldots,w_n)] \tag{3.16}$$

We report in Table 3.14, for all the different numerical simulations performed in section 3.5, the computational time (in minutes) spent to produce the output files necessary to determine the optimal solution of the optimization problem.

To determine the optimal solution of the optimization problem written in Equation 3.16 in a two-asset universe, for a given initial wealth and initial target withdrawal rates, accounting (or not) for life events, on a weight grid with a 1% step, it requires to run 10 000 Monte Carlo simulations for each of the 101 points of grid in order to evaluate the welfare function value on each of the 101 points of the grid. Since we consider a grid with a 1% step for the admissible weights, w_1 and $w_2 = 100\% - w_1$ vary between 0% and 100% with a 1% step and with the constraint that $w_1 + w_2 = 1$. We then have to evaluate a function with one variable on 101 points to determine its optimum. For each point of the grid, we use a code written on the R programming language that produces a file where all the necessary outputs to calculate the welfare function value on this point are stored.

To generate the results for the base case we consider that the individual (a 65-year-old female) is endowed with \$500,000 and that she has access to a simple investment universe that contains a balanced fund with a 50%/50% stock/bond annually rebalanced mix, and an immediate annuity with a 2% cost-of-living adjustment (SPIA-COLA). In this base case analysis, we systematically test three levels of initial target withdrawal rates, namely 3%, 4% and 5%. For each of the three initial target withdrawal rates, we have to evaluate the welfare function on the 101 points of the grid. We generate $3 \times 101 = 303$ files for the base case without accounting for life events and also 303 files for the base case accounting for life events. We assess that the average computation time per file is significantly higher when we account for life events (0.191 minutes) than when we do not (0.027 minutes). In the case where we account for life events, we need to account for stochastic target withdrawals in the code: this can explain this additional computation time.

For first the robustness check (RC1), we studied the impact of changes in the initial wealth level with three possible values of \$250,000, \$500,000 and \$1,000,000. We still consider that the individual (a 65-year-old female) has access to a simple investment universe that contains a balanced fund with a 50%/50% stock/bond annually rebalanced mix, and an immediate annuity with a 2% cost-of-living adjustment (SPIA-COLA). We assume an initial target withdrawal rate of 4%. We already generated the files we need for an initial wealth level of \$500,000 when we

generated the results for the base case. Moreover, when we do not account for life events, the optimal allocation will be independent of initial wealth since the replacement income needs of the individual will represent the same deterministic percentage of her initial wealth. As such, we do not need not generate any additional files for RC1 when we do not account for life events. However, when life events are taken into account, their relative costs will have an impact, which will be stronger for individuals with a lower initial wealth. We need to generate 101 files for an initial wealth level of \$250,000 and also 101 files for an initial wealth level of \$1,000,000. The average computation time per file is 0.204 minutes which is very close and consistent with the average computation time per file of the base case. We studied the impact of changes in the investment universe in a second robustness check (RC2) by considering three alternative two-asset universes in addition to the base case universe made of a SPIA-COLA and a balanced fund, for an individual with a \$500,000 initial wealth and a 4% initial target withdrawal rate. We need to generate 101 files for each of the three other three alternative two-asset universes when we do not account for life events and repeat the exercise when we account for life events. We find computation times that are consistent with those of the base case: the average computation time per file is still significantly higher when we account for life events (0.182 minutes) than when we do not (0.028 minutes).

We finally enlarged the base case universe by adding a variable annuity product in the first robustness check (RC3) where we considered a 65-year-old female individual with a \$500,000 initial wealth and a three-asset universe made of a SPIA with a 2% COLA, a variable annuity and a balanced fund. We performed the simulations for an initial target withdrawal rate of 4%.The optimal solution of the optimization problem written in Equation 3.16 in a three-asset universe made of a SPIA-COLA, a VA and a BF, for a given initial wealth and initial target withdrawal rates, accounting (or not) for life events, on a weight grid with a 1% step requires to run 10 000 Monte Carlo simulations for each of the $102 \times 101 \times \frac{1}{2} = 5151$ points of grid in order to evaluate the welfare function value on each of the 5151 points of the grid. Indeed if we assume that the number of assets in the universe is 3 and consider a grid with a 1% step for the admissible weights, w_1, w_2 and $w_3 = 100\% - w_1 - w_2$ vary between 0% and 100% with the constraint that $w_1 + w_2 + w_3 = 1$. We then have to evaluate a function with 2 variables on $\sum_{i=1}^{101} i \times (101 - i) = 102 \times 101 - \frac{102 \times 101}{2} = \frac{102 \times 101}{2} = 5151$ points to determine its optimum. As for the previous cases described above, the average computation time per file is significantly higher when we account for life events (0.184 minutes) than when we do not (0.028 minutes).

Overall we note that (1) for the base case and all the robustness checks the average computation time per file is significantly higher when we account for life events, (2) when we do not account for life events the average computation time per file is roughly the same for the base case and all the robustness checks and (3) the average computation time per file is significantly higher when we account for life events and is roughly the same for the base case and all the robustness checks.

We acknowledge that, when accounting for life events, generating all the files needed to determine the optimal portfolio can become burdensome in term of computation time as soon as the number of assets in the universe is higher than or equal to 3 and for a thin grid with a grid step lower than or equal to 1%. Two possible extensions to speed up the computation time are (1) to take the randomness of life events more efficiently in the code and (2) to develop an algorithm that can find the optimal solution or a weight vector that is close to the optimal solution on the whole grid without having to run Monte Carlo simulations on all the points of the grid. We leave these extensions for further research.

3.6 Conclusion and Extensions

In the presence of longevity risk, the annuitization problem is best solved by purchasing a life annuity. Our analysis shows, however, that the costly reversibility of annuitization decisions leads individuals facing life event uncertainty to lower their optimal demand for annuities. This chapter is primarily concerned with the question of optimal investment decisions in decumulation taking as given a stream of income withdrawal cash-flows. A useful extension of the framework would relate to a joint optimization of investment and withdrawal decisions. The original model of Merton (1971), which addresses the joint optimization of investment and consumption decisions, predicts that the optimal consumption level is a function of both current wealth and market conditions, so that the agent must adjust consumption to changes in wealth and to changes in parameters such as interest rates, volatilities and risk premia. This analysis, however, abstracts away from the presence of minimum and/or target levels of replacement income, and cannot be directly used to develop an actionable decision making process for individuals in decumulation. Many heuristic withdrawal rules exist such as the 3% (or 4%) rule, amongst others (see Suri et al. (2020)). In this context, it would be desirable to analyze whether rule-based state-dependent investment and withdrawal strategies can lead to substantial welfare improvements with respect to simpler static rules. In practice, one key source of complexity in this problem is the presence of accounts with multiple tax regimes, the presence of other sources of income that have an indirect impact on the tax treatment of the managed wealth, as well as the presence of relocation decisions in retirement that may impact the tax efficiency of investment and withdrawal strategies. Another avenue for future research would be to repeat the optimization exercise by introducing a financial asset that provides a fixed income without the constraints of annuities to see if the presence of life events has consequences on its optimal weight. In particular it has been argued that retirement goal-hedging liquid bond portfolios can be used to generate replacement income for a fixed-period in retirement. Such retirement goal-hedging bond portfolios, or retirement bonds in short, do not exist as off-the-shelves fixed-income products.[19] On the other hand, they can be replicated by standard dynamic factor exposure matching strategies (typically duration-matching but also duration/convexity-matching or duration/slope matching strategies) and recent research (Kobor and Muralidhar (2018) and Martellini et al. (2019)) show substantial benefits from replacing standard bond holdings by retirement bond holdings in investor portfolios. We leave these developments for further research.

3.A Pricing of Liabilities and Insurance Contracts

In this Appendix, we ask what is the minimum capital that should be invested to hedge given liabilities or a given insurance contract. The answer depends on the chosen hedging criterion, and we examine two criteria: hedging on average, and hedging almost surely.

3.A.1 Notations

Uncertainty is represented by a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{P}$ being the physical probability measure, which models investors beliefs. The finite time span is the range $[0, T]$, where T is an integer number. Liabilities are represented by a sequence of payments occurring at dates $1, 2, ..., T$, with values $\ell_1, \ell_2, \ldots, \ell_T$. We require that for all $t = 1, \ldots, T$, ℓ_t is nonnegative almost surely

[19] A series of recent articles in the financial and general press have made a case for the issuance by governments and other public or semi-public institutions of retirement bonds, sometimes also called SeLFIES for Standard of Living indexed, Forward-starting, Income-only Securities (see Muralidhar (2015); Muralidhar et al. (2016); Merton and Muralidhar (2017); Martellini et al. (2018); Kobor and Muralidhar (2018)). These bonds would enjoy the following two main characteristics: (1) payments are deferred to the retirement date, and (2) interest payment and capital amortization are spread over time in such a way that the annual income paid by the bond is constant or preferably cost-of-living-adjusted.

(a.s.), and takes on positive values with non-zero probability. The value of assets at time t just after a payment, is denoted with A_t, and the value just before the payment is denoted with A_{t-}. On a payment date, asset value jumps downwards and we have

$$A_t = A_{t-} - \ell_t$$

Examples include:

- Pension liabilities for a defined-benefit pension fund with a finite horizon T . Then, payments are typically either constant or indexed on some inflation metrics
- An annuity contract with payments conditional on an individual being alive. Then, each payment has the form
$$\ell_t = c_t \times \mathbb{1}_{\{\tau > t\}}$$
where τ is the time of death and c_t is the payment, to be made only if the individual is alive. T is the maximum longevity, beyond which all individuals are assumed to pass away.

Assets are invested in some (self-financing) fund, whose return between dates s and t is denoted with $r_{s,t}$. We have, for each $t = 1, \ldots, T$,

$$A_t = A_{t-1}[1 + r_{t-1,t}] - \ell_t$$

Induction then shows that the value of assets at time t after the payment is given by

$$A_t = A_0[1 + r_{0,t}] - \sum_{u=1}^{t} \ell_u[1 + r_{u,t}] \tag{3.17}$$

3.A.2 Hedging Liabilities on Average

The first hedging criterion that we consider is to have an expected value of zero for assets after the very last payment:

$$\mathbb{E}[A_T] = 0 \tag{3.18}$$

Substituting Equation (3.17) in this condition and rearranging terms, we obtain the following equivalent condition:

$$A_0 = \sum_{u=1}^{T} \frac{\mathbb{E}[\ell_u[1 + r_{u,T}]]}{\mathbb{E}[1 + r_{0,T}]}$$

This equation simplifies if the payments are independent (in the probabilistic sense) of the returns of the underlying investment. In the case of an annuity with fixed payments, this condition is equivalent to the condition that the death time should be independent from the returns. Under the independence assumption, Equation (3.18) is equivalent to

$$A_0 = \sum_{u=1}^{T} \mathbb{E}[\ell_u] \frac{\mathbb{E}[1 + r_{u,T}]}{\mathbb{E}[1 + r_{0,T}]} \tag{3.19}$$

Define now the annualized expected return between dates s and t as

$$\begin{aligned} \rho_{s,t} &= \mathbb{E}[1 + r_{s,t}]^{\frac{1}{t-s}} - 1 \text{ if } s < t, \\ \rho_{t,t} &= 0 \end{aligned}$$

Then, Equation (3.19) can be rewritten as

$$A_0 = \sum_{u=1}^{T} \mathbb{E}[\ell_u] \frac{[1+\rho_{u,T}]^{T-u}}{[1+\rho_{0,T}]^T} \tag{3.20}$$

To simplify further this equation, define $k_{u,T}$ as

$$[1+\rho_{0,T}]^T = [1+k_{u,T}]^u [1+\rho_{u,T}]^{T-u}$$

Equation (3.20) can be rewritten as

$$A_0 = \sum_{u=1}^{T} \frac{\mathbb{E}[\ell_u]}{[1+k_{u,T}]^u}$$

In the special case where $\rho_{u,T}$ equals $\rho_{0,T}$ for each u ranging from 1 to $T-1$, this equation becomes

$$A_0 = \sum_{u=1}^{T} \frac{\mathbb{E}[\ell_u]}{[1+\rho_{0,T}]^u} \tag{3.21}$$

This equation provides some justification for the calculation of the present value of liabilities by discounting expected payments (or the payments themselves if they are nonstochastic) with the expected return on the asset portfolio. But it is derived by imposing a weak hedging criterion, which is that the expected value of terminal assets should be zero. We consider in what follows a much stronger criterion.

3.A.3 Hedging Liabilities Almost Surely

We now require that final assets should be nonnegative with 100% probability:

$$A_T \geq 0 \text{ a.s.} \tag{3.22}$$

To calculate the minimum capital to invest to have this condition satisfied, we make the following two assumptions:

1. For each $t = 1, \ldots, T$, there exists a zero-coupon bond that pays ℓ_t at date t. Let $d_{s,t}$ denote the price of this bond at date s for $s \leq t$
2. There are no arbitrage opportunities.

A consequence of the second assumption that we will use below is that if S_1 and S_2 are two securities whose returns between two dates s and t satisfy

$$\frac{S_{1t}}{S_{1s}} \geq \frac{S_{2t}}{S_{2s}} \text{ a.s.}$$

then they actually have equal returns, i.e.,

$$\frac{S_{1t}}{S_{1s}} = \frac{S_{2t}}{S_{2s}} \text{ a.s.}$$

Equation (3.22) implies that

$$\frac{A_{T-}}{A_{T-1}} \geq \frac{\ell_T}{d_{T-1,T}} \times \frac{d_{T-1,T}}{A_{T-1}}$$

Suppose for a while that $A_{T-1} < d_{T-1,T}$. Then, we have

$$\frac{A_{T-}}{A_{T-1}} \geq \frac{\ell_T}{d_{T-1,T}} \quad \text{a.s.}$$

and the inequality is strict with positive probability because ℓ_T is positive with positive probability. So, the asset portfolio outperforms the last zero-coupon bond almost surely, and outperforms it strictly with positive probability. This violates the assumption of no arbitrage, so it must be the case that

$$A_{T-1} \geq d_{T-1,T}$$

It follows that $A_{[T-1]-} \geq \ell_{T-1} + d_{T-1,T}$
hence that

$$\frac{A_{[T-1]-}}{A_{T-2}} \geq \frac{\ell_{T-1} + d_{T-1,T}}{d_{T-2,T-1} + d_{T-2,T}} \times \frac{d_{T-2,T-1} + d_{T-2,T}}{A_{T-2}}$$

The first factor in the right-hand side is the return between dates $T-2$ and $T-1$ of the basket that contains the zero-coupons maturing respectively at dates $T-1$ and T. By the same no-arbitrage argument as above, the asset portfolio cannot outperform this basket almost surely, so it must be the case that

$$A_{T-2} \geq d_{T-2,T-1} + d_{T-2,T}$$

By induction on k, it is shown that for all k ranging from 0 to T , we have

$$A_{T-k} \sum_{i=0}^{k-1} d_{T-k,T-i}$$

In particular, we have

$$A_0 \geq \sum_{t=1}^{T} d_{0,t} \tag{3.23}$$

By Equation (3.23), the minimum capital to invest in order to hedge liabilities almost surely is the sum of the prices of the zero-coupon bonds that deliver the payments. This quantity is called the present value of liabilities and is denoted with L_0. When payments are non-stochastic, the zero-coupon bonds are nominal zero-coupon bonds, so we can rewrite the present value of liabilities in the more familiar form

$$L_0 = \sum_{t=1}^{T} \ell_t b_{0,t} \tag{3.24}$$

It is the sum of future payments discounted at the nominal zero-coupon rates. For an annuity, payments are conditional on an individual being alive. Because individual longevity risk is (in principle) not hedgeable with financial securities, we are in an incomplete market case, and the zero-coupon bonds with payoffs matching the payments do not exist.

3.A.4 Actuarial Value of Liabilities

We define the actuarial value of liabilities as

$$L_0^{\text{act}} = \sum_{t=1}^{T} \mathbb{E}[\ell_t] D_{0,t} \tag{3.25}$$

where $D_{0,t}$ is a discount factor for the maturity t known at date 0. It can be seen that the actuarial value of liabilities equals the present value of liabilities when the payments are non-stochastic (so that they equal their expectations) and the discount factors are taken equal to the prices of nominal zero-coupon bonds. The actuarial value also coincides with the minimum capital to invest to hedge liabilities on average provided the discount factors are taken to be

$$D_{0,T} = \frac{1}{[1+\rho_{0,T}]^t} \tag{3.26}$$

where $\rho_{0,T}$ is the expected return on the portfolio in which assets are invested (see Equation (3.21)).

3.B Solution of the Continuous-Time VAR Model

We define $Z_t = \begin{bmatrix} \log S_t \\ \log B_t \\ r_t \end{bmatrix}$ and consider the continuous-time VAR model defined as

$$dZ_t = (A + BZ_t)dt + \sigma' dW_t \tag{3.27}$$

with $A = \begin{bmatrix} \sigma_S \lambda_S - \frac{\sigma_S^2}{2} \\ \sigma_B \lambda_B - \frac{\sigma_B^2}{2} \\ ab \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -a \end{bmatrix}$ and σ such that $\sigma'\sigma = \Sigma = \begin{bmatrix} \sigma_S^2 & \sigma_{SB} & \sigma_{Sr} \\ \sigma_{SB} & \sigma_B^2 & \sigma_{Br} \\ \sigma_{Sr} & \sigma_{Br} & \sigma_r^2 \end{bmatrix}$

where Σ is the covariance matrix of the random vector Z_t. σ is determined by Cholesky decomposition.

We multiply each side of Equation (3.27) by $\exp(-tB)$ where for a given matrix M of size $(n \times n)$, $\exp(M)$ is defined as:

$$\exp(M) = \sum_{k=0}^{+\infty} \frac{M^k}{k!}$$

We have:

$$\exp(-tB)\, dZ_t - \exp(-tB)\, BZ_t dt = \exp(-tB)\left(Adt + \sigma' dW_t\right) \tag{3.28}$$

We define the function f as:

$$\begin{aligned} f : \mathbb{R}^+ \times \mathbb{R}^3 &\to \mathbb{R}^3 \\ (t, z) &\mapsto f(t,z) = \exp(-tB)\, z \end{aligned}$$

We apply the three-dimensional version of the Ito lemma to $f(t, Z_t)$:

$$d\left(\exp(-tB)\, Z_t\right) = (-B)\exp(-tB)Z_t dt + \exp(-tB)dZ_t \tag{3.29}$$

We use Equation (3.29) in Equation (3.28) and integrate between t and $t+h$:

$$\exp(-(t+h)B)\, Z_{t+h} - \exp(-tB)\, Z_t = \int_t^{t+h} \exp(-sB)Ads + \int_t^{t+h} \exp(-sB)\sigma' dW_s \tag{3.30}$$

By multiplying each side of Equation (3.30) by $\exp((t+h)B)$, we obtain:

$$Z_{t+h} = \exp(hB)\, Z_t + \exp((t+h)B) \left[\int_t^{t+h} \exp(-sB) A ds + \int_t^{t+h} \exp(-sB) \sigma' dW_s \right] \tag{3.31}$$

$$= \exp(hB)\, Z_t + \int_t^{t+h} \exp((t+h-s)B) A ds + \int_t^{t+h} \exp((t+h-s)B) \sigma' dW_s \tag{3.32}$$

$$= \exp(hB)\, Z_t + \int_0^{h} \exp(uB) A du + \int_0^{h} \exp(uB) \sigma' dW_u \tag{3.33}$$

where the last equality is obtain via the change of variable $t + h - s = u$.

We finally obtain the discretized equation of the continuous-time VAR model:

$$Z_{t+h} = \Psi_0 + \Psi_1 Z_t + \varepsilon_{t+h}$$

where $\Psi_0 = \left(\int_0^h \exp(sB) ds \right) A$, $\Psi_1 = \exp(hB)$ and $\varepsilon \sim \mathcal{N}\left(\mathbf{0}, \int_0^h \exp(sB)\, \sigma' \sigma \exp(sB')\, ds \right)$.

Bibliography

Amblard, P. O. and J. F. Coeurjolly (2011). Identification of the multivariate fractional Brownian motion. *IEEE Transactions on Signal Processing 59*(11), 5152–5168.

Amblard, P. O., J. F. Coeurjolly, F. Lavancier, and A. Philippe (2010). Basic Properties of the Multivariate Fractional Brownian Motion. arXiv preprint arXiv:1007.0828.

Amenc, N., F. Goltz, L. Martellini, and P. Retkowsky (2011). Efficient Indexation: An Alternative to Cap-Weighted Indices. *Journal of Investment Management 9*(4), 1–23.

Ameriks, J., A. Caplin, S. Laufer, and S. Van Nieuwerburgh (2011). The Joy of Giving or Assisted Living? Using Strategic Surveys to Separate Public Care Aversion from Bequest Motives. *The Journal of Finance 66*(2), 519–561.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2009). High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence. *Journal of Financial Economics 91*(1), 1–23.

Arnott, R. and R. Lovell (1993). Rebalancing: Why? when? how often ? *Journal of Investing 2*(1), 5–10.

Asmussen, S. (1998). *Stochastic simulation with a view towards stochastic processes*. Stochastic simulation with a view towards stochastic processes. University of Aarhus. Centre for Mathematical Physics and Stochastics (MaPhySto)[MPS].

Bachelier, L. (1900). *Théorie de la spéculation*. Gauthier-Villars Paris.

Bacinello, A., E. Biffis, and P. Millossovich (2009). Pricing Life Insurance Contracts with Early Exercise Features. *Journal of Computational and Applied Mathematics 233*(1), 27–35.

Bacinello, A., E. Biffis, and P. Millossovich (2010). Regression-Based Algorithms for Life Insurance Contracts with Surrender Guarantees. *Quantitative Finance 10*(9), 1077–1090.

Bacinello, A., P. Millossovich, A. Olivieri, and E. Pitacco (2011). Variable Annuities: a Unifying Valuation Approach. *Insurance: Mathematics and Economics 49*(3), 285–297.

Balduzzi, P. and W. Lynch (1999). Transaction Costs and Predictability: Some Utility Cost Calculations. *Journal of Financial Economics 52*(1), 47–78.

Bauer, D., D. Bergmann, and R. Kiesel (2010). On the Risk-Neutral Valuation of Life Insurance Contracts with Numerical Methods in View. *Astin Bulletin 40*(1), 65–95.

Bauer, D., A. Kling, and J. Russ (2008). A Universal Pricing Framework for Guaranteed Minimum Benefits in Variable Annuities. *Astin Bulletin 38*(2), 621–651.

Bernstein, W. and D. Wilkinson (1997). Diversification, rebalancing, and the geometric mean. *Frontier*. Research manuscript.

Biagini, F., Y. Hu, B. Øksendal, and T. Zhang (2008). *Stochastic calculus for fractional Brownian motion and applications*. Springer Science and Business Media.

Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy 81*(3), 637–654.

Bodie, K., A. Kane, and A. Marcus (1999). *Investments*. McGraw-Hill, 4th Edition.

Booth, D. and E. Fama (1992). Diversification Returns and Asset Contributions. *Financial Analysts Journal 48*(3), 37–57.

Bouchey, P., V. Nemtchinov, A. Paulsen, and D. Stein (2012). Volatility harvesting: Why does diversifying and rebalancing create portfolio growth? *Journal of Wealth Management 15*(2), 26–35.

Brigo, D. and F. Mercurio (2001). *Interest rate models: theory and practice (Vol.2)*. Berlin: Springer.

Brown, J. R. and A. Finkelstein (2011). Insuring Long-Term Care in the United States. *Journal of Economic Perspectives 25*(4), 119–42.

Brown, J. R., O. S. Mitchell, and J. M. Poterba (2001). The Role of Real Annuities and Indexed Bonds in an Individual Accounts Retirement Program. In *Risk Aspects of Investment-Based Social Security Reform*, pp. 321–370. University of Chicago Press.

Brown, J. R. and M. J. Warshawsky (2001). Longevity-Insured Retirement Distributions From Pension Plans: Market and Regulatory Issues. Technical report, National Bureau of Economic Research.

Brown, R. (1828). A brief account of microscopical observations on the particles contained in the pollen of plants, and on the general existence of active molecules in organic and inorganic bodies.

Chambers, D. and J. Zdanowicz (2014). The limitations of diversification return. *Journal of Portfolio Management 40*(4), 65–76.

Chapman, D. A., J. B. Long, and N. D. Pearson (1999). Using proxies for the short rate: When are three months like an instant? *Review of Financial Studies 12*(4), 763–806.

Charupat, N. and M. A. Milevsky (2002). Optimal Asset Allocation in Life Annuities: A Note. *Insurance: Mathematics and Economics 30*(2), 199–209.

Cheridito, P. (2003). Arbitrage in fractional brownian motion models. *Finance and Stochastics 7*, 533–553.

Choueifaty, Y. and Y. Coignard (2008). Toward Maximum Diversification. *Journal of Portfolio Management 35*(1), 40–51.

Craigmile, P. F. (2003). Simulating a class of stationary Gaussian processes using the DaviesHarte algorithm, with application to long memory processes. *Journal of Time Series Analysis 24*(5), 505–511.

Cuthbertson, K., S. Hayley, N. Motson, and D. Nitzche (2015). Diversification returns, rebalancing returns and volatility pumping. *Working Paper*.

Davidoff, T., J. R. Brown, and P. A. Diamond (2005). Annuities and Individual Welfare. *American Economic Review 95*(5), 1573–1590.

Davies, R. B. and D. S. Harte (1987). Tests for Hurst effect. *Biometrika 74*(1), 95–101.

De Nardi, M., E. French, and J. B. Jones (2010). Why Do the Elderly Save? The Role of Medical Expenses. *Journal of Political Economy 118*(1), 39–75.

DeMiguel, V., L. Garlappi, F. Nogales, and R. Uppal (2009). A Generalized Approach to Portfolio Optimization: Improving Performance By Constraining Portfolio Norms. *Management Science 55*(5), 798–812.

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.

Dichtl, H., W. Drobetz, and M. Wambach (2014). Testing rebalancing strategies for stock-bond portfolios across different asset allocation. HFRC Working Paper Series No.4-5.

Dickson, D., M. R. Hardy, and H. R. Waters (2013). *Actuarial Mathematics for Life Contingent Risks*. Cambridge University Press.

Duffie, D. (2001). *Dynamic Asset Pricing Theory*. Princeton: Princeton University Press.

Duncan, T. E., Y. Hu, and B. Pasik-Duncan (2000). Stochastic calculus for fractional Brownian motion I. Theory. *SIAM Journal on Control and Optimization 38*(2), 582–612.

Elliott, R. J. and J. Van Der Hoek (2003). A general fractional white noise theory and applications to finance. *Mathematical Finance 13*(22), 301–330.

Erb, C. B. and C. R. Harvey (2006). The strategic and tactical value of commodity futures. *Financial Analysts Journal 62*(2), 69–97.

Feldstein, M. and E. Ranguelova (2001). Individual Risk in an Investment-Based Social Security System. *American Economic Review 91*(4), 1116–1125.

Fernholz, R. (2002). *Stochastic Portfolio Theory*. Springer.

Fernholz, R. and B. Shay (1982). Stochastic portfolio theory and stock market equilibrium. *Journal of Finance 37*(2), 615–624.

French, E. and J. B. Jones (2004). On the Distribution and Dynamics of Health Care Costs. *Journal of Applied Econometrics 19*(6), 705–721.

Gabay, D. and D. Herlemont (2007). Benchmarking and rebalancing. Working Paper.

Gaillardetz, P. and J. Y. Lakhmiri (2011). A New Premium Principle for Equity-Indexed Annuities. *Journal of Risk and Insurance 78*(1), 245–265.

Geweke, J. and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis 4*(4), 221–238.

Hallerbach, W. (2014). Disentangling rebalancing return. *Journal of Asset Management 15*, 301–316.

Harjoto, M. and F. Jones (2006). Rebalancing strategy for stocks and bonds asset allocation. *Journal of Wealth Management 9*(1), 37–44.

Horneff, W. J., R. H. Maurer, O. S. Mitchell, and M. Z. Stamos (2009). Asset Allocation and Location over the Life Cycle with Investment-Linked Survival-Contingent Payouts. *Journal of Banking and Finance 33*(9), 1688–1699.

Horneff, W. J., R. H. Maurer, O. S. Mitchell, and M. Z. Stamos (2010). Variable Payout Annuities and Dynamic Portfolio Choice in Retirement. *Journal of Pension Economics and Finance 9*(2), 163–183.

Horneff, W. J., R. H. Maurer, and M. Z. Stamos (2008). Life-Cycle Asset Allocation With Annuity Markets. *Journal of Economic Dynamics and Control 32*(11), 3590–3612.

Hosking, J. R. (1984). Modeling persistence in hydrological time series using fractional differencing. *Water resources research 20*(12), 1898–1908.

Hu, Y. and B. Øksendal (2010). Fractional white noise calculus and applications to finance. *Infinite dimensional analysis, quantum probability and related topics 6*(01), 1–32.

Huang, D., M. Grove, and T. Taylor (2012). The Efficient Income Frontier: A Product Allocation Framework for Retirement. *The Retirement Management Journal*, 9–22. Link to RIIA presentation: http://riia-usa.org/pdfs/fc11-web/D-EIF-MattGrove.pdf.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American society of civil engineers 116*(1), 770–799.

Jaconetti, C., F. Kinniry, and Y. Zilbering (2010). Best practices for portfolio rebalancing. *Vanguard*, 1–17.

Jagannathan, R. and T. Ma (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *Journal of Finance 58*(4), 1651–1683.

Jordan, B. and T. Miller (2008). *Fundamentals of Investments*. McGraw-Hill, 4th Edition.

Kelly, J. (1956). A new interpretation of information rate. *Bell System Technical Journal 35*, 917–926.

Kobor, A. and A. Muralidhar (2018). How a New Bond Can Greatly Improve Retirement Security. Working paper.

Koijen, R., S. Van Nieuwerburgh, and M. Yogo (2016). Health and Mortality Delta: Assessing the Welfare Cost of Household Insurance Choice. *The Journal of Finance 71*(2), 957–1009.

Koijen, R. S., T. E. Nijman, and B. J. Werker (2011). Optimal Annuity Risk Management. *Review of Finance 15*(4), 799–833.

Kolmogorov, A. (1940). Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *C.R.(Doklady) Acad. URSS (N.S) 4*, 115118.

Latané, H. (1959). Criteria for choice among risky ventures. *Journal of Political Economy 67*, 144–155.

Lavancier, F., P. A. . S. D. (2009). Covariance function of vector self-similar processes. *Statistics & probability letters 79*(23), 2415–2421.

Lintner, J. (1965a). Security Prices, Risk, and Maximal Gains from Diversification. *The Journal of Finance 20*(4), 587–615.

Lintner, J. (1965b). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics 47*(1), 13–37.

Luenberger, D. (1997). *Investment Science*. Oxford University Press.

Maillard, S., T. Roncalli, and J. Teiletche (2010). The Properties of Equally Weighted Risk Contribution Portfolios. *Journal of Portfolio Management 36*(4), 60–70.

Mandelbrot, B. and J. Wallis (1968). Noah, joseph, and operational hydrology. *Water Resources Research 4*(5), 909–918.

Mandelbrot, B. B. and J. W. Van Ness (1968). Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review 10*(4), 422437.

Mantilla-Garcia, D. (2016). Maximizing the volatility return: A risk-based strategy for homogeneous groups of assets. *Working Paper*.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance 7*(1), 77–91.

Markowitz, H. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly 3*(1-2), 111–133.

Martellini, L., R. Merton, and A. Muralidhar (2018). Pour la création « d'obligations retraite ». *Le Monde,* samedi 7 avril 2018.

Martellini, L., V. Milhau, and J. Mulvey (2019). Flexicure Retirement Solutions: A Part of the Answer to the Pension Crisis? *Journal of Portfolio Management 45*(5), 136–151.

Maurer, R., O. S. Mitchell, R. Rogalla, and V. Kartashov (2013). Lifecycle Portfolio Choice with Systematic Longevity Risk and Variable InvestmentLinked Deferred Annuities. *Journal of Risk and Insurance 80*(3), 649–676.

Merton, R. (1969). Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *Review of Economics and Statistics 51*(3), 247–257.

Merton, R. (1971). Optimal Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory 3*(4), 373–413.

Merton, R. (1973a). An Intertemporal Capital Asset Pricing Model. *Econometrica 41*(5), 867–887.

Merton, R. (1973b). Rational Theory of Option Pricing. *Bell Journal of Economics and Management Science 4*(1), 141–183.

Merton, R. (1980). On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics 8*(4), 323–361.

Merton, R. and A. Muralidhar (2017). Time for Retirement 'SelFIES'? Working paper.

Milevsky, M. (1998). Optimal Asset Allocation Towards The End of the Life. Cycle: To Annuitize or Not to Annuitize? *The Journal of Risk and Insurance 65*(3), 401–426.

Milevsky, M. A. (2013). Life Annuities: An Optimal Product for Retirement Income. *CFA Institute*.

Milevsky, M. A. and A. Macqueen (2015). *Pensionize Your Nest Egg: How to Use Product Allocation to Create a Guaranteed Income for Life*. Wiley, Hoboken, NJ.

Mitchell, O. S., J. M. Poterba, and M. J. Warshawsky (1999). New Evidence on the Money's Worth of Individual Annuities. *The American Economic Review 89*(5), 1299–1318.

Muralidhar, A. (2015). New Bond Would Offer a Better Way to Secure DC Plans. Technical report, Pensions and Investments.

Muralidhar, A., K. Ohashi, and S. H. Shin (2016). The Most Basic Missing Instrument in Financial Markets: The Case for Forward Starting Bonds. *Journal of Investment Consulting 47*(2), 34–47.

Pang, G. and M. Warshawsky (2010). Optimizing the Equity-Bond-Annuity Portfolio in Retirement: The Impact of Uncertain Health Expenses. *Insurance: Mathematics and Economics 46*(1), 198–209.

Pang, G. and M. Warshawsky (2012). *Comparing Strategies for Retirement Wealth Management: Mutual Funds and Annuities*. MIT Press.

Pashchenko, S. (2013). Accounting for Non-Annuitization. *Journal of Public Economics 98*, 53–67.

Peijnenburg, K., T. Nijman, and B. J. Werker (2017). Health Cost Risk: A Potential Solution to the Annuity Puzzle. *The Economic Journal 127*, 1598–1625.

Peng, C., S. Buldyrev, S. Havlin, M. Simons, H. Stanley, and A. Goldberger (1994). Mosaic organization of dna nucleotides. *Physical Review E 49*(2), 1685–1689.

Pfau, W. (February 2013). A Broader Framework for Determining an Efficient Frontier for Retirement Income. *Journal of Financial Planning*, 44–53.

Plerou, V., P. Gopikrishnan, B. Rosenow, L. Nunes Amaral, and H. Stanley (1999). Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters 83*(7), 1471–1474.

Plyakha, Y., R. Uppal, and G. Vilkov (2012). Why Does an Equal-Weighted Portfolio Outperform Value-and Price-Weighted Portfolios? Technical report, EDHEC-Risk Institute Publication.

Politis, D. and J. Romano (1994). The stationary bootstrap. *Journal of American Statistical Association 89*, 1303–1313.

Qian, E. (2012). Diversification Return and Leveraged Portfolios. *Journal of Portfolio Management 38*, 14–25.

Qian, E. (2014). To rebalance or not to rebalance: A statistical comparison of terminal wealth of fixed-weight and buy-and-hold portfolios. Working Paper.

Reichling, F. and K. Smetters (2015). Optimal Annuitization With Stochastic Mortality and Correlated Medical Costs. *American Economic Review 105*(11), 3273–3320.

Roncalli, T. (2013). *Introduction to Risk Parity and Budgeting*. London: Taylor & Francis.

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *Journal of Finance 19*(3), 425–442.

Spinu, F. (2015). Buy-and-hold vs. constantly rebalanced portfolios: A theoretical comparison. *Journal of Asset Management 16*(2), 79–84.

Suri, A. M., N. Vrdoljak, Y. Liu, and R. Zhang (2020). Beyond the 4% Rule-Determining Sustainable Retiree Spending Rates. Merrill Lynch Internal Publication.

Tobin, J. (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies 25*(2), 65–86.

Tsai, C. (2001). Rebalancing diversified portfolios of various risk profiles. *Journal of Financial Planning 14*(10), 104–110.

Turra, C. M. and O. S. Mitchell (2008). The Impact of Health Status and Out-of-Pocket Medical Expenditures on Annuity Valuation. In *Recalibrating Retirement Spending and Saving*, pp. 227–250. Oxford University Press.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics 5*(2), 177–188.

Vernon, S. (2014). The Role of Annuities in Retirement. *The Journal of Retirement 1*(3), 63–76.

Vernon, S. (2015). Optimal Retirement Income Solutions in DC Retirement Plans. *Stanford Center on Longevity*.

Weron, R. (2002). Estimating long-range dependence: finite sample properties and confidence intervals. *Physica A 312*, 285–299.

Willenbrock, S. (2011). Diversification return, portfolio rebalancing, and the commodity return puzzle. *Financial Analysts Journal 67*(4), 42–49.

Wise, A. (1996). The Investment Return From a Portfolio With a Dynamic Rebalancing Policy. *British Actuarial Journal 2*(4), 975–1001.

Yaari, M. E. (1965). Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *The Review of Economic Studies 32*(2), 137–150.