



HAL
open science

Locality in time of the European insurance regulation "risk-neutral" valuation framework, a pre-and post-Covid analysis and further developments

Fabrice Borel-Mathurin, Nicole El Karoui, Stéphane Loisel, Julien Vedani

► To cite this version:

Fabrice Borel-Mathurin, Nicole El Karoui, Stéphane Loisel, Julien Vedani. Locality in time of the European insurance regulation "risk-neutral" valuation framework, a pre-and post-Covid analysis and further developments. 2020. hal-02905181

HAL Id: hal-02905181

<https://hal.science/hal-02905181v1>

Preprint submitted on 23 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Locality in time of the European insurance regulation “risk-neutral” valuation framework, a pre- and post-Covid analysis and further developments

Fabrice BOREL-MATHURIN * Nicole EL KAROUI † Stéphane LOISEL ‡
Julien VEDANI §¶

July 23rd, 2020

Abstract

The so-called *market-consistency* of the European life insurance valuation as shaped by regulation guidelines embeds numerous theoretical and practical misstatements. Since El Karoui et al. (2017) the manipulation risk induced by the framework imprecision and, in particular, its high dependency to regulatory and non-regulatory calibration data is clear. In this paper we update some results and analysis of El Karoui et al. (2017) using data from a more recent “classical” year (2017) and from an exceptional year (first quarter of 2020, with Covid-19 effects), and test additional sensitivities. Based on the updated values we obtain up to -45% in the *VIF* estimates values depending on the swaption implied volatilities matrix used to calibrate the interest rates model. Then trying different calibration sets we obtain up to 105% difference. In parallel, we see that using 3-month averages to calibrate Economic Scenario Generators do not make effects of crises like Covid-19 disappear. We then address the “simulation seed” setting issue, and the interest and limits of keeping the same seed when estimating and comparing economic valuations, be it on horizontal (comparing valuations through time) or vertical (studying sensitivities at the same date) analysis. We finally open our study to propose various tools for a better risk management of economic scenarios and valuation, through a better understanding of Asset-Liability Management models.

*Autorité de Contrôle Prudentiel et de Résolution, Paris, France, Email: Fabrice.borel-mathurin@acpr.banque-france.fr

†Université Pierre et Marie Curie - Paris 6, Laboratoire de Probabilité, Statistique et Modélisation, Paris, France, Email: nicole.el_karoui@upmc.fr

‡Université de Lyon, Université Claude Bernard Lyon 1, Laboratoire de Science Actuarielle et Financière, Institut de Science Financière et d’Assurances, Lyon, France, Email: stephane.loisel@univ-lyon1.fr

§Milliman, Paris, France, Email: julien.vedani@milliman.com

¶Université de Lyon, Université Claude Bernard Lyon 1, Laboratoire de Science Actuarielle et Financière, Institut de Science Financière et d’Assurances, Lyon, France, Email: julien.paul.vedani@gmail.com

Key words Risk-neutral valuation, economic valuation, market-consistency, European regulation, simulation seed.

Introduction

Along with the development of financial markets, since the 70s, the market price notion has been widely employed and considered as a fair trading price. Theory has also helped this evolution thanks to the development of risk-neutral valuation for financial options (Black & Scholes (1973), Merton (1973),...), works that overlap with economy general equilibrium theory (Arrow & Debreu (1954)) and portfolio theory (Markowitz (1952)).

Since the mid-90s the fair value concept has been widely used by accounting practices (Nelson (1996), Barth & Landsman (1995),...), especially through International Financial Reporting Standards (IFRS, the most recent international accounting norms). It is commonly assumed that the fair value of an asset or liability can be seen as the price agreed upon by a willing buyer and seller, with equivalent full knowledge on the traded product. Obviously, as soon as the asset or liability is traded on a complete financial market, its fair value (accounting speaking) should be its market value.

Finally, a 2002 European Union Parliament regulation (European Parliament (2002)) states that, for fiscal years starting after January 2005, the IFRS norms apply on every European listed company, and all balance sheets should be displayed in fair value. However, in practice, fair valuation is not that easy to apply to full balance sheets and the IFRS norms have evolved a lot since their initial texts. For example, concerning IFRS 13 *Fair Value Measurement*, see International Accounting Standards Board (2011), three drafts have been released before the final issue in May 2011 which has already been amended twice. One of the main difficulty is to fair price life insurance liabilities (mixing IFRS 4: *Insurance contracts*, International Accounting Standards Board (2004), and IFRS 9: *Financial Instruments*, International Accounting Standards Board (2014)), an issue not clearly resolved yet.

Directive Solvency II (European Commission (2009)), applied in Europe since 2016, adapts the fair valuation to life insurance regulation by introducing the economic valuation framework, an accounting paradigm where every items are estimated under a *market-consistent* and so-called *risk-neutral* measure (see Groupe Consultatif Actuariel Européen (2012), Wüthrich et al. (2008), Wüthrich & Bühlmann (2010)). The valuation scheme progressively implemented from the Directive and European Insurance and Occupational Pensions Authority (EIOPA) guidelines leads to numerous theoretical and practical issues, some of them having been raised year ago (see *e.g.* Vesa & Lasse (2007), Huerta de Soto (2009)), and more recently in El Karoui et al. (2017). In this last paper, the authors underline various relatively easy manipulation processes, mostly in line with the high complexity level of financial models used by actuarial practitioners, disconnected from their original applications in finance. These processes lead to a framework excessively difficult to manage, control and regulate.

In this paper we first zoom on the European insurance regulator texts framing the *market-consistency* of the valuation scheme introduced by directive Solvency II. Considering the limits of this valuation framework, in particular its lack of stability and strong dependency to calibration parameters, we then focus on one of El Karoui et al. (2017)) propositions to stabilise the obtained values. The considered proposition (presented in Subsection 2.2.3 of previous paper) was tested on end-of-year 2014 data with very impacting results. In Section 2 we update this implementation and the analysis to end-of-year 2017 data. Then we test additional sensitivities for the updated method in order to refine our analysis and conclusions. We also investigate, using 2020 data, whether events like Covid-19 would be "erased" by the 3-month smoothing technique if it occurred at the end of a calendar year. We also focus on the

interest of keeping similar simulation seeds when calculating and comparing economic valuations (a fact already raised in Borel-Mathurin & Vedani (2019)) through empirical tests. Then, in Section 4, we address deeper the impact of the simulation seed conservation from a more theoretical viewpoint. In final Section 4 we enlarge our study to propose different risk management tools to better understand (and question) the impact of economic scenarios on the Asset-Liability Management (ALM) model used by practitioners when achieving an economic valuation.

1 Life insurance valuation in Solvency II - from regulation theory to practice

1.1 Quick reminder on risk-neutrality in theory and practice

Solvency II regulation imposes the market price as the reference value for the different balance sheet elements of insurance companies operating in the European Union. This fair valuation of the balance sheet is explicitly mentioned in the Solvency II Directive, in the recitals and in article 76 & 77¹. On the asset side this represents only minor difficulties since quoted prices represent most of the securities booked by insurance companies. Their investments are hardly done outside products listed on active markets. On the liability side this is another story: prices cannot be based on exchanges like on the asset side. Due to their legal specificity, insurance contracts are not meant to be transferred nor hedged², even in theory. In this context, those products cannot be valued, and provide any fair price, in a risk-neutral framework, be it in theory or in practice. In absence of a no-arbitrage pricing measure, one needs to use expedients to produce these liability prices in these incomplete markets. A market-consistent insurance premium can for example be based on the utility indifference as in Malamud et al. (2008). However, those product "prices" can neither be really considered as fair nor, more obviously, risk-neutral. For all these reasons the economic values should better (in all generality) be considered as insurance regulation-specific than risk-neutral. As a consequence, instead of "market-consistency" with its broad meaning, a better wording for the setup here described would be "insurance market-consistency".

Despite the different caveats above, we still use the risk-neutral denomination to talk about economic valuation to fit the European actuarial regulation and operational wording.

1.2 Valuing insurance provisions as a financial option

Based on the IFRS fair-value first insights, directive Solvency II tries to frame the definition of a *market value of insurance technical provisions* and introduces the *Best Estimate*, a Monte-Carlo estimation of the *option-value* of these provisions. From Article 77 - Calculation of technical provisions - paragraph 2,

“The best estimate shall correspond to the probability-weighted average of future cash-flows,

¹Recital 54 of the Solvency II Directive requests that "The calculation of technical provisions should be consistent with the valuation of assets and other liabilities, market consistent [...]"

²Which is highly required when using risk-neutral schemes.

taking account of the time-value of money (expected present value of future cash-flows), using the relevant risk-free interest rate term structure.”

This lays the foundations of economic valuation, an insurance-specific adaptation of financial option valuation theory. Indeed, the price of a financial option is divided into two parts, its intrinsic value, the value the option would provide if exercised today, and its time-value (or extrinsic value), which reflects the potential future gains, for the owner, if exercised later. The time-value is the price of the optionality associated to an option contract. In a similar fashion it is possible to see the time value of insurance technical provision as their optionality price.

Article 77 paragraph 2 then continues with,

“The calculation of the best estimate shall be based upon up-to-date and credible information and realistic assumptions and be performed using adequate, applicable and relevant actuarial and statistical methods.”

This sentence is in line with Article 76 - General provisions - paragraph 3,

“The calculation of technical provisions shall make use of and be consistent with information provided by the financial markets and generally available data on underwriting risks (market consistency).”

The underlying idea here is that the economic values (of provisions, but also own funds) should be consistent with financial markets data. This also introduces (from practitioners and regulators interpretation) the idea of risk-neutral valuation, that is assumed “financial market standards”. This market-consistency is still today the main (though quite loose) criteria used to value and control economic valuation efficiency.

However, this criteria is also limited by the idea of *relevant risk-free interest rate term structure*, which leaves the possibility to introduce a regulatory-specific, non market-based, yield curve.

1.3 The regulatory interest rates term structure

The relevant risk-free interest rates term structure has long been an issue for both practitioners and regulators. If not regulatory-specified the market-consistent yield curve would have been an open door to various in-homogeneous practices, among European countries but also between companies of the same country. First, the concept of *risk-free* is unclear for European insurance companies (EURIBOR, EONIA, other LIBOR maturities, etc.). Even financial markets are complex to reconcile around the same interest rates (see *e.g.* Rebonato & Pogudin (2011)).

Various yield curves have been tested until the idea of using a swap (Euroswap) curve plus an *illiquidity premium* in the 5th Quantitative Impact Study (European Insurance and Occupational Pensions Authority (2010a)) of the European Insurance and Occupational Pensions Authority (EIOPA, former CEIOPS, Committee of European Insurance and Occupational Pensions Supervisors). Finally, the relevant risk-free interest rates term structure, used by all European countries, is regulatory-fixed. It is

adjusted from the Euroswap term structure, as of end-of-year of each yearly valuation, and published by the EIOPA (with additional curves after each end of month, for intermediary calculations, between years N and $N + 1$).

As can be seen in El Karoui et al. (2017), the regulator curve is based on the market swap curve as at end-of-day with additional adjustments (see European Insurance and Occupational Pensions Authority (2017) and Figure 1 for the impact of these adjustments on the 12/31/17 yield curve):

- minus a few bps due to the Credit Risk Adjuster (“CRA” =-10bps applied on swap rates, in 2017)
- plus a few bps due to the Volatility Adjuster (“VA” =+4bps applied on Zero-Coupon rates, in 2017),
- then, from maturity 20 years (the “Last Liquid Point” of the swap rates) to 60 years, the regulator uses a Smith-Wilson algorithm for the 1-year forward rates to converge towards an “Ultimate Forward Rate” (“UFR” =4.2%, in 2017).

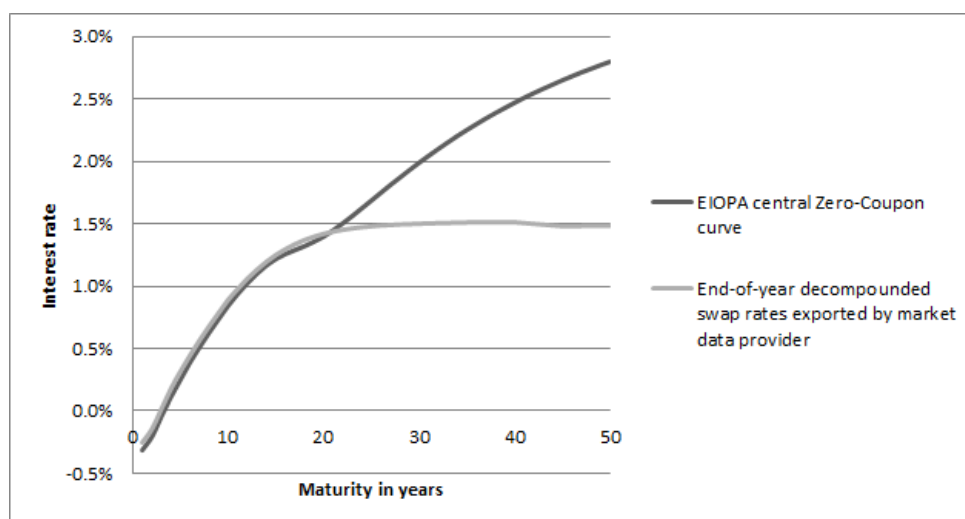


Figure 1: Yield curves as at end-of-year 2017: EIOPA ZC yield curve vs. Euroswap decomposed curve

Note that, in reality, two different yield curves are used by practitioners depending on which balance sheet part is economically valued. The liability best estimate is valued using economic scenarios simulated with the *with VA* EIOPA curve but the asset part is valued using scenarios simulated with the *without VA* EIOPA curve. Eventually, two different economic scenarios tables must be used to assess economic own funds.

1.4 Market-consistency from the insurance regulation point of view

There is one crucial problem with the use of a non market-based term structure. Market prices (in particular those of interest rates options and derivatives) are assessed conditionally to these rates. So it is

impossible to stay consistent with both market implied data and optionality pricing. Interest rates models are generally calibrated based on at-the-money receiver swaption implied volatilities (IV) but with the regulatory term-structure, it is impossible to keep both IV and market prices data.

This problem was already raised in European Insurance and Occupational Pensions Authority (2010b), with the following practitioner's question (question 76 of section *Technical provisions Best estimate*),

“The answer to [question 21] asks us to include the appropriate liquidity premium for both projecting and discounting. In this way, assets roll up and get discounted at the same rate. *The answer does not give any indication of whether or not we accept that option prices will change.*”

With the following answer from EIOPA,

“Your understanding of the answer to question 21 is correct: the appropriate illiquidity premium for the valuation of the liabilities has to be included both for projecting and for discounting the assets, thus assuring that the assets are rolled up and discounted with the same rates.

- According to [QIS 5], asset models should be calibrated to a risk-free interest rate curve that includes an illiquidity premium - and thus differs from the risk-free term structure implicit in the market price of some options.
- The convention in the over-the-counter option market is to use swaps as risk-free rates. As QIS5 is based on a different relevant risk-free rate, market option prices and market IV can no longer be replicated simultaneously.
- ***The asset models should nevertheless be market-consistent*** and comply with [QIS 5].
- The market-consistency of the asset models that no longer reproduce observable market prices can be demonstrated in a two stage approach. In the first stage relatively simple closed form solutions can be parameterized to match the market value of observable options using the swap rate, i.e. the market implied discount rate. These closed form solutions and the same parameters [*i.e.* IV] should then be reused with the relevant QIS 5 risk-free rate to establish ***theoretical market values*** consistent with the definition of risk-free used in the valuation of the liabilities in QIS 5. These theoretical market values can then be used ***to validate the market consistency*** of the liability valuation approach by confirming that the liability approach adequately reproduces those theoretical market values.”

This notion of “theoretical market values” is ambiguous. The underlying idea is to estimate regulatory based prices by coupling the EIOPA interest rates and market swaption IV (under the Black/lognormal or Bachelier/normal paradigm) to obtain alternative (“theoretical”) market values through (the respective) closed-form formulas (see *e.g.* Brigo & Mercurio (2006) for swaptions Black formula).

This specific market-consistency understanding seem to be undertaken by the regulator. The conclusion is that the optionality pricing of the insurance technical provisions has a very complex and unpredictable link with true market option prices.

1.5 Market-consistent calibration in practice - the interest rates model case

Today, the European life insurance market has integrated the market consistency as a criteria (see Kemp (2009) for a definition, adapted to insurance, Sheldon & Smith (2004), van Bragt et al. (2010), Malamud et al. (2008) or Wüthrich & Bühlmann (2010) for theoretical uses) to quantify the efficiency of its valuations. Practitioners consider a calibration process of their stochastic models that allows to reprice market IV through Monte-Carlo simulation (see *e.g.* Devineau et al. (2017) for some theoretical calibration issues and El Karoui et al. (2017) for a practical preview and scheme manipulability evidences). The stochastic models are then used to simulate various financial drivers (equity indexes, Zero-Coupon bonds, Inflation-linked bonds, spreads, etc.) through long-term horizons and compute economic scenarios tables. Asset-Liability Management models valuating the economic balance sheet of life insurance companies then use these market-consistent tables as inputs and estimate each simulation x projection year balance sheet item cash-flow. an economic value (of assets, liabilities or own fund) is then obtained as an intrinsic part plus a discounted average of these future cash-flows.

Concerning interest rates, the European life insurances use various stochastic models to simulate economic scenarios. In line with the recent negative short term swap rates in Europe, the models that seem to be favoured by the European regulator support the simulation of negative interest rates, such as the Displaced Diffusion Libor Market Model (DDLMM), with or without stochastic volatility (see Rebonato (2002), Joshi & Rebonato (2003), Joshi et al. (2003), etc.), the Displaced Diffusion Black Karasinski (direct displacement of the Black-Karasinski model from Black & Karasinski (1991)), the Hull-White model (see Hull & White (1994)), etc.

The market consistent calibration of a stochastic (risk-neutral) interest rate model consists often in the choice of a couple of:

- one interest rate term structure (“the relevant risk-free interest rate term structure.” Directive Solvency II, article 77, paragraph 2) and
- one set of market swaption IV (often matrices as of end-of-day 12/31 of the calculation year).

Then, according to the interest rates model chosen, its parameters are calibrated so as to replicate the calibration swaption volatility set conditionally to the regulatory-term structure.

2 Use of a monthly averaged IV matrix for interest rate calibration

We have chosen now to focus on a specific tool proposed by El Karoui et al. (2017) to improve the stability of prices obtained through insurance regulation valuations, and decrease the manipulability of the scheme: the use of monthly-averaged IV sets/matrices when calibrating interest rates models.

2.1 The El Karoui et al. (2017) calibration process - Results as of end-of-year 2014

In Subsection 2.2.3 of their paper, El Karoui et al. shape an “adapted market-consistent constraint” for interest rate models calibration aiming to provide more stable and robust values. They test four

calibrations as of end-of-year 2014 (*i.e.* with the corresponding regulator interest rates term-structure) on different swaption IV.

- The IV averaged on the whole month of October 2014. Only the receiver swaptions of maturity 5, tenors 1 to 10, and the receiver swaptions of tenor 5, maturities 1 to 10.
- The IV averaged on the two months of October and November 2014. Only the receiver swaptions of maturity 5, tenors 1 to 10, and the receiver swaptions of tenor 5, maturities 1 to 10.
- The IV as of end-of-year 2014. Only the receiver swaptions of maturity 10, tenors 1 to 10, and the receiver swaptions of tenor 10, maturities 1 to 10 (v2).
- And the IV as of end-of-year 2014. Only the receiver swaptions of maturity 5, tenors 1 to 10, and the receiver swaptions of tenor 5, maturities 1 to 10 (v2).

The authors add that both calibrations as of end-of-year “can be considered as market-consistent”.

The underlying idea is that averaged data can improve the stability of the valuations (compared to the use of one single day matrix) and reduce the dependency to the end-of-year huge trades made for accounting reasons (“turn-of the year effect”, well-known by traders, see *e.g.* Ritter (1988)). But, as the regulator yield curve is built from a market curve as of end-of-year, the “no arbitrage opportunity” has long been a reason to use the corresponding end-of-year swaption IV matrix, and a justification not to consider the El Karoui et al. matrix. This issue is discussed in Borel-Mathurin & Vedani (2019).

The four calibrated then simulated tables embed 1’000 economic scenarios (a standard number among practitioners, though probably not enough), without any use of variance reduction techniques, obtained from the same simulation seed. The interest rates model considered is a Libor Market Model, often used by French practitioners at that time. They are used to value three standard saving products³ economic own funds⁴, with the following results,

Table 1: Comparison between the obtained economic own funds values - 2014 results

	October 14’	Oct. & Nov. 14’	12/31/14 v1	12/31/14 v2
portfolio nb 1	16’898	15’614	7’046	10’000
portfolio nb 2	12’826	12’283	9’517	10’000
portfolio nb 3	12’553	12’073	6’050	10’000

As it can be seen here, the results are very volatile and strongly depend on the calibration swaption IV. We can see a gap of about 140% between the end-of-year v1 calibration and the October 14’ averaged IV calibration for portfolio 1. This underlined the great difficulty to assess stable values in the European economic valuation framework.

³See the paper for more information about the portfolios tested.

⁴The economic own funds valuation presents the same issues as for the economic liabilities Best Estimate as the market value of the Asset part, mostly traded on financial markets, in the economic balance sheet is quite obvious.

2.2 Update of the results to 2017 valuation

The update of the El Karoui et al. results needs additional adjustments. Indeed, the regulator term-structure now embeds negative zero-coupon interest rates for short term maturities (see Figure 1). Therefore, the Libor Market Model is not adapted to replicate this term-structure. We therefore use a Displaced Diffusion Libor Market Model (DDLMM, see Joshi & Rebonato (2003)) with a shift parameter of 2.5%. The LMM assumes the forward rates are log-normal. The DDLMM assumes the displaced forward (forward rate +2.5% here, with a displacement factor *a priori* fixed but not calibrated⁵) is log-normal, so that simulated interest rates can reach negative values until -2.5%. In addition, to achieve a better convergence of the estimated economic values, the simulated tables embeds 2'000 economic scenarios that integrate antithetic variables. Concerning the asset-mix of this 2017 ALM model the proportions are the following,

Table 2: Initial asset-mix of the ALM model

Asset	Proportion
Stock	15.4%
Obligations	78.2%
Inflation-linked obligations	3.9%
Cash	2.5%

The results displayed in Table 1 consider complete economic Own Funds. In practice the key-variable to be considered should rather be the Value of In-Force (*VIF*), that is the stochastic part, the *time value* of the economic own funds. We therefore analyse here only *VIF* results

Note that the savings product used here is similar to the product behind portfolio nb 1 with updated characteristics. It is still a representative product of the French market. Considering the economic scenarios we simulate zero-coupon bonds (ZCB, DDLMM) but also a stock index (Black-Scholes with deterministic by term volatility) and Inflation Linked Zero-Coupon Bonds (ILZCB, 2-Factor Vasicek). For all these models see Brigo & Mercurio (2006) for theoretical and Institut des Actuaire (2019) for practical (actuarial) insights. These model choices are very standard among practitioners.

Finally, considering the DDLMM interest rates model, we consider five different calibrations.⁶

- One on the IV averaged on the whole month of October 2017. Only the receiver swaptions of maturity 5, tenors 1 to 10, and the receiver swaptions of tenor 5, maturities 1 to 10.
- One on the IV averaged on the whole month of November 2017. Only the receiver swaptions of maturity 5, tenors 1 to 10, and the receiver swaptions of tenor 5, maturities 1 to 10.
- One on the IV averaged on the whole month of December 2017. Only the receiver swaptions of maturity 5, tenors 1 to 10, and the receiver swaptions of tenor 5, maturities 1 to 10.

⁵This model and shift value are quite common among practitioners.

⁶We keep the same calibration data, standard as at end-of-year 2017 for the two other stochastic models. Keeping the same simulation seed we know the only difference between our *VIF* estimations will be due to our different DDLMM calibration choice (and not to any sample bias on stock or ILZCB in particular).

- One on the IV averaged on the whole two months of October and November 2017. Only the receiver swaptions of maturity 5, tenors 1 to 10, and the receiver swaptions of tenor 5, maturities 1 to 10.
- And one on the IV as of end-of-year 2017. Only the receiver swaptions of maturity 5, tenors 1 to 10, and the receiver swaptions of tenor 5, maturities 1 to 10.

The Values of In-Force estimates are provided in Table 3, with the Monte-Carlo 95%-confidence intervals.

Table 3: Comparison between the obtained *VIF* values - 2017 results

	October 17'	November 17'	December 17'	Oct. & Nov. 17'	12/29/17 v2
VIF	6'118 ($\pm 2'069$)	10'186 ($\pm 1'873$)	9'484 ($\pm 1'932$)	5'508 ($\pm 2'103$)	10'000 ($\pm 1'894$)

First, it is relevant to note the low level of convergence of our results. Though we have doubled the number of scenarios and used antithetic variables the valuations stay highly unstable. One of the reasons of such instability is that we consider *VIF* values, the pure stochastic part of economic own fund, not including the revalued net asset (the intrinsic part of the option price) which would add some inertia to the own funds values. We underline here the necessity for practitioners to question the number of simulations they consider and to further analyse the confidence intervals of their results. This raises a crucial issue (particularly for control authorities) of such valuation scheme, the choice of the simulation seed and corresponding potential manipulations. Anyway, remind the same seeds are used for all tables, which should somehow lead to reliable differences between the results.

This updated implementation provides interesting results. In particular, we see that, contrary to 2014, the averaged IV approaches for October 17' and the 2-months period October and November 17' are much lower than the standard results using end-of-year IV (up to -82%) . In addition, the 2-months, when separated, lead to higher *VIF* values. This underlines the unpredictability of such approach impact and its non-linearity.

It also provides a wide illustration of the valuation framework lack of stability. Many different calibrations can be implemented, with unpredictable impact, not talking about diffusion model changes impact. Another complexity lies in the fact that each company applies its own subjective choices, leading to a global tricky control and no comparability of the assessed values, be it economic liabilities, own funds, but also required capital and solvency ratios. The scheme is simply too complex and lacks of homogeneity among undertakings.

To conclude, we acknowledge we base these results and differences on a single example and they can depend on the product characteristics. However, the El Karoui et al. paper has shown a relative stability of observations when tested on other products.

2.3 Enlarging the analysis to different squared IV sets

In practice, actuaries often prefer to consider a higher number of IV when calibrating their model, though using models embedding a limited number of parameter to calibrate leads to instability in the calibra-

tion (overparametrization). Thus we try to test the same implementation as in Subsection 2.2 but now comparing three sets of calibration swaptions IV (see Figure 2 for easier visualisation),

- set 1: only the receiver swaptions of maturity 5, tenors 1 to 10, and the receiver swaptions of tenor 5, maturities 1 to 10 (same set as in Subsection 2.2),
- set 2: the full receiver swaptions squared matrix of maturity 1 to 10 / tenors 1 to 10,
- set 3: the full receiver swaptions squared matrix of maturity 1 to 10, 15 and 20 / tenors 1 to 10, 15 and 20.

maturities \ tenors	1	2	3	4	5	6	7	8	9	10	15	20
1	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border: 1px solid black; padding: 5px;">set 2</div> <div style="border: 1px solid black; padding: 5px;">set 1</div> <div style="border: 1px solid black; padding: 5px;">set 3</div> </div>											
2												
3												
4												
5												
6												
7												
8												
9												
10												
15												
20												

Figure 2: The three calibration IV sets

The Values of In-Force estimates are provided in Table 4, with the Monte-Carlo 95%-confidence intervals.

Table 4: Comparison between the obtained *VIF* values - 2017 results for various calibration swaption IV sets

	October 17'	November 17'	December 17'	Oct. & Nov. 17'	12/29/17
Set 1	6'118 (±2'068)	10'186 (±1'873)	9'484 (±1'932)	5'507 (±2'102)	10'000 (±1'893)
Set 2	9'418 (±1'935)	9'581 (±1'943)	10'189 (±1'922)	9'400 (±1'946)	10'718 (±1'851)
Set 3	10'609 (±1'862)	10'916 (±1'850)	10'971 (±1'861)	10'768 (±1'853)	11'369 (±1'838)

These results provide two essential information.

First switching from set 1 to set 2 drives to much higher stability in valuation. Probably due to more inertia in the calibrated parameters chosen so as to fit an overall squared matrix rather than a “cross” that enables any movements of the non-replicated values for each corner of the matrix. But also leading to less dependency in time of the IV. The final results are then more “purely” dependent to the used yield curve.

Then, the 11% increase on average between set 2 and set 3 is also relevant. The major differences between market swaption prices and *regulation-specific* prices are captured by set 3 but not by set 2 (see Figure 3 from Borel-Mathurin & Vedani (2019)). We therefore assume the use of such IV, associated

to swap rates modified by the UFR convergence, drives the values upward. The impact of the UFR on simulated yield curves (through long term horizons, here 30 years) seem evident ⁷ but this underlines the *impact on calibration* the UFR convergence can have.

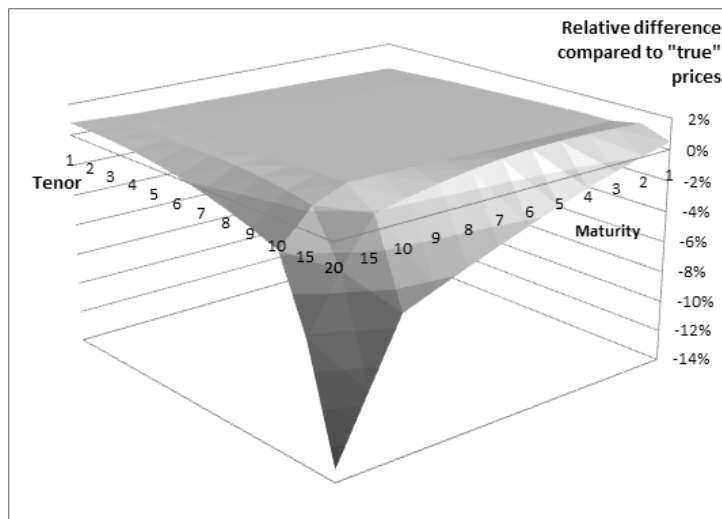


Figure 3: Differences between market and *regulatory-specific* swaption prices

Finally, practitioners are incited in practice to calibrate their models so as to stabilise their estimates. According to our results, they should be encouraged both to use the averaged approach, for less dependency in time of the market-consistent valuations and a large enough calibration IV matrix.

2.4 Quantification of the regulator term-structure adjustments (CRA+VA+UFR) impact on valuation

As it is aforementioned, the simulation impact of using the UFR-adjusted regulator term-structure seem evident and its impact through calibration is also significant. In this Section, we try to quantify the overall impact of the regulator term structure adjustments, including the CRA, VA and UFR, on valuation, compared to the use of the crude swap yield curve.

We therefore apply the same process used in Subsection 2.3, testing both set 2 and set 3 for calibration IV, but using the crude Euroswap term-structure (on which the regulator rates term-structure is built). We obtain the results displayed in Table 5, with the Monte-Carlo 95%-confidence intervals.

Table 5: Comparison between the obtained *VIF* values - use of the crude swap yield curve

	October 17'	November 17'	December 17'	Oct. & Nov. 17'	12/29/17 v2
Set 2	5'231 ($\pm 2'103$)	5'830 ($\pm 2'067$)	11'617 ($\pm 1'778$)	5'730 ($\pm 2'067$)	7'373 ($\pm 1'949$)
Set 3	4'994 ($\pm 2'074$)	5'091 ($\pm 2'057$)	5'682 ($\pm 2'032$)	5'266 ($\pm 2'052$)	5'633 ($\pm 2'041$)

⁷In particular, between $t=20$ and $t=30$ years the simulated curves, using a risk neutral model, swing around the forward curve, stable and equal to the UFR value, 4.2%, for high maturities. This naturally has a strong impact on valuations.

These results are astonishing. On the first hand we tend to obtain relatively stable lower values around 6'000 ($\pm 2'000$), which is in line with what was expected from not considering the UFR convergence. These low values, when compared to those obtained using the regulatory term-structure underline the huge gains induced by the regulation yield curve, both lying in the calibration and simulation process. On the other hand, we have obtained an *a priori* outlier in the 'Set1 - December 17', 11'617 ($\pm 1'778$) value, the only one twice bigger than others. We have first thought of a mistake in our implementation and used a second valuation table generated using a different seed but we obtain a similar price, 11'152 ($\pm 1'803$).

Finally, we must admit this framework is even more surprising than we thought. We probably have reached a calibration abnormality leading to a specifically high *VIF*. Such events may happen in particular with such over-parameterized calibration (minimisation) programs. It is almost impossible to predict such occurrences (assumed quite rare in practice) but it is interesting to note such potential behaviour of the estimated economic values.

2.5 A recent update of our implementation on Covid-19 crises data

In order to complete our testing of the averaged matrixes calibration method we have focused on very recent data as of 03/31/2020 with various averaging time ranges. The scenarios table embeds a stock index simulated with very high implied volatilities as at 03/31/20, Eventually the *VIF* estimated (standard market way) is negative : $-8'385 (\pm 2'322)$. This is highly conjectural, the 12/31/19 was close to $+19'000...$ The values in 6 are obtained for different interest rates model calibrations (but keeping the high stock volatilities and other simulation parameters).

Table 6: Comparison between the obtained *VIF* values as at 03/31/20 - IR model calibrated on different averaged 20x20 swaption implied volatilities matrixes)

time ranges	01/20	02/20	01/20 to 02/20	01/20 to 03/20
<i>VIF</i> value	$-6'095 (\pm 2'172)$	$-6'105 (\pm 2'202)$	$-5'640 (\pm 2'165)$	$-7'085 (\pm 2'242)$

Some interesting facts can be deduced from these results. First thing is the major impact of this crises on the value. This underlines the high dependency of market-consistency on financial markets and the high risk induced by making the whole market solvency dependent of one market day. Imagine if end of March was end-of-year.

Then note the relative impact of the smoothed data, with only a reduced decrease of the crisis impact, in particular the 3-month smoothing. This test is only applied to interest rate model calibration, with still an impact on ALM parameters and other models calibration. This shows the smoothing technique still keeps the impact of significant market event while lowering their effect. This is therefore a reasonable counter-cyclical tool to enforce the stability of the solvency assessment while keeping a consistency with market data and movements.

2.6 Back to 2017 - Testing the simulation seed impact on *VIF* differences

It is clear at first that our presented results suffer from great volatility (high confidence intervals), which tends to question the interest of the observed differences between the values, and to some extent our analysis. The answer often operationally provided to this problem is that we used the same simulation seed for all our valuations. Therefore, a sample bias exists but it should be very close from one estimation to another. Finally, the observed differences between two *VIF* estimators “should” be relatively stable and significant.

In this Subsection, we have tried to empirically challenge this explanation. To do so we have focused on the two most extreme values obtained previously. Remarkably, these two values have been obtained using the same calibration yield curve (crude swap rates) and matrix set (set 2), one calibrated on the averaged October 10x10 matrix (5231 (± 2103)) and the other on the averaged December 10x10 matrix (11617 (± 1778)). According to the 95%-confidence intervals, the difference between both values, 6386, may as well be higher than 10000 or lower than 3000 in reality.

Let us now evaluate the 95%-confidence interval of the difference value Monte-Carlo estimator. We now obtain 6386 (± 500). As expected, we obtain a very low confidence interval for the two values obtained with the same simulation seed. And if we now check this asymptotic confidence interval (in particular the asymptotic assumption efficiency) by considering, this time, 50000 scenarios to evaluate the *VIF* instead of 2000, we obtain the following results (see table 7).

Table 7: Comparison between the obtained *VIF* values - use of the crude swap yield curve (50000 scenarios estimators with same sim. seed)

	October 17'	December 17'	Difference
Set 2	3'993 (± 440)	10'674 (± 371)	6'681 (± 101)

These results are in line with the previous estimators, with much lower confidence intervals due to the larger number of Monte-Carlo simulations, which tends to validate our initial assumption: even with a limited number of scenarios and limited efficient Monte-Carlo estimators, the estimated difference between estimators may be rather efficient, *as soon as the simulation seed used to generate the economic scenarios is the same*. Let us now test the same calculation but on 25'000 simulations using different simulation seeds (see results in table 8).

Table 8: Comparison between the obtained *VIF* values - use of the crude swap yield curve (25000 scenarios estimators with same vs. different seeds)

	October 17'	December 17'	Difference
Set 2 same seed	4'029 (± 621)	10'731 (± 523)	6'703 (± 145)
Set 2 different seeds	4'029 (± 621)	10'617 (± 526)	6'589 (± 815)

This is an example of what the confidence interval for the *VIF* difference would be if different seeds were used.

A last interesting point is to simply focus on the relatively small difference between the *VIF* values obtained for the calibration as at end-of-year 2017 using the EIOPA curve and calibration sets 2 and 3.

If we just focus on the *VIF* confidence intervals it may look very similar, even equal, but in fact the difference between the two values is very stable (see Table 9)

Table 9: Comparison between the obtained *VIF* values - use of the EIOPA yield curve (2000 scenarios estimators calibrated on set 2/set 3)

	set 2	set 3	Difference
<i>VIF</i> value at end-of-year 17' calibrated on EIOPA curve	10'718 (± 1851)	11'369 (± 1838)	651 (± 134)

Note that such seed impact efficiency is probably highly linked to the use of the same ALM algorithm, applied to the same product. The ALM rules and product characteristics being similar this leads to strong correlation of the Monte-Carlo estimated Net Present Values used to assess the *VIF*.

The interest of keeping the same simulation seed for more efficient comparability of estimations is well-known in many fields of simulation theories and practices (see *e.g.* one of the multiple application of seed setting in Bates et al. (2014)) but its efficiency for economic scenarios has great implications.

3 Advantages and drawbacks of keeping the same simulation seed - a theoretical viewpoint

Keeping the same seed corresponds in applied probability to so-called common random numbers or common random paths techniques. If we had only one risk driver and one time period, we could estimate the difference $f(X) - g(Y)$ between the Value of In Force (or another quantity of interest) in a central scenario and the one in a stressed scenario, or the difference between the VIFs in two models calibrated with two different datasets, thanks to Monte Carlo approach. As noted by Glasserman & Yao (1992), the variance of the estimator of $f(X) - g(Y)$ can be reduced if $f(X)$ and $g(Y)$ are positively correlated. In particular, if f and g are increasing functions, say, then it is relevant to choose the same "seed" to generate comonotonic versions of (X, Y) thanks to c.d.f. inversion.

In our context, we are interested in differences between two VIF which involve sample paths for which one generates some vector of increments or innovations. Even if there is no guarantee that variance is reduced in general, as noted by Glasserman & Yao (1992), there are theoretical cases in which the variance reduction is guaranteed. One of them corresponds to sensitivity analysis, when we want to compare the values of the output before and after a small change. In finance, physics and insurance risk management, such "common random path techniques" are widely used. Fournié et al. (1999), Tezuka (1998), Benhamou (2003) and Chen & Glasserman (2007), among many others have used it for the analysis of Greeks in mathematical finance, in relation to Malliavin calculus. Common random paths are also useful in physics (see for example Hall et al. (2016) for applications to generalised Langevin dynamics. Some actuarial models might include jumps, which can be handled with similar techniques (see for example Loisel & Privault (2009) for sensitivity analysis of ruin probabilities and integration by parts, or Loisel et al. (2009) who make the link between the sensitivity of ruin probabilities with respect to the claim intensity and the difference (that can be computed thanks to common random paths) between the

classical ruin probability and the associated ruin probability for a partly shifted risk process, which features one additional jump with the same jump size distribution as the other ones, occurring at a random instant uniformly distributed between time 0 and the (finite) time horizon.

For our problem, even if there is no theoretical guarantee that using common random paths reduces the variance of the studied difference, we have several empirical observations in favour of the method: first, sometimes, we may be in the small change context, particularly if we compare situations when using closely related calibration datasets or models. Second, as noted by Chauvigny et al. (2011), some quantities of interest may often present monotonicity and/or convexity with respect to some risk drivers like stock returns, interest rate levels, mortality rates, cumulated claim amounts or inflation. However, these characteristics may disappear in unusual times or if the complex hedging strategy and product guarantees offset them. Third, we tested the method on many types of insurance portfolios, asset mixes and market conditions, and the common random path method always reduced the variance of the estimator of the considered difference by an important factor.

Another advantage of keeping the same seed is the reproducibility of results, for Value of In Force for example. This makes financial communication easier, and it makes it easier for Executive Committee members and for members of the Administration, Management or Supervisory Board (AMSB) to become more familiar with the model outputs and their interpretation.

However, it also breeds manipulation risk. Actuaries or financial analysts could, with good or bad intentions, choose some seed that corresponds, from their point of view, to an "average" or "favourable" set of scenarios. It may also trigger suspicion even if nothing goes wrong: members of the AMSB or supervisors could suspect that the seed was chosen in an inappropriate way.

Another drawback corresponds to the fact that top management and board members could become less aware of the random nature of the outputs and of the uncertainty around the presented results. Executives receiving results produced with different seeds would notice immediately that the standard deviation of the estimated output is too large and that one has to find a way to carry out more simulations or to reduce variance with another way.

Fixing the seed corresponds in certain cases to fixing the set of scenarios. When the number of scenarios is small (hundreds or thousands) in large dimensions, this might trigger some inappropriate, dangerous optimisation strategy that is too much dependent on the seed. In reinsurance, a similar experience occurred in the 1990's and early 2000's for hurricane risk on the East coast of the United States of America: some brokers and tied agents started to underwrite houses that were in "blindspots", i.e. in zones that were not covered by any hurricane path in the few climate events database.

Using same seed and common random path make sense when using the same model. However, one would need different models for central and extreme scenarios, or a unique model that is particularly designed for accommodating extreme scenarios, which is not enough the case presently. Therefore, one

could rephrase the question about the seed addressed in this Subsection into "Should one keep the same (conditional) model for central and extreme scenarios?".

4 Open developments on risk management

During this whole section, we focus on the *VIF* calculation we estimated on standard insurance market calibration (IV as at end-of-year 2017 for the DDLMM model) and simulation (using the EIOPA regulatory yield curve as at end-of-year 2017) we have estimated in Subsection 2.3 based on the 20×20 IV matrix (Set 3). We chose to focus on this valuation because it is a standard regulation-specific one and because it seems the wider the calibration matrix we use the more stable the valuation. Recall the obtained *VIF* is $11'369 (\pm 1'838)$. However, a simple valuation is not sufficient to understand the underlying risk processes. We now want to enlarge the analysis to get more information on this value, on the reasons under this estimation, the way the underlying risks (interest rates, stock, real rates) are taken into account by the ALM model. In this Section, we want to focus on solvency regulation-adapted Enterprise Risk Management tools. It could be interesting to consider geometric quantiles of the risk drivers (seen as a random vector) or related concepts (see Chauvigny et al. (2011) for more details), but for the sake of concision we only consider here (1-dimensional, classical) quantiles of random variables which are functions of several risk drivers.

4.1 Understanding better the link between economic scenarios and ALM

Basically, this study background lies on the disconnection actuaries make between the Economic Scenarios Tables and the final estimated *VIF* values. From the basic definition of Enterprise Risk Management (ERM) its objective is to understand the way a company reacts to economic movements, both to figure out and quantify in which ways its own funds decrease (for risk management) but also increase (for strategic choices). In the book of Lam (2014) we note as soon as in its preface:

"Risk management is about balancing risk and rewards. Interestingly, the Chinese characters for risk are actually the combination of the characters for danger and opportunities."

The underlying idea is both sides are useful to understand the mechanic of one business.

As we consider a way to assess own funds using a Monte-Carlo estimation it seems very interesting to focus on the *NPV* (remember it is the interest variable on which every Monte-Carlo valuation is assessed).

Eventually it seems very interesting to focus on the way the ALM model produces the *NPV* (1-dimension indexes) and reacts to the various movements induces by the economic scenarios. We raise here the necessity for an efficient risk management not only to assess the *VIF* value (necessary to assess the regulatory own funds) but also to understand how this value is obtained. What is the impact of economic scenarios on the *NPV*.

Of course it is clear that this impact will depend on the ALM model choices, the model points, the valuation time, the EIOPA yield curve choices (UFR in particular). However, it seems we can easily study some scenarios impacts, read them, and artificially build sensitivities.

Before focusing on sensitivities let us come back to the variable NPV and to its link with its underlying economic scenarios. Each of our scenarios is a simulation through 30 years of

- a stock index (Black-Scholes – > 1 Brownian motion),
- ZCB curves on 30 maturities (DDLMM – > 2 Brownian motions),
- ILZCB curves on 30 maturities (2-factors Vasicek – > 2 Brownian motions).

So we can summarised each scenario by its 5 random sources through 30 years (as the stochastic drivers are discretized yearly this leads to using $5 \times 30 = 150$ correlated standard Gaussian outcomes - Brownian motion increments) - let us denote this 150 sample vector ε - but also by both its initial economic conditions, that is the calibration data, the initialisation ZCB curve (be it the EIOPA regulatory curve or the crude Euroswap curve, as used in Subsection 2.4 and 2.6) - let us denote this set of information ζ - and by the ALM model functions and parameters designed by the practitioners - let us denote this set of information ξ . Finally we can write any NPV as a function $f(\varepsilon, \zeta, \xi)$. This is a general setting and, in most cases (our Section's included), ζ and ξ are fixed so we write below $NPV(\varepsilon)$.

4.2 Working on extreme scenarios - first interpretations

We have assessed our estimation of $VIF = 11'369$ based on 2'000 economic scenarios so we have obtained 2'000 NPV . Let us now focus on the extreme NPV values, see which underlying scenarios lead to these values and try to interpret them. We present below the NPV sorted by increasing order and the corresponding economic scenarios.

Scenario	NPV	Scenario	NPV
1298	-194'551
1920	-179'370	885	71'111
153	-167'278	1206	71'600
206	-164'832	1677	72'880
502	-160'295	1502	75'299
...	...	1891	76'338

We now know which scenarios to focus on. Some other statistical moments of the NPV empirical distribution: standard deviation=37'048, skewness=-1'851. The heavy negative skewness is standard for such heavy left tail distributions (see Figure 4).

Let us now examine the 10 spotted scenarios. First, we see the relatively limited importance of the RIR impact which is mainly explained by the little part of inflation-linked Obligation in our product asset-mix (3.9%).

Concerning the worse scenarios we obtain Figure 5 showing the grand lines (stock and nominal rates, up and down risks, for more complete lines see the graphs Figure 10 and 11 in Appendix) allowing an easier interpretation. One directly sees the worse scenarios are associated to a balance between great and quick decrease of the stock index and very low, even negative, nominal rates. This enables to quantify the impact of this link between low IR and stock decrease on the variable of interest, NPV , level (up to $-19 \times$ the estimated VIF).

Concerning the top scenarios we obtain Figure 6. We see here one issue of using risk neutral simulations on so long time horizon (never done by market banking practitioners), this leads to unrealistic possibilities, which have little interest to assess solvency. In each of these top scenarios, we can see a

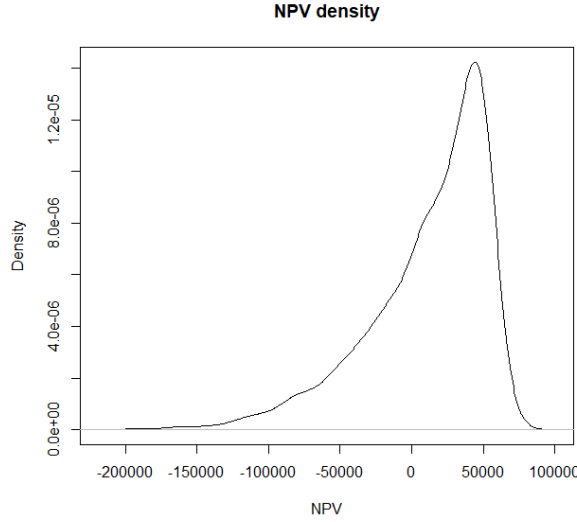


Figure 4: Density of the 2'000 sample empirical distribution

balance between very high and quick increase of the stock index and of nominal rates⁸. As expected the corresponding *NPV* are very high (almost 7 times the *VIF* value).

This is first interesting but general interpretation and any risk management specialist could quantify in depth, study the returns, stock scores and estimate liability Greeks, relevant option durations associated to these scenarios. This is a first easy tool to underline the interest studying the economic scenarios table more deeply.

4.3 Test on sensitivities scenarios based on the random sources

In this Subsection, we focus on *NPV* sensitivities by building artificial sensitivities scenarios, to quantify the impact of large movements of the drivers on the *NPV* valuations when they are inputted in the ALM model. A first naive idea is to focus on ε , the set of 150 standard Gaussian outcomes (yearly Brownian motion increments). Basically, one can write with evident notation,

$$\varepsilon = \left(\varepsilon_1^{stock}, \dots, \varepsilon_{30}^{stock}, \varepsilon_1^{IR_1}, \dots, \varepsilon_{30}^{IR_1}, \varepsilon_1^{IR_2}, \dots, \varepsilon_{30}^{IR_2}, \varepsilon_1^{RIR_1}, \dots, \varepsilon_{30}^{RIR_1}, \varepsilon_1^{RIR_2}, \dots, \varepsilon_{30}^{RIR_2} \right),$$

where the $(\varepsilon_i^{stock})_{i \in \llbracket 1;30 \rrbracket}$, $(\varepsilon_j^{IR_1})_{j \in \llbracket 1;30 \rrbracket}$, $(\varepsilon_j^{IR_2})_{j \in \llbracket 1;30 \rrbracket}$, $(\varepsilon_k^{RIR_1})_{k \in \llbracket 1;30 \rrbracket}$ and $(\varepsilon_k^{RIR_2})_{k \in \llbracket 1;30 \rrbracket}$ are the stock, nominal interest rates and real interest rates $\mathcal{N}(0, 1)$ outcomes, correlated when $k = j = i$ using the 5×5 Economic Scenarios Generator (ESG) correlation matrix, we call Σ below.⁹

⁸Which is nice for two reasons: first the 2017 yields are very low so an increase of nominal rates renders the minimum guaranteed rate easier to serve, and second, as the simulation measure is risk neutral, the nominal rate drives the stock index level. It must however not increase too much this would lead to higher lapse risk, etc.

⁹Note that, if we simply denote $(\varepsilon^s)_{s \in \llbracket 1;S \rrbracket}$ the S sets associated to a table of S economic scenarios, the choice to keep the same simulation seed for this table and for a second S scenarios table (if the models used are the same, but with no restriction on the calibration data used to parameterized the two tables) simply means the sets associated to the second table are the same,

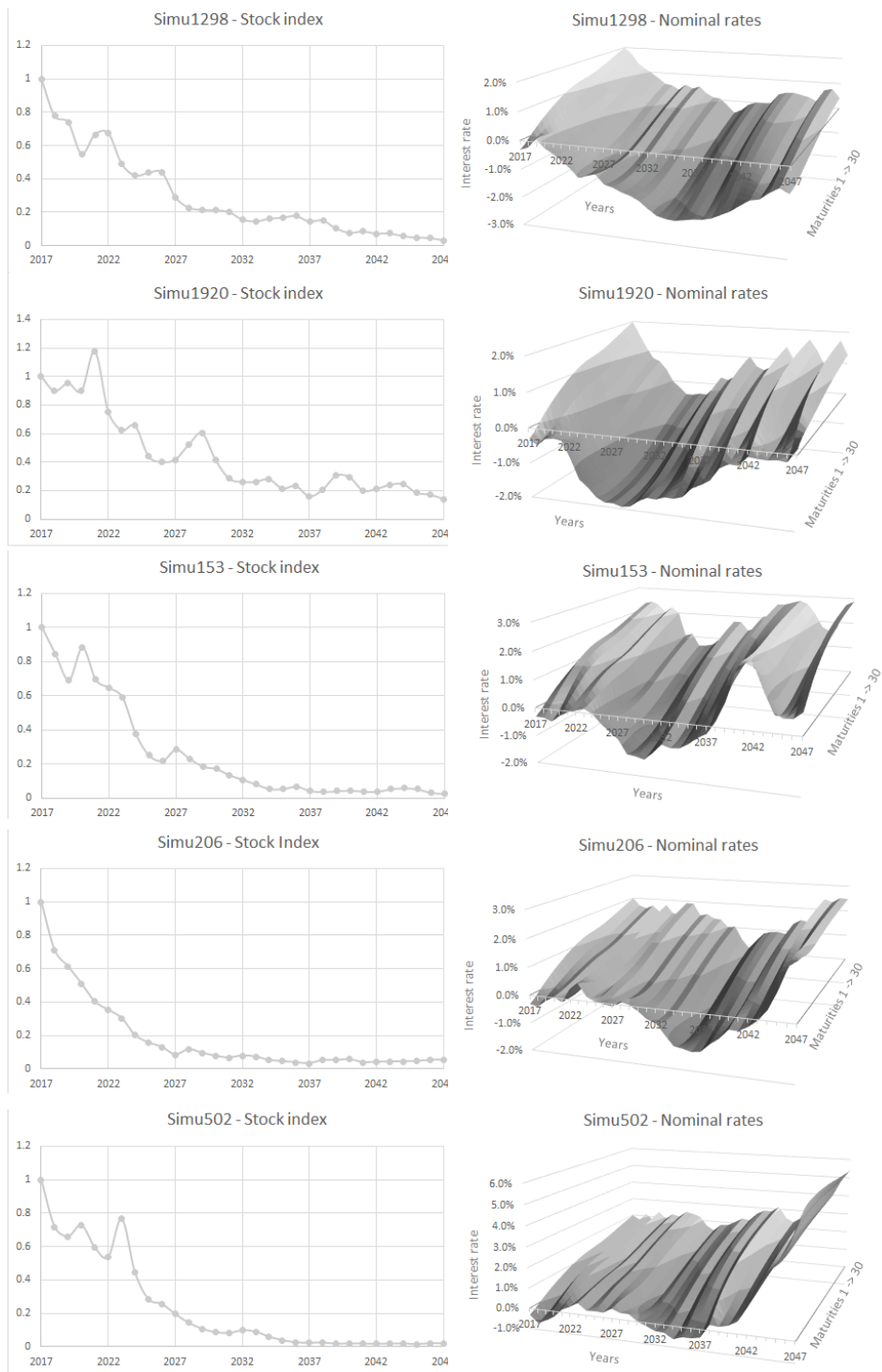


Figure 5: 5 worse scenarios ("SimuN" meaning simulation/scenario number N) information

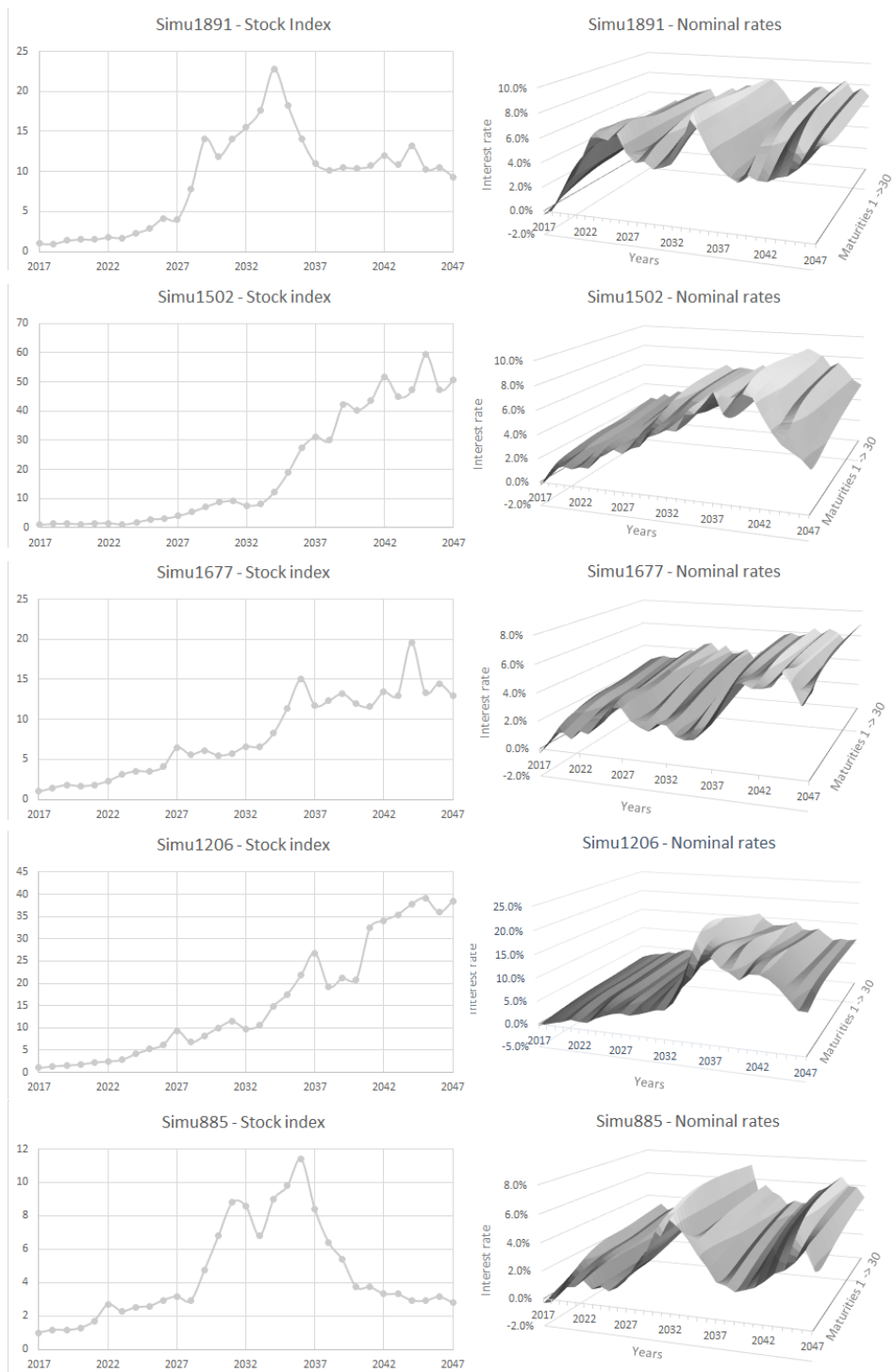


Figure 6: 5 best scenarios ("SimuN" meaning simulation/scenario number N) information

A first possibility to test extreme impacts on the NPV could be to use quantiles of the standard Gaussian outcomes. For example here we mix the following outcomes,

- for the stock index, $\left(\varepsilon_i^{stock,sensi.down} = qtl_{25\%}(\mathcal{N}(0,1))\right)_{i \in \llbracket 1;30 \rrbracket}$,
 $\left(\varepsilon_i^{stock,mid} = qtl_{50\%}(\mathcal{N}(0,1))\right)_{i \in \llbracket 1;30 \rrbracket}$, $\left(\varepsilon_i^{stock,sensi.up} = qtl_{75\%}(\mathcal{N}(0,1))\right)_{i \in \llbracket 1;30 \rrbracket}$,
- for the nominal IR, $\left(\varepsilon_j^{IR_1,sensi.down} = qtl_{25\%}(\mathcal{N}(0,1)), \varepsilon_j^{IR_2,sensi.down} = qtl_{25\%}(\mathcal{N}(0,1))\right)_{j \in \llbracket 1;30 \rrbracket}$,
 $\left(\varepsilon_j^{IR_1,mid} = qtl_{50\%}(\mathcal{N}(0,1)), \varepsilon_j^{IR_2,mid} = qtl_{50\%}(\mathcal{N}(0,1))\right)_{j \in \llbracket 1;30 \rrbracket}$,
 $\left(\varepsilon_j^{IR_1,sensi.up} = qtl_{75\%}(\mathcal{N}(0,1)), \varepsilon_j^{IR_2,sensi.up} = qtl_{75\%}(\mathcal{N}(0,1))\right)_{j \in \llbracket 1;30 \rrbracket}$,
- and for the RIR, $\left(\varepsilon_k^{RIR_1,sensi.down} = qtl_{25\%}(\mathcal{N}(0,1)), \varepsilon_k^{RIR_2,sensi.down} = qtl_{25\%}(\mathcal{N}(0,1))\right)_{k \in \llbracket 1;30 \rrbracket}$,
 $\left(\varepsilon_k^{RIR_1,mid} = qtl_{50\%}(\mathcal{N}(0,1)), \varepsilon_k^{RIR_2,mid} = qtl_{50\%}(\mathcal{N}(0,1))\right)_{k \in \llbracket 1;30 \rrbracket}$,
 $\left(\varepsilon_k^{RIR_1,sensi.up} = qtl_{75\%}(\mathcal{N}(0,1)), \varepsilon_k^{RIR_2,sensi.up} = qtl_{75\%}(\mathcal{N}(0,1))\right)_{k \in \llbracket 1;30 \rrbracket}$,

So we get $3 \times 3 \times 3 = 27$ sensitivities scenarios. This leads to artificial and really extreme sensitivities (the up and down are associated to 0.01% and 99.99% (\mathcal{N})(0, 30) quantiles) but the objective of this test is to get extreme impacts and see how the ALM model reacts to such heavy movements of the economy. Some of the relevant features associated to these sensitivities scenarios are given in Figures 7 and 8. We obtain the sensitivities results in Tables 10, 11 and 12, and shown in Figure 9.

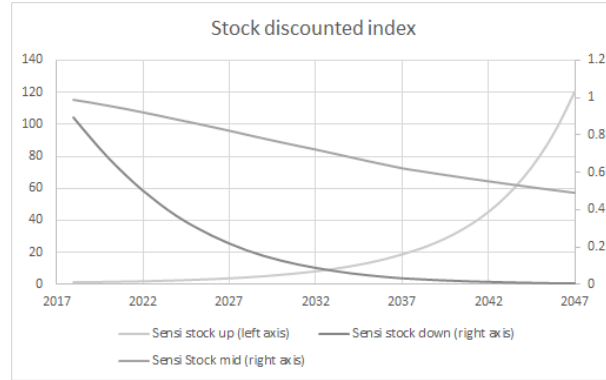


Figure 7: down/mid/up stock discounted returns

Note first that the NPV associated to $\varepsilon = 0$ is 32'503, much higher than the averages VIF of 11'369, so that $\mathbb{E}[NPV(\varepsilon)] < NPV(\mathbb{E}[\varepsilon])$ which is expected, associated to the concavity of the NPV risk measure. The scenario associated to $\varepsilon = 0$ could be seen as a *forward-type* scenario (though still integrating some stochastic diffusion parameters like the Ito parameter in the Black-Scholes diffusion). Such scenarios are useful when estimated regularly on central VIF assessment in order to quantify the distortion of the NPV distribution through time.

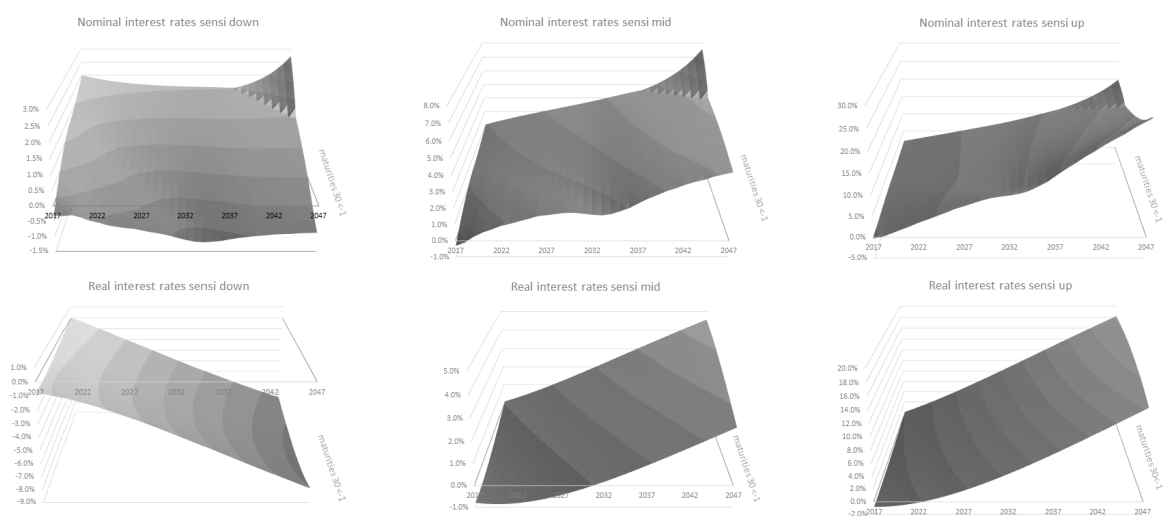


Figure 8: down/mid/up yield curves associated to the sensitivities

Table 10: NPV sensitivities - RIR down

Risk	sensitivities								
IR	down	down	down	mid	mid	mid	up	up	up
stock	down	mid	up	down	mid	up	down	mid	up
NPV	-199'520	-43'008	66'395	-61'019	32'872	70'452	-3'802	34'722	66'037

Table 11: NPV sensitivities - RIR mid

Risk	sensitivities								
IR	down	down	down	mid	mid	mid	up	up	up
stock	down	mid	up	down	mid	up	down	mid	up
NPV	-210'808	-48'707	65'070	-65'436	32'503	71'808	-5'325	34'583	66'095

Table 12: NPV sensitivities - RIR up

Risk	sensitivities								
IR	down	down	down	mid	mid	mid	up	up	up
stock	down	mid	up	down	mid	up	down	mid	up
NPV	-214'314	-51'187	63'628	-67'718	31'539	71'227	-6'988	34'211	67'668

These sensitivities are interesting from a risk management perspective and for regulatory check-ups but the number of underlying risks is a great limitation. Here we have only 3 underlying risks but with 4 (e.g. adding real estate or even credit risk) the number of sensitivities scenarios to build and launch would rise up to $3 \times 3 \times 3 = 81$ and even more for a more precise study. So we need to focus on more general “quantile” scenarios.

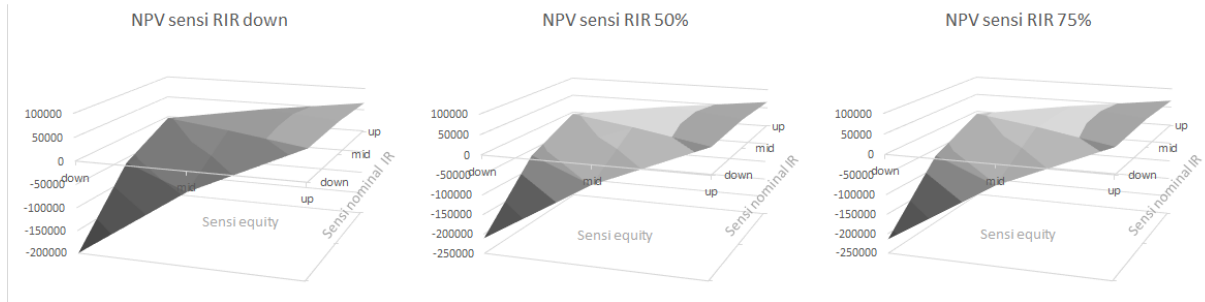


Figure 9: 27 extreme sensitivities impact

4.4 Developments on quantile scenarios - basics and implementations

Basically, two possibilities can be considered to define any “ $x\%$ ”-quantile scenario.

On one hand it is interesting to look for scenarios that would lead to the NPV quantiles. In this case, one wants to find the space $\Omega_1 \subset \mathbb{R}^{150}$ such that $\forall \tilde{\varepsilon} \in \Omega_1, NPV(\tilde{\varepsilon}) = qt_{x\%}(NPV)$.

On the other hand it is also interesting to study scenarios associated to the quantiles of ε . In this case, one focuses on the space $\Omega_2 \subset \mathbb{R}^{150}$ such that $\forall \tilde{\varepsilon} \in \Omega_2, F(\tilde{\varepsilon}) = x\%$. With F the cumulative function of ε .

In all generality $\Omega_1 \neq \Omega_2$. Ω_2 is the level set (hyperplan) of a 150-dimensions correlated Gaussian distribution associated to the $x\%$ -quantile.¹⁰

Considering Ω_1 it is generally impossible to get any theoretical assessment due to the high level of complexity of function $NPV(\varepsilon)$. However we already know one estimated/empirical quantile scenario thanks to our empirical distribution of NPV .

4.5 Focus on marginal quantile scenarios to understand the time-impact of risks on the ALM model

Another possibility is to assess 1-risk factor (say $r \in (stock, IR_1, IR_2, RIR_1, RIR_2)$) sensitivity, that is to consider the space Ω_2' obtained by limiting the $x\%$ -quantile to the sub 30-tuple of outcomes $(\varepsilon_s^r)_{s \in [1;30]}$ and considering the other risks 120 outcomes as equal to 0 (deterministic with no impact on the simulation). This allows to focus on the marginal impact of risk r on the NPV through the ALM model calculation.

Remember the true objective here is to better understand the ALM model and quantify its movements to “realistic” potential risk-neutral trajectories. A first interesting idea could be to focus on Ω_2^{stock} and to focus on the temporal impact of the major risk that is equity decrease. We propose to test the 25%-quantile scenarios associated to different periods, that is 1 to 10 years, 11 to 20 years and 21 to 30 years. We note the studied periods $p_i, i \in [1;3]$ ($p_1 = [1;10]$, $p_2 = [11;20]$ or $p_3 = [21;30]$). The scenarios are therefore defined as follow, for each period $p_i, i \in [1;3]$,

¹⁰It is easy to characterise Ω_2 . Indeed it is well-known that $\sum_{s=1}^{30} (\varepsilon_s^{stock} + \varepsilon_s^{IR_1} + \varepsilon_s^{IR_2} + \varepsilon_s^{RIR_1} + \varepsilon_s^{RIR_2}) \sim \mathcal{N}(0, \eta^2)$ for a given scalar η depending on Σ . This relation is sufficient to define the whole hyperplan.

- for any year $y \in \llbracket 1; 30 \rrbracket$, $\varepsilon_y^{IR_1, p_i} = \varepsilon_y^{IR_2, p_i} = \varepsilon_y^{RIR_1, p_i} = \varepsilon_y^{RIR_2, p_i} = 0$,
- for any year $y \in \llbracket 1; 30 \rrbracket$ out of p_i $\varepsilon_y^{stock, p_i} = 0$,
- for any year $y \in p_i$, $\varepsilon_y^{stock, p_i} = \frac{\sqrt{30}}{10} qtl_{25\%}(\mathcal{N}(0, 1)) = \frac{1}{10} qtl_{25\%}(\mathcal{N}(0, 30))$,

The idea behind these three scenarios is to reach Ω_2^{stock} associated to the 25%-quantile level but, depending on i , to focus on a different section of the 30-years time horizon (considering only the stock decrease risk). The resulting *NPV* are given in Table 13.

Table 13: *NPV* time-sensitivities

Period	p_1	p_2	p_3	case $\varepsilon = 0$
<i>NPV</i>	2'708	17'697	30'189	32'503
comparison with the case $\varepsilon = 0$	-91.7%	-45.6%	-7.1%	\

These *NPV* allow to quantify the time impact of the stock decrease risk, as at constant ALM model and economic situation. It is normal to see a stronger impact when considering decreasing period sensitivities for different reasons. The stock shock, when applied early, has an impact on all future dates through asset management and ALM rules. In addition, the calculations are made assuming a closed population of policyholders (run-off diffusion) so that the shock on p_3 impacts fewer contracts than p_2 (some have ended before 21 years, other policyholders have exercised their lapse option, others are dead, etc.) and the shock on p_2 impacts fewer contracts than p_1 (the impact is $2 \times$ higher) for the same reasons. The relative impacts must be compared to the *NPV* associated to $\varepsilon = 0$ and not to the *VIF*. Indeed we are only considering specific sets of cash-flows, which is not the case through the mitigation of risks induced by the *VIF* average. The sensitivities on *VIF* values are comparable when estimating different averages, e.g. obtained on scenarios tables calibrated on different economic situations, instantaneously shocked liabilities or ALM assumptions, market data (such sensitivities are used to assess marginal capitals when using the Standard Formula to estimate the Solvency Capital Requirement).

This analysis shows the strong relative time-impact induced by the ALM model and the initial economic situation on the way the stock risk is undertaken by the studied product, under the economic valuation assumptions (risk-neutral measure, chosen market-consistency assumption - including the regulatory yield curve - and simulation models). Ideally the division into 3 periods, only for the stock decrease risk, could be enlarged to yearly periods and to every risk (IR up, IR down, RIR up, RIR down here) and even to cross-sensitivities (stock down + IR up, etc.). Once again, the number of underlying risks is a great limitation but it would be interesting to focus more precisely on highest risks or higher probability risks (according to ERM practitioners). One must balance limiting the number of sensitivities between less complexity and increasing relevance. In particular it seems possible that this balance could change during time and should be well justified.

4.6 About using a risk-neutral valuation framework

The marginal-quantile *NPV* values obtained above also introduce some questions on the interest of using a risk-neutral measure and models to value insurance accounting quantities. If one accepts the lack

of realism of the simulations (a point often discarded by regulation and practitioners based on the fact one only aims at estimating a price/an average value) them one must undertake the fact that many of the *NPV* assessed on the economic scenarios table can be negative (based on these results one can only say “probably more than 20%” as here we only focus on one single risk - 30.8% of the 2'000 scenarios used to assess the *VIF*)¹¹. Anyway the risk neutral measure is supposed to lead to fair market prices due to the martingality of the variable of interest (e.g. discounted option cash-flows for financial markets). The *NPV* is a very complex function of martingale underlyings here but it is not built to be a martingale, so the averaged prices obtained are absolutely no fair prices according to risk neutral theory. In addition the ALM rules (lapse and mortality assumptions, asset optimisation, rate credited to policyholders, rate assumed to be expected by policyholders, etc.) put inside the valuation model are calibrated based on realistic potential movements of the economy. So the application of such parameters to unrealistic scenarios seem counter-intuitive and may even lead to non-adapted movements and algorithmic choices (e.g. in asset reallocation, unrealistic/in-adapted - extreme or too low - lapse rates). For these reasons we have tested more “realistic” marginal quantile-style “v2” scenarios associated to a stock crises event then a rise again: 3 years of 10% (negative) shock followed by 2 years of 70% (positive) shock on the 5 first years of each period p_1 , p_2 and p_3 . Finally, for each period $p_i, i \in \llbracket 1; 3 \rrbracket$,

- for any year $y \in \llbracket 1; 30 \rrbracket$, $\varepsilon_y^{IR_1, p_i^{y^2}} = \varepsilon_y^{IR_2, p_i^{y^2}} = \varepsilon_y^{RIR_1, p_i^{y^2}} = \varepsilon_y^{RIR_2, p_i^{y^2}} = 0$,
- for any year $y \in \llbracket 1; 30 \rrbracket$ out of $\llbracket (i-1) \times 10 + 1; (i-1) \times 10 + 5 \rrbracket$ $\varepsilon_y^{stock, p_i^{y^2}} = 0$,
- for any year $y \in \llbracket (i-1) \times 10 + 1; (i-1) \times 10 + 3 \rrbracket$, $\varepsilon_y^{stock, p_i^{y^2}} = qtl_{10\%}(\mathcal{N}(0, 1))$,
- and for years $y \in \llbracket (i-1) \times 10 + 4; (i-1) \times 10 + 5 \rrbracket$, $\varepsilon_y^{stock, p_i^{y^2}} = qtl_{70\%}(\mathcal{N}(0, 1))$

The resulting *NPV* are given in Table 14.

Table 14: *NPV* time-sensitivities “v2”

Period	p_1	p_2	p_3	case $\varepsilon = 0$
<i>NPV</i> ^{v2}	10'218	14'438	24'435	32'503
comparison with the case $\varepsilon = 0$	-68.6%	-55.6%	-24.8%	\

These scenarios are also associated to a Ω_2^{stock} space but for $\sim 30\%$ -quantile scenarios. The values still present the time-impact character, which seems normal, though with reduced impact. But more important, the higher realism given to the economic scenario has little impact on the interrogations raised above about using a risk-neutral measure and models.¹²

¹¹In practice, the management should probably react promptly in such an occurrence. But such reactions are not necessarily taken into account in the model because the associated scenarios are not realistic enough. Though such negative *NPV* could be considered as over-weighted in true practice, the regulatory values are used to quantify solvency in real life.

¹²Note that in reality the *risk-neutral* denomination is not present in the directive but only used in practice as an European Union wide operational interpretation of the idiom “market-consistent”. In fact, this term has never been formally framed by regulation but it is supposed to, at least, set aside the regulation difficulties induced by real-world or more realistic valuations (estimation/validation/control of risk premiums in particular).

To conclude on this “ERM” Section, we wanted here to focus on the economic scenarios, often forgotten by practitioners who generally only focus on realistic evolution of the *VIF* and the Solvency regulation quantities through realistic-type sensitivities scenarios. This last use is indeed the most important due in particular to regulatory constraints (*e.g.* Own Risk and Solvency Assessment / ORSA process). However, it is also very important to understand in depth the impact of the risk-neutral scenarios and the different risks diffused under such a specific measure on the developed ALM model. This objective mixes ALM and ERM in a regulatory-specific valuation scheme, adapting standard strategic ERM to solvency regulation management constraints. After focusing on and analysing the top and worse *NPV* scenarios, we have proposed various economic sensitivity scenarios definitions that enables practitioners to assess a complete analysis based on risk quantification, comparison and time impact of these risks on the ALM model. Finally, the values obtained have led us to reflect upon the practical interest of risk neutrality to estimate insurance accounting quantities for solvency assessment purposes. After these reflections we must warn practitioners and regulators on the ill-adapted but also hazardous character (not even talking about the manipulability risks) of this framework to quantify solvency.

Conclusion

In this paper, we have first tried to synthesised the European life insurance valuation process as shaped by directive Solvency II, the EIOPA texts, and used by practitioners.

Then, after recalling the results and conclusions obtained by El Karoui et al. for 2014, we update them to 2017, still observing strong differences in valuations. We also test new parameters for the method to estimate sensitivities. This analysis shows the high impact of the UFR convergence on calibration, simulation and valuation. We conclude on two main clues for a better stability of economic valuations. First, we underline the efficiency of using averaged matrices to tackle some issues due to the dependency in time of the market-consistency calibration process. Second, we incite practitioners to consider large sets of calibration IV for more stability. We have also shown, using 2020 data, that the impact of a financial crises event like Covid-19 would not be “erased” by the 3-month smoothing technique if it occurred at the end of a calendar year. This underlines the interest of such smoothing technique for a more stable market-consistent scheme. We have also empirically underlined the interest of setting a fixed simulation seed when comparing economic valuations. Even with high confidence intervals in the valuations, we show the higher stability of the valuations differences, which directly underline the robustness of our previous comparison of values under various calibration sets.

This led us to replace the “same seed” practitioners wording into its literature and theory of common random numbers and paths, widely known and used, in particular in finance.

Then, in a final Section, we have focused on alternative ERM practices based on analysing economic scenarios and their impact on the ALM model used by practitioners. We have proposed different fashions to select scenarios and estimate sensitivities, depending on the objective followed by companies risk management. These propositions enable users to focus on in-depth economic scenarios simulation and impact. Our tests have eventually led us to raise the risk of using a risk-neutral measure and simulation models to quantify solvency from a regulation viewpoint.

This study breeds different research perspectives. First we aim at testing ALM valuation on averaged

matrices in further years and continue sensitivity tests on the results presented in El Karoui et al. (2017). Indeed, this kind of study is specifically interesting if regularly updated. In particular, the UFR value will evolve in next years, affecting both arbitrage opportunities and economic valuations. The follow-up of our results will be of interest for applied actuarial research. In addition, new sensitivities should be tested in particular on the interest rates term structure used in the calibration *e.g.* to quantify the UFR impact (on calibration and simulation distinctly). Finally, it would be interesting to apply the same tests on other interest rates models to evaluate the impact of model changes. Finally, our alternative ERM propositions require further developments to focus more precisely on ALM model fingerprints. This field has been barely pointed out but it is clear that the wide freedom left to practitioners to build and parameterized their model has a deep impact on valuation and can lead to additional level of manipulability. Further studies will be necessary to quantify and propose solutions to these issues. In addition, we have introduced various seemingly legitimate questions concerning the hazard of using such risk-neutral valuation scheme to frame the solvency regulation. This risk must be better quantified and managed to improve the scheme stability and regulation efficiency.

Finally, it is interesting to loop on the subject of the manipulability of this valuation scheme, introduced by El Karoui et al. One major problem lies in the non-unicity of the risk-neutral valuation measure. No-arbitrage is one of the conditions for risk neutral measure and for prices unicity. Though financial markets are not fully arbitrage-free, the high liquidity of trades ensures a relative fairness of prices. In this actuarial field, there is no trade nor liquidity of insurance liabilities. The adaptability of a risk-neutral valuation is at least questionable and there are high risks of valuation instability, control issues and regulation difficulties. Basically, if the regulator does not force models (simulation and ALM) and parameters to be used by companies no efficient stability and comparability of economic values can be achieved. Even the integration of stochastic scenarios-based valuation instead of quota-share provisions (one of the major assumed *improvement* between Directives Solvency I and Solvency II) is, with the current guidelines, more than questionable. Another manipulability problem lies in the simulation seed choice made by companies. The number of simulations to assess economic own funds seem currently too low. This choice used to be justified by algorithmic complexity but with the everyday higher computer capacities, this explanation seems questionable today.

Appendix

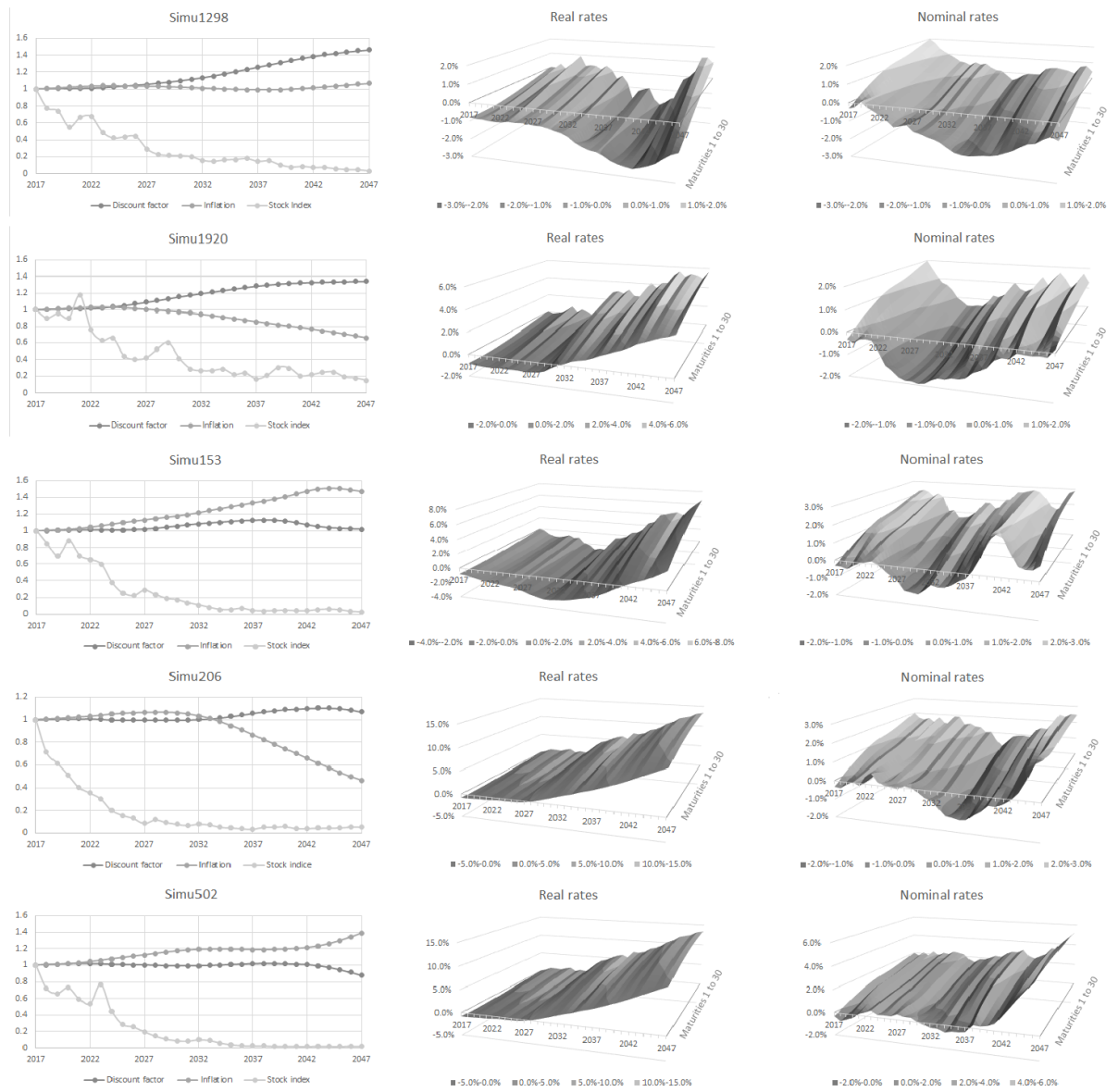


Figure 10: 5 best scenarios ("SimuN" meaning simulation/scenario number N) full information

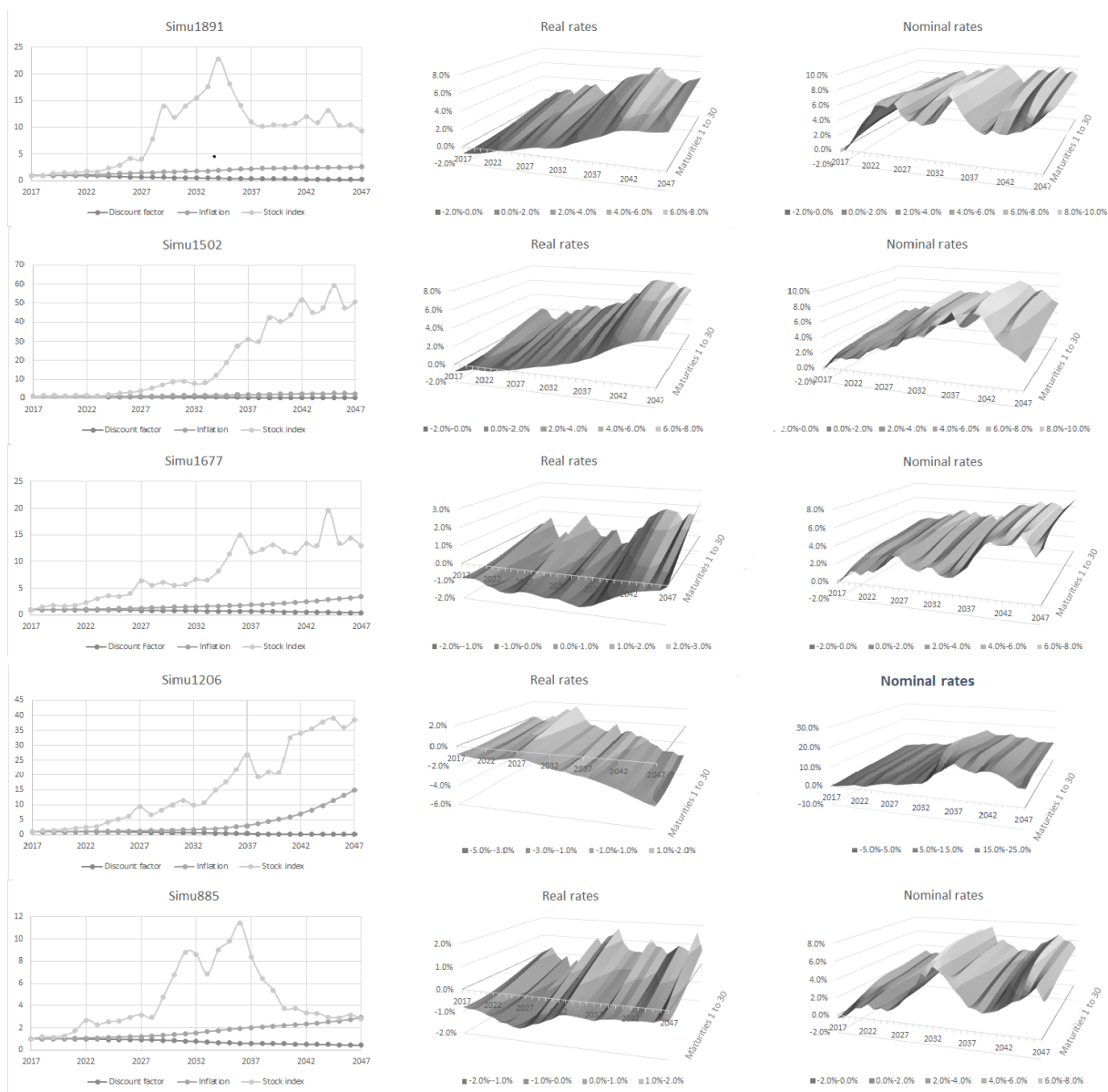


Figure 11: 5 best scenarios ("SimuN" meaning simulation/scenario number N) full information

References

- Arrow, K. J. & Debreu, G. (1954), 'Existence of an equilibrium for a competitive economy', *Econometrica: Journal of the Econometric Society* **22**(3), 265–290.
- Barth, M. E. & Landsman, W. R. (1995), 'Fundamental issues related to using fair value accounting for financial reporting', *Accounting Horizons* **9**(4), 97.

- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2014), ‘Fitting linear mixed-effects models using lme4’, *arXiv preprint arXiv:1406.5823* .
- Benhamou, E. (2003), ‘Optimal malliavin weighting function for the computation of the greeks’, *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* **13**(1), 37–53.
- Black, F. & Karasinski, P. (1991), ‘Bond and option pricing when short rates are lognormal’, *Financial Analysts Journal* **47**(4), 52–59.
- Black, F. & Scholes, M. (1973), ‘The pricing of options and corporate liabilities’, *The journal of political economy* **81**(3), 637–654.
- Borel-Mathurin, F. & Vedani, J. (2019), ‘Market-consistent valuation: a step towards calculation stability’.
- Brigo, D. & Mercurio, F. (2006), *Interest rate models-theory and practice: with smile, inflation and credit*, Vol. 2, Springer.
- Chauvigny, M., Devineau, L., Loisel, S. & Maume-Deschamps, V. (2011), ‘Fast remote but not extreme quantiles with multiple factors: applications to solvency ii and enterprise risk management’, *European Actuarial Journal* **1**(1), 131–157.
- Chen, N. & Glasserman, P. (2007), ‘Malliavin greeks without malliavin calculus’, *Stochastic Processes and their Applications* **117**(11), 1689–1723.
- Devineau, L., Arrouy, P.-E., Bonnefoy, P. & Boumezoued, A. (2017), ‘Fast calibration of the libor market model with stochastic volatility and displaced diffusion’, *arXiv preprint arXiv:1706.00263* .
- El Karoui, N., Loisel, S., Prigent, J.-L. & Vedani, J. (2017), ‘Market inconsistencies of market-consistent european life insurance economic valuations: pitfalls and practical solutions’, *European Actuarial Journal* **7**(1), 1–28.
- European Commission (2009), ‘SOLVENCY II – EC directive 2009/138’.
- European Insurance and Occupational Pensions Authority (2010a), ‘Quantitative Impact Study 5 - technical specifications’.
- European Insurance and Occupational Pensions Authority (2010b), ‘Quantitative Impact Study 5 Questions & Answers’.
- European Insurance and Occupational Pensions Authority (2017), ‘Technical documentation of the methodology to derive eiopas risk-free interest rate term structures’.
- European Parliament (2002), ‘Regulation no 1606/2002 of the european parliament and of the council of 19 july 2002 on the application of international accounting standards’, *Official Journal of the European Communities* **45**(243), 1–4.
- Fournié, E., Lasry, J.-M., Lebuchoux, J., Lions, P.-L. & Touzi, N. (1999), ‘Applications of malliavin calculus to monte carlo methods in finance’, *Finance and Stochastics* **3**(4), 391–412.

- Glasserman, P. & Yao, D. D. (1992), 'Some guidelines and guarantees for common random numbers', *Management Science* **38**(6), 884–908.
- Groupe Consultatif Actuariel Européen (2012), 'An actuarial view of market-consistency'.
- Hall, E. J., Katsoulakis, M. A. & Rey-Bellet, L. (2016), 'Uncertainty quantification for generalized langevin dynamics', *The Journal of chemical physics* **145**(22), 224108.
- Huerta de Soto, J. (2009), 'The fatal error of solvency ii', *Economic Affairs* **29**(2), 74–77.
- Hull, J. & White, A. (1994), 'Numerical procedures for implementing term structure models i: Single-factor models', *Journal of derivatives* **2**(1), 7–16.
- Institut des Actuaire (2019), 'Document d'Orientation - Les Générateurs de scénarios économiques', https://www.institutdesactuaire.com/global/gene/link.php?doc_id=15033&fg=1.
- International Accounting Standards Board (2004), 'IFRS 4: Insurance contracts'.
- International Accounting Standards Board (2011), 'IFRS 13: Fair value measurement'.
- International Accounting Standards Board (2014), 'IFRS 9: Financial instruments'.
- Joshi, M. & Rebonato, R. (2003), 'A stochastic-volatility, displaced-diffusion extension of the libor market model', *Quantitative Finance* **3**(6), 458–469.
- Joshi, M. S., Rebonato, R. et al. (2003), 'A displaced-diffusion stochastic volatility libor market model: motivation, definition and implementation', *Quantitative Finance* **3**(6), 458–469.
- Kemp, M. (2009), *Market Consistency: Model Calibration in Imperfect Markets*, John Wiley & Sons.
- Lam, J. (2014), *Enterprise risk management: from incentives to controls*, John Wiley & Sons.
- Loisel, S., Mazza, C. & Rullière, D. (2009), 'Convergence and asymptotic variance of bootstrapped finite-time ruin probabilities with partly shifted risk processes', *Insurance: Mathematics and Economics* **45**(3), 374–381.
- Loisel, S. & Privault, N. (2009), 'Sensitivity analysis and density estimation for finite-time ruin probabilities', *Journal of computational and applied mathematics* **230**(1), 107–120.
- Malamud, S., Trubowitz, E. & Wüthrich, M. V. (2008), 'Market consistent pricing of insurance products', *ASTIN Bulletin-Actuarial Studies in Non Life Insurance* **38**(2), 483.
- Markowitz, H. (1952), 'Portfolio selection', *The journal of finance* **7**(1), 77–91.
- Merton, R. C. (1973), 'Theory of rational option pricing', *The Bell Journal of economics and management science* **4**(1), 141–183.
- Nelson, K. K. (1996), 'Fair value accounting for commercial banks: An empirical analysis of sfas no. 107', *Accounting Review* pp. 161–182.

- Rebonato, R. (2002), *Modern pricing of interest-rate derivatives: The LIBOR market model and beyond*, Princeton University Press.
- Rebonato, R. & Pogudin, A. (2011), 'Is it possible to reconcile the caplet and swaption markets? evidence from the us-dollar market', *Journal of Derivatives* **19**(2), 8.
- Ritter, J. R. (1988), 'The buying and selling behavior of individual investors at the turn of the year', *The Journal of Finance* **43**(3), 701–717.
- Sheldon, T. J. & Smith, A. D. (2004), 'Market consistent valuation of life assurance business', *British Actuarial Journal* **10**(3), 543–605.
- Tezuka, S. (1998), Financial applications of monte carlo and quasi-monte carlo methods, in 'Random and quasi-random point sets', Springer, pp. 303–332.
- van Bragt, D., Steehouwer, H. & Waalwijk, B. (2010), 'Market consistent ALM for life insurers steps toward Solvency II', *The Geneva Papers on Risk and Insurance-Issues and Practice* **35**(1), 92–109.
- Vesa, R. & Lasse, K. (2007), 'Topical modelling issues in solvency ii', *Scandinavian Actuarial Journal* **2007**(2), 135–146.
- Wüthrich, M. V. & Bühlmann, H. (2010), *Market-consistent actuarial valuation*, Springer.
- Wüthrich, M. V., Bühlmann, H. & Furrer, H. (2008), *Market-consistent actuarial valuation*, Springer.