Mémoire présenté le :

pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA
et l'admission à l'Institut des Actuaires

Par : Florian BOLLOTTE

Titre Impact du calibrage d'un Générateur de Scénarios Economiques et tests de scénarios risque neutre

Confidentialité : ■ NON ☐ OUI (Durée : ☐ 1 an ☐ 2 ans)

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

Membre présents du jury de l'Institut des Actuaires	signature	*Entreprise :* GALEA & Associés 25 rue de Choiseul 75002 PARIS Tél. 01 43 22 11 11 R.C.S. Paris - 492 379 839
		Nom : GALEA & Associés
		Signature :
Membres présents du jury de l'ISFA		*Directeur de mémoire en entreprise :*
		Nom : Pierre THEROND
		Signature :
		Invité :
		Nom :
		Signature :
		***Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels** (après expiration de l'éventuel délai de confidentialité)*
		Signature du responsable entreprise
		Signature du candidat

Page 1 De 1

Résumé

Dans le cadre des évaluations prudentielles exigées par la directive Solvabilité 2, le recours à un Générateur de Scénario Economique (GSE) est devenu indispensable en assurance vie. Un GSE va permettre de simuler un ensemble de trajectoires de variables financières suffisamment représentatif pour mettre en place une méthode de Monte Carlo, nécessaire à l'évaluation de la juste valeur des engagements d'un assureur vie caractérisés notamment par leurs multiples sources d'optionalité. Pour pouvoir générer un scénario économique cohérent avec la réalité du marché et adapté au profil de risque de l'assureur, un GSE doit reposer sur une théorie économique et financière robuste. La directive Solvabilité 2 impose donc plusieurs exigences vis-à-vis des GSE pour s'assurer que les scénarios économiques générés soient suffisamment robustes face aux évaluations *Best Estimate* à mener, notamment le respect de la propriété *market-consistency*.

Malgré ces exigences réglementaires, les assureurs conservent une certaine liberté de modélisation dans les différentes étapes de développement d'un GSE. Il revient tout d'abord à l'assureur la responsabilité de retenir les modèles stochastiques qui lui semblent les plus adaptés pour projeter les variables macro-économiques du GSE. Le calibrage de ces modèles représente ensuite une étape très délicate du développement au sens où l'assureur doit également choisir les données de marché et la méthode de calibrage à utiliser. Enfin, il doit implémenter des tests pour vérifier que les scénarios générés respectent bien les propriétés théoriques des modèles et les exigences réglementaires. Toutes ces étapes de modélisation sont autant de responsabilités pour la fonction actuarielle qui doit être en mesure de justifier les choix retenus au regard du profil de risque de l'assureur.

Ce mémoire propose une étude des enjeux mentionnés supra. On commencera dans une première partie par essayer d'apprécier l'impact du choix du modèle de taux du GSE et de son calibrage sur les scénarios générés et les évaluations *Best Estimate market-consistent*. A partir d'un GSE risque-neutre et d'un outil ALM déjà existants, nous tenterons de tirer des conclusions quant à la sensibilité de ces deux outils au choix du modèle de taux et des données de marché utilisées pour le calibrer. Dans une seconde partie, on portera attention aux tests pouvant être utilisés pour auditer les scénarios générés par un GSE risque-neutre, en particulier aux tests statistiques de martingalité dont nous présentons une revue détaillée. Bien que nous n'ayons trouvé que peu de travaux faisant référence à ces tests statistiques dans le cadre d'un GSE, nous avons pu constater qu'ils pouvaient se révéler très utiles du point de vue de la fonction actuarielle, en permettant par exemple de détecter des erreurs d'implémentation du GSE. Leur implémentation peut néanmoins soulever de nombreuses interrogations pour la fonction actuarielle, et nous conclurons donc ce mémoire en proposant des recommandations d'utilisation de ces tests au sein d'un GSE risque-neutre.

Mots clés : Générateur de Scénario Economique, calibrage, G2++, Libor Market Model, martingale, test statistique de martingale, Solvabilité 2, fonction actuarielle.

Abstract

In the context of the prudential assessments required by the Solvency 2 directive, the use of an Economic Scenario Generator (ESG) has become essential in life insurance. An ESG allows to simulate a set of trajectories of financial variables that is sufficiently representative to implement a Monte Carlo method, necessary to assess the fair value of a life insurer's commitments characterized especially by their multiple sources of optionality. In order to generate an economic scenario consistent with the market reality and adapted to the insurer's risk profile, an ESG must rely on a solid economic and financial theory. The Solvency 2 Directive therefore imposes several requirements for ESG to ensure that the generated economic scenarios are robust enough to the Best Estimate evaluations to be conducted, in particular the respect of the market consistency criterion.

Despite these regulatory requirements, insurers retain a certain amount of freedom in several stages of the ESG development. The insurer must first select the most appropriate models to project the macro-economic variables of the ESG. The calibration of these models is also a delicate step in the sense that the insurer has to choose the market data and the calibration methodology to use. Finally, the insurers must implement tests to check that the generated scenarios respect the theoretical properties of the models and the regulatory requirements. All these modeling steps represent responsibilities for the actuarial function which must be able to justify the choices made based on the insurer's risk profile.

We propose of study of the several issues mentioned above. In a first part, we will begin by trying to measure the impact of the choice of the interest rate model of the ESG and of its calibration on the generated scenarios and the Best Estimate market consistent evaluations. Using a risk-neutral ESG and an ALM tool already developed, we will try to draw conclusions about the sensitivity of these two models to the choice of the interest rate model and the market data used to calibrate it. In a second part, we will deal with the tests that can be used to audit the scenarios generated by a risk neutral ESG, with a particular attention to statistical martingale tests of which we present a detailed review. Despite the fact that we have not found many studies that refer to these statistical tests in the framework of ESG, we have noticed that they can be very useful from the point of view of the actuarial function, allowing for example to detect implementation errors of the ESG. Nevertheless, their application can raise many questions for the actuarial function, and we will thus conclude this study by proposing some recommendations for the use of these tests within a risk neutral ESG.

Keywords: Economic Scenario Generator, calibration, G2++, Libor Market Model, martingale, statistical martingale test, Solvency 2, actuarial function.

Remerciements

En premier lieu, je tiens à remercier Florence Chiu, consultante du pôle Normes et Modèles, pour le temps qu'elle a passé à guider mes pas et tout ce qu'elle m'a appris au cours de mon stage de M1 puis de mon alternance. Je lui suis reconnaissant de m'avoir laissé prendre beaucoup d'initiatives dans mon travail et de m'avoir responsabilisé sur mes missions.

De même, je souhaite remercier mon tuteur Pierre Thérond, associé en charge du pôle Normes et Modèles, qui m'a guidé tout au long de ce mémoire et a suivi avec l'intérêt l'évolution de mon travail.

Mes remerciements s'adressent également à Christian Robert, mon tuteur universitaire pour son attention portée à mon mémoire.

Enfin, je remercie Norbert Gautron, président du cabinet GALEA & Associés, de m'avoir accueilli au sein de son entreprise.

Contents

Chapter 1

General introduction

1.1 Regulatory framework

1.1.1 Introduction to the Solvency 2 Directive

The Solvency 2 Directive has redefined the requirements applicable to insurance companies in terms of their overall solvency[2] and risk management. It is based on principles rather than rules in order to allow insurers to better adapt to requirements and standards according to their risk profile. The Directive sets out broad principles and gives insurance companies a certain amount of freedom to apply these principles (with the exception of the standard SCR formula). The Solvency 2 Directive is based on three pillars:

Pillar 1: Quantitative requirements
This pillar defines the quantitative rules to be respected in terms of technical provisions and minimum capital threshold to be held. It especially aims at introducing a forward vision of risks and harmonizing at European level the standards and methodologies for valuing balance sheet items so that it is easier to compare insurance companies. In particular, it requires insurers to draw up an economic balance sheet based on the valuation of their assets at market value and their liabilities at transfer value. In addition, it defines two capital requirements:

- The Minimum Capital Requirement which corresponds to the minimum capital level below which the insurer would no longer be able to meet its commitments towards its policyholders and which would lead to the automatic intervention of the supervisory authority;
- The Solvency Capital Requirement which is the amount of funds to protect against economic ruin for the coming year with a 99.5% probability.

Pillar 2: Qualitative requirements
The second pillar deals with qualitative requirements for risk management, control and internal audit to complete the quantitative requirements of pillar 1. These measures must allow insurers to manage all their risks (material or not, quantifiable or not) in a prudent manner over a horizon longer than the one considered in the pillar 1, through a system of governance. This system of

[2]Solvency refers to the ability of an insurer to meet its long-term commitments towards its policyholders.

governance must be an adequate and transparent organizational structure that facilitates supervision. The Directive defines it in a very broad sense (risk management, actuarial function, internal audit, etc.) but requires a strict level of formalization. Insurance companies are thus encouraged to develop an Enterprise Risk Management approach, so that they can themselves assess and measure their risks, and to set up an internal risk and solvency assessment process, called Own Risk and Solvency Assessment, adapted to the specificities of the company and including both a quantitative and a qualitative part.

Pillar 3: Disclosure and transparency requirements
The third pillar deals with the transparency obligation of insurers. It defines the information insurers have to regularly publish to enable supervisory authorities to assess their risk management (prudential reporting) and to provide policyholders and financial actors a level of information about the economic and financial situation of the company (public reporting).

The use of ESG is especially motivated by the first pillar, through the need of valuing the insurer's assets and liabilities in accordance with the requirements of pillar 1.

1.1.2 The pillar 1 of Solvency 2

The Solvency 2 Directive encourages insurers to better manage their risks by considering them in an approach consistent with the current economic environment. One of the most noteworthy reforms imposed by the first pillar has been the recasting of the balance sheet assessment. The latter must now be carried out in the form of a prudential balance sheet according to an economic approach and not an accounting one. This methodology requires a more realistic vision of the company's accounting management through a more consistent identification and estimation of risks. The economic approach of the balance sheet relies especially on the criterion of market consistency, that is to say an evaluation that uses as much as possible market data and hypothesis. Despite being today the standard approach for asset and liability valuation in the insurance framework, the market-consistent methodology has been criticized by the academic community (see for example El Karoui et al. (2015)).

Under Solvency 2, assets have to be valued at market value, that is at the price at which they could be traded on the market. As concerns liabilities, the problem is more complex since there is generally no market to obtain the value at which an insurer's liability could be exchanged. Hence, a specific method, based on the Best Estimate and the Risk Margin, is used to assess the economic value of technical provisions.

Technical provisions

Technical provisions correspond to an anticipation of the insurer's future obligations and must allow it to settle all its expected commitments towards the policyholders. Under Solvency 2, they must be valued at their current transfer value, as defined by Article 76 of the Directive. More precisely, technical provisions are estimated as the sum of the Best Estimate of Liabilities and the Risk Margin. Moreover, the calculation of technical provisions has to be carried out in a manner consistent with the market, to respect the market consistency principle.

Best Estimate of Liabilities

The Best Estimate (BE), defined in Article 77 of the Directive, corresponds to the expected current value of future cash-flows required to settle the insurer commitments towards policyholders. The cash inflows and outflows considered must include the premiums, the insurer's benefits and all types of expenses due to the insurance contracts. These cash flows are discounted using either the risk-free yield curve provided by the EIOPA or the market swap rate curve.
An important notion in Solvency 2, especially for the BE of premiums, is the boundary of a recognized (re)insurance contract, defined in the Technical Specification for the Preparatory Phase (Part 1) published by the EIOPA in 2014, because it allows to identify the insurer's commitment to be taken into account in the calculation of the BE. Under Solvency 2, commitments must be taken into account when the insurer no longer has the option to unilaterally act on the terms of the contract, which requires to consider more contracts, particularly a broader basis of premiums than under the current French accounting standards.

The Best Estimate of Liabilities (BEL) is the best estimation of the total technical provisions that must take into account every option inherent to the contract, as mentioned in Article 32 and Article 272 of the Commission Delegated Regulation of 10 October 2014. These options are financial guarantees or rights granted to the policyholder by regulation or by contractual clauses that aim at making the contract more attractive on the market. Even though they are identified, these options are not separately provisioned and therefore, must be included in the Best Estimate, which complicates its calculation.
If we consider a life insurance or retirement savings contract, the best examples of options and financial guarantees inherent in the contract are:

Surrender: Surrender is a right granted to the policyholder. During the term of the contract, the policyholder has rights on the mathematical provisions of the contract through the surrender mechanism. He may claim his right on it during the contract and thus terminate it before the initial settlement date. The policyholder is the only actor who is able to exercise this right of surrender. This choice can be motivated by the difference between the rate paid by the insurer and those of competing insurers.

Minimum guaranteed rate: When subscribing to a savings or retirement insurance contract, the insurer undertakes to guarantee a minimum revaluation of the accumulated savings, through a guaranteed minimum rate. It is a rate fixed at the subscription that must be respected by the insurer during all the life of the contract, regardless of the evolution of its portfolio. Prudential regulation can impose some limitation on the guaranteed minimum rate, by imposing for example a maximum value, in order to secure the solvency of insurance companies.

Profits sharing: According to the French Insurance Code (L.331-3), life insurers have to share with their policyholders a portion of their technical and financial benefits. This distribution is carried out via the profit sharing mechanism, which revalues the insurer's commitments towards the policyholders, depending on the evolution of the assets held by the insurer, and therefore, directly impacts the BE. The profit sharing mechanism can be contractual or discretionary too (distribution according to decisions of the insurer).

Risk Margin

The Risk Margin is the amount required to be added to the Best Estimate to cover the insurance or reinsurance obligations over the entire life of the portfolio, taking into account the capital cost. It guarantees that the total amount of technical provisions is equal to the amount the company would have to pay if it transferred its commitments to another insurer.

Solvency Capital Requirement

The SCR corresponds to the amount of funds that an insurer must hold today so that its probability of economic ruin[1] for the coming year is less than 0.5%. It has to be calculated annually and notified to the supervisory authority. If the risk profile of the company is significantly changed during the course of a year, then the SCR must be re-evaluated and notified to the regulator.
The EIOPA offers insurers two main methods to compute their SCR, by using either the standard formula, based on accounting data and known indicators, or an internal model that must be validated by the supervisory authority. An internal model allows insurers to better adapt the regulation to their risk profile. It should be noted that there are also several alternative approaches, through the use of hybrid models. For example, an insurer may use its own parameters, called "Undertaking Specific Parameters", in the standard formula instead of the given parameters. We can also mention the existence of partial internal models, characterized by the fact that some significant risks are not modeled.

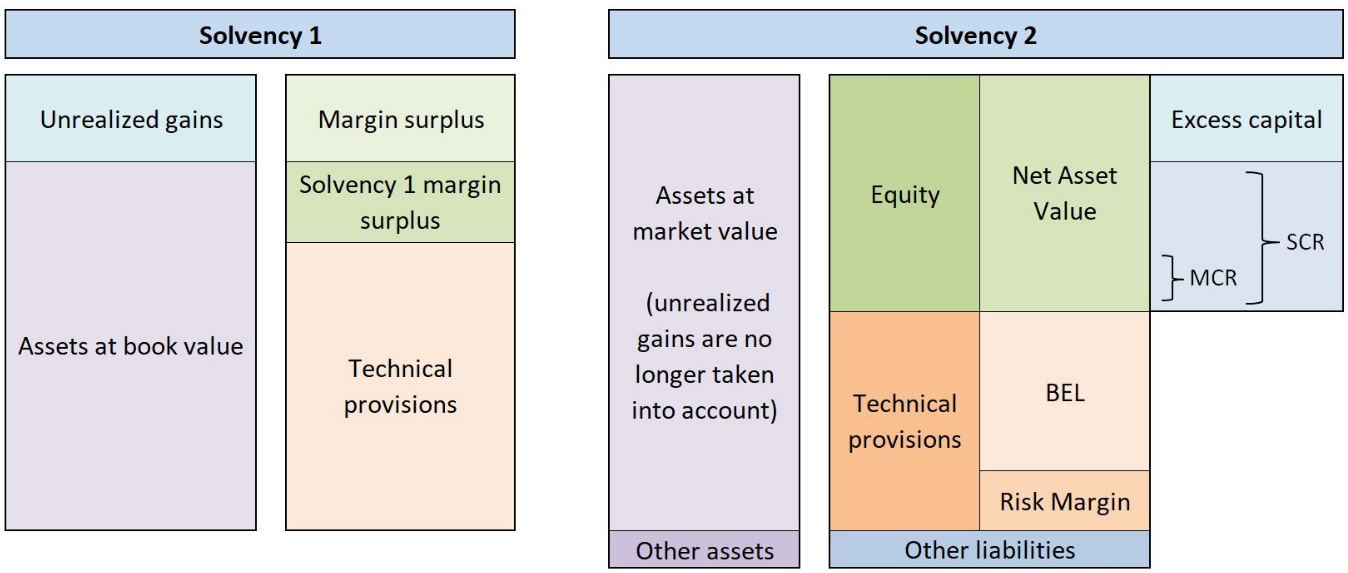


Figure 1.1 – Review of the recasting of the balance sheet assessment under Solvency 2

[1]Economic ruin refers to the situation where the insurer's assets, valued at market value, are worth less than the economic value of its liabilities.

1.2 Economic Scenario Generator and ALM model

1.2.1 Economic Scenario Generator under Solvency 2

As part of the pillar 1 requirements of Solvency 2, insurers have to evaluate their liabilities towards the policyholders at their current execution value, which is the best valuation, consistent with market observations, of the current value of future cash-flows required to settle insurer commitments towards the policyholders, taking into account every contractual option and financial guarantee included in the contract. The complexity of the estimation of these inherent options, especially for a life insurance contract, requires the use of a simulation approach, based on the Monte Carlo method. In particular, the profit sharing scheme, distributed to policyholders, indexed to the financial and technical results and mainly resulting from a discretionary process (at least when the minimal guaranteed interest rate is not too high) that is added to the compulsory profit sharing, requires to model both the policyholder behavior and the one of the life insurance company management, in order to compute the Best Estimate. To estimate the price of the uncertainties inherent to the multiple sources of optionality of a life-insurance product, a deterministic model is not adequate, especially because of the non-linear structure of the financial guarantees. Therefore, life insurers resort to stochastic models to make a sufficiently robust evaluation that takes into account the peculiar risks implied by the volatility over time of the market and underlying assets.

Economic Scenario Generators (ESG) address the need to perform stochastic calculus for balance sheet valuation in accordance with the principle of consistency with market prices and observations. An ESG is a simulation tool that models and projects macroeconomic variables in an economic environment consistent with the market. Planchet et al. (2009) define an economic scenario as a projection of economic and financial variables over a horizon. An ESG projects the economic environment of an institution (in particular an insurance company) by modeling the different economic and financial risks to which it exposes itself. To do so, an ESG simulates a set sufficiently representative of future economic scenarios that depend on the evolution of the risk factors proper to the insurer (term structure of interest rates, share return, obligation return, etc.). The latter are modeled over time with stochastic models that take into account their interdependence and, in the case of a market consistent ESG, are calibrated to market data that should reflect the market situation at the evaluation date.

The scenarios generated by an ESG are used as inputs for an ALM model that models the interactions between assets and liabilities, asset and liability management strategies and the policyholders behavior. The combination of these tools then allows to evaluate the economic value of the future cash flows relative to the insurer obligations in order to perform a Best Estimate valuation consistent with the market. As explained previously, the association of these two tools is essential in life insurance to take account of the interaction between asset and liability implied by the profit sharing scheme, and more generally by the different contractual options (surrender, etc.). Moreover, they also enable to compute the SCR and measure the impact of an instantaneous change in the economic and financial environment, as required by European regulation (stress-test scenarios used in the SCR evaluation).

Furthermore, since the regulator requires insurers to value their assets and liabilities in accordance with observable and measurable market prices and risks, the Solvency 2 Directive describes ESGs as essential tools to perform market consistent evaluation and projection of life insurance

products. Concretely, this translates into the fact that the generated economic scenarios have to be consistent with the market prices observed at the evaluation date.
It is important to note that the importance of ESG is not only due to the coming into effect of Solvency 2. The implementation of MCEV (Market Consistent Embedded Value) models and provision calculations methods under IFRS already required the use of stochastic models for the valuation of an insurer's balance sheet. ESGs allow insurance companies to manage their investments and their activity with greater precision in the current context of ever-increasing accounting standards and prudential requirements.

In addition to the interests previously given, the development of an ESG presents some other advantages for an insurer (several are taken from Qureshi (2013)):

The policyholder behavior can be modeled more accurately:
The policyholder behavior is the result of a complex process that depends on the overall economic environment, the financial situation, the insurer's product compared to the competition and the effectiveness of its financial communication. This complexity is found especially in the surrender issue. Policyholders compare the rates paid by their insurer to those of the competition and can then decide to surrender their policy, which impacts the insurer's balance sheet through realization of potential unrealized losses. The dynamic of surrender and their financial consequences in the short and long term can be modeled in a prospective insurance model.

It allows to hedge more efficiently against interest rate variation:
Interest rate variations can impact the financial result of an insurer through the rate guarantees towards the policyholders. Indeed, a decrease in the interest rates generally leads to an increase in the medium and long term financial liabilities, even if it may also lead to short-term unrealized gains beneficial to financial communication. Conversely, when interest rates increase, the rates paid by the insurer may not be as high as those of the market or competition and thus, cause a wave of surrenders. An ESG allows the insurer to better direct its asset and liability management in order to foresee market evolutions and their potential impact on the balance sheet and be able to set up hedging strategies against these rate variations.

It offers independence from the market:
Economic scenarios are necessary for the evaluation of the contractual options and financial guarantees inherent to life insurance product. However, in order to avoid that the majority of insurers rely on the same scenarios to compute their technical provisions and regulatory capital, the prudential supervisory authority does not provide them. Insurers can have recourse to ESG provided by specialized companies, but the acquisition may be expensive and the ESG obtained less suitable for the risk profile of the insurer than an internally developed ESG.

Remark 1.1. *An ESG enables insurers to model the close link between assets returns and revaluation of accumulated savings. Nonetheless, regulatory and contractual clauses for profit-sharing are defined in relation to the financial balance, which is based on the accounting standards that apply to insurers. Therefore, the elements that form the financial balance of an insurer (such as discount amortization) have to be accounted under social standards. Consequently, in order to be able to simulate future profit-sharing schemes and also compute all the income statement flows that have to be taken into account in the Best Estimate evaluation, it is necessary to model not only the evolution of the market value of the assets but also their book value, which makes the models particularly more complex.*

One can schematize the use of an ESG as follows:

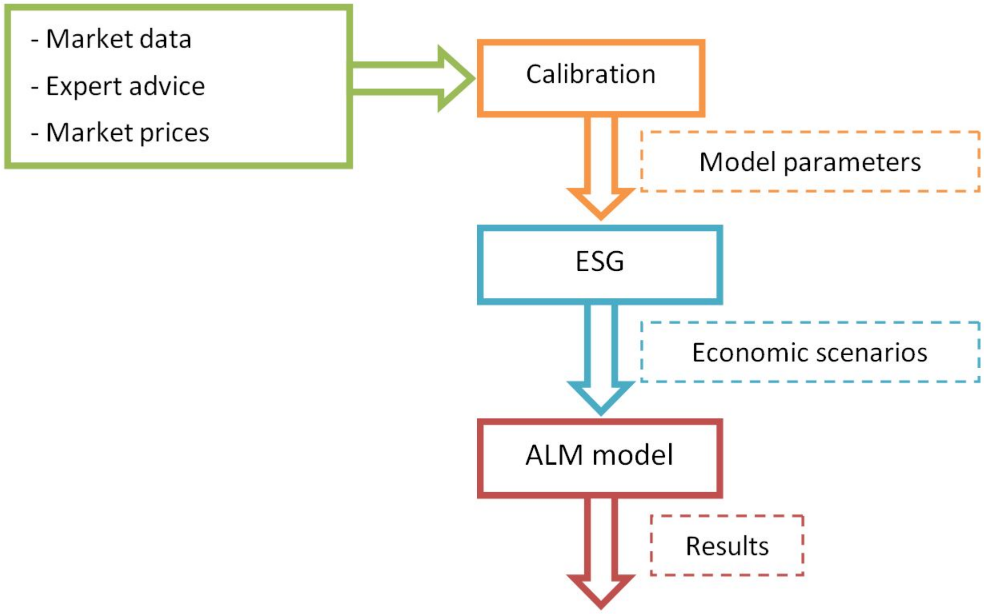


Figure 1.2 – Diagram summarizing the use of an ESG

1.2.2 Global structure of an ESG

The development of an ESG is divided into several essential stages:

1. Identification of risk sources: choice of the economic and financial variables to project in the ESG (interest rate, inflation rate, share return, etc.).
2. Determination of the projection method (real world / risk-neutral approach).
3. Choice of the stochastic models used to model the ESG economic and financial variables.
4. Choice of the dependence structure between risk sources: this step is decisive because the dependence structure between asset classes partly determines the coherence of the generated scenarios. The use of copula is the ideal approach for modeling dependent financial variables, but the simpler use of correlation matrices is allowed.
5. Model calibration: estimating the models parameters is a very delicate step. The estimation methodology must be adapted to the use and constraints, especially operational ones, of the ESG. The resort to expert advice might also be necessary.
6. Model validation: this raises the question of auditing the scenarios generated by the ESG. The validation of the choices made and the model quality can be based on several complementary criteria: statistical arguments about the model suitability to the data used, backtesting, consistency of the generated scenario with market data, etc.

1.2.3 Risk-neutral / Real world approach

In the context of valuation of liabilities in the insurance framework, two approaches are possible with an ESG:

Risk-neutral approach

With this approach, we suppose that there is no arbitrage opportunity and that all economic agents are risk-neutral. As a result, all risk premiums are null, which avoids the need of determining their complex form, and valuations can be conducted by discounting future cash flows at the risk-free rate. The neutrality assumption for economic agents is widely used for the pricing of derivative product in finance and of option derivative included in an insurance policy because it simplifies valuation calculations. Indeed, under the risk-neutral probability, the expected returns of financial assets is equal to the risk-free rate and their discounted prices are martingales. It is important to note that this approach is based on an evaluation principle and that the generated scenarios do not aim at reflecting the future evolutions expected by the insurer. For example, 5 year-long market consistent scenarios do not reflect the expectations of the insurer about the economy in 5 years, but can be used to estimate the market price of a derivative with maturity 5 years.

The use of the risk-neutral probability is also recommended by the first pillar of Solvency 2 since it facilitates market consistent evaluation of the insurer's liabilities, in particular contracts with financial options and guarantees. Models have to be calibrated to market prices quoted on the evaluation date, so that the generated scenarios respect the property of consistency with the observed market prices.

Real world approach

With this approach, we aim at replicating the historical behavior of the market data. The simulations project economic variables under the real world probability as closely as possible to the observable market environment. Models have to be calibrated to historical data representative of the economy over the projection horizon in order to project the actual behavior of these variables as observed in the historical data. Note that models can also be calibrated to anticipations of the future economic environment, so that the projected variables replicate these anticipations. This approach considers the real returns and therefore involves risk premium in assets returns, i.e. a return in excess of the risk-free rate, which greatly complicates assets valuation since risk premiums are in practice very difficult to model.
A real-world ESG allows to generate valid real world distribution for all the risk factors. The generated scenarios are then especially used for optimizing the strategic asset allocation, performing calculations in the ORSA framework or computing the SCR.

Evaluations based on an ESG can be conducted either with the risk-neutral approach or the real world one provided that a deflator is used to ensure the consistency between the two projection methods. A deflator is a stochastic discounting function that includes a time component and measures the market value of a risk from a projection of the flows under the real world probability.

1.2.4 Different types of ESG

To model the different asset classes and their dependence structure, ESG can be divided into two broad categories:

- Composite models that first separately model the dynamics of each asset class before aggregating them with a dependence structure between asset classes and obtaining an overall description of the portfolio,
- Integrated model that model asset classes based on an explanatory variable.

Composite models

The main advantage of composite models lies in their simplicity of implementation, since a correlation matrix alone can synthesize the dependence structure between asset classes. Nevertheless, with this approach, it may be difficult to simulate a consistent global economic environment. In particular, as a variable that affects all asset classes, inflation is difficult to include in this type of model. On the contrary, integrated models allow to develop a more coherent structure that takes into account the global impact of the inflation.

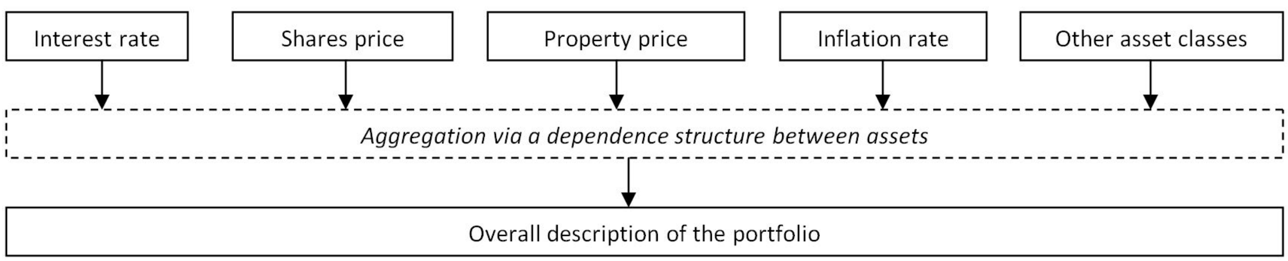


Figure 1.3 – Structure of a composite model

Integrated models

Because of its impact on the premiums paid and the benefits provided, inflation has a major importance in the insurance framework and thus, the majority of integrated models use this key variable as an explanatory variable. We present in the following the most famous integrated models within the actuarial community, considering that the ESG includes the following asset classes: obligation, share, property and monetary. The structural schemes of the following models are taken from Planchet et al. (2009) and presented in appendix A.1.

Wilkie model:
Wilkie model was introduced for the first time in 1985 and still remains the best known integrated model. Its main characteristics are modeling shares prices based on dividend rates and amounts and including the inflation rate as well as long-term interest rates. The inflation rate plays a major role in this model since it impacts many economic parameters. However, Wilkie model may appear

quite summary in the modeling of each asset class and is not well-suited for the calculation of prices according to a market consistent approach.

Brennan and Xia model:
Brennan and Xia model, presented in 2000, relies on stochastic models of real interest rates, inflation and shares to derive a stochastic consumer price index and a stochastic real discount factor. The inflation and the real interest rate play a major role in the modeling structure. Nonetheless, this model is more restrictive than Wilkie and Ahlgrim models because it only considers real interest rates, inflation and share. This feature may seem questionable in the sense that property is not modeled while it can represent up to 40% of the portfolio of an insurance company.

Ahlgrim model:
Ahlgrim model was introduced for the first time in 2005 in order to offer an alternative to the models given previously and their limits while emphasizing the accessibility of the approach. The model integrates, amongst others, interest rate, share return, inflation rate and property return with the ability of using stochastic model developed for each asset class. Moreover, it emphasizes the interdependence between the different economic and financial variables considered. However, the model implementation can be complex since the projection of short and long-term rates depend on that of inflation.

1.2.5 Constructing a risk-neutral ESG

When constructing a risk-neutral ESG, regulation gives insurers a certain amount of freedom for developing a model that complies with regulatory requirements. First, the ESG must take as an input the EIOPA zero-coupon curve so that the interest rate model of the ESG exactly replicates the current risk-free term structure to discount cash flows. Secondly, as a market consistent ESG, the ESG models have to be calibrated to market prices or volatilities, quoted on the evaluation date, of financial instrument that can reflect the nature and term of the insurer's liabilities. Those are the two principal constraints imposed by the regulator for constructing a risk-neutral ESG.

The insurer has then the choice on how to reach to the desired risk-neutral ESG, which will have to check two main theoretical properties:

- The market consistency criterion: the aim of market consistent scenarios is to reproduce market prices and implied volatilities;
- The martingale criterion: if there is no arbitrage opportunity, the discounted prices of financial assets are martingales under the risk-neutral probability.

The freedom of the insurer is found in several stages. First of all, the implementation of a risk-neutral ESG begins with the selection of the most appropriate models to project macro-economic variables of the market (interest rate, inflation etc.). The insurer must question the characteristic and market phenomena (yield curve shape for example) that it wants to be reproduced by the models since they do not have the same theoretical properties. It must also put these choices in perspective with the current economic environment. For example, interest rate models that can generate negative rates were until recently criticized for this characteristic. However, in the current context of low interest rates, the use of these models may seem more attractive to insurers, and even be explicitly required by the regulator[1]. This reflection is a major step in the development of an ESG because its results largely depend on these initial choices that form the key components of the ESG.

The calibration of an ESG represents another very delicate step in the development process, and will greatly influence the correct market consistent valuation of the portfolio. The insurer must first select which market data to use to calibrate the ESG models, ensuring that they reflect in a certain manner the market situation at the evaluation date and that they are adapted to the considered models. The imperfection of stochastic models means that the latter can not simultaneously reproduce all the observed market prices. Hence, it is necessary to resort to an optimization problem, which nonetheless depends heavily on operational constraints, especially the available computing power. The results obtained can be very sensitive to the method retained (optimization algorithm) and to the market data used to calibrate the models parameters.

Finally, it is necessary to check the reliability of the generated scenarios by verifying that the theoretical properties of the underlying models and the regulatory requirements for ESG are respected. In the case of a risk-neutral ESG, we can especially test the market consistency of the generated scenarios and the martingality of the discounted simulated processes. The market consistency criterion is the easiest one to check because the test simply consists in repricing financial instrument whose price or implied volatility is quoted on the market and comparing the resulting model value to the market one. The martingale criterion is more difficult to check since different methodologies could be proposed to test the martingale property of the discounted simulated asset prices.

This freedom given to the insurer emphasizes the importance of the actuarial function which has to justify the solutions retained for each step of the development process, and possibly arbitrate some modeling problems. Nevertheless, the Solvency 2 requirements in terms of ESG must not be achieved at the expense of an overly complex model. Even though a more complex model would allow to perform projections more consistent with the real market conditions, the actuarial function must focus more on respecting a parsimony criterion in order to favor simple but well-controlled models, whose assumptions can be clearly justified based on the insurer's risk profile, rather than models that are as much complex as they are unadapted to operational constraints.

We propose a study of several issues mentioned above. We will study in a first part the impact of the interest rate model of a risk-neutral ESG and its calibration on the generated scenarios and the Best Estimate market consistent evaluations. This first part will consist in particular in observing, by means of two families of interest rate models, the consequences of the choice of model (with identical market data to calibrate them) on the evaluations. A specific attention will also be paid to the choice of the market data used to calibrate the model. In a second part, we will pay attention to the tests that can be employed to audit the economic scenarios generated by a risk-neutral ESG, in particular to statistical martingale tests. To summarize, the overall issue of this study is to reach to a more precise knowledge of these different stages and to provide some answers to legitimate questions from the point of view of the actuarial function, in order to result in a better global mastery of ESG.

[1] Jérémie Courbon [2017] "Conférence ACPR du 16 juin 2017", *GALEA et Associés*, *http://www.galea-associes.eu/2017/08/conference-acpr-16-juin-2017/*

1.3 Data used in the study

This study is based on an existing ALM model developed by GALEA & Associés. The objective of this tool is to model the decision of purchase and sell of assets up to a given horizon, according to the cash-flows of liabilities, the minimum guaranteed rates and the revaluation rates of the insurance contracts. In the end, the tool computes the Best Estimate of the cash flows of liabilities and the present value of future accounting results. The ALM tool is composed of an ALM model, a composite risk-neutral ESG and a liability projection tool that projects the deterministic cash flows of liabilities of the insurer.
The calculations performed with the ALM model in this study are based on a modified portion of a life insurance company's asset and liability portfolio. We consider individual supplementary pension contracts, with individual subscription. Each contract is characterized by a technical interest rate that defines the minimum revaluation of the associated cash flows towards the policyholder. In addition to this minimum revaluation, these contracts are affected by the profit-sharing mechanism.

Part I

Impact of a risk-neutral ESG Calibration

Chapter 2

A Brief History of Interest Rate Modeling

2.1 Introduction

The interest rate model is the key component of an Economic Scenario Generator (ESG). It estimates the interest rates for different maturities, which especially enables to generate the zero-coupon bond prices and construct the yield curves. A stochastic interest rate model allows to deform the yield curve in a generated scenario and therefore represents the first necessary step to estimate the value of an insurance company's assets and liabilities in an economic scenario. Moreover, in a risk-neutral ESG, the short rates represent the drifts of the other asset classes. Nevertheless, the purpose of an interest rate model is neither to exactly explain past evolution of interest rates, nor to perfectly predict any future interest rate. It rather aims to generate a variety of plausible interest rate scenarios to which an insurance company might be exposed in the future.

Ideally, an stochastic interest rate model should feature the following properties:

- the model is complex enough to capture the various aspects of the yield curve dynamics observable in the market and to correctly reproduce the current term structure of interest rates and the market volatility curves,
- it implies closed form formulas for zero-coupon bond prices and interest rate derivatives (cap, floor, etc.), especially to ease the calibration process,
- it does not imply arbitrage opportunities,
- it is mean reverting, i.e. the interest rate tends to move to its average value over time,
- it is simple enough to ease its implementation and calibration in practice, especially through a simple discretization,
- because of the current context of low interest rates, it can generate negative rates.

In this chapter, we will introduce the main evolutions that occurred in the history of interest rate modeling. To do so, we will explain the reasons and approaches that led to the introduction of the most important existing interest rate models. This presentation does not aim to be exhaustive and much more information can be found in the voluminous literature dedicated to interest rate

modeling theory (see Brigo & Mercurio (2006) for example). Especially, we will pay more attention to the G2++ model and the Libor Market Model since we consider both these models in our study. We use the following notations:
- r_t is the instantaneous rate, usually referred to as instantaneous spot rate or briefly as spot rate,
- $P(t,T)$ denotes the value at time $t < T$ of a T-maturity zero-coupon bond,
- $R(t,T)$ is the spot interest rate prevailing at time t for the maturity T. We have: $r_t = \lim_{T \to t^+} R(t,T)$,
- $F(t,T,S)$ is the forward interest rate prevailing at time t for the expiry $T > t$ and maturity $S > T$.

2.2 One factor term-structure models

Black & Scholes (1973) approach to price options had a large impact on the theory of interest-rate modeling. Most interest rate models constructions were based on a similar approach, using the argument of arbitrage-free market and the construction of a suitable locally-riskless portfolio. The argument of arbitrage-free market is the fundamental assumption of the theory of interest-rate modeling. Indeed, this assumption allows to assume the existence of a risk-neutral measure (the uniqueness is guaranteed if the market is complete) such that, under this measure, every discounted price is a martingale.
The first interest rate models used the instantaneous short rate as fundamental quantity, which offered a large liberty in choosing the related dynamics. These models are known as *equilibrium models*. The pioneering approach to interest-rate modeling was proposed by Vasicek (1977). He used the basic Black & Scholes arguments to obtain an arbitrage-free price for any interest rate derivative. Cox, Ingersoll and Ross (Cox et al. (1985)) then presented a short rate model which ensures that the instantaneous short rate is always positive, contrary to the Vasicek model. Both of these time-homogeneous[1] models present several interesting properties, such as explicit formulas for useful quantities related to the interest rate world, simple simulation, etc. However, these models are *endogenous*, meaning that the current term structure of rates is an output rather than an input of the model. Thus, no matter how the model parameters are calibrated, the initial term structure of rates does not necessarily match the one observed in the market, which makes endogenous term-structure models unqualified for a market-consistent ESG.

Then, exogenous one-factor term-structure models, also known as *arbitrage free models*, were introduced in order to solve the poor fitting of the initial term structure of interest rates implied by endogenous models. Ho & Lee (1986) were the first to propose an exogenous short-rate model. Their pioneering approach showed how an arbitrage-free interest rate model can be designed so that it is automatically consistent with any specified initial term structure. Nevertheless, their model was based on the assumption of a binomial tree governing the evolution of the entire term structure of rates and therefore, the model (and its continuous time equivalent derived by Dybvig (1988) and Jamshidian (1988)) has the disadvantage that it incorporates no mean reversion in the short-rate dynamics. In addition, negative rates are possible and the model is also known to be computationally quite time consuming.
In order to exactly fit to the currently-observed yield curve, Hull & White (1990) introduced time-varying parameters in the Vasicek model. They considered that the market expectations about future interest rates involve time-dependent parameters because of the cyclical nature of the economy and of the expectations concerning the future trends in macroeconomic variables or the

[1] An interest rate model is *time-homogeneous* if the assumed short rate dynamics depends only on constant coefficients.

future impacts of monetary policies. Therefore, the drift rate and volatility of the instantaneous short rate should be functions of time. Hull and White presented an extended Vasicek model and an extended CIR one, which were both consistent with both the current term structure of interest rates and either the current volatilities of all spot rates or the current volatilities of all forward rates. Moreover, despite a Gaussian distribution, negative short rates are rarer in these models than in the Vasicek one.

Lognormal short rate models have also been considered to face the drawback of negatives rates since they can keep the rate away from zero. Furthermore, it may seem reasonable to assume a lognormal distribution for the instantaneous short-rate process given that market formulas for cap and swaption pricing (Black's formulas) are based on the assumption of lognormal rates. Black, Derman and Toy (Black et al. (1990)) first proposed a binomial-tree model for a lognormal short-rate process. Nevertheless, this model has several drawbacks, especially it presents no analytic properties and its tree construction implies a time-dependence in two ways.
However, Hull and White pointed out that a normal (or lognormal) model with mean reversion can depend on time in three ways. Following this idea, they created a general lognormal model that depends on time in three ways by dropping the tie between the volatility parameter and the reversion rate parameter in Black, Derman and Toy' model. Black and Karasinski then used a special case of the general model developed by Hull and White. Black & Karasinski (1991) assumed that the short rate process evolves as the exponential of an Ornstein-Uhlenbeck process with time dependent coefficients, where the reversion rate and volatility are determined independently of each other, contrary to Black, Derman and Toy' model. The Black-Karasinski model presents a rather good fitting quality to market data, especially to the swaption volatility surface. However, this model has no analytic tractability and the expected value of the money-market account is infinite, no matter which maturity is considered, because of the assumed lognormal distribution.

In 2001, Brigo and Mercurio described a procedure to extend any time-homogeneous spot rate model in such a way that the extended model exactly fits to any observed term structure of rates and preserves the possible analytical tractability of the original model. The most relevant extended model is the CIR++ model derived in Brigo & Mercurio (2006) by applying this approach to the original CIR model. Especially, the CIR++ model features analytical formulas and can guarantee positive rates.
Moreover, we can note that dynamics with jumps have also been presented to model the short rate.

2.3 The Heath-Jarrow-Morton (HJM) Framework

Choosing the instantaneous short rate as the fundamental quantity to model the interest-rate evolution gives a large freedom in choosing the related dynamics. In particular, when using one-factor short-rate models, one can choose the drift and instantaneous volatility coefficient in the diffusion dynamics as one deems fit. However, short-rate model have some clear disadvantages:

- Most involve only one factor, and therefore only one source of uncertainty,
- The user does not have the complete freedom in choosing the volatility structure,
- An exact calibration to the initial curve of discount factors is difficult to achieve if the model is not analytically tractable.

By modeling the evolution of the entire yield curve in a binomial-tree setting, the Ho and Lee model (Ho & Lee (1986)) was the first historically important alternative to short-rate models. A new methodology emerged with Ho and Lee (1986) and Heath et al. (1992*a*). Their approach used the martingale measure approach to provide arbitrage-free prices that depend on an exogenous specified initial forward rate curve rather than the market price for risk. Heath, Jarrow and Morton (Heath et al. (1990)) then published a paper that described the no-arbitrage conditions that must be satisfied when modeling the yield curve. They translated Ho and Lee' intuition in continuous time but instead of focusing upon bond prices as Ho and Lee, they concentrated on forward rates. Their key result was the link between the drift and the diffusion coefficient of the instantaneous forward rate. Precisely, they proved that, in an arbitrage-free framework, the drift in the forward rate dynamics is a suitable transformation of the diffusion coefficient in the same dynamics. In other words, when modeling the instantaneous forward rate, one does not have any liberty in selecting the drift of the process because it is fully determined by the volatility coefficient. Hence, in their arbitrage-free framework, whose fundamental quantities are the instantaneous forward rates, the forward rate dynamics is fully specified through their instantaneous volatility structures. This is a major difference with arbitrage free one factor short-rate dynamics. In 1992, Heath, Jarrow and Morton (Heath et al. (1992*b*)) introduced the HJM model which is the best example of this framework. The main advantage of the HJM theory is the fact that any exogenous term-structure interest-rate model can be derived within this framework. Indeed, the current term structure of rates is by construction an input of the HJM model. Nonetheless, only a restricted class of volatilities implies a Markovian short-rate process.

2.4 Multi-Factor Short-Rate Models

The major drawback of one-factor models is that they induce perfect correlation between instant rates for all maturities in the yield curve. Furthermore, one-factor models are quite likely to badly reproduce some typical shapes that can be observed in the market. So, when the correlation plays a major role or when a higher precision is needed, multi-factor models may represent a better solution to provide a realistic evolution of the zero-coupon curve. To choose the number of factors, one must find a compromise between the capability of the model to represent realistic correlation structures and numerically-efficient implementation. Some historical analysis of the whole yield curve (for example Jamshidian & Zhu (1997)) have shown that two components can explain 85% to 90% of variations in the yield curve.

The first multi-factor models proposed to model the dynamics of the short and long-term rates. For example, Brennan & Schwartz (1982) introduced a two-factor model which models both the spot interest rates and the long term rates. However, this kind of model has been dropped for several reasons. Firstly, unlike one-factor short-rate models, explicit formulas can not be derived for zero-coupon bonds. Secondly, these models are in contradiction with the result proved by Dybvig et al. (1996) who explained that in an arbitrage-free framework, the long-term rates can not decrease. This is why most of the multi-factor models used nowadays do not include long-term rates but are based on the assumption that the short-rate dynamics depends on several factors. Many two-factor short-rate models are simply generalization of well-known one-factor short-rate models. There also exist models that model the dynamics of the short rate and the inflation rate (for example, the Richard model in Richard (1978)).

2.4.1 The G2++ Model

Brigo & Mercurio (2006) introduced a two-factor additive Gaussian model called G2++. One can prove the equivalence between the G2++ model and the Hull-White two-factor model. They assumed that the dynamics of the instantaneous-short-rate process under the risk-adjusted measure is given by:

$$r_t = x_t + y_t + \varphi(t), \quad r(0) = r_0,$$

where the processes x_t and y_t satisfy:

$$dx_t = -ax_t dt + \sigma dW_x(t), \quad x(0) = x_0,$$
$$dy_t = -by_t dt + \eta dW_y(t), \quad y(0) = y_0,$$

with (W_x, W_y) a two-dimensional Brownian motion with instantaneous correlation ρ such that:

$$dW_x(t)dW_y(t) = \rho dt,$$

and where r_0, a, b, σ, η are positive constants and $-1 \leq \rho \leq 1$. The function φ is deterministic and chosen so as to exactly fit the current term structure of discount factors:

$$\varphi(T) = f^M(0,T) + \frac{\sigma^2}{2a^2}\left(1 - e^{-aT}\right)^2 + \frac{\eta^2}{2b^2}\left(1 - e^{-bT}\right)^2 + \rho\frac{\sigma\eta}{ab}\left(1 - e^{-aT}\right)\left(1 - e^{-bT}\right),$$

with $f^M(0,T)$ the instantaneous forward rate at time 0 for a maturity T of a given term structure.

Brigo & Mercurio (2006) introduced this model for several reasons. First of all, the use of two factors allows a better description of the actual variability of market rates. Secondly, the Gaussian distribution makes the model quite analytically tractable in that explicit formulas can be derived for discount bonds and a number of financial instruments, such as European swaption, cap and floor. Furthermore, unlike some two-factor models (especially the CIR2++ model), the G2++ model stays analytically tractable even if the correlation between the two Brownian motions is not null, which implies a more flexible correlation structure and "a more precise calibration to correlation-based products like European swaptions"(Brigo & Mercurio (2006), p.142). Moreover, the G2++ model is easier to implement in practice than some other two-factor models (for example the CIR2++) and there are only five parameters to calibrate.
Nonetheless, the G2++ model has several drawbacks: because of the Gaussian distribution, the spot rate process can assume negative values with positive probabilities; the model does not fit precisely to the volatility surface of interest rate derivatives; the interpretation of the model and its parameters may be difficult.

2.5 The LIBOR Market Models

Markovian term structure models (with normally distributed interest rates or log-normally distributed bond-prices) that imply closed-form solutions for interest rate derivatives have been introduced by many authors. However, these models imply negative interest rates with positive probabilities and are therefore not arbitrage-free. Alternatively, models with log-normally distributed interest rates have been proposed to avoid negative interest rates. Unfortunately, such rates explode with positive probabilities (as shown by Morton (1988) and Hogan & Weintraub

(1993)) and thus, they also imply arbitrage opportunities. Moreover, no closed-form solutions were known for these models for a long time. Yet, market practice was to price caps and floors assuming that the underlying forward rate process was lognormally distributed with zero drift. This assumption allowed to use Black's formula to price the inherent options of these derivatives. Still, the academic community was skeptic toward this market practice because, in an arbitrage-free setting, forward rates can not all be lognormal under one arbitrage-free measure since they are related to one another over consecutive time intervals.

2.5.1 The Standard LIBOR Market Model

Before the LIBOR Market Model was introduced, there was no interest-rate dynamics compatible with Black's formulas. Miltersen, Sandmann and Sondermann in Miltersen et al. (1997) and Brace, Gatarek and Musiela in Brace et al. (1997) first proved that this market practice could be consistent with an arbitrage-free term structure model. They derived an unified term structure model that gives closed form solutions for cap and floor, consistent with Black's formulas used by market practitioners, within the context of a log-normal interest rate model.
The first main key of their approach was the result shown by Sandmann & Sondermann (1994). They proved that the problem of exploding rates disappears if one assumes that simple interest rates are log-normally distributed over a fixed finite period instead of assuming that the continuously compounded interest rates are log-normally distributed. Indeed, Brace et al. (1997) explained that consecutive quarterly or annual forward rates can all be lognormal in an arbitrage-free framework if each rate is log-normal under the forward arbitrage-free measure rather than under one spot arbitrage-free measure.
The resolution of the term-structure puzzle came through many stages. For example, Sandmann and Sondermann were the first to propose the change of focus from the instantaneous continuously compounded rates to the instantaneous effective annual rates, while Brace, Gatarek and Musiela proved the existence of a unique non-negative solution of the stochastic differential equation (under suitable regularity conditions) that describes the dynamics of the forward rates in this framework. Jamshidian (1997) also contributed significantly to the development of the model.
Finally, under the assumption of deterministic volatility and lognormal distribution for forward rates (in the appropriate measure), Miltersen et al. (1997) and Brace et al. (1997) derived, with different approaches, a no-arbitrage term structure model of the HJM type where the stochastic behavior of the term structure of interest rates is determined by the forward rates and where simple closed form solutions for pricing cap and floor can be derived. This model is called the LIBOR Market Model, also known as the Brace-Gatarek-Musiela model or the "Log-normal Forward-LIBOR Model"(Brigo & Mercurio (2006)). It can be seen as a collection of Black models that would be simultaneously evolved under a specified measure.

We shall consider:

- A sequence of dates $0 = T_0 < T_1 < ... < T_M$. The pairs (T_{i-1}, T_i) represent the expiry-maturity pairs of dates for a family of spanning forward rates.
- τ_i denotes the year fraction associated with the expiry-maturity pair (T_{i-1}, T_i), i.e. $\tau_i = T_i - T_{i-1}$.
- The generic forward rate $F_k(t) = F(t; T_{k-1}, T_k)$, $k = 1, ..., M$. We have $F_k(T_{k-1}) = L(T_{k-1}, T_k)$ with $L(u, s)$ the LIBOR rate[1] at time u for maturity s.

[1] The London Interbank Offered Rate is a benchmark rate that is charged between some of the world's leading

- Q^k denotes the forward measure for the maturity T_k, that is to say the probability measure associated with the numeraire $P(., T_k)$.

The **LIBOR Market Model** (LMM) models forward LIBOR rates F_k through a log-normal distribution under the relevant measure. Indeed, it assumes the following driftless dynamics for F_k under Q^k:

$$dF_k(t) = \sigma_k(t) F_k(t) dZ_k(t), \quad t \leq T_{k-1}, \tag{2.1}$$

with $\sigma_k(t)$ the instantaneous volatility at time t for the forward LIBOR rate F_k and $Z(t)$ an M-dimensional Brownian motion under Q^k with instantaneous covariance $\rho = (\rho_{i,j})_{i,j=1,\ldots,M}$ such that:

$$dZ_i(t) Z_j(t) = \rho_{i,j} dt.$$

The LMM has become an industry standard for interest rates modeling. Indeed, unlike short rate models previously discussed, the LMM captures the dynamics of the entire curve of interest rates by choosing dynamic LIBOR forwards as its fundamental quantity. The model is intrinsically multi-factor, thus it can capture with accuracy various aspects of the curve dynamics (parallel shifts, steppenings and flattenings, etc.). Furthermore, the LMM allows to choose an approximately stationary volatility and correlation structure of LIBOR forwards. This offers a solution to the problem faced by pre-LMM term structure models that tend to produce unrealistic volatility structures of forwards rates and therefore, lead to overvaluation of instruments depending on forward volatility. Moreover, as a log-normal model, it can only produce positive rates. However, the LMM is not Markovian and is far less tractable than the Hull-White model for example. In addition, because of the lognormal distribution, rates can explode with positive probabilities. This is a major drawback of the LMM in the actual context of low interest rates.

2.5.2 The Displaced Diffusion LIBOR Market Model

The typical shape of a term structure of volatility is a hump shape, but empirical observations show that the volatility does not simply monotonically decrease, it is also smiling. Therefore, an interest rate model should reproduce the volatility smile that can be observed on the equity markets. However, observations of the different shapes of the implied volatility surfaces obtained from the standard LMM show that this model does not feature the previous property. Indeed, the lognormal distribution of forward rates implies that the LMM model produces an implied volatility surface that is flat across strikes.
A displaced diffusion model was proposed to face this problem of constant volatility across strikes. Instead of modeling the forward rate as a diffusion, the Displaced Diffusion LMM models the forward LIBOR rate plus some constant δ called the displacement coefficient. The dynamics for F_k under Q^k becomes:

$$d\left[F_k(t) + \delta\right] = \sigma_k(t)\left[F_k(t) + \delta\right] dZ_k(t), \quad t \leq T_{k-1}.$$

If the diffusion coefficient δ is positive, the implied volatility produced by the model is decreasing across strikes, contrary to the standard LMM. Rebonato (2002) proved that for reasonable values of δ, rates can assume negative values with rather small probabilities. Furthermore, all the results of the standard LMM linked to the lognormal distribution are retained in this model.

banks for short-term loans. It is the primary benchmark, along with the Euribor, for short-term interest rates around the world. LIBOR rates are calculated for 5 currencies and 7 borrowing periods ranging from overnight to one year.

2.5.3 The Stochastic Volatility Displaced Diffusion LIBOR Market Model

Joshi & Rebonato (2003) proposed another extension of the standard LMM, the Stochastic Volatility Displaced Diffusion LMM which allows for stochastic instantaneous volatilities of the forward rates in a displaced diffusion setting. According to them, the unprecedented changes in cap and swaption implied volatilites that occurred in 1998 (in the aftermath of the Russia crisis) showed that a deterministic smile surface, even if it is non-flat (as implied by the Displaced Diffusion LMM), can not reproduce the complex shapes that can be assumed by the implied volatility smile. Hence, they proposed to complete the standard LMM with a “combined displaced-diffusion stochastic-instantaneous-volatility approach”(Joshi & Rebonato (2003)) to capture these changes in the shape of the implied volatility.
In the Stochastic Volatility Displaced Diffusion LMM, the instantaneous volatility evolves according to a diffusion process driven by a Brownian motion:

$$\begin{aligned} dF_k(t) &= a_k(t)\varphi\left(F_k(t)\right)\left[V(t)\right]^{\gamma} dZ_k(t), \\ dV_t &= a_V dt + b_V dW(t), \end{aligned}$$

where a_k and φ are deterministic functions, $\gamma \in \{1/2, 1\}$, a_V and b_V are adapted processes, and Z_k and W are possibly correlated Brownian motions.
This model can accurately fit the market smile surface and the caplet market. Moreover, the European swaption matrix can be approximated with a closed-form formula which does not need to carry out a Monte Carlo simulation.

Chapter 3

Libor Market Model Implementation

In this chapter, we will focus on the implementation of the standard Libor Market model in the ESG of GALEA & Associés. We did not implement one of its extensions because in this study, we especially deal with the impact of the derivative used in the calibration process. Since the standard LMM features more closed-form formulas for financial instruments, we chose to use the standard version, despite its clear weakness of not being able to generate negative rates, in particular in the current economic environment.

3.1 The dynamics of the Libor Market Model

This section is mainly inspired by Brigo & Mercurio (2006) and Schatz (2011). In the following, we will use the notation introduced in the last part (concerning the Libor Market Models). First of all, recall that, in the standard Libor Market model (LMM), the dynamics of each forward rate F_k under its individual forward measure Q^k is given by:

$$dF_k(t) = \sigma_k(t)F_k(t)dZ_k(t), \quad t \leq T_{k-1}, \quad \forall k = 1, ..., M. \tag{3.1}$$

To price a general contingent claim and for simulation purposes, the behavior of all forward rates has to be modeled simultaneously under one single measure. One can model (see Brigo & Mercurio (2006), p.213) the behavior of all forward rates simultaneously under the forward-adjusted measure Q^i in the three cases $i < k$, $i = k$ and $i > k$:

$$\begin{aligned}
&i < k, \quad t \leq T_i : dF_k(t) = \sigma_k(t)F_k(t) \sum_{j=i+1}^{k} \frac{\rho_{j,k}\tau_j\sigma_j(t)F_j(t)}{1+\tau_j F_j(t)} dt + \sigma_k(t)F_k(t)dZ_k^i(t),\\
&i = k, \quad t \leq T_{k-1} : dF_k(t) = \sigma_k(t)F_k(t)dZ_k^i(t),\\
&i > k, \quad t \leq T_{k-1} : dF_k(t) = -\sigma_k(t)F_k(t) \sum_{j=k+1}^{i} \frac{\rho_{j,k}\tau_j\sigma_j(t)F_j(t)}{1+\tau_j F_j(t)} dt + \sigma_k(t)F_k(t)dZ_k^i(t),
\end{aligned}$$

where Z^i is a M-dimensional correlated Brownian motion under Q^i (Z_k^i is the k^{th} component) with:

$$dZ_i(t)dZ_j(t) = \rho_{i,j}dt, \quad \forall i, j = 1, ..., M.$$

The deterministic diffusion term $\sigma_k(.)$ is commonly interpreted as the instantaneous volatility at time t for the LIBOR forward rate F_k. If these functions $\sigma(.)$ are bounded, then the equations above admit a unique solution (see Brigo & Mercurio (2006), p.209).

Remark 3.1. *When deriving the risk-neutral dynamics, some residual appear in the change of measure which lead to the appearance of continuous term in the dynamics, which is an undesirable property. Moreover, if we choose the terminal measure Q^M for simulation purposes, some rates will be more biased than others because the bias coming from the discretized drift term is distributed differently on different rates. In order to improve these two situations, we can consider a discretely rebalanced bank account numeraire B_d as an alternative to the continuously rebalanced bank account $B(t)$ whose dynamics is given by $dB(t) = r_t B(t) dt$.*

Consider the numeraire asset:

$$B_d(t) = \frac{P(t, T_{\beta(t)-1})}{\prod_{j=0}^{\beta(t)-1} P(T_{j-1}, T_j)},$$

with $\beta(t)$ such that $\beta(t) = m$ if $T_{m-2} < t \leq T_{m-1}, m \geq 1$. If we choose B_d as numeraire, we obtain the dynamics of the LMM under the spot LIBOR measure Q^d which is the measure associated with B_d:

Proposition 1. *The dynamics of forward LIBOR rates in the LMM under the spot LIBOR measure Q^d is:*

$$dF_k(t) = \sigma_k(t) F_k(t) \sum_{j=\beta(t)}^{k} \frac{\tau_j \rho_{j,k} \sigma_j(t) F_j(t)}{1 + \tau_j F_j(t)} dt + \sigma_k(t) F_k(t) dZ_k^d(t). \tag{3.2}$$

3.1.1 An equivalent specification of the LMM dynamics

In the following, we consider M forward rates which follow the dynamics under their respective individual forward measure Q^k given by:

$$d \ln F_k(t) = -\frac{\sigma_k(t)^2}{2} dt + \sigma_k(t) dZ_k(t), \quad t \leq T_{k-1}, \tag{3.3}$$

where $Z(t)$ is a M-dimensional correlated Brownian motion with $dZ_i(t)dZ_j(t) = \rho_{i,j}dt$ and $\sigma_i : [0, T_{i-1}] \to \mathbb{R}^+$ are one dimensional volatility functions with $i = 1, ..., M$. M is the number of factors (or dimension) of the model.

If the vector $\gamma_i(t) \in \mathbb{R}^M$ is such that:

$$\sigma_i(t) := |\gamma_i(t)| = \sqrt{\sum_{k=1}^{M} \gamma_{i,k}^2(t)}, \quad 0 \leq t \leq T_{i-1}, \quad 1 \leq i \leq M,$$

and

$$\rho_{i,j}(t) = \frac{\gamma_i^\intercal(t)\gamma_j(t)}{|\gamma_i(t)|\,|\gamma_j(t)|}, \quad 0 \leq t \leq min\left(T_{i-1}, T_{j-1}\right), \quad 1 \leq i, j \leq M,$$

then (3.3) is equivalent to:

$$d \ln F_k(t) = -\frac{\gamma_k^\intercal(t)\gamma_k(t)}{2} dt + \gamma_k^\intercal(t) dW_k(t), \quad t \leq T_{k-1},$$

with $Z(t) = (Z_1(t), ..., Z_M(t))$ a standard M-dimensional Brownian motion.

For the drift we get:

$$\begin{aligned}\mu_i(t) &= \sigma_i(t) \sum_{j=i+1}^{k} \frac{\tau_j \rho_{i,j}(t) \sigma_j(t) F_j(t)}{1 + \tau_j F_j(t)} dt, \\ &= \sum_{j=i+1}^{k} \frac{\tau_j F_j(t)}{1 + \tau_j F_j(t)} \gamma_j^\intercal(t) \gamma_i(t) dt.\end{aligned}$$

Since we want to obtain an equivalent model, we set:

$$\gamma_i(t) = \sigma_i(t) e_i(t), \quad e_i \in \mathbb{R}^M, \tag{3.4}$$

with:

$$\rho_{i,j}(t) = e_i^\intercal(t) e_j(t) = \sum_{k=1}^{M} e_{i,k}(t) e_{j,k}(t),$$

where e_k are unit vectors in $\mathbb{R}^M$, i.e. $|e_i(t)| = 1$, and $\sigma_i : [0, T_{k-1}] \to \mathbb{R}^+$. We can easily see that:

$$|\gamma_i(t)| = |\sigma_i(t) e_i(t)| = \sigma_i(t).$$

Therefore, there exists a $M \times M$ matrix $E = (e_{i,j})_{i,j=1,...,M}$ such that:

$$dZ_i(t) = \sum_{k=1}^{M} e_{i,k}(t) dW_k(t), \quad dZ(t) = E(t) dW(t).$$

Moreover, if we multiply E by an orthonormal matrix $P \in \mathbb{R}^{M \times M}$, with $PP^\intercal = I_M$, then we obtain the same scalar volatility.

This equivalent specification (especially the formulation (3.4)) shows that the covariance between forward rates can be decomposed into the volatility of each forward rate and the correlation structure, without loss of generality. As a result, it enables separate calibration method to both the caps and swaptions markets in the sense that σ_i only influences the caplets prices while e_i specifies the correlation structure. Indeed, as we will see later on, the correlation between forward rates does not afflict the payoff of a caplet. On the contrary, the value of a swaption is influenced by the joint distribution of forward rates. So, volatility parameters σ_i can first be calibrated to the caps market while the remaining parameters stay free for swaption calibration, in order to contain the market information about the correlation structure between forward rates.

3.2 Specification of the instantaneous Volatility and Correlation structures of Forward Rates

The specification of the functional forms for the instantaneous volatility σ_i and the instantaneous correlation structure ρ of forward rates is a crucial modeling step when implementing the Libor Market Model. Indeed, given the high dimensionality of the model (i.e. the dimension of the Brownian motion Z), one must find a balance between the number of parameters to calibrate and the model flexibility to reproduce the shapes of the term structure of volatilities and to calibrate to the given market data.

3.2.1 Volatility Structure

There exist many different specifications of the instantaneous volatility σ_i. First of all, it is often assumed that the instantaneous volatility is piecewise-constant, meaning that it is constant in "expiry-maturity"time interval where the forward rate is alive.

Table 3.1 – Piecewise-constant parametrization of the instantaneous volatility

Instantaneous volatility	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	...	$(T_{M-2}, T_{M-1}]$
$F_1(t)$	$\sigma_{1,1}$	Dead	Dead	...	Dead
$F_2(t)$	$\sigma_{2,1}$	$\sigma_{2,2}$	Dead	...	Dead
...	...	...	...	...	...
$F_M(t)$	$\sigma_{M,1}$	$\sigma_{M,2}$	$\sigma_{M,3}$	...	$\sigma_{M,M}$

Still, in table 3.1, the model is over-parametrized and the number of volatility parameters to calibrate may be too high. If we assume stationarity or time-homogeneity of the volatilities, then the volatilities of the forward rates depend only on the time to maturity $T_k - T_{\beta(t)-1}$ (for example, $\beta(t) = 1$ if $t \in (0, T_0]$) rather than on time t and maturity T_k separately and therefore, the number of parameters σ is significantly reduced. With this assumption, the above table becomes:

Table 3.2 – Piecewise-constant parametrization of the instantaneous volatility assuming stationarity

Instantaneous volatility	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	...	$(T_{M-2}, T_{M-1}]$
$F_1(t)$	η_1	Dead	Dead	...	Dead
$F_2(t)$	η_2	η_1	Dead	...	Dead
...	...	...	...	...	...
$F_M(t)$	η_M	η_{M-1}	η_{M-2}	...	η_1

The time-homogeneity assumption is motivated by observations of the term structures of forward rates. Indeed, in general, the term structure shape does not change significantly over time. Rebonato (2002) proved that time homogeneity of the volatility term structure can be preserved if the volatility is only a function of time to maturity.
This piecewise constant specification of the instantaneous volatility can be perfected. In particular, the specification in table 3.2 is not suitable for accurate calibration of the LMM as it does not take into account the well-known financial phenomenon of mean reversion of long term rates. Lesniewski

(2008) proposes to introduce a volatility kernel which would take into account "the deviation of the instantaneous volatility form from the purely stationary model".

Along with piecewise-constant models, parametric forms for instantaneous volatilities have also been introduced. These parametric specifications forms may be preferred to the piecewise-constant ones in case of a long projection horizon since the number of parameters to calibrate may be too high with a piecewise-constant specification.
Let's first consider a widely used parametric form, due to Rebonato, and often called the abcd formula:

$$\sigma_i(t) := [a + b\,(T_{i-1} - t)]\, e^{-c(T_{i-1}-t)} + d. \tag{3.5}$$

Its popularity is in particular due to the fact that it is "sufficiently flexible to allow an initial steep rise followed by a slow decay, and that its square has an analytical integral"(Joshi (2008)). In other words, the volatility term structure is "hump shaped". Rebonato (2002) gives an explanation of this hump shape by analyzing the caplet market over different maturities. Therefore, a suitable parametric form should be able to both reproduce this hump shape and respect the time-homogeneity restriction. In addition, this specification possesses an analytical integral, presented in the appendix A.2, which avoids computationally expensive numerical integration schemes in the calibration process. Nevertheless, Brigo & Mercurio (2006) warns that its flexibility is not sufficient to jointly calibrate the LMM to both the caps and swaptions markets.

One must add the following constraints on the parameters in the calibration process:

1. $a + d > 0$,
2. $d > 0$,
3. $c > 0$,
4. $b > 0$.

These constraints can be easily interpreted because they allow to preserve the short and long time behavior and the humped shape of the term structure of volatilities. If we consider $t \to T_{i-1}$ and $T_{i-1} \to \infty$, we easily see that the first two constraints insure that the instantaneous volatility stays positive in these two extreme cases. We also note that $a + d$ and d can be interpreted as the volatilities of the forward rates with respectively the shortest and longest maturity. Furthermore, $(b - ca)/cb$ is the location of the extremum of the humped curve and $b > 0$ constrains the extremum to be a maximum.

The formulation (3.5) can be perfected into a richer parametric form by introducing factors ϕ_i that measure the extent to which time homogeneity is lost, such that:

$$\sigma_i(t) := \phi_i \left([a + b\,(T_{i-1} - t)]\, e^{-c(T_{i-1}-t)} + d \right).$$

If the "abcd"formulation is already able to reproduce markets prices, then the factors ϕ_i should be close to 1. These additional factors ϕ's add flexibility to the parametric form in order to improve the joint calibration of the LMM to the caps and swaptions markets. In particular, they can be parametrized such that the model fits perfectly to the market caplet volatilities (see Brigo & Mercurio (2006), p.224-225 for more details). However, to calibrate this parametric form, you have to have as many quoted caplets as there are parameters ϕ_i and therefore, as there are forward rates to project, which is impossible if the projection horizon is long. One could extrapolate market data over the projection horizon, but it may produce inconsistent results if the horizon is too long.

3.2.2 Correlation Structure

In the previous section, we presented several possible volatility structures for the LMM. We will now focus on possible formulation for instantaneous correlation. This is a central issue if one wants to calibrate the model to both the cap and swaption markets in a consistent manner.
Without any specification, the correlation matrix associated with the LMM is a $M \times M$ full-rank matrix, with M the number of forward rates. Given symmetry and the ones in the diagonal, this matrix is characterized by $M(M-1)/2$ entries which have to be calibrated. This number is too high in practice and thus, parsimonious parametric forms have to be used for ρ in order to reduce the number of parameters. We can distinguish two kinds of parametric forms for the correlation matrix: full rank parametrization and reduced rank formulation.

We will also focus on rank reduction techniques which are frequently used to facilitate the implementation of a multi-factor LMM. Indeed, without any restriction, the number of stochastic factors, i.e. the number of Brownian motions driving the LMM dynamics, is equal to the number of forwards, which leads to slow performance when simulating LIBOR forward rates. Hence, one has to consider reducing the rank of the LMM, i.e. reducing the number of factors in the dynamics.

Reduced rank Formulation

These correlation structures are based on Rebonato's angles parametrization to directly obtain a reduced-rank correlation matrix. Indeed, since ρ is a positive definite symmetric matrix, linear algebra proves that:

$$\rho = PHP^{\intercal},$$

with H a diagonal matrix of the positive eigenvalues of ρ and P a real orthogonal matrix whose columns are the eigenvectors of ρ. We also have:

$$\rho = AA^{\intercal}, \tag{3.6}$$

with $A = P\Lambda$ and Λ the diagonal matrix of the square roots of the entries of H.

The problematic of reduced-rank formulations is finding a suitable n-rank $M \times n$ matrix B such that $BB^{\intercal}$ is an n-rank correlation matrix which can approximate the decomposition $\rho = AA^{\intercal}$, with $n << M$. Reduced-rank formulations are widely used because they allow to use as new noise a standard n-dimensional Brownian motion W which replaces the original M-dimensional random shocks $dZ(t)$ by $BdW(t)$. Thus, the correlation structure moves from $dZdZ^{\intercal} = \rho dt$ to $BdW\,(BdW)^{\intercal} = BB^{\intercal}dt$. Put differently, we assume that only n independent Brownian motions drive the process.

The resulting correlation matrix is: $\rho^B = BB^{\intercal}$. The issue is to choose a suitable parametric form for the matrix B. Rebonato (1999) suggests a general spherical form for the i-th row of the matrix B:

$$\begin{aligned}
b_{i,1} &= \cos(\theta_{i,1}), \\
b_{i,k} &= \cos(\theta_{i,k}) \sin(\theta_{i,1}) \ldots \sin(\theta_{i,k-1}), \quad 1 < k < n, \\
b_{i,n} &= \sin(\theta_{i,1}) \ldots \sin(\theta_{i,n}),
\end{aligned}$$

for $i = 1, ..., M$. The instantaneous volatility structure is then said to be written in terms of spherical coordinates. Rebonato (1999) shows that using spherical coordinates allows for a more robust minimization scheme in the calibration process. Moreover, this formulation guarantees the independence between the total instantaneous volatility of the i-th forward and the different forward rates correlation.
With this parametrization (based on $M \times (n-1)$ parameters $\theta_{i,k}$), the resulting correlation matrix ρ^B is positive semidefinite and its diagonal terms are ones. The main advantage of this formulation is that we are sure that once the parameters $\theta_{i,k}$ have been calibrated, we can directly perform simulations of a n factor LMM with an n-dimensional independent shocks structure dW by using the M-dimensional correlated shocks BdW, without any other transformation.

Because of the desirable properties of the resulting correlation matrix, the spherical parametrization is widely used today. However, if M is large, it might be difficult to calibrate the resulting large number of parameters $\theta_{i,k}$. Brigo & Mercurio (2006) suggests in this case selecting a subparameterization for the θ's such that $\theta_k = f(k)$ with f a function depending on a small number of parameters, even though they explain that it may complicate the calibration process.

Full rank Parametrization

Full rank parametrizations maintain a full-rank correlation matrix but characterized by a smaller number of parameters than $M(M-1)/2$. With this kind of parametric forms, a Principal Component Analysis (PCA) is usually applied in order to reduce the rank of the resulting correlation matrix.

Schoenmakers & Coffey (2000) proposed several convenient full rank parametrizations that lead to correlation matrix with appealing qualities (such as increasing along sub-diagonals, positive semi-definite), see Brigo & Mercurio (2006), p247 for some examples. Rebonato (1999) also proposed some parametric forms, such as:

$$\rho_{i,j} = \rho_\infty + (1 - \rho_\infty)\, e^{-\beta|i-j|}, \quad \beta \geq 0, \tag{3.7}$$

with ρ_∞ the asymptotic level of correlations and β the decay rate of correlations. Brigo & Mercurio (2006) points out that this formulation does not necessarily lead to positive semi-definite correlation matrix, contrary to Schoenmakers and Coffey' parametrizations. However, Rebonato's formulation is often preferred because the parameters are easier to calibrate and the property of positive semi definitiveness is hardly violated in practice by the resulting correlation matrix. We also note the more flexible parametric form proposed by Lutz (2011) to more accurately fit to historical correlation matrices.

Nonetheless, in contrast to reduced rank formulations, full rank parametrizations present the undesirable feature of full dimensionality, that is to say, despite a parametric specification, you need a M-dimensional random shocks structure Z to perform simulations. To alleviate this problem, one has to find a reduced rank correlation matrix $\rho^{(n)}$ which would be the best approximation of the full M-rank correlation matrix ρ, with typically $n << M$. We would then obtain a n factor LMM (usually $n = 2$ or 3).

One possible approach consists in performing a Principal Component Analysis (PCA). As in the previous section, we want to mimic the decomposition $\rho = AA^\intercal$ with a suitable n rank $M \times n$

matrix B such that $BB^{\intercal}$ is a n rank correlation matrix which can accurately approximate the M rank correlation matrix ρ. Previously, the B matrix was defined as an angle parametrized matrix. It is now defined as follows (more details can be found in Brigo & Mercurio (2006), p253). Recall that with simple linear algebra, ρ can be written as $\rho = PHP^{\intercal}$. Consider now:

- the $n \times n$ diagonal matrix $\Lambda^{(n)}$ obtained from Λ, defined in (3.6), by only keeping its n highest diagonal elements (i.e. the eigenvalues) and shrinking the matrix correspondingly,
- The $M \times n$ matrix $P^{(n)}$ obtained from matrix P by keeping the columns (i.e. the eigenvectors) corresponding to the n highest eigenvalues, i.e. the diagonal elements of Λ^{n},
- The $M \times n$ matrix $B^{(n)} = P^{(n)}\Lambda^{(n)}$.

The resultant reduced rank correlation matrix obtained with this method is:

$$\bar{\rho}^{(n)} = B^{(n)} \left(B^{(n)}\right)^{\intercal}.$$

However, even if the matrix $\bar{\rho}^{(n)}$ is positive semi definite, it does not necessarily features ones in the diagonal, contrary to the correlation matrix directly obtained with spherical form in the previous section, owing to the eigenvalues taken away previously. That is why Brigo & Mercurio (2006) proposes to interpret $\bar{\rho}^{(n)}$ as a covariance matrix, so that we can derive the associated correlation matrix by defining:

$$\rho_{i,j}^{(n)} = \frac{\bar{\rho}_{i,j}^{(n)}}{\sqrt{\bar{\rho}_{i,i}^{(n)}\bar{\rho}_{j,j}^{(n)}}}.$$

$\rho^{(n)}$ is now an n-rank correlation matrix (with ones in the diagonal) which can approximate the original correlation matrix ρ.
Tanimura (2006) proposes another method to obtain ones in the diagonal of the resulting correlation matrix. He suggests to let the ones in the diagonal be one objective of the calibration process by minimizing the sum of squared errors between unit vectors and the diagonal components of the resultant rank-reduced correlation matrix.

So far, we have presented two methods to reduce the number of factors (i.e. the dimension of the random shocks structure) of the LMM. Note that another approach was proposed by Grubisic (2002) to reduce the dimension of the Wiener process governing the dynamics, called the "Projection Method".

3.2.3 Parametrizations retained for the implementation of the LMM

The Libor Market Model has been implemented in the ESG of the ALM model of GALEA & Associés. Recall that this ALM model projects retirement agreements. Thus, the projection horizon has to be sufficiently long to perform a projection consistent with the long duration of the cash flows involved in this kind of life insurance liabilities. As a matter of fact, the projection horizon usually retained is 80 years.

Thus, to choose which specifications of the volatility and correlation structures we shall use, we had to take into account the fact that we could not implement parametrizations whose number of parameters would depend on the projection horizon. In particular, piecewise constant specifications of the volatility or reduced rank parametrization of the correlation structure based on spherical coordinates could not be used since the number of parameters to calibrate would have been much too high in practice.

Hence, for the instantaneous volatility structure, we have chosen to use the abcd parametrization defined in (3.5), without the additional parameters ϕ. As concerns the correlation structure, we retained the parametrization in (3.7) and use a Principal Component Analysis to implement a two factor LIBOR market model (as the correlation specification depends on two parameters).

3.3 Monte Carlo Simulation

The LMM dynamics admit no known transition densities. Thus, we can not simulate a vector of LIBOR forward rates from time t to a horizon T in one step. Neither does it allow for an implementation based on recombining trees. Therefore, in order to perform simulations of the joint evolution of forward rates, we have to discretize time from t to T, with a sufficiently but not too small time step Δt.

Consider the dynamics of the LMM under the spot LIBOR measure given in (3.2). By Ito's formula and taking logs, we obtain:

$$d\ln F_k(t) = \sigma_k(t)\sum_{j=\beta(t)}^{k}\frac{\tau_j\rho_{j,k}\sigma_j(t)F_j(t)}{1+\tau_jF_j(t)}dt - \frac{\sigma_k(t)^2}{2}dt + \sigma_k(t)dZ_k(t).$$

Since the diffusion coefficient in the last equation is deterministic, the Euler discretization scheme coincides with the Milstein scheme. The discretization is therefore:

$$\begin{aligned}\ln F_k^{\Delta t}(t+\Delta t) = \ln F_k^{\Delta t}(t) + \sigma_k(t)\sum_{j=\beta(t)}^{k}\frac{\tau_j\rho_{i,j}(t)\sigma_j(t)F_j^{\Delta t}(t)}{1+\tau_jF_j^{\Delta t}(t)}\Delta t - \frac{\sigma_k(t)^2}{2}\Delta t \\ + \sigma_k(t)\left[Z_k(t+\Delta t) - Z_k(t)\right],\end{aligned}$$

with $Z_k(t+\Delta t) - Z_k(t) \sim \sqrt{\Delta t}\mathcal{N}(0,\rho)$. Since the projection horizon used in our case is long, the equation is discretized annually (i.e $\Delta t = 1$ year). This represents a first drawback of this study because of the resulting lack of precision of the discretization.

Forward rates simulation consists in implementing a diffusion triangle as follows:

Table 3.3 – Diffusion triangle of the LIBOR forward rates

T_0	T_1	...	T_i	...	T_M
$F_0(T_0)$	Dead	Dead	Dead	Dead	Dead
$F_1(T_0)$	$F_1(T_1)$	Dead	Dead	Dead	Dead
...	...	...	Dead	Dead	Dead
$F_i(T_0)$	$F_i(T_1)$	...	$F_i(T_i)$	Dead	Dead
...	...	...	...	...	Dead
$F_M(T_0)$	$F_M(T_1)$	...	$F_M(T_i)$	...	$F_M(T_M)$

The forward rate $F_i(T_j)$ (red cell) is derived with $F_k(T_{j-1})$, $j \le k \le i$ (green cells). The forward rates in the first column can easily be derived from the given EIOPA zero-coupon curve. We then have to compute every forward rate in each column before moving to the next one.

Once the forward rates have been computed with a discretization process, one can easily derive the associated zero coupon prices at time $0 \leq t < T_k$ through:

$$P\left(t, T_k, T_{k+m}\right) = \prod_{i=1}^{m} \frac{1}{1 + \delta F(t, T_k, T_{k+i})}.$$

Remark 3.2. *Discretizing the continuous time dynamics of LMM does not lead to arbitrage free LMM in discrete time. In other words, bond prices implied by a discretization of LMM dynamics are not martingales when expressed under the relevant numeraires. Besides, Glasserman & Zhao (2000) proposed a discretization scheme that maintains the martingale property required by no arbitrage in discrete time. Nevertheless, the violation of the arbitrage free condition can be considered as negligible if the discretization step Δt is small enough (which is generally the case in practice).*

Chapter 4

Libor Market Model and G2++ Calibration

4.1 Introduction

4.1.1 Founding Principles

In order to generate meaningful and market-consistent economic scenarios, an interest rate model must be able to replicate market prices and observed volatility and correlation surfaces. That is why calibration to given market data is such a crucial step when it comes to modeling. We can define calibration as the process of estimating model parameters consistently with market information. Depending on which type of projection one wants to implement, we distinguish two different approaches:

- Risk neutral (or endogenous, implied) calibration where the model parameters are derived from market liquid products like cap or swaption, in order to reproduce market prices and implied volatility or correlation quoted on the evaluation date,
- Historical (or exogenous) calibration where we fit the model parametrization to empirically observed market data (or to anticipations of the future economic environment).

The purpose of market consistent calibration is to reproduce the prices liabilities would be traded at if they were really traded on the market. Accordingly, models should be calibrated to financial instruments that can reflect the nature and term of these liabilities. In practice, we calibrate (in general with a least squares minimization problem) ESG models to market prices of assets which have similar characteristic to insurers liabilities. Therefore, to ease the calibration process, one must try to find closed-form formulas to evaluate the prices of assets whose market prices can be easily recovered (with a professional market data provider, such as *Bloomberg*). Indeed, even if it is in general possible to evaluate theses prices using Monte Carlo simulations, the presence of closed-form formulas (giving exact or approximated prices) can considerably reduce the computation time of the calibration process. In the end, if the risk neutral calibration process was correctly executed, a market consistent ESG should be able to replicate market prices.

As concerns historical calibration, the purpose is to estimate parameters such that the model replicates an empirically observed history of market prices. This introduces a new problem compared to risk neutral calibration: the choice of the initial date and the horizon of the historic. In this case, to calibrate the parameters, statistical methods like Moments Matching or Maximum Likelihood are commonly used.

In our case, we will only focus on risk neutral calibration process. When calibrating a market consistent model to given market data, we seek for the parameters that minimize the difference (or error) between the prices obtained with our model and market prices. Precisely, this problem can be reduced to a least squares minimization problem:

$$\theta^* = \arg\min_{\theta} \sum_{i} \left[\omega_i \left(\text{Model Price}\,(i,\theta) - \text{Market Price}(i)\right)\right]^2 .$$

Depending on which function we want to minimize, we can choose weights ω's to be:

- all equal to 1. We then consider squared absolute errors, which gives more weight to instruments with high prices,
- equal to $1/MarketPrice(i)$. We then consider squared relative errors, which gives the same weight to all instruments,
- equal to specified values chosen by the user in order to give more weight to specific instruments.

We are then left with two problems: choosing the derivatives to calibrate the model, and choosing the optimization algorithm for finding the best fit to market data.

4.1.2 Optimization algorithm

We can distinguish two families of optimization methodologies: deterministic (or local) and stochastic algorithms. Once again, one must find a balance between time and efficiency. Indeed, on the one hand, deterministic algorithms do not guarantee a global optimum (only a local one) but are much faster. On the other hand, stochastic algorithms have a higher probability of convergence to the global optimum but are much slower.

Firstly, when using a local algorithm, a preliminary choice is required because these algorithms need an initial point to start their routines, unlike most stochastic algorithms. Moreover, even if a deterministic algorithm always converges to a minimum, if we start from a random initial point, it will most likely get a local minimum. Therefore, we have to find starting values near the region of the global minimum in order to converge to the desired global minimum. In this case, one can either guess starting values (by taking the latest calibrated parameters for example) or generate a large number of initial points and then run the optimization process from different best guesses (see Gilli & Schumann (2010) for more details). Nevertheless, even with this approach, the probability that the deterministic algorithm converges to a local minimum (rather than the global one) remains high. In fact, the main weakness of local algorithms is that at each iteration, they always directly move to states that reduce the value of the objective function. Thus, in presence of local minimums, they have high chances of getting stuck into one of them.

We now give a few examples of local optimization algorithms. First, many local algorithms (such as Newton method or Quasi-Newton method) use gradient descent to find local minimum.

However, in this case, the objective function has to be differentiable, amongst other properties, otherwise the algorithm depends heavily on the starting values. Secondly, the Nelder-Mead (or Downhill Simplex) algorithm is a commonly applied local optimization method in multidimensional space which does not require the objective function to be differentiable, as it only needs evaluations of the target function. It is based on a simplex[1] that can be expanded, contracted or shrunk at each iteration depending on the function evaluations. Still, to be efficient, the feasible region has to be small enough so that the algorithm can test a sufficiently high number of different combination of parameters. In addition, the Nelder-Mead method is known not to be well adapted to optimization problems with saturated constraints (i.e. when optimum parameters can be equal to their corresponding extrema).

Local optimization algorithms may be useful when minimizing smooth functions with appealing properties (monotonic, convex). Otherwise, stochastic algorithms should be preferred to perform global optimization, despite their longer computation time. Indeed, contrary to deterministic ones, stochastic algorithms introduce some randomness in the optimum research, which increases the probability of convergence to the global optimum, and can even guarantee it if we suppose that the algorithm runs infinitely long. Especially, as concerns interest rate model calibration, stochastic algorithms appear particularly useful because the corresponding objective functions generally present many strong local minimums (due to their complexity and the high number of parameters) and are highly sensitive to small variation in input data (market prices, initial zero-coupon curve).

Simulated annealing and genetic algorithm are two examples of stochastic algorithms. We also note differential evolutionary algorithms (recommended by Di Francesco (2012) in an article in which he calibrated a G2++ model) and algorithms based on random initialization and multiple restarts. First, a simulated annealing algorithm is inspired by thermodynamic theory and works as follows:

- At each iteration, it generates a random new point whose position from the previous one is determined by a probability distribution, known as the acceptance probability and given by the Boltzmann distribution (which can be specified by the user). This acceptance probability has to be parametrized such that points that minimize the objective function have high probabilities and vice versa.
- Just like a deterministic algorithm, it converges to new points that reduce the value of the objective function. However, it can also move to states that increase the value of the target function with a specific probability (also specified by the user). As a result, the probability of being stuck into a local minimum is reduced since the algorithm is able to consider more combinations of parameters. As a matter of fact, if the distributions specified present some specific properties, the convergence to the global minimum is theoretically guaranteed.

Still, a simulated annealing algorithm requires an initial point to start. On the contrary, genetic algorithm avoids this need of starting values. Genetic algorithm is inspired by the process of natural selection and based on natural evolution theories to solve optimization problem. It belongs to the larger class of evolutionary algorithms. However, it is known to be very computationally intensive and it does not guarantee a convergence to the global minimum, not even to a local one.

Eventually, a best practical method for complex optimization problem would be to first call a stochastic algorithm to at least converge to the region of the global minimum, and then use a local

[1] a polytope which generalizes the notion of triangle to arbitrary dimensions

deterministic algorithm in order to refine the solution. In this study, we used this method as much as possible. Still, some optimization problems were too much time consuming to use a stochastic algorithm and therefore, we had to use local algorithms for them. In those cases, we used the method previously suggested, that is running the algorithm from different initial points to be sure that we found a solution that can reasonably be considered as a global optimum.

4.1.3 Input Market Data

In this section, we first introduce the two main derivatives (cap and swaption) used for calibration purposes. Since these instruments are very liquid assets and considered vanilla by market participants, their market price can be reasonably supposed to be correct.

Cap and Floor

Before defining a cap (floor), we have to introduce caplets (floorlet) since a cap (floor) contract can be decomposed as a portfolio of caplets (floorlets). A caplet is a call option on a forward-rate agreement[1] (FRA), or equivalently on a forward LIBOR rate. The discounted payoff of a caplet which resets at time T_{i-1} (i.e. the payoff is fixed in T_{i-1}), pays at T_i, whose nominal and strike are respectively N and K, is:

$$D\left(t,T_i\right) N\tau_i \left[L\left(T_{i-1},T_i\right) - K\right],$$

with $D\left(t,T_i\right)$ the discount factor and $L\left(T_{i-1},T_i\right)$ the LIBOR rate from T_{i-1} to T_i.

If we consider the individual measure for each caplet, then the underlying LIBOR forward rate is log-normally distributed. We can then evaluate each caplet payoff with the Black's formula:

$$\mathbf{Cpl}^{Black}\left(0,T_{i-1},T_i,N,K\right) = P\left(0,T_i\right) N\tau_i \mathrm{Bl}\left(K,F_i(0),\upsilon_i\right),$$

where:

$$\begin{aligned}
\mathrm{Bl}\left(K,F_i(0),\upsilon_i\right) &= F_i(0)\Phi\left[d_1\left(K,F_i(0),\upsilon_i\right)\right] - K\Phi\left[d_2\left(K,F_i(0),\upsilon_i\right)\right],\\
d_1\left(K,F,\upsilon\right) &= \frac{\ln(F/K) + \upsilon^2/2}{\upsilon},\\
d_2\left(K,F,\upsilon\right) &= \frac{\ln(F/K) - \upsilon^2/2}{\upsilon},
\end{aligned}$$

with Φ the standard Gaussian cumulative distribution function, $\upsilon_i = \sqrt{T_{i-1}}\upsilon_{T_{i-1}-caplet}$ and $\upsilon_{T_{i-1}-caplet}$ the T_{i-1}-caplet volatility.

As mentioned above, a cap (floor) is a collection of caplets (floorlet) with a common strike K. Let $\mathcal{T} = \{T_\alpha, ..., T_\beta\}$ where $T_{\alpha+1}, ..., T_\beta$ and $T_\alpha, ..., T_{\beta-1}$ are respectively the payment and reset dates, with $\tau = \{\tau_{\alpha+1}, ..., \tau_\beta\}$ the associated year fractions. A cap can be considered as successive call options on forward LIBOR rates. In other words, the discounted payoff of a cap is equal to:

$$\sum_{i=\alpha+1}^{\beta} D\left(t,T_i\right) N\tau_i \left(L\left(T_{i-1},T_i\right) - K\right)^+. \tag{4.1}$$

[1]A FRA is a contract which allows its holder to lock-in the interest rates between two fixed future maturities at a desired value (the strike), see Brigo & Mercurio (2006) p.11 for more details.

Still, the market practice to price a cap is to apply a sum of Black's formulas:

$$\mathbf{Cap}^{Black}\left(0, \mathcal{T}, N, K\right) = N \sum_{i=\alpha+1}^{\beta} \tau_i P\left(0, T_i\right) \mathrm{Bl}\left(K, F_i(0), \sqrt{T_{i-1}} \upsilon_{\mathcal{T}-cap}\right).$$

The market then quotes caps in Black volatilities $\upsilon_{\mathcal{T}-cap}$.

Remark 4.1. *The risk neutral expectation of a cap discounted payoff is given by:*

$$N \sum_{i=\alpha+1}^{\beta} P\left(0, T_i\right) \tau_i E^i\left[\left(F_i\left(T_{i-1}\right) - K\right)^+\right],$$

with Q^i the forward-adjusted measure. Notice that the joint distribution of forward rates is not involved in the payoff of a cap. As a result, only the marginal distributions of single forward rates are necessary to compute the expectation of a cap payoff while the correlation between forward rates does not afflict it.

Remark 4.2. *A cap is said to be at-the-money (ATM) if and only if the strike K is equal to the forward swap rate, i.e.:*

$$K_{ATM} := S_{\alpha,\beta}(0) = \frac{P\left(0, T_\alpha\right) - P\left(0, T_\beta\right)}{\sum_{i=\alpha+1}^{\beta} \tau_i P\left(0, T_i\right)}.$$

A cap is in-the-money if $K < K_{ATM}$ and out-of-the-money if $K > K_{ATM}$ (and vice versa for a floor).

In a calibration process, one often has to consider cap instead of caplets because, unlike caplets that are not quoted, caps are traded liquidly in the market. Indeed, the market quotes the flat volatilities of caps for a number of different maturities. The flat volatility, also called forward volatility, is derived such that if we apply it to each caplet with Black's formula, we obtain the cap market price.

We have extracted flat volatilities for EUR market ATM caps for which the reference rate is 6 month EURIBOR on the 31/12/2016 with Bloomberg. Note that cap black volatilities were not quoted before a maturity of 9 years, which therefore does not give us much market data to use for the calibration process.

Table 4.1 – ATM caps volatility at 31/12/2016

Maturity (in year)	9	10	12	15	20	25	30
Black ATM cap volatility	152,35%	115,17%	85,78%	70,27%	63,27%	64,16%	67,06%

Moreover, we can easily notice a large increase in the caps volatilities if we compare 31/12/2016 values to older ones, as can be seen in table 4.2 which gives market caps volatilities quoted on 31/02/2001 (given by Brigo & Mercurio (2006)) and on 18/08/2014 (given by Ferranto (2015)).

Table 4.2 – ATM caps volatility quoted on 31/02/2001 and 18/08/2014

Maturity	1	3	5	7	10	15	20	30
Volatility 31/02/2001	15,2%	16,4%	16,05%	15,55%	14,75%	13,50%	12,60%	
Volatility 18/08/2014	50,97%	58,73%	67,81%	60,85%	49,04%	40,02%	35,06%	32,05%

Swaption

We first need to introduce interest rate swap. An interest rate swap (IRS) is a "contract that exchanges fixed payments for floating ones at previously fixed time dates from a future time instant"(Schatz (2011). Consider a set of previously specified time $T_{\alpha+1}, ..., T_\beta$, a fixed interest rate K, a nominal N and year fractions τ_i. At every payment date, the fixed leg payment is equal to $N\tau_i K$ whereas the floating leg one is $N\tau_i L(T_{i-1}, T_i)$. The floating leg rate resets at $T_\alpha, ..., T_{\beta-1}$ and pays at $T_{\alpha+1}, ..., T_\beta$. There are two types of IRS. If the fixed leg is paid and the floating one is received, the IRS is a Payer IRS. In the other case, it is a receiver IRS. The discounted payoff of a payer IRS is:

$$N \sum_{i=\alpha+1}^{\beta} D(t, T_i) \tau_i (L(T_{i-1}, T_i) - K).$$

The fixed leg that makes fair an IRS contract fair at time t is called the forward swap rate $S_{\alpha,\beta}(t)$ and is given by:

$$S_{\alpha,\beta}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum_{i=\alpha+1}^{\beta} \tau_i P(t, T_i)} = \frac{1 - \prod_{j=\alpha+1}^{\beta} \frac{1}{1+\tau_j F_j(t)}}{\sum_{i=\alpha+1}^{\beta} \tau_i \prod_{j=\alpha+1}^{i} \frac{1}{1+\tau_j F_j(t)}}.$$

A swaption (or termed swap option) is an option on an interest rate swap. We also have to distinguish two types of swaption. A payer (receiver) swaption gives the holder the right (and no obligation) to enter a payer (receiver) IRS, at a prespecified time (the swaption maturity) and a prespecified strike. The length of the underlying IRS is called the tenor of the swaption. The swaption maturity often coincides with the first reset date of the underlying IRS.

The discounted payoff of a payer swaption is given by:

$$ND(t, T_\alpha) \left(\sum_{i=\alpha+1}^{\beta} P(T_\alpha, T_i) \tau_i (F_i(T_\alpha) - K) \right)^+. \tag{4.2}$$

Unlike caps, a swaption payoff cannot be decomposed easily. Hence, one has to consider the joint distribution of forward rates, and thus their correlation, in order to value swaption.

The market practice to value swaption is to use a Black like formula. The price of a payer swaption is then given by:

$$N.\mathrm{Bl}\left(K, S_{\alpha,\beta}(0), \sqrt{T_\beta}\sigma_{\alpha,\beta}\right) \sum_{i=\alpha+1}^{\beta} \tau_i P(0, T_i),$$

with $\sigma_{\alpha,\beta}$ is the Black volatility of the swaption quoted on the market.

Remark 4.3. *A swaption (either payer or receiver) is said to be at-the-money (ATM) if and only if the strike K is equal to the forward swap rate $S_{\alpha,\beta}(0)$. A payer swaption is in-the-money if $K < S_{\alpha,\beta}(0)$ and out of the money if $K > S_{\alpha,\beta}(0)$, and vice versa for a receiver swaption. Note that if we consider an ATM swaption, the Black like formula for swaption pricing becomes very simple given that $K = S_{\alpha,\beta}(0)$.*

Mat/Tenor	1	2	3	4	5	7	10	12	15	20	25	30
1					1,985	1,104	0,785	0,692	0,627	0,591	0,584	0,583
2				1,627	1,196	0,859	0,684	0,625	0,584	0,559	0,555	0,557
3			1,385	1,035	0,867	0,709	0,616	0,568	0,543	0,528	0,528	0,531
4	2,043	1,166	0,903	0,774	0,694	0,608	0,558	0,530	0,508	0,503	0,503	0,507
5	1,031	0,812	0,701	0,636	0,591	0,541	0,512	0,500	0,478	0,482	0,481	0,486
6	0,725	0,639	0,585	0,547	0,521	0,500	0,491	0,479	0,464	0,466	0,467	0,473
7	0,588	0,548	0,517	0,494	0,480	0,472	0,475	0,462	0,452	0,456	0,456	0,462
8	0,512	0,494	0,475	0,462	0,460	0,450	0,465	0,451	0,449	0,452	0,455	0,460
9	0,468	0,461	0,450	0,452	0,446	0,449	0,458	0,452	0,448	0,452	0,454	0,460
10	0,437	0,439	0,446	0,441	0,431	0,451	0,453	0,455	0,448	0,450	0,453	0,459
12	0,445	0,440	0,428	0,445	0,452	0,459	0,474	0,471	0,463	0,458	0,458	0,465
15	0,494	0,498	0,490	0,484	0,478	0,507	0,521	0,513	0,494	0,476	0,472	0,478
20	0,610	0,629	0,628	0,623	0,620	0,632	0,621	0,598	0,560	0,522	0,526	0,540
25	0,681	0,737	0,731	0,724	0,710	0,722	0,700	0,672	0,623	0,601	0,642	0,635

Figure 4.1 – ATM swaption Black volatilities at 31/12/2016

We have extracted in figure 4.1 Black swaption volatilities for EUR market ATM swaptions for which the reference rate is 6 month EURIBOR on the 31/12/2016 with Bloomberg.

We can once more notice in table 4.3 a large increase between swaption volatilities in 4.1 (quoted on 31/12/2016) and values given by Brigo & Mercurio (2006) (quoted on 13/02/2001) or by Ferranto (2015) (quoted on 18/08/2014). As already explained by Ferranto (2015), the large increase in the market volatility surface level can be partially explained by the extremely low levels of current interest rates, "which greatly influences the discount factor curve and the lognormal volatility"(Ferranto (2015)).

Table 4.3 – Differences between volatility surface levels

Date	13/02/2001	18/08/2014	31/12/2016
Lowest swaption volatility	8.4%	30.1%	42.83%
Highest swaption volatility	16.4%	66.05%	204.26%

EIOPA risk-free zero-coupon curve

A market consistent ESG should use as an input the EIOPA risk-free interest rate term structure. Still, to implement either the LMM or the G2++ model, we need some discount factors that cannot be derived from the given EIOPA term structure because some are not on the stripped grid. In order to estimate these missing rates, one possible approach for building the needed whole zero coupon curve is to fit a parametric form to the given EIOPA zero coupon curve. A parametric form will guarantee the continuity of the zero coupon curve and the analytical tractability of the forward rate. The model, which was already implemented in the ESG of GALEA & Associés, is the Nelson, Siegel and Svensson method. In this model, the spot rate at time t is expressed as follows:

$$R(0,t) = \beta_1 + \beta_2 \left(1 - \frac{e^{-t/\alpha_1}}{t/\alpha_1}\right) + \beta_3 \left(1 - \frac{e^{-t/\alpha_1}}{t/\alpha_1} - e^{-t/\alpha_1}\right) + \beta_4 \left(1 - \frac{e^{-t/\alpha_2}}{t/\alpha_2} - e^{-t/\alpha_2}\right).$$

Once the model has been calibrated at the evaluation date, we obtain the following zero-coupon curve in figure 4.2:

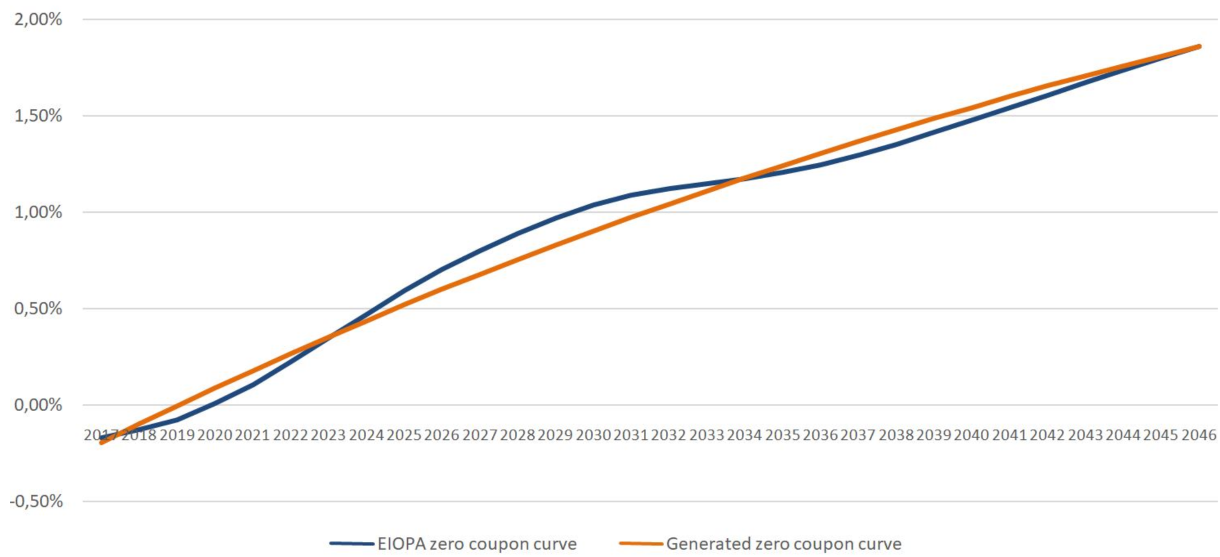


Figure 4.2 – Generated zero-coupon curve vs EIOPA zero-coupon curve at 31/12/2016

4.2 Libor Market Model Calibration

4.2.1 Theoretical Calibrations

We have implemented two different methods to calibrate the standard Libor Market Model. Indeed, as explained before, instantaneous correlation between forward rates afflicts the payoff of a swaption but not the one of a cap. Consequently, one can first calibrate the volatility parameters of the LMM to the caps market and then use the correlation parameters, which are the only parameters left, to fit the swaptions market. This first approach can be seen as a joint calibration to the caps and swaptions markets. The second method simply consists in calibrating in one step the volatility and correlation parameters to the swaptions market, without considering caps. Both approaches have been implemented, so that we can analyze the impact of these two different methods.

Joint calibration to the caps and swaptions markets

We will first focus on the joint calibration of the LMM to the caps and swaptions markets (i.e. the two-steps calibration approach). Recall that with the volatility and correlation parametrizations that we retained in section 3.2.3, we have six parameters to calibrate: a, b, c, d (volatility parameters) and β, ρ_∞ (correlation parameters).
Since correlation between forward rates does not afflict cap payoff, a cap volatility (or payoff) can be expressed in the LMM as a function of parameters a, b, c and d. Therefore, we will first try to fit as much as possible the given market cap prices to calibrate the four volatility parameters. Afterwards, we will use a closed-form approximation for swaption volatility to calibrate the remaining correlation parameters β and ρ_∞.

First of all, remember that the LMM allows to recover Black's formula, which is the standard formula for pricing European-style option on the market rate, such as cap. As the LMM was developed in particular to be consistent with market cap prices, one has to be particularly demanding with the fitting of the LMM to the caps market. Recall the Black's formula used to price caplet:

$$\mathbf{Cpl}^{Black}\left(0, T_{i-1}, T_i, N, K\right) = P\left(0, T_i\right) N \tau_i \mathrm{Bl}\left(K, F_i(0), \upsilon_i\right),$$

with $\upsilon_i^2 = T_{i-1} \upsilon_{T_{i-1}-caplet}^2$. In the LMM, we have:

$$\upsilon_{T_{i-1}-caplet}^2 := \frac{1}{T_{i-1}} \int_0^{T_{i-1}} \sigma_i(t)^2 dt,$$

where σ_i is the volatility parametrized by the user for the LIBOR forward rates. In our case (with the "abcd"formula), we obtain a closed-form formula for the squares of caplet volatilities (multiplied by time T_{i-1}):

$$\upsilon_i^2 = \int_0^{T_{i-1}} \left(\left[a + b\left(T_{i-1} - t\right)\right] e^{-c\left(T_{i-1}-t\right)} + d\right)^2 dt,$$

by using the formula in the appendix A.2.

Nonetheless, the market does not quote caplet volatilities but only cap ones, and assumes flat volatility among caplets that compose the caps. That is why several stripping or bootstrap algorithms have been proposed in order to recover implied caplet volatilities from quoted cap volatilities. Actually, we have distinguished three different methods to calibrate the volatility parameters to the caps market.

Method 1:
Brigo & Mercurio (2006) (p.226) proposed a simple stripping algorithm that allows to recover caplet volatilities by using Black's formula for caplet. Indeed, remember that caps are basically sums of individual caplets and that quoted cap volatilities are computed such that they allow to find the price of the associated cap with:

$$\mathbf{Cap}^{MKT}\left(0,\mathcal{T}_j,K\right)=\sum_{i=1}^{j}\tau_i P\left(0,T_i\right)\mathrm{Bl}\left(K,F_i(0),\sqrt{T_{i-1}}\upsilon_{\mathcal{T}_j-cap}\right),$$

with $\mathcal{T}_j=[T_0,...,T_j]$. Note that the market cap volatility $\upsilon_{\mathcal{T}_j-cap}$ is assumed for all caplets and that the strike K is the forward swap rate because we consider ATM caps. As there is a one-to-one correspondence between caplet price and caplet implied volatility, Brigo & Mercurio (2006) proposed to recover the caplet volatilities $\upsilon_{T_{i-1}-caplet}$ such that:

$$\sum_{i=1}^{j}\tau_i P\left(0,T_i\right)\mathrm{Bl}\left(K,F_i(0),\sqrt{T_{i-1}}\upsilon_{\mathcal{T}_j-cap}\right)=\sum_{i=1}^{j}\tau_i P\left(0,T_i\right)\mathrm{Bl}\left(K,F_i(0),\sqrt{T_{i-1}}\upsilon_{T_{i-1}-caplet}\right).$$

Then, we can compute caplet volatilities based on this equation for $j=1,2,3,...$ and using numerical non-linear equation solving. Once we have obtained the caplet volatilities, we can calibrate the volatility parameters with a least square optimization problem. Note that as cap volatilities are quoted annually, we first need to interpolate given market cap volatilities to obtain cap volatilities at each caplet maturity. For example, if all reset dates are six-months spaced, we need the cap volatilities for all these reset dates (i.e. every six months starting from the first reset date $T_0=6$ months), and not just the annual ones in order to apply this method. However, one main weakness of this approach is that we need to have cap volatilities for all reset dates, especially for the first ones. Still, in our case, we have seen in 4.1 that we only have at our disposal cap volatilities whose maturities are superior to 9 years. Since an estimation of earlier cap volatilities would probably be inaccurate, we did not implement this approach.

Method 2:
Hull & White (1999) proposed a method very similar to the first one but which follows a piecewise linear approach and does not require cap volatilities for all caplet maturities. First, we still need to perform interpolation to get the flat cap volatility for all reset dates. However, instead of using Black's formula to derive caplet prices, we compute the fair value of the caplets composing the caps by taking the difference between the adjacent cap prices:

$$\mathbf{Cpl}^{Black}\left(0,T_{i-1},T_i,K\right)=\mathbf{Cap}^{Black}\left(0,\mathcal{T}_i,K\right)-\mathbf{Cap}^{Black}\left(0,\mathcal{T}_{i-1},K\right).$$

We can then estimate caplet volatilities by inverting Black's formula. We tried to apply this method. However, the estimated caplet volatilities were too irregular (see the appendix A.3) to be reasonably used in the calibration process. In fact, we tried to calibrate the volatility parameters to these estimated caplet volatilities but we did not manage to obtain a satisfactory calibration (the resulting volatility was not hump-shaped or was even negative sometimes). Furthermore, Andersen & Piterbarg (2010) presented some possible problems with this method, especially when caps are quoted in a different strike range. Thus, we have decided to use a third method which does not need neither to estimate caplet volatilities nor to inverse Black's formula.

Method 3:
Andersen & Piterbarg (2010) and West (2010) introduced a different approach which allows to circumvent the problems of bootstrapping caplet volatilities and interpolating cap prices. We now directly consider caplet prices as functions of the volatility parameters a, b, c, d.
Let's denote $\mathbf{Cap}^{MKT}(\mathcal{T}_i)$ the market price of the i-th cap, $\mathbf{Cpl}^{LMM}(T_i, K_i, a, b, c, d)$ the model price of the i-th caplet (evaluated with Black's formula) with depends on the volatility parameters, N the number of different caps and $n_i (n_1 < ... < n_N)$ the number of caplets in the i-th cap. The objective function to minimize is:

$$\mathcal{F} = \sum_{i=1}^{N} \left(\sum_{j=1}^{n_i} \mathbf{Cpl}^{LMM}(T_j, K_i, a, b, c, d) - \mathbf{Cap}^{MKT}(\mathcal{T}_i) \right)^2 .$$

We thus consider the differences between the market cap prices and the model ones, computed as sums of caplet premiums calculated with Black's formula. The calibrated parameters are solutions to the following optimization parameters:

$$\underset{a,b,c,d}{\arg\min}\, \mathcal{F} \quad \text{subject to} \quad a + d > 0, b > 0, c > 0, d > 0.$$

Using this third method, we obtained the volatility parameters in table 4.4. Numerically, we used an algorithm based on random initialization and multiple restarts (R package *gosolnp*) and the Nelder-Mead algorithm to refine the solution previously obtained (R function *constrOptim*). We note in table 4.5 that with these parameters, the fitting to the market caps prices is good, except for the first maturities.

Table 4.4 – Calibrated LMM volatility parameters using caps volatilities

a	b	c	d
0.5542	1.8253	1.1416	0.4567

Table 4.5 – Comparison between market caps prices and LMM cap prices obtained with calibrated parameters in table 4.4

Maturity (in year)	9	10	12	15	20	25	30
Market cap price	0.054	0.063	0.081	0.109	0.168	0.249	0.348
LMM cap price	0.044	0.054	0.077	0.112	0.176	0.253	0.342
Relative error	-19.9%	-14.1%	-4.90%	2.53%	5.11%	1.93%	-1.67%

Remark 4.4. *When jointly calibrating the LMM to the caps and swaptions market, one can face the problem that the cap market forward rates are semi-annual whereas the forward swap rates are annual rates. Brigo & Mercurio (2006) (p.300) proposed a methodology to reconcile volatilities of semi-annual forward rates and those of annual forward rates, in order to implement a consistent joint calibration to the caps and swaptions markets. Indeed, they proposed formulas that enable to treat caplets as if they were on annual forward rates, just like forward swap rates. However, as we retained the method 3 previously introduced, which only used annual cap volatilities to calibrate the LMM to the caps market, we did not have to consider this issue.*

As concerns swaptions, the market does not quote directly swaption prices but rather quotes swaption volatilities such that when inserting into the Black-like formula for swaptions, we obtain the option premium. The major assumption of the Black-like formula is that forward swap rates are log-normally distributed. However, forward rates and forward swap rates cannot be log-normally distributed at the same time. As a result, since the LMM supposes that forward rates are log-normally distributed, there exists no exact closed form formula for swaption price in the LMM. Even it is possible to estimate swaption price by performing Monte Carlo simulations, this solution is computationally not viable. Nonetheless, approximate closed form formulas have been proposed to compute swaption prices, which gives us an ideal method to calibrate the LMM to the swaptions market since the accuracy of these approximations is recognized. The demonstrations of these approximations can be found in Brigo & Mercurio (2006) (p.281-284).

Consider a swaption whose floating leg rate resets at $T_\alpha, ..., T_{\beta-1}$ and pays at $T_{\alpha+1}, ..., T_\beta$. The first approximation is Rebonato's formula, from Rebonato (1998), for swaption volatility:

Proposition 2. *The (squared) swaption volatility (multiplied by T_α) can be approximated by:*

$$(\upsilon_{\alpha,\beta})^2 = \sum_{i,j=\alpha+1}^{\beta} \frac{w_i(0)w_j(0)F_i(0)F_j(0)\rho_{i,j}}{S_{\alpha,\beta}(0)^2} \int_0^{T_\alpha} \sigma_i(t)\sigma_j(t)dt, \tag{4.3}$$

where $S_{\alpha,\beta}(0)$ is the forward swap rate at time 0 and:

$$w_i(t) = \frac{\tau_i \prod_{j=\alpha+1}^{i} \frac{1}{1+\tau_j F_j(t)}}{\sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}}.$$

Hull & White (1999) proposed a more sophisticated and accurate approximation for swaption volatility:

Proposition 3. *The (squared) swaption volatility (multiplied by T_α) can be better approximated by:*

$$(\bar{\upsilon}_{\alpha,\beta})^2 = \sum_{i,j=\alpha+1}^{\beta} \frac{\bar{w}_i(0)\bar{w}_j(0)F_i(0)F_j(0)\rho_{i,j}}{S_{\alpha,\beta}(0)^2} \int_0^{T_\alpha} \sigma_i(t)\sigma_j(t)dt, \tag{4.4}$$

where:

$$\bar{w}_h(t) = w_h(t) + \sum_{i=\alpha+1}^{\beta} F_i(t) \frac{\partial w_i(t)}{\partial F_h},$$

$$\frac{\partial w_i(t)}{\partial F_h} = \frac{w_i(t)\tau_h}{1+\tau_h F_h(t)} \left[\frac{\sum_{k=h}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}}{\sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}} - 1_{\{i \geq h\}} \right].$$

Let's get back to our joint calibration process. Recall that we now only have to calibrate the correlation parameters β and ρ_∞. Since the volatility parameters have already been calibrated to the caps market, we can now express swaption volatility as a function of the only free parameters left, that is β and ρ_∞, and calibrate these parameters to the swaption volatility table 4.1. As it is known to be more accurate, we have retained Hull and White's formula for swaption volatility. We also tested Rebonato's formula to check if we obtained similar results, which was the case.

We have obtained the following correlation parameters. Note that these calibrated parameters seem inconsistent because of the value of ρ_∞ which is supposed to be the asymptotic level of correlations. Still, we have met the same problem in the other calibrations of the LMM, which are presented in the following.

Table 4.6 – Calibrated correlation parameters using the swaption volatilities and the volatility parameters in table 4.4

a	b	c	d	β	ρ_∞
0.5542	1.8253	1.1416	0.4567	0.4218	0.9999

The matrix of relative errors is presented below in figure 4.3. We notice that the fitting to the swaptions market is rather good though our calibrated LMM does not manage to reproduce all the fluctuations present in the market swaption volatility matrix. Especially, in the resulting model swaptions matrix, volatilities generally decrease when the swaption maturity or the tenor increase while the market matrix presents more fluctuations (for example, for a fixed tenor, the volatilities are first decreasing and then increasing with the swaption maturity).

Mat/Tenor	1	2	3	4	5	7	10	12	15	20	25	30
1					-72,1%	-54,5%	-39,7%	-33,0%	-27,5%	-25,0%	-25,7%	-27,5%
2				-59,6%	-50,2%	-38,2%	-28,2%	-23,4%	-20,3%	-19,1%	-20,6%	-22,9%
3			-47,4%	-37,4%	-31,3%	-24,3%	-19,1%	-14,6%	-13,2%	-13,5%	-15,7%	-18,4%
4	-56,0%	-32,5%	-23,2%	-18,9%	-15,9%	-12,1%	-10,4%	-8,1%	-6,6%	-8,7%	-11,0%	-14,0%
5	-19,1%	-8,7%	-5,4%	-4,3%	-3,2%	-2,3%	-2,7%	-2,7%	-0,7%	-4,5%	-6,7%	-10,0%
6	8,5%	10,5%	9,3%	8,3%	7,5%	4,5%	0,9%	1,2%	2,2%	-1,2%	-3,8%	-7,3%
7	27,4%	23,6%	19,8%	17,0%	14,8%	9,7%	3,8%	4,7%	4,7%	1,1%	-1,4%	-5,0%
8	40,6%	32,7%	27,1%	22,8%	17,8%	14,0%	5,5%	6,7%	5,2%	1,8%	-1,1%	-4,6%
9	48,7%	38,3%	31,2%	23,4%	19,9%	13,2%	6,6%	6,2%	5,1%	1,7%	-1,1%	-4,5%
10	54,5%	41,7%	29,8%	24,5%	22,6%	12,0%	7,3%	5,2%	4,9%	1,9%	-0,8%	-4,5%
12	44,8%	35,9%	31,1%	20,5%	14,8%	8,6%	1,6%	0,9%	1,1%	-0,2%	-2,3%	-5,7%
15	23,9%	15,1%	10,7%	7,9%	6,3%	-3,2%	-8,4%	-8,1%	-5,9%	-4,3%	-5,3%	-8,5%
20	-5,4%	-13,0%	-16,7%	-18,7%	-20,0%	-23,6%	-23,9%	-21,8%	-17,6%	-13,1%	-15,3%	-19,0%
25	-18,5%	-28,1%	-30,2%	-31,3%	-31,2%	-33,8%	-33,0%	-30,8%	-26,2%	-24,7%	-30,6%	-30,9%

Figure 4.3 – Relative errors between the swaptions volatilities derived from the LMM jointly calibrated to the caps and swaptions markets and the market ATM swaptions volatilities 4.1

Calibration of the LMM to the swaptions market only

We will now deal with a second method to calibrate the LMM, which consists in calibrating the model parameters to the swaptions market only, without considering caps. Using either Rebonato's formula 4.3 or Hull and White's one 4.4, we can express swaption Black volatility as a function of the LMM volatility and correlation parameters in order to calibrate them to the swaptions market. From there, we can either directly calibrate the LMM to the market swaptions Black volatilities or calibrate the model to the swaptions market prices by using one of these approximations and then putting the resulting quantities in Black's formula for swaption to compute approximated swaptions prices.
We have implemented these two different methods to see if they led to the same parameters, as could be obviously expected. In the previous calibration method, we have only considered a calibration of the correlation structure to the market swaptions Black volatilities because we have noticed that the correlation parameters had a far less important impact on the resultant model than the volatility structure, and therefore we have decided that it was not relevant enough in this case to implement two different calibrations to the swaptions market.
In the following, we will also study the impact on the calibrated parameters of the objective function retained in the minimization problem, by considering both the squared absolute and the squared relative errors.

Calibration to the market swaptions Black volatilities:
First of all, we consider a calibration of the LMM to the market swaptions Black volatilities. If we consider the squared absolute errors in the least squares minimization problem, we obtain the following volatility and correlation parameters:

Table 4.7 – Calibrated LMM parameters using only market swaption Black volatilities and minimizing squared absolute errors

a	b	c	d	β	ρ_∞
-0.4665	1.9654	0.8141	0.4666	3.5919	0.9947

If we consider the squared relative errors, we obtain the following LMM parameters:

Table 4.8 – Calibrated LMM parameters using only market swaption Black volatilities and minimizing squared relative errors

a	b	c	d	β	ρ_∞
-0.4953	1.0009	0.9222	0.4953	0.2998	0.9999

Except for the parameter a and β, these calibrations are quite close to the parameters previously obtained in tables 4.6 with the joint calibration, especially if we consider the squared absolute errors (notice the close values computed for parameters b and c). The relative errors matrices are presented below in figures 4.4 and 4.5. Note that if we consider squared relative errors, the swaption volatilities (and consequently the forward rates volatilities see figure 4.8) are lower than the ones obtained when considering squared absolute errors, as expected since we give less weight to higher values in the minimization problem.

Furthermore, we easily notice that even when calibrating the model to the swaptions market only, the LMM does not manage to reproduce all the fluctuations present in the market swaption volatility matrix. The global trends of the resulting swaption volatility matrix are the same as in the joint calibration.

Mat/Tenor	1	2	3	4	5	7	10	12	15	20	25	30
1					-58,8%	-39,0%	-27,5%	-22,7%	-19,6%	-19,5%	-21,7%	-24,3%
2				-45,9%	-33,8%	-22,1%	-16,0%	-13,3%	-12,5%	-13,8%	-16,7%	-19,7%
3			-33,4%	-19,2%	-12,1%	-7,4%	-7,1%	-4,4%	-5,5%	-8,3%	-11,7%	-15,1%
4	-52,0%	-19,4%	-3,9%	2,6%	5,2%	5,3%	1,7%	1,8%	1,0%	-3,5%	-7,0%	-10,6%
5	-10,3%	8,9%	17,0%	18,9%	18,7%	15,2%	9,3%	7,0%	6,8%	0,7%	-2,7%	-6,5%
6	20,8%	31,1%	33,6%	32,6%	29,9%	21,6%	12,4%	10,6%	9,5%	3,9%	0,3%	-3,6%
7	41,8%	45,8%	44,7%	41,3%	36,8%	26,2%	14,7%	13,7%	11,8%	6,2%	2,8%	-1,2%
8	56,2%	55,4%	52,0%	46,5%	38,7%	29,9%	15,9%	15,4%	12,0%	6,8%	3,1%	-0,7%
9	64,8%	60,9%	55,4%	45,8%	39,8%	28,0%	16,5%	14,4%	11,7%	6,7%	3,1%	-0,4%
10	70,7%	63,8%	52,4%	45,7%	41,7%	25,8%	16,9%	13,1%	11,2%	6,8%	3,6%	-0,2%
12	59,0%	55,4%	51,9%	39,1%	31,0%	20,7%	10,0%	8,0%	7,0%	4,8%	2,3%	-1,1%
15	35,0%	29,9%	26,2%	22,5%	19,5%	6,5%	-1,4%	-2,1%	-0,5%	0,6%	-0,5%	-3,5%
20	2,1%	-3,4%	-6,8%	-9,3%	-11,6%	-17,0%	-18,7%	-17,0%	-13,0%	-8,5%	-10,5%	-14,0%
25	-12,7%	-21,1%	-22,9%	-24,4%	-24,9%	-28,7%	-28,8%	-26,8%	-22,2%	-20,6%	-26,6%	-26,5%

Figure 4.4 – Relative errors between the swaptions volatilities derived from the LMM calibrated to the swaptions markets (considering squared absolute errors) and the market ATM swaptions Black volatilities

Mat/Tenor	1	2	3	4	5	7	10	12	15	20	25	30
1					-69,3%	-49,5%	-33,6%	-26,4%	-20,9%	-18,6%	-19,8%	-21,9%
2				-61,5%	-50,2%	-35,4%	-23,6%	-18,2%	-14,7%	-13,5%	-15,2%	-17,6%
3			-53,8%	-41,0%	-32,5%	-22,4%	-15,3%	-9,9%	-8,1%	-8,2%	-10,5%	-13,2%
4	-68,2%	-44,9%	-31,4%	-23,2%	-17,4%	-10,4%	-6,7%	-3,7%	-1,6%	-3,5%	-5,7%	-8,6%
5	-38,8%	-23,3%	-14,1%	-8,6%	-4,6%	-0,4%	1,2%	1,9%	4,4%	0,9%	-1,3%	-4,4%
6	-15,3%	-5,2%	0,6%	4,3%	6,6%	6,7%	5,0%	6,0%	7,5%	4,4%	1,9%	-1,4%
7	2,0%	8,1%	11,6%	13,6%	14,3%	12,4%	8,1%	9,6%	10,3%	6,9%	4,6%	1,2%
8	15,0%	17,7%	19,7%	20,1%	17,9%	17,1%	10,1%	11,9%	10,9%	7,7%	5,0%	1,9%
9	23,8%	24,3%	24,6%	21,5%	20,6%	16,6%	11,5%	11,6%	11,0%	7,9%	5,3%	2,2%
10	30,7%	28,9%	24,3%	23,2%	23,9%	15,7%	12,5%	10,8%	10,9%	8,3%	5,9%	2,6%
12	25,7%	26,0%	27,4%	20,6%	17,0%	12,8%	6,9%	6,7%	7,3%	6,6%	4,9%	1,9%
15	10,8%	9,1%	9,3%	9,3%	9,4%	1,4%	-3,0%	-2,3%	0,5%	2,8%	2,4%	-0,3%
20	-12,3%	-15,4%	-16,2%	-16,4%	-16,7%	-19,4%	-18,9%	-16,3%	-11,4%	-5,9%	-7,6%	-10,9%
25	-22,6%	-28,8%	-28,9%	-28,7%	-27,8%	-29,7%	-28,2%	-25,6%	-20,3%	-18,0%	-23,9%	-23,6%

Figure 4.5 – Relative errors between the swaptions volatilities derived from the LMM calibrated to the swaptions markets (considering squared relative errors) and the market ATM swaptions Black volatilities

Calibration to the market swaptions prices:
Secondly, we have calibrated the LMM to the market swaptions prices using Hull and White' approximation and the Black-like formula for swaption pricing. If we consider the squared absolute errors in the least squares minimization problem, we obtain the following parameters:

Table 4.9 – Calibrated LMM parameters using only market swaption prices and minimizing squared absolute errors

a	b	c	d	β	ρ_∞
-0.294	0.267	0.353	0.484	0.715	0.999

If we consider the squared relative errors, we obtain the following LMM parameters:

Table 4.10 – Calibrated LMM parameters using only market swaption prices and minimizing squared relative errors

a	b	c	d	β	ρ_∞
-0.487	1.0466	0.844	0.487	0.451	0.999

When considering squared relative errors, the resulting calibration is consistent with the one obtained when calibrating to the market swaptions Black volatilities, in the sense that the calibrated parameters are close. On the other hand, it is not the case when considering squared absolute errors, which seems a little odd. This might be due to an absence of convergence of our optimization program to the global minimum, despite the retained stochastic algorithm, which would represent a major weakness of the calibration processes.

The relative errors matrices are presented below in figures 4.6 and 4.7. Note that, at first sight, the model seems to better fit the market when considering squared relative errors in the minimization problem. Moreover, the relative errors tendencies are very similar to the ones previously obtained when calibrating to the market swaptions Black volatilities.

Remark 4.5. *Other methods have been introduced to approximate swaption price in the LMM. For example, Brace (1996) proposed a rank-one analytical formula to approximate swaption price. He extended in Brace (1997) his rank-one analytical approximation to the rank-r case. However, these approximated formulas are more complex and difficult to implement than Rebonato's or Hull and White' approximations, and present the undesirable feature of requiring an explicit eigenvalues computation. More details can be found in Brigo & Mercurio (2006) (p.271-281).*

Mat/Tenor	1	2	3	4	5	7	10	12	15	20	25	30
1					-62,3%	-40,1%	-22,7%	-16,3%	-12,7%	-13,5%	-16,5%	-19,6%
2				-53,8%	-42,0%	-24,8%	-12,4%	-8,1%	-6,7%	-8,3%	-11,7%	-15,1%
3			-46,5%	-33,7%	-23,5%	-10,9%	-4,0%	0,2%	-0,2%	-3,1%	-6,9%	-10,5%
4	-54,0%	-39,1%	-25,6%	-15,4%	-7,6%	1,5%	4,5%	6,1%	6,0%	1,5%	-2,2%	-5,9%
5	-33,6%	-19,4%	-8,3%	-0,4%	5,5%	11,3%	11,9%	11,1%	11,5%	5,5%	2,0%	-1,8%
6	-13,0%	-2,1%	6,2%	12,3%	16,3%	17,8%	14,9%	14,4%	13,9%	8,5%	4,9%	1,0%
7	3,0%	10,7%	16,7%	21,0%	23,3%	22,4%	17,1%	17,1%	15,8%	10,5%	7,2%	3,2%
8	15,0%	19,7%	24,1%	26,6%	26,0%	25,8%	18,0%	18,4%	15,8%	10,9%	7,3%	3,7%
9	23,1%	25,5%	28,1%	27,2%	27,5%	24,3%	18,3%	17,2%	15,1%	10,6%	7,2%	3,9%
10	28,8%	29,1%	27,2%	27,9%	29,4%	22,4%	18,3%	15,7%	14,4%	10,5%	7,4%	4,0%
12	24,0%	25,6%	28,3%	23,8%	21,4%	18,0%	11,9%	10,8%	10,2%	8,2%	6,0%	3,0%
15	11,2%	10,7%	11,8%	12,5%	12,9%	6,7%	2,5%	2,3%	3,5%	4,3%	3,3%	0,9%
20	-5,0%	-6,3%	-6,4%	-6,3%	-6,4%	-8,2%	-8,7%	-7,7%	-5,3%	-2,6%	-4,1%	-6,5%
25	-9,5%	-11,6%	-11,5%	-11,4%	-11,2%	-12,2%	-12,3%	-11,6%	-9,7%	-9,1%	-12,3%	-12,4%

Figure 4.6 – Relative errors between the swaptions prices derived from the LMM calibrated to the market swaptions prices (considering squared absolute errors) and the market ATM swaptions prices

Mat/Tenor	1	2	3	4	5	7	10	12	15	20	25	30
1					-62,9%	-45,9%	-31,8%	-25,4%	-20,6%	-18,8%	-20,2%	-22,4%
2				-51,9%	-43,3%	-31,2%	-21,5%	-17,0%	-14,1%	-13,5%	-15,4%	-17,8%
3			-42,5%	-33,0%	-26,3%	-18,4%	-13,2%	-8,8%	-7,6%	-8,3%	-10,6%	-13,4%
4	-46,9%	-33,4%	-23,2%	-16,8%	-12,4%	-7,3%	-5,2%	-3,0%	-1,6%	-3,8%	-6,1%	-8,9%
5	-27,6%	-15,6%	-8,2%	-4,0%	-1,1%	1,5%	1,8%	1,9%	3,8%	0,2%	-2,0%	-5,0%
6	-8,7%	-0,4%	4,2%	6,9%	8,4%	7,5%	5,0%	5,4%	6,4%	3,2%	0,9%	-2,2%
7	5,4%	10,4%	13,1%	14,4%	14,6%	12,0%	7,4%	8,4%	8,6%	5,3%	3,2%	0,0%
8	15,7%	17,9%	19,3%	19,3%	17,0%	15,5%	8,8%	10,1%	8,9%	5,9%	3,4%	0,6%
9	22,4%	22,7%	22,7%	19,7%	18,7%	14,7%	9,6%	9,5%	8,7%	5,9%	3,6%	0,9%
10	27,3%	25,7%	21,7%	20,5%	20,8%	13,4%	10,2%	8,6%	8,4%	6,0%	4,0%	1,2%
12	21,7%	21,9%	22,8%	17,2%	14,1%	10,3%	5,2%	4,9%	5,2%	4,4%	3,0%	0,5%
15	9,0%	7,7%	7,7%	7,4%	7,3%	1,1%	-2,4%	-2,1%	-0,3%	1,3%	0,9%	-1,1%
20	-6,4%	-8,2%	-8,9%	-9,3%	-9,8%	-11,6%	-11,8%	-10,5%	-7,7%	-4,6%	-5,7%	-7,9%
25	-10,5%	-13,0%	-13,3%	-13,6%	-13,5%	-14,6%	-14,4%	-13,5%	-11,4%	-10,5%	-13,4%	-13,4%

Figure 4.7 – Relative errors between the swaptions prices derived from the LMM calibrated to the swaptions markets (considering squared relative errors) and the market ATM swaptions prices

A plot comparing the instantaneous volatilities resulting from the calibrated "abcd"parameters in tables 4.4, 4.7, 4.8, 4.9 and 4.10 can be found in figure 4.8. Notice that we obtain the desired humped-shape. This plot shows that the five calibrations are consistent between each other as concerns the value of the parameter d (i.e. the volatility of the forward rate with the longest maturity) and that considering squared absolute errors in the calibration process lead to higher instantaneous volatilities than when using squared relative ones. In the legend of the figure, "Swaption volatility - absolute errors"for example means that the model has been calibrated to the market swaptions Black volatilities considering squared absolute errors in the optimization problem. However, the volatilities obtained with these parameters are very high, which will make these calibration unusable in practice with the ALM model as we will see later on.

Moreover, to test the consistency of these new calibrations with the one obtained with the joint

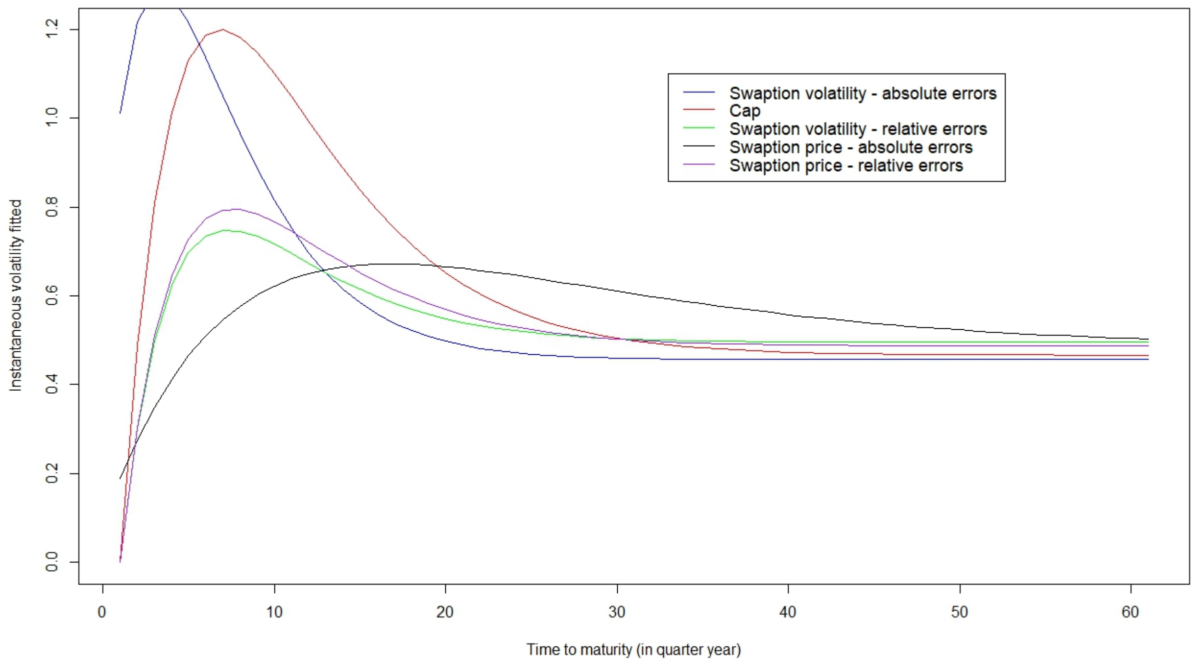


Figure 4.8 – Comparison between the instantaneous volatilities resulting from the "abcd"parameters of the five calibrations of the LMM

calibration to both the caps and swaptions markets, we have computed in table 4.11 the model caps prices resulting from the "abcd"parameters calibrated to the market swaptions Black volatilities in tables 4.7 and 4.8. We note that the LMM obtained when calibrating to the swaptions market only and considering squared absolute errors in the minimization problem fits rather well the caps market even though we did not use any information from it in the calibration process.

Table 4.11 – Relative errors between market caps prices and LMM caps prices obtained with parameters calibrated to market swaptions volatilities (tables 4.7 and 4.8)

Maturity (in year)	9	10	12	15	20	25	30
Relative errors with table 4.7	-17.70%	-11.50%	-1.74%	6.08%	8.67%	5.18%	1.25%
Relative errors with table 4.8	-31.20%	-26.10%	-17.70%	-10.20%	-6.10%	-7.32%	-9.29%

Remark 4.6. *There exist other methods to calibrate the Libor Market Model. For example, Schatz (2011) implements a calibration to the CMS spread option market because it is known to be a liquidly traded product which heavily depends on correlation. Indeed, as explained by Schatz (2011) (p.49), when using swaptions to calibrate the LMM, we face the problem that different volatility parametrizations imply decorrelation independent of the instantaneous correlation. However, swaptions are influenced by terminal correlation but not directly by instantaneous correlation, which can make difficult to derive correlation information from swaption quotes.*

4.2.2 LMM Calibrations used in practice

In the previous section, we have obtained calibrations for the LMM with the only concern of being as much market-consistent as possible. However, these “theoretical calibrations”imply high instantaneous volatilities for the LIBOR forward rates (see figure 4.8), which make them unusable in practice with an ALM model. Indeed, recall that GALEA & Associés's ALM model projects cash flows associated with retirement agreements. Thus, the projection horizon usually retained is 80 years, in order to perform an estimation of the insurer's liabilities consistent with the long duration of the cash flows involved. Otherwise, if we retained a horizon of 20 years for example, the technical provisions at the end of the projection would be still very high, meaning that the insurer would still have substantial liabilities towards its policyholders present in the portfolio at the evaluation date, and that the evaluation would not estimate with enough accuracy the cash flows involved in these long-term liabilities.

If we used calibrated parameters that imply high volatility for the forward rates and thus for all variables involved in the ESG (especially zero-coupon prices), the results of the ALM model (the resulting Best Estimate or asset values for example) would be aberrant and very unstable. Therefore, in order to obtain calibrated parameters that could be used in practice, we had to add some constraints to the minimization problems such that the volatility of the resulting forward rates could be decreased.

A major problem is that, as we have seen in section 4.1.3, market volatilities are very high nowadays, which consequently makes the LMM parameters tend to be high too in order to fit the market data. Thus, we have to find the best balance between market consistency and the additional constraints applied to the model parameters, since the more constrained the parameters are in the calibration process, the less market consistent the resultant model is. Tests have been performed to determine the minimum additional constraints that we should apply to the calibration processes. Indeed, we did not want to over-constrain the model because, despite these additional constraints, we still wanted to obtain a model as much market consistent as possible. To determine these minimum constraints, our measure was the relative *convergence gap*[1] of the ALM model, which we wanted to be inferior to 1%.

Furthermore, in order not to over-penalize our study, we decided to consider a projection horizon equal to 30 years, thus much smaller than the one usually retained. Indeed, in this study, we do not aim to compute the rightest Best Estimate but rather analyze the impact of the interest rate model calibration on the ALM model results. This restriction enables us to put less restrictive constraints on the model parameters since obviously the longer the projection horizon is, the more you have to constrain the parameters. We have managed to find calibrated parameters that meet the above criteria, and which will be presented in the following. However, this choice of adding constraints to the parameters represents another weakness of this study because the market consistency of the ALM model may be more questioned now. In addition, it also questions our study of the impact of the derivative used in the calibration process on the generated scenarios and the ALM model results given that the additional constraints give less freedom to the parameters which are as a result less affected by the derivative used to calibrate them. The above problem might represent a global issue for the projection of any long-term insurer liability in the current economic environment.

[1] The convergence gap can be seen as the difference between the projected wealth and the initial one. Formally, it is equal to: VIF + BEL - Assets Market Value.

As we will see later on, we have faced the same problem with the G2++ calibration. Hence, the standard LMM and the G2++ model may not be well-adapted models to capture high market volatilities without generating aberrant and unstable economic scenarios. Nonetheless, as concerns the LMM, the parametrizations retained for the volatility and correlation structure were quite classic ones. If we used more sophisticated or recent specifications, the impact of the practical constraints would maybe be less important because the parametrizations would be more appropriate to the current economic environment or the freedom of the parameters would be less reduced (with a parametrization that would depend on a higher number of parameters for example).

By jointly calibrating the LMM to the caps and swaptions markets with the additional constraints, we have obtained the following parameters:

Table 4.12 – Calibrated LMM parameters using both the caps and swaptions markets and adding practical constraints

a	b	c	d	β	ρ_∞
-7.003e-09	0.331	0.281	2.132e-08	3.987	0.499

When calibrating the LMM to the market swaptions Black volatilities only and considering squared relative errors, the resulting calibrated parameters are:

Table 4.13 – Calibrated LMM parameters using only market swaption Black volatilities and adding practical constraints

a	b	c	d	β	ρ_∞
-0.0349	0.324	0.298	0.0350	4.154	0.5

whereas when calibrating to the market swaptions prices, considering squared relative errors:

Table 4.14 – Calibrated LMM parameters using only market swaption prices and adding practical constraints

a	b	c	d	β	ρ_∞
-0.0350	0.324	0.297	0.0350	4.98	0.5

We note that the obtained calibrations are quite close (especially the parameters b,c and ρ_∞). However, it is now harder to conclude that the different derivatives used in the calibration process lead to the same parameters since it might be partly due to the new "practical"constraints, which give less freedom to the parameters. Moreover, notice the very low, and thus odd values of the parameters a and d calibrated to the caps market, especially if we compare them to the ones derived from the swaptions market. This may illustrate the remark of Brigo & Mercurio (2006) about the "abcd"formulation, who explain that "its flexibility is not sufficient for practical purposes when jointly calibrating to the caps and swaptions markets." (Brigo & Mercurio (2006), p.212).

We have plotted the instantaneous volatility of the forward rates in figure 4.9 obtained with the "abcd"parametrization calibrated in table 4.13. Notice how the volatilities of the forward rates have been decreased, compared to figure 4.8, in order to make this calibration usable in the ALM model.

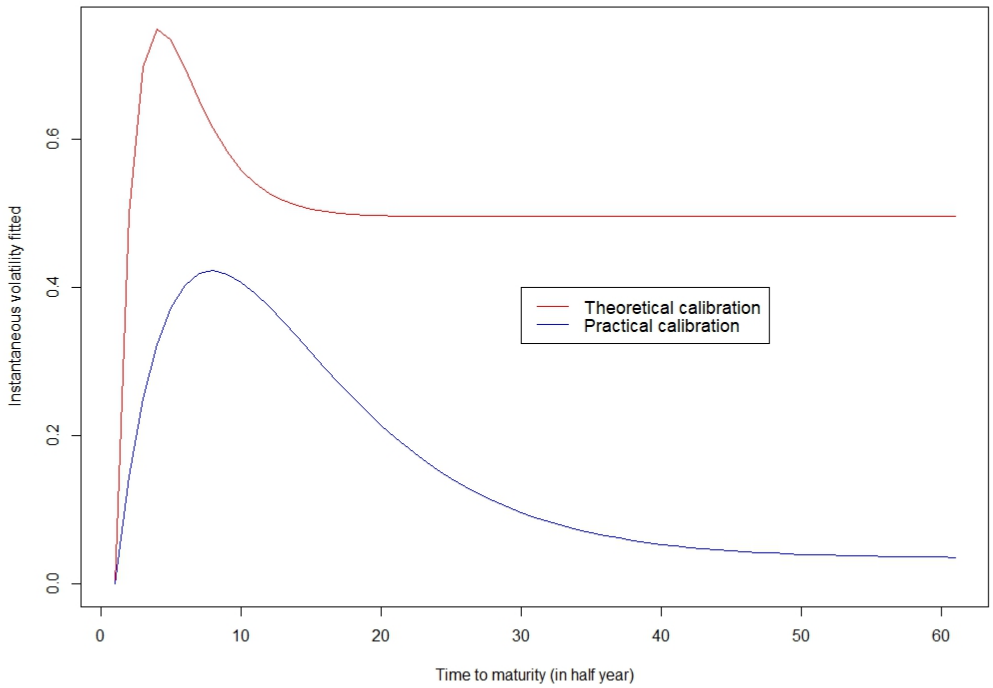


Figure 4.9 – Instantaneous volatility obtained when calibrating the LMM to the swaptions market and adding practical constraints (parameters in table 4.13)

However, this volatility decrease has an undesirable cost on the market-consistency of the resulting model. For example, in table 4.15, we see that the fitting of the LMM to the caps market is far less satisfactory than the one previously obtained in table 4.5. The same problem can be noticed for the fitting of the model to the swaptions market.

Table 4.15 – Relative errors between market caps prices and LMM cap prices obtained with calibrated parameters in table 4.12

Maturity (in year)	9	10	12	15	20	25	30
Relative errors	-45.9%	-42.9%	-38.3%	-35.4%	-35.9%	-39.2%	-42.0%

4.3 G2++ Calibration

4.3.1 Theoretical Calibrations

The G2++ model was already implemented in the ESG of GALEA & Associés before this study. Thus, we will not focus on its implementation (see for example Brigo & Mercurio (2006) p.142) but only on different approaches to calibrate the five model parameters: a, b, σ, η and ρ. Actually, we have implemented three different calibration methodologies for the G2++: calibration to the caps market, to the market ATM swaptions Black volatilities and to the market ATM swaptions prices. We recall the G2++ dynamics:

$$\begin{aligned} r_t &= x_t + y_t + \varphi(t), \quad r(0) = r_0, \\ dx_t &= -ax_t dt + \sigma dW_x(t), \quad x(0) = x_0, \\ dy_t &= -by_t dt + \eta dW_y(t), \quad y(0) = y_0, \end{aligned}$$

where (W_x, W_y) is a two-dimensional Brownian motion with instantaneous correlation ρ such that:

$$dW_x(t)dW_y(t) = \rho dt.$$

Calibration of the G2++ to the caps market

The G2++ model allows to derive an explicit exact formula for cap and floor pricing. However, as already explained, the correlation between forward rates does not afflict caps prices. Therefore, in a calibration process, as explained by Brigo & Mercurio (2006) (p.167), the correlation parameter ρ often happens to be close to 1 or -1 because a simpler Gaussian one-factor model is sufficient to fit the caps market in many situations, even if such a model implies perfectly correlated rates. We have also noticed this phenomena in our calibration as we will see later on.

This leads us to a major problematic for interest rate model calibration. On the one side, the presence of closed-form formulas for pricing derivatives is a desirable feature for an interest rate model (especially in term of computational optimization) but on the other side, one has to pay attention to the theoretical properties (for example which financial phenomena can a derivative allow to reproduce with the model) and possible limits of the considered derivatives in order not to obtain aberrant results. In the case of the G2++, one should therefore prefer the swaptions market to calibrate the model without obtaining a trivial value for ρ, given that swaptions prices depend on the forward rates correlation.

Still, we have implemented a calibration of the G2++ to the caps market. Consider $\mathcal{T} = \{T_0, T_1, ..., T_n\}$ the set of a cap payment dates with T_0 the first reset date, $\tau = \{\tau_1, ..., \tau_n\}$ the set of the corresponding year fractions. In the G2++ model, the no-arbitrage value of a caplet which resets at time T_{i-1}, pays at time T_i, with nominal N and strike K is given by the following closed-form formula:

$$\mathbf{Cpl}\left(t,T_{i-1},T_i,N,K\right) = -N\left(1+K\tau_i\right)P\left(t,T_i\right)\Phi\left(\frac{\ln\frac{NP(t,T_{i-1})}{N(1+K\tau_i)P(t,T_i)}}{\Sigma\left(t,T_{i-1},T_i\right)} - \frac{1}{2}\Sigma\left(t,T_{i-1},T_i\right)\right) + P\left(t,T_{i-1}\right)N\Phi\left(\frac{\ln\frac{NP(t,T_{i-1})}{N(1+K\tau_i)P(t,T_i)}}{\Sigma\left(t,T_{i-1},T_i\right)} + \frac{1}{2}\Sigma\left(t,T_{i-1},T_i\right)\right),$$

where:

$$\Sigma\left(t,T,S\right) = \frac{\sigma^2}{2a^3}\left[1-e^{-a(S-T)}\right]^2\left[1-e^{-2a(T-t)}\right] + \frac{\eta^2}{2b^3}\left[1-e^{-b(S-T)}\right]^2\left[1-e^{-2b(T-t)}\right] + 2\rho\frac{\sigma\eta}{ab\left(a+b\right)}\left[1-e^{-a(S-T)}\right]\left[1-e^{-b(S-T)}\right]\left[1-e^{-(a+b)(T-t)}\right].$$

Consequently, since a cap price is the sum of the underlying caplets prices, the cap price is:

$$\mathbf{Cap}\left(t,\mathcal{T},\tau,N,K\right) = \sum_{i=1}^{n}\left[-N\left(1+K\tau_i\right)P\left(t,T_i\right)\Phi\left(\frac{\ln\frac{P(t,T_{i-1})}{(1+K\tau_i)P(t,T_i)}}{\Sigma\left(t,T_{i-1},T_i\right)} - \frac{1}{2}\Sigma\left(t,T_{i-1},T_i\right)\right) + P\left(t,T_{i-1}\right)N\Phi\left(\frac{\ln\frac{P(t,T_{i-1})}{(1+K\tau_i)P(t,T_i)}}{\Sigma\left(t,T_{i-1},T_i\right)} + \frac{1}{2}\Sigma\left(t,T_{i-1},T_i\right)\right)\right].$$

Using the above formula and Black's formula to compute the caps market prices with the volatilities in table 4.1, we have calibrated the following G2++ parameters. Notice that, as mentioned previously, the value of ρ is close to -1. Numerically, we used a simulated annealing algorithm, followed by a Nelder-Mead algorithm to refine the solutions previously obtained.

Table 4.16 – Calibrated G2++ parameters using using market caps volatilities in table 4.1

a	b	σ	η	ρ
0.0953	0.0060	0.2570	0.1363	-0.9665

Table 4.17 shows that the fitting of the G2++ to the caps market is almost perfect. Especially, it is better than the one obtained with the LMM (see table 4.5), which is noteworthy since the LMM was in particular developed in order to be an interest rate model that could perfectly replicate the caps market.

Table 4.17 – Comparison between market caps prices and G2++ caps prices obtained with the calibrated parameters in table 4.16

Maturity (in year)	9	10	12	15	20	25	30
Market cap price	0.054	0.063	0.081	0.109	0.168	0.249	0.348
G2++ cap price	0.055	0.063	0.081	0.109	0.168	0.248	0.348
Relative errors	0.42%	-0.16%	-0.16%	-0.14%	0.21%	-0.10%	0.02%

Calibration to the swaptions market

There exists an exact formula to perform swaption pricing (see Brigo & Mercurio (2006) p.158) with the G2++ model. However, the formula is complex, difficult to implement and quit time consuming. Thus, one often rather considers closed-form approximations for swaption volatility (for example, see Collin Dufresne & Goldstein (2002) or Singleton & Umantsev (2002)) which are known to be quite accurate. Especially, the approximation proposed by Schrager & Pelsser (2005) allows to derive a closed-form formula for the volatility and the price of a swaption.

Proposition 4. *Consider an ATM swaption with nominal N, strike K and whose underlying swap resets at $T_\alpha, ..., T_{\beta-1}$ and pays at $T_{\alpha+1}, ..., T_\beta$. Schrager & Pelsser (2005) proved that the value of this swaption can be approximated by:*

$$ES\left(0, T_\alpha, T_\beta, N, K = S_{\alpha,\beta}(0)\right) = BPV_{\alpha,\beta}(0)\frac{\sigma_{\alpha,\beta}}{\sqrt{2\pi}}N,$$

where the swaption volatility $\sigma_{\alpha,\beta}$ is given by:

$$\sigma_{\alpha,\beta} = \sqrt{\sigma^2\left(C^{(a)}_{\alpha,\beta}\right)^2\left[\frac{e^{2aT_\alpha}-1}{2a}\right] + \eta^2\left(C^{(b)}_{\alpha,\beta}\right)^2\left[\frac{e^{2bT_\alpha}-1}{2b}\right] + 2\rho\sigma\eta C^{(a)}_{\alpha,\beta}C^{(b)}_{\alpha,\beta}\left[\frac{e^{(a+b)T_\alpha}-1}{a+b}\right]},$$

with:

$$C^{(a)}_{\alpha,\beta} = \frac{1}{a}\left[e^{-aT_\alpha}P^{BPV}\left(0,T_\alpha\right) - e^{-aT_\beta}P^{BPV}\left(0,T_\beta\right) - S_{\alpha,\beta}(0)\sum_{i=\alpha+1}^{\beta}\tau_i e^{-aT_i}P^{BPV}\left(0,T_i\right)\right],$$
$$C^{(b)}_{\alpha,\beta} = \frac{1}{b}\left[e^{-bT_\alpha}P^{BPV}\left(0,T_\alpha\right) - e^{-bT_\beta}P^{BPV}\left(0,T_\beta\right) - S_{\alpha,\beta}(0)\sum_{i=\alpha+1}^{\beta}\tau_i e^{-bT_i}P^{BPV}\left(0,T_i\right)\right],$$

$$P^{BPV}\left(t,T_i\right) = \frac{P\left(t,T_i\right)}{BPV_{\alpha,\beta}\left(t\right)},$$
$$BPV_{\alpha,\beta}\left(t\right) = \sum_{i=\alpha+1}^{\beta}\tau_i P\left(t,T_i\right).$$

Therefore, the price or volatility of a swaption can be approximately expressed as a function of the five parameters a, b, σ, η, ρ. From there, two methods can be used to calibrate the G2++:

- Calibrate the model to the market swaptions Black volatilities by first computing the model swaption prices with Schrager and Pelsser' approximation and then, derive the implied swaption Black volatilities by inverting the Black-like formula for swaption pricing. Indeed, it is important to note that the above formula for swaption volatility does not give the implied swaption Black volatility.
- Directly use the above approximation for swaption pricing to calibrate the G2++ to the market swaptions prices, derived from the quoted swaptions Black volatilities and the Black-like formula for swaption pricing.

Using the first method (i.e. calibrating the G2++ model to the market swaptions Black volatilities in table 4.1), we have obtained the following G2++ parameters. Numerically, we have used the same optimization method as for the calibration to the caps market.

Table 4.18 – Calibrated G2++ parameters using using market swaptions Black volatilities in table 4.1

a	b	σ	η	ρ
0.0116	0.00083	0.108	0.102	-0.998

As we have already seen with the LMM calibration, calibrated parameters tend to imply high volatility (see the low value of b and the high values of σ and η) due to the high values of market implied volatilities. Therefore, we have decided to use the squared relative errors in the G2++ calibration in order to reduce the weight of the highest market values.
The matrix of relative errors is presented below in figure 4.10. Notice that with this calibration, the G2++ generally fits rather well the market volatilities of swaptions, except for long swaption maturity and long tenor. Compared with the caps calibration, it is more difficult to tell which model between the LMM and the G2++ can best fit the market swaptions volatilities. Nevertheless, it seems that once again, the G2++ model can better fit the market ATM swaptions Black volatilities than the standard LMM.

Mat/Tenor	1	2	3	4	5	7	10	12	15	20	25	30
1					26,2%	6,4%	-4,8%	-6,7%	-7,6%	-6,5%	-5,5%	-5,8%
2				15,8%	5,8%	-2,3%	-8,5%	-8,7%	-8,1%	-4,9%	-3,2%	-3,1%
3			1,1%	0,0%	-2,0%	-6,7%	-10,8%	-8,1%	-6,4%	-2,3%	0,0%	0,2%
4	-15,4%	-4,3%	-2,5%	-3,5%	-4,7%	-7,2%	-9,6%	-7,2%	-3,2%	0,9%	3,8%	4,2%
5	-8,1%	-4,1%	-3,4%	-4,1%	-4,6%	-6,0%	-6,6%	-4,9%	1,1%	4,6%	8,0%	8,3%
6	-1,6%	-1,3%	-1,6%	-1,9%	-2,3%	-4,5%	-5,3%	-2,3%	3,7%	8,1%	11,4%	11,5%
7	2,1%	0,7%	0,3%	0,2%	-0,2%	-2,0%	-3,1%	1,2%	6,8%	11,2%	14,6%	14,6%
8	5,2%	2,9%	2,6%	2,4%	0,7%	1,4%	-0,7%	4,2%	8,5%	13,0%	15,7%	15,8%
9	7,7%	5,0%	4,5%	2,3%	2,4%	1,5%	1,7%	5,4%	10,1%	14,4%	16,9%	16,8%
10	10,8%	7,4%	3,8%	4,0%	5,9%	2,1%	4,4%	6,4%	11,7%	16,1%	18,5%	18,1%
12	6,0%	6,1%	8,6%	4,6%	3,4%	3,4%	3,2%	6,3%	11,6%	17,1%	19,7%	19,2%
15	-2,3%	-3,0%	-1,1%	0,9%	2,9%	-1,2%	-0,9%	2,5%	9,4%	17,4%	20,9%	20,4%
20	-15,1%	-17,1%	-16,4%	-15,1%	-13,9%	-13,9%	-9,8%	-4,7%	4,4%	16,0%	18,1%	16,7%
25	-17,9%	-23,4%	-22,0%	-20,3%	-17,8%	-17,2%	-11,4%	-5,4%	5,7%	15,9%	13,7%	19,3%

Figure 4.10 – Relative errors between the swaptions volatilities derived from the G2++ calibrated to the market swaption volatilities and the market ATM swaptions Black volatilities

When calibrating the G2++ to the market swaptions prices using Schrager and Pelsser' approximation and the Black-like formula for swaption pricing, the resulting parameters are:

Table 4.19 – Calibrated G2++ parameters using using market swaptions prices

a	b	σ	η	ρ
3.593e-08	2.976e-03	0.3259	0.3310	-0.9998

The matrix of relative errors between market swaptions prices and the ones computed with the G2++ can be found below in figure 4.11. We easily notice that the fitting of the model to the market swaptions prices is very satisfactory.
However, it is noteworthy to see that we do not find the same calibrated parameters whether we use the Black volatilities or the prices of the market ATM swaptions to calibrate the G2++. Still, the two calibrations are not inconsistent between each other because the fitting of the model to the market swaptions volatilities with the parameters derived from the market prices (or inversely) is quite good.

Furthermore, note that whatever derivative we used in the calibration process, the correlation parameter ρ was always very close to -1, which means that the G2++ degenerated to a simpler one-factor model (even if we used swaptions). As explained by Ferranto (2015) (p.77), the shape of the swaption volatility surface may explain the degeneration of the G2++ to a Gaussian one-factor model. Indeed, in addition to the high volatilities across the whole volatility matrix, we notice that if we consider only one row or column, the corresponding section of the volatility surface is rather flat. Thus, the optimization algorithm can not generate enough convexity in the volatility surface.

Mat/Tenor	1	2	3	4	5	7	10	12	15	20	25	30
1					14,5%	11,6%	0,8%	-2,8%	-6,6%	-8,3%	-7,6%	-6,6%
2				12,6%	11,1%	5,6%	-2,8%	-5,1%	-7,3%	-6,8%	-5,1%	-3,5%
3			8,2%	9,1%	7,6%	1,9%	-5,5%	-5,2%	-6,2%	-4,4%	-2,0%	-0,1%
4	2,7%	6,6%	8,2%	7,0%	5,1%	0,5%	-5,5%	-5,4%	-4,0%	-1,8%	1,6%	3,8%
5	4,3%	7,0%	6,9%	5,3%	3,7%	-0,1%	-4,2%	-4,5%	-1,1%	1,1%	5,2%	7,5%
6	8,2%	7,8%	6,5%	5,1%	3,5%	-0,9%	-4,7%	-3,6%	0,2%	3,7%	8,0%	10,3%
7	9,2%	7,1%	5,7%	4,4%	3,0%	-0,9%	-4,5%	-2,0%	1,9%	6,0%	10,6%	12,9%
8	9,2%	6,2%	4,9%	3,7%	1,2%	-0,1%	-4,0%	-0,6%	2,5%	7,1%	11,4%	13,9%
9	8,5%	5,3%	3,9%	1,1%	0,3%	-1,9%	-3,2%	-0,6%	3,2%	8,0%	12,3%	14,8%
10	8,3%	4,7%	0,9%	0,3%	1,2%	-3,0%	-2,0%	-0,5%	4,1%	9,3%	13,6%	15,8%
12	0,1%	-0,4%	1,1%	-2,3%	-3,5%	-3,8%	-3,8%	-1,1%	3,9%	10,3%	14,8%	16,8%
15	-8,9%	-9,3%	-8,0%	-6,5%	-4,9%	-7,2%	-5,9%	-2,7%	3,4%	11,4%	16,2%	18,1%
20	-14,1%	-14,7%	-13,8%	-12,6%	-11,5%	-10,4%	-6,7%	-2,9%	3,5%	12,3%	15,2%	16,0%
25	-9,0%	-10,4%	-9,3%	-8,1%	-6,7%	-5,3%	-1,7%	1,5%	7,1%	12,6%	12,6%	15,6%

Figure 4.11 – Relative errors between the swaptions prices derived from the G2++ with parameters in table 4.19 and the market ATM swaptions prices

4.3.2 G2++ Calibrations used in practice

When calibrating the G2++ model, we have faced the same problem as with the LMM calibration, that is to say the calibrations previously given could not be used in the ESG of the ALM model because they implied too much volatility and aberrant results. Indeed, whatever derivative we used, the calibrated volatility parameters σ and η were quite high (and parameters a and b quite low), especially if we compare them to the same calibrated parameters obtained by Brigo & Mercurio (2006) on the 13/02/2001 (p.168) or by Ferranto (2015) on the 18/08/2014 (p.76) (which were close to 1%). This increase in model parameters can be once more explained by the increase in the market caps and swaptions Black volatilities.

Once again, we had to add some additional constraints in the calibration processes in order to obtain calibrations that could be used in practice. To determine theses constraints, we used the same method as for the LMM calibration, that is to say constrain the optimization problems such that the relative convergence gap of the ALM model was inferior to 1%. However, the impact of those additional constraints on the market-consistency of the G2++ model was worse than for the LMM calibration. As a matter of fact, we had to put some really restrictive constraints in the minimization problems which limit the freedom of the G2++ parameters to calibrate to the given market data. Especially, the G2++ "practical"constraints were more restrictive than the ones used for the LMM calibration, which might make the LMM a more adapted model when performing a projection with a very long horizon from the market-consistency point of view, despite its clear disadvantage of not being able to generate negative rates (at least if we consider the basic formulation), particularly in the current economic environment.

The resulting calibrated parameters of the three different calibration processes with the additional "practical"constraints are given in the following tables. Note that the calibrated volatility parameters σ and η are now quite low and that the correlation parameter ρ stays close to 1.

Table 4.20 – Calibrated G2++ parameters using market caps volatilities and practical constraints

a	b	σ	η	ρ
0.4567	0.3294	0.008203	0.007246	0.9552

Table 4.21 – Calibrated G2++ parameters using market swaptions volatilities and practical constraints

a	b	σ	η	ρ
0.5321	0.3140	0.006101	0.007684	0.9594382

Table 4.22 – Calibrated G2++ parameters using market swaptions prices and practical constraints

a	b	σ	η	ρ
0.2910	0.4141	0.008779	0.005652	0.8209

Despite their low values, the calibrated volatility parameters are not aberrant, especially if we compare them to the G2++ parameters calibrated by Brigo & Mercurio (2006) (p.167), whose volatility parameters were also close to 1%. Nonetheless, Brigo and Mercurio derived these parameters 15 years ago, in a completely different economic environment. In other words, our resulting calibrations are not theoretically aberrant but may be inconsistent with the current market reality.

To see how the market-consistency of the model has been weakened, consider the table 4.23 which shows how the fitting of the LMM to the caps market is now far less satisfactory. We could also say the same about the fitting to the swaptions market.

Table 4.23 – Relative errors between market caps prices and G2++ cap prices obtained with calibrated parameters in table 4.20

Maturity (in year)	9	10	12	15	20	25	30
Relative errors	-57.35%	-57.54%	-58.52%	-61.67%	-67.42%	-71.56%	-72.96%

To conclude the calibration chapter, the LMM seems to be potentially more market-consistent than the G2++ model in the current economic environment when used in a practical situation that imposes restrictive constraints on the model parameters. On the contrary, if we remove these additional constraints and only try to fit the market data, the G2++ was more appropriate, which may be due to its ability of generating negative rates which makes it a more suitable model in the current economic environment. Nevertheless, the fact that the G2++ model can better fit market data than the standard LMM is quite surprising since the implemented LMM depends on six parameters and has therefore one more degree of freedom than the G2++ to calibrate to the given market data.

Chapter 5

Impact of the interest rate calibration on the ESG and the ALM model

5.1 Introduction

In the previous chapter, we considered the impact of both the calibration and the interest rate model on the pricing of the derivatives used in the calibration process. In this chapter, we will firstly consider their impact on some quantities generated by the ESG and then, on the results of the ALM model.
Comparing the LMM to the G2++ model allows to compare a very popular market model which can not generate negative rates to a model that is capable of doing it, which gives the G2++ an advantage in the current economic environment where the supervisors can demand insurers' ESG to be able to generate negative rates. Furthermore, using the standard LMM and not one of its extensions enables to directly compare a forward rate approach to a spot rate one, without any additional property possessed by the LMM extensions (shifted dynamics, stochastic volatility).

5.2 Impact of the interest rate model and of its calibration on some ESG quantities

5.2.1 Impact on one-year rates

The first considered quantity of the ESG is the forward rate whose maturity is 1 year, i.e. the quantity $F(t,t,t+1) = R(t,t+1)$ which can be seen as the "one-year rate".

We first plotted its evolution over time (see figure A.5 in appendix A.4) using the G2++ model and its theoretical calibration to the market caps volatilities in table 4.16. Note that since these calibrated parameters imply too much volatility for the G2++ model, the one-year rate $t \to R(t,t+1)$ explodes and gets aberrant values.

In contrast, if we consider the "practical"calibrations of the LMM and G2++, the resulting one-year rates are more satisfactory. However, probably because of the restrictive constraints applied on the parameters, the derivative used in the calibration process seems to have a negligible impact on the evolution over time of the resulting one-year rate, as can be seen in the figure 5.1.
Nevertheless, we notice that the two models imply two different shapes for the one-year rate. Actually, for low maturities, the G2++ curves are always higher than the LMM ones. Still, at specific times, the LMM curves continue to increase while the G2++ curves tend to take a more concave shape. Especially, this assessment is even clearer with the LMM calibration derived from the swaptions market only, because of the higher calibrated parameter d (the long-term volatility of forward rates). This evolution of the one-year rates implied by the LMM might be partly due to one of the weaknesses of the model due to its lognormal distribution: exploding rates. Indeed, as we will see later on, despite an upper bound of 50% put on forward rates, a minority of forward rates generated by the LMM tend to explode for high maturities.

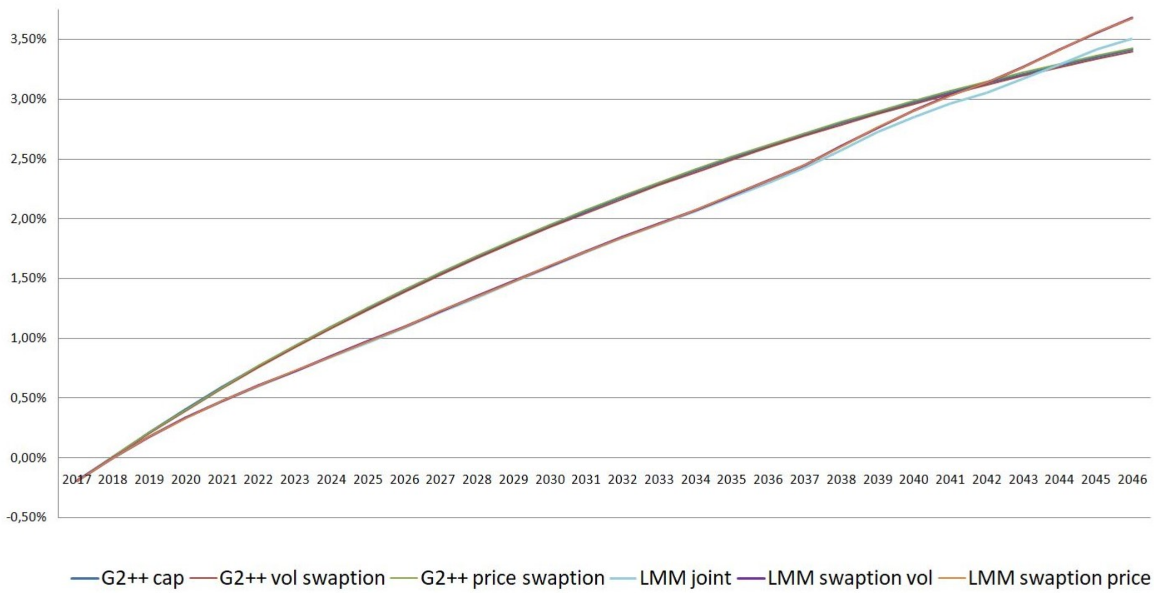


Figure 5.1 – Evolution of the one-year rate $t \to R(t, t+1)$ derived from the different calibrations of the G2++ and LMM

5.2.2 Impact on share returns

Shares market values are projected in the ESG through a Black, Scholes and Merton dynamics with dividend. Under the historical probability, the shares dynamics (S_t) is then given by:

$$dS_t = S_t \left[(\mu_S - q)\, dt + \sigma_S dW_S(t) \right],$$

with q the dividend rate (in this study, $q = 0.5\%$), μ_S the drift rate and σ_S the volatility of the process. Using Girsanov theorem, one can prove that under the risk neutral probability Q, the shares dynamics becomes:

$$dS_t = S_t \left((r_t - q)\, dt + \sigma_S dW_S^Q(t) \right).$$

Given the above dynamics, the only parameter to calibrate is σ_S. The shares model is calibrated using the Black and Scholes' formula for European call and assuming that the dividend rate is null. The model is then calibrated to European call prices to derive the market implied volatility, using Newton-Raphson algorithm. The calibrated implied volatility parameter obtained with the different calibrations of the LMM and G2++ was always close to $\sigma_S = 0.165$.

We have plotted in figure 5.2 the evolution of the mean of share returns obtained with the different calibrations of the G2++ and LMM. In a risk-neutral ESG, share returns should be equal to one-year rates. It is approximately the case here since the share returns curves have similar shapes to the one-year rate curves, presented previously in figure 5.1, in terms of global evolution even if they are far more volatile.

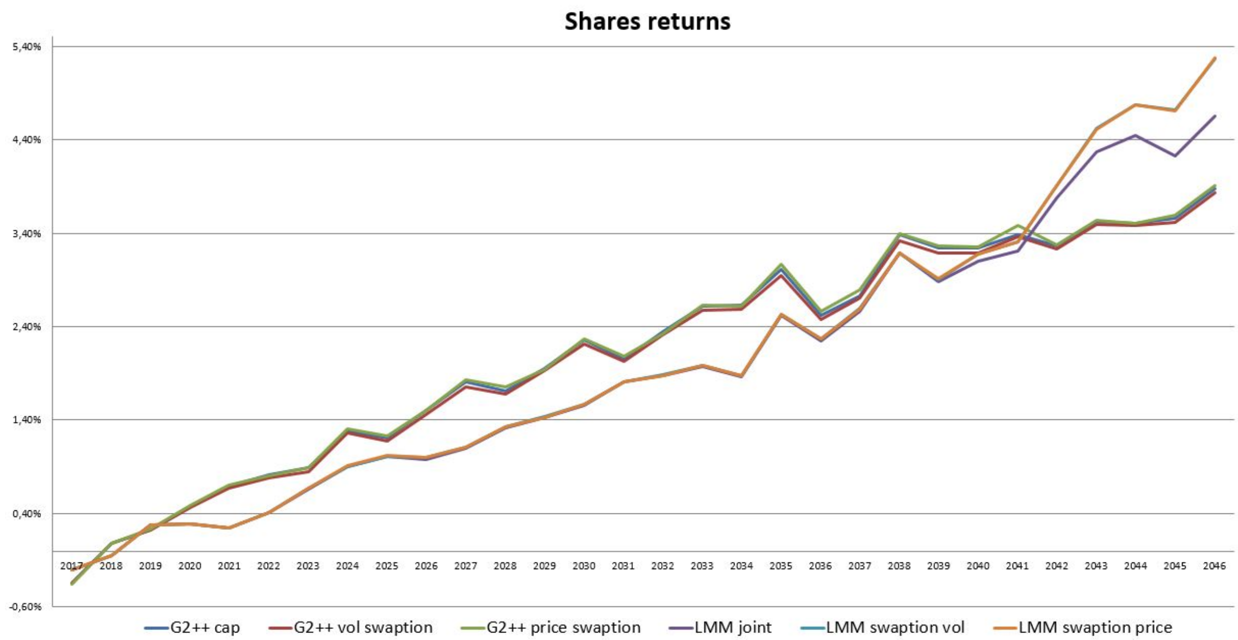
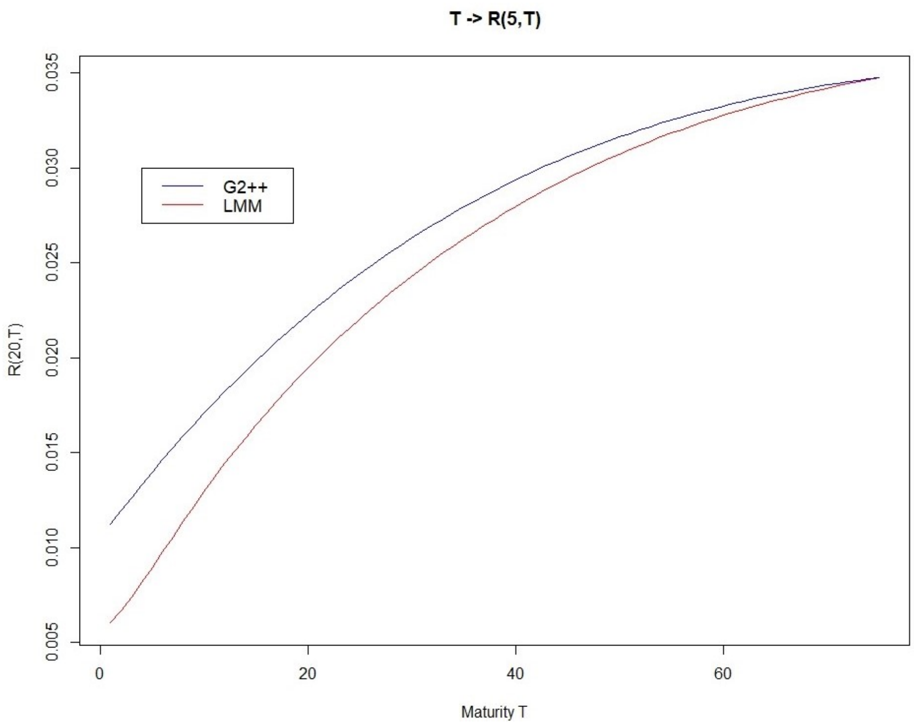


Figure 5.2 – Evolution of the mean of share returns derived from the different calibrations of the G2++ and LMM

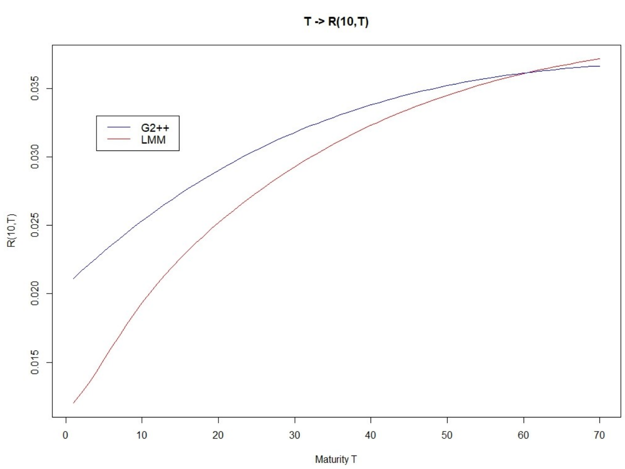
5.2.3 Impact on spot rate curves

We then plotted the spot rate curves in 5, 10 and 20 years to see the impact of the LMM and the G2++ on these curves, using parameters calibrated to the market swaptions Black volatilities. We easily notice that the more distant from now the spot rate curves are, the more different they become. Indeed, if we consider low maturities, the G2++ spot rate curve always takes higher values than the LMM yield curve. Still, in the three cases, at a specific time, the LMM spot rate curve continues to increase while the G2++ curve tends to stabilize itself around a stationary value. When considering the spot rate curve in 20 years, we note that the LMM and G2++ curves have completely different shapes. This absence of stabilization of the LMM spot rate curve might be explained by the exploding forward rates generated by the model. Put differently, the analysis of the spot rate curves is quite similar to the one-year rate one in section 5.2.1.
Furthermore, since we have decided to retain a projection horizon of 30 years for the ALM model, we can expect the Best Estimates derived from the G2++ model to be lower than the ones obtained from the LMM in the ESG because of the higher values of the G2++ spot rate curve for small

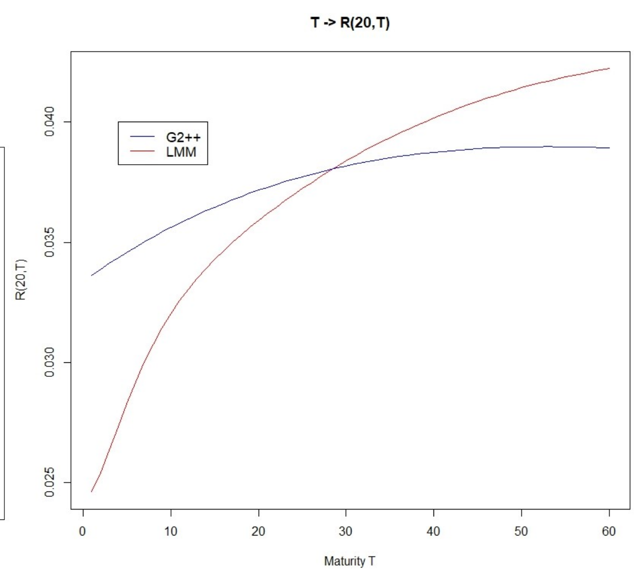
maturities.



(a) $t = 5$



(b) $t = 10$



(c) $t = 20$

Figure 5.3 – Spot rate curve $T \rightarrow R(t, T)$ in t years using the G2++ and LMM parameters calibrated to the market ATM swaptions Black volatilities

5.2.4 Impact on zero-coupon prices

We now consider the distribution of the zero-coupon price of maturity 10 years in 5 years in figure 5.4. Analyzing the densities allows to see better the difference implied by the different models and derivatives used in the calibration processes. As could be guessed from the spot rate curve in 5 years in figure 5.3, the zero-coupon prices derived from the G2++ are lower than the ones derived from the LMM. The derivatives used in the LMM calibration process seems to have a negligible impact, contrary to the G2++ model in which the variance of the zero-coupon price is quite influenced by the choice of the derivative. As a matter of fact, the difference between the zero-coupon price distributions derived from the G2++ could be guessed from the calibrated parameters. Indeed, note that the speeds of reversion parameters are the lowest when calibrating to the market swaptions prices (notice that it was already the case for the "theoretical"calibrations), therefore implying more volatility than the other two G2++ calibrations. Not surprisingly, the variance of the ZC price derived from this calibration is the highest of the three. Also note that with the G2++ model, the probability of the zero-coupon price to be higher than 1 is not null, unlike the LMM, because of the negative rates that can be generated by this model.

In the legend of figure 5.4, "G2pp"stands for G2++, "LMM joint"means that the LMM was jointly calibrated to the caps and swaptions market.

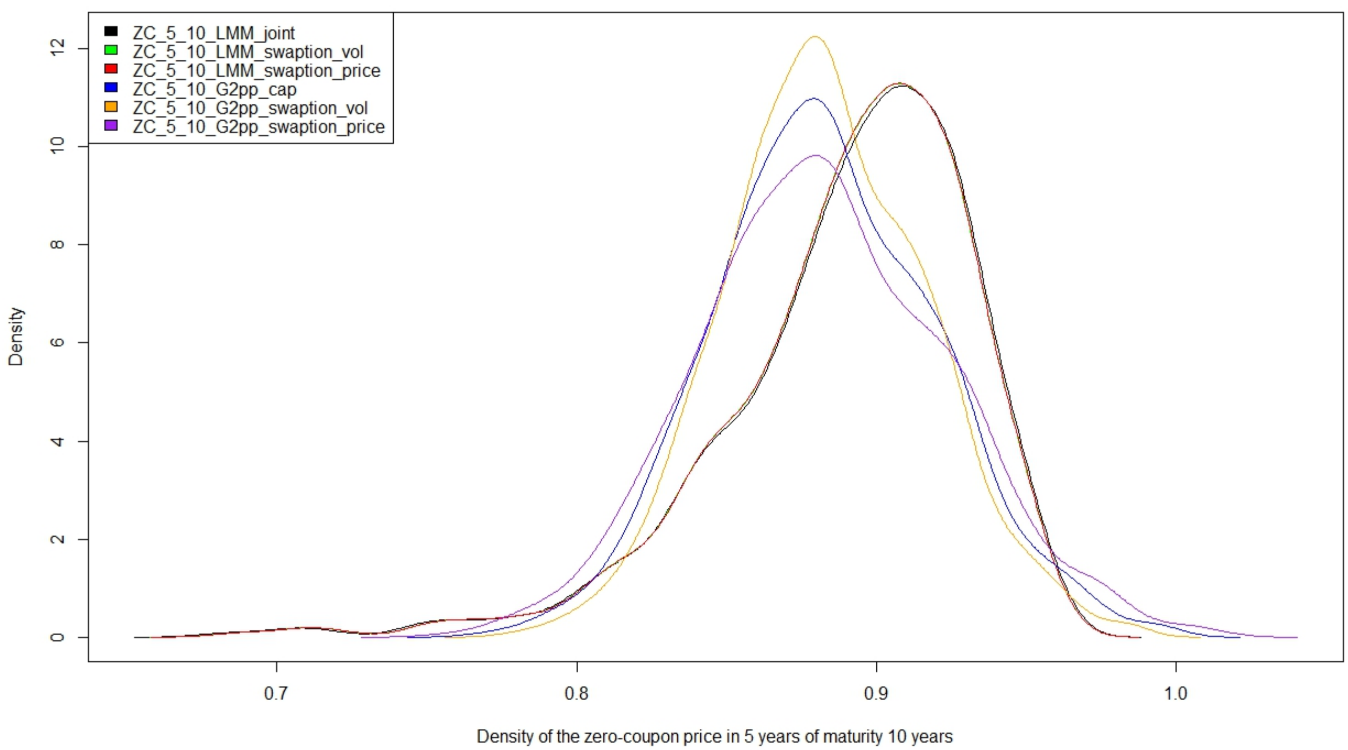


Figure 5.4 – Density of the zero-coupon price of 10 years maturity in 5 years obtained with the different calibrations of the LMM and G2++

5.3 ALM modeling

An ALM model does not aim to perfectly predict the future but rather to derive in the best way possible the value of the Best Estimate of an insurance company, given the retained hypothesis on the assets and liabilities and their interaction. Thus, the Best Estimate may be the best quantity to consider to analyze the impact of both the interest rate model of the ESG and the derivative used in the calibration process. Hypotheses must be retained such that the ALM model is sufficiently robust to the generated cash flows required to compute the Best Estimate. The results of the ALM model depend heavily on the hypotheses and inputs that are used in each generated economic scenario.

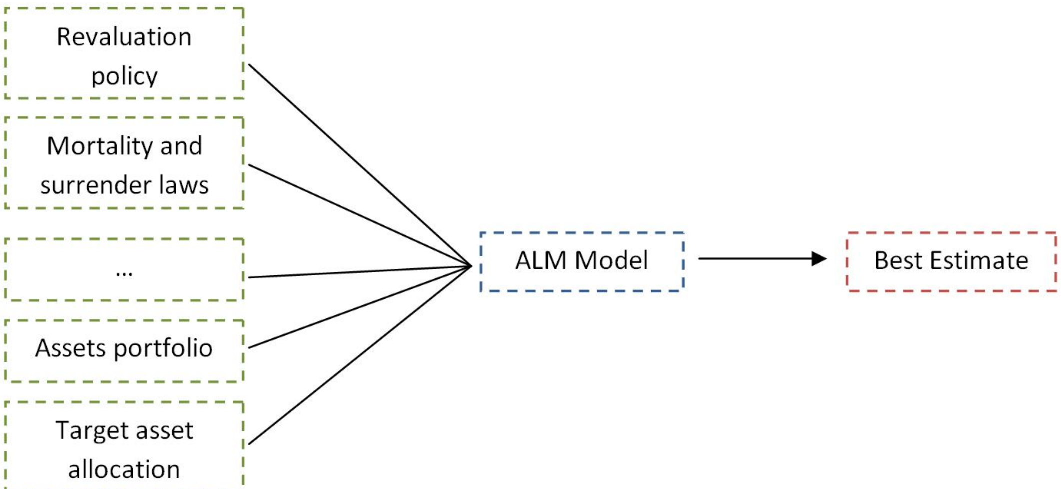


Figure 5.5 – Main inputs and hypotheses of an ALM model

An ALM model should present the following properties:

- The policyholders portfolio should not mostly be alive at the end of the projection. Indeed, the BE relies on the evaluation of all future cash flows associated with the policies on the projection horizon. Therefore, if most of the policyholders are still alive at the end of the projection (which could be detected through a high residual mathematical provision), then cash flows posterior to the projection horizon may not be negligible, but won't be taken into account in the BE since they are not modeled.
- On the contrary, the policyholders portfolio should not go extinct too early (due to a wrong parametrization of the law of mortality or surrender) because we would then estimate numerous null cash flows.
- Mathematical provisions should not take aberrant values caused by an imbalance between asset and liability. Recall that a major goal of an insurer is to find an equilibrium between its assets and liabilities. If the assets values get too high, the mathematical provision may explode and take aberrant and unstable values, especially because of the profit sharing mechanism. This is what we wanted to avoid with the practical calibrations of the G2++ and LMM in chapter 4.
- On the contrary, assets portfolio should not go extinct too fast. This might be caused by too high revaluation rates that would make the insurer's liabilities too high and the assets portfolio run out before the end of the projection.

We will now briefly describe the retained hypotheses of the ALM model (and some of the ESG that were not described previously) used in this study.

5.3.1 Asset

- Bonds market values are easily derived during the projection from the interest rate model. The default risk of bonds is modeled with Longstaff, Mithal and Neis model which allows to generate a yield curve that takes into account the probability of default of bond.
- As explained previously in section 5.2.2, shares market values are projected through a Black, Scholes and Merton dynamics with a dividend rate of $q = 0.5\%$ and an implied volatility parameter σ_S equal to 0.165.
- Property returns are modeled through a Black, Scholes and Merton dynamics with rental. Similarly to the shares market values, the dynamics of property returns under the risk neutral dynamics is given by:
$$dI_t = I_t \left((r_t - l)\, dt + \sigma_I dW_I^Q(t) \right),$$
with l the rental rate (in this study, $l = 0.5\%$) and σ_I the volatility of the process. However, the Property model is not calibrated to market data because of the low number of available data for this kind of asset. The volatility parameter σ_I is thus fixed arbitrary to 0.05, based on expert judgment.
- The instantaneous correlations between asset classes and the interest rates model are fixed arbitrary (i.e. not calibrated to market data), based on expert judgment. These Gaussian correlations are implemented by using Cholesky decomposition to generate correlated Brownian motions, that will drive the dynamics of the assets. The correlation matrix used in this study is given by (the symbol “/”means that it depends on the model and calibration retained):
$$\begin{array}{c} \\ \text{Rates 1st factor} \\ \text{Rates 2nd factor} \\ \text{Shares} \\ \text{Property} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Rates 1st factor} & \text{Rates 2nd factor} & \text{Shares} & \text{Property} \end{array} \\ \begin{pmatrix} 1 & / & 0.2 & 0.4 \\ / & 1 & 0.2 & 0.4 \\ 0.2 & 0.2 & 1 & 0 \\ 0.4 & 0.4 & 0 & 1 \end{pmatrix} \end{array}$$
- At the beginning of the projection, the assets portfolio allocation is 75.9% of bonds, 14.1% of property, 4.1% of shares and 5.9% of cash. The investment or disinvestment are computed each year in order to reevaluate the liability cash flows, given the returns generated by the different asset classes (bond, shares, property) and the target assets allocation.

5.3.2 Liability

- We only project cash flows of annuities and death benefits.
- The mortality of the policyholders is derived from the mortality tables TGH/TGF05.
- We consider that there is no new policyholder, no contribution and no surrender during the projection.
- Charge rates are constant during the projection.
- Each policy is characterized by a technical interest rate which defines the minimum revaluation of the associated cash flows towards the policyholder. In addition to this minimum revaluation, the profit-sharing mechanism is implemented, which also reevaluate the liabilities

depending on the insurer technical and financial results, to a target revaluation rate. In this study, the target revaluation rate is 0.5%.

- There are 54 666 policies in the portfolio at the beginning of the projection. This portfolio is divided into model points, each one associated to one technical interest rate. Since we consider two different technical interest rates in this study (0.25% and 0.5%), there are only two model points.

5.4 Impact on the Best Estimate

To study the potential impact of both the interest rate model and its calibration on the ALM model results, we now consider the value of the Best Estimate since it has become the key quantity to value an insurance policy and can be easily interpreted. 3100 simulations have been performed to compute these Best Estimates, with a projection horizon of 30 years.

Table 5.1 – Resulting Best Estimates using the different calibrations of the G2++ and LMM

Model	Calibration	Best Estimate
G2++	Cap	81 674 203
	Swaption volatility	81 465 599
	Swaption price	81 868 172
LMM	Cap & Swaption	82 110 818
	Swaption volatility	82 578 324
	Swaption price	82 898 463

The resulting Best Estimates in table 5.1 are quite close, though we notice that the BE's obtained with the LMM are always higher than the ones derived with the G2++ model. However, this first observation has to be questioned because of the practical constraints that have been implemented in the calibration processes and which may impose the models to produce similar results. The higher BE's derived with the LMM can be partly explained by the higher volatility implied by the calibrated parameters of the LMM than the G2++ ones, because of the less restrictive constraints put on the LMM parameters. Hence, in our case, the LMM allows to generate more extreme scenarios, which obviously increase the value of the BE, while being more market consistent than the G2++. On the contrary, the G2++ model may lead to underestimate the BE compared to the LMM especially because the calibrated G2++ parameters does not allow to generate enough volatile rates with the restrictive imposed constraints. The difference might also be partly due to convergence error, to a lesser extent because of the high number of simulations.

The volatility of an interest rate model can be seen as a desirable feature in the current market situation where economic scenarios are unstable. Generate too many central scenarios may harm the resulting BE, while a higher variance of the technical provisions enables insurers to anticipate more diverse scenarios. In the current context of high market volatilities, to reduce this risk, one should use an interest rate model that can generate scenarios volatile enough. Thus, we may recommend to use a forward rate model rather than a spot rate one because of the current high market volatilities, but also the increasing complexity of financial derivatives traded on the market.

The derivative used in the calibration process of the LMM seems to have a negligible impact on the BE, as could be expected given that the parameters calibrated from different derivatives

were very close. Nevertheless, we can specify that the highest value of the BE derived from the calibration of the LMM to the market swaptions prices seems reasonable because of the highest value of the long-term volatility of the forward rates (i.e. the parameter d) which should imply more volatility and thus, more extreme scenarios.
As concerns the G2++, the difference between the BE's also seems negligible, because of the close calibrated parameters. Just like for the zero-coupon price in section 5.2.4, the difference between the BE's can be partly guessed from the calibrated parameters, since the BE derived from the G2++ calibration to the market swaptions prices is the highest of the three.

Afterwards, we have considered the distribution of the resultant technical provisions for each calibration in figure 5.6. At first sight, it may be surprising to obtain a higher Best Estimate with the LMM than with the G2++ since the G2++ technical provisions are concentrated around a higher value than the LMM ones. As a matter of fact, the higher BE derived with the LMM is due to some extreme scenarios, which have been truncated to plot the densities, generated by the LMM and which increase the BE value. In other words, the G2++ tail is actually finer than the LMM one, as can be seen on the graph.

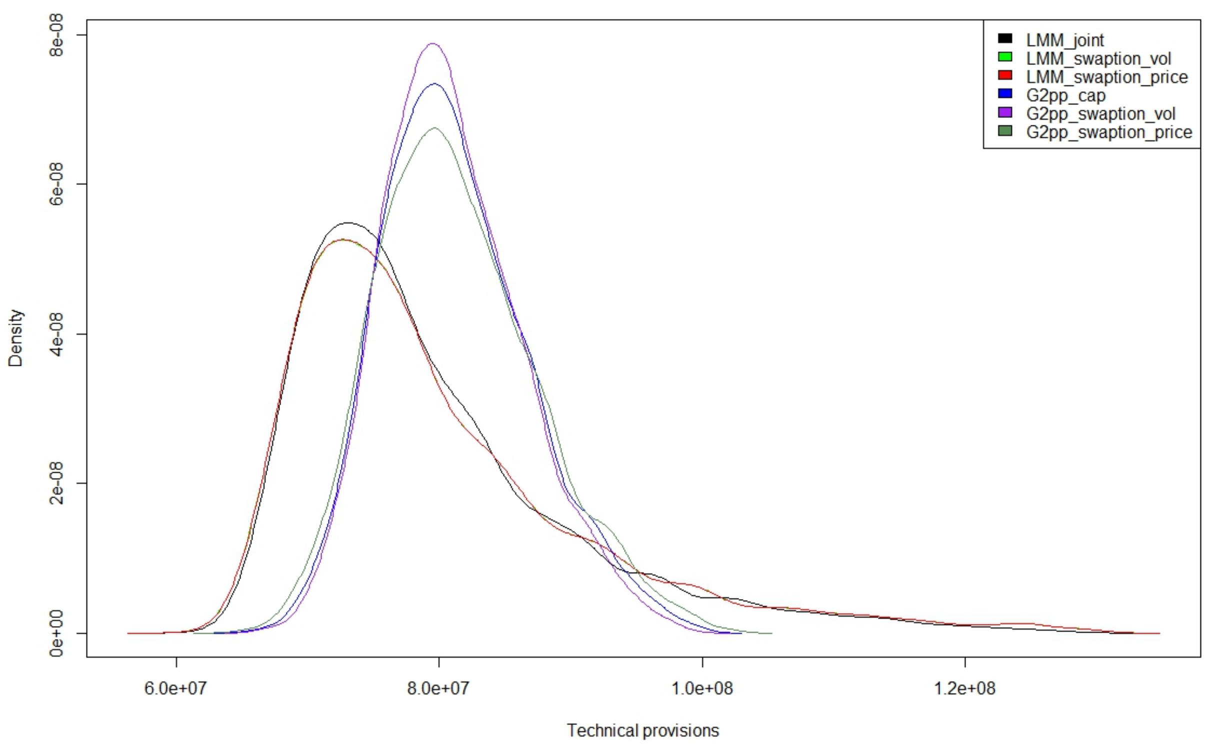


Figure 5.6 – Technical provisions densities derived from the different calibrations of the G2++ and LMM

We also tried to see if the evolution over time of the technical provisions was impacted by the interest rate model or the calibration used in the ESG. However, this comparison did not give any valuable information since the resulting curves were very similar and their difference negligible.

The presence of extreme scenarios in the LMM projections leads us to briefly consider the problematic of the cap put on the forward rates generated by the LMM. Indeed, the 0.5 cap already applied previously did not prevent the occurrence of extreme scenarios. To analyze the impact on the resulting BE's, we have plotted in figure 5.7 the evolution, as a function of the number of simulation, of the BE resulting from the LMM calibration to the market swaptions volatilities for three different caps put on forward rates (0.5, 0.3 and 0.1). Obviously, the lower the cap is, the lower the BE becomes. Moreover, we can notice that only one extreme scenario can have a noticeable impact on the final BE (see the jump provoked by the simulation around 2100). Still, setting a cap on the generated forward rates leads to truncate some economic scenarios, which weakens the market consistency of the projection. Hence, we have decided to keep this cap of 0.5, given that we already put some restrictive constraints on the LMM parameters in the calibration processes.

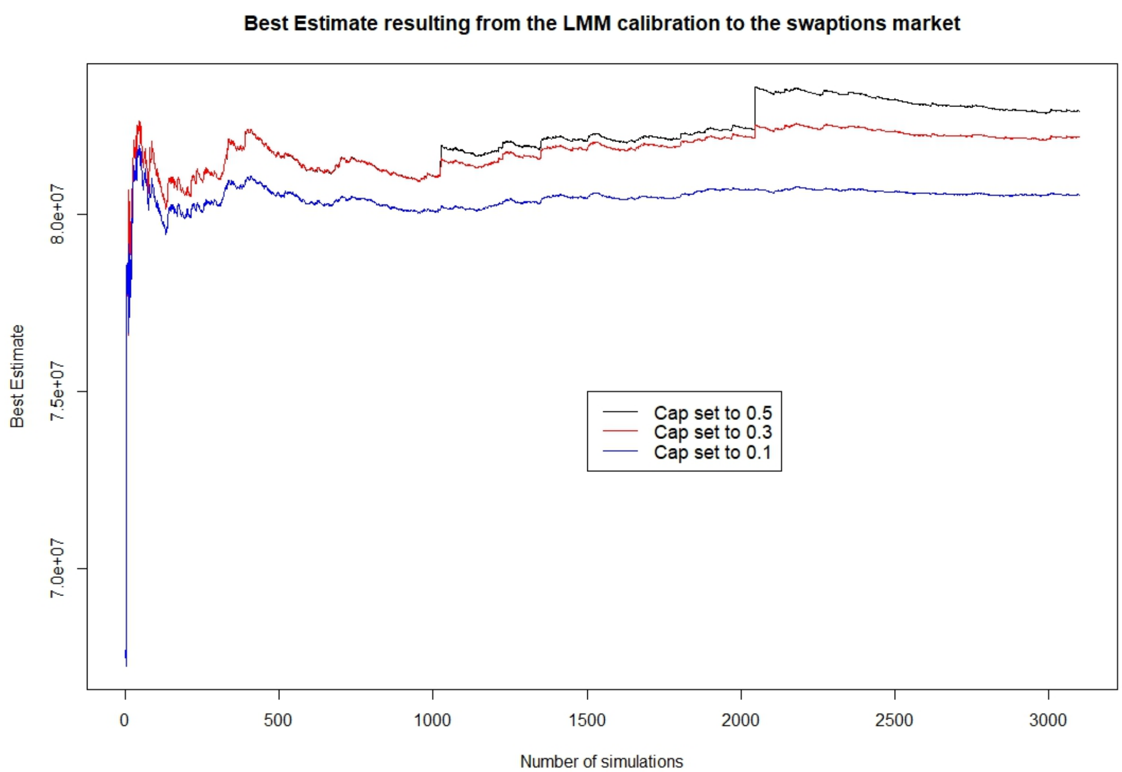


Figure 5.7 – Evolution of the BE resulting from the LMM calibration to the swaptions market, for three different caps put on forward rates

Chapter 6

Calibration to older market data

6.1 Introduction

We have seen previously in chapter 4 that recent market data weaknesses our study because of the high market implied volatility. Therefore, it could be interesting to calibrate our models to older market data in order to once again study the impact of the calibration on the ESG, hoping that in this case, we would not have to impose too much restrictive constraints on the model parameters, as we had to do with market data quoted on December 31, 2016.

We now consider a calibration to an imaginary market environment, based on the EIOPA zero-coupon curve quoted on December 31, 2013 and market volatilities quoted on February 13, 2001 (taken from Brigo & Mercurio (2006), p.167-168). The ATM EUR caps market volatilities used in this chapter are given in table 6.1 and the ATM EUR swaption volatilities in figure 6.1.

Table 6.1 – ATM cap volatility at 13/02/2001

Maturity	1	2	3	4	5	7	10	15	20
Cap volatility	15.2%	16.2%	16.4%	16.3%	16.05%	15.55%	14.75%	13.5%	12.6%

Mat/Tenor	1	2	3	4	5	6	7	8	9	10
1	16,4%	15,5%	14,3%	13,1%	12,4%	11,9%	11,6%	11,2%	11,0%	10,7%
2	16,0%	15,0%	13,9%	12,9%	12,2%	11,9%	11,6%	11,3%	11,0%	10,8%
3	15,7%	14,5%	13,4%	12,4%	11,9%	11,5%	11,3%	11,0%	10,8%	10,6%
4	14,8%	13,6%	12,6%	11,9%	11,4%	11,2%	10,9%	10,7%	10,5%	10,3%
5	14,0%	12,8%	12,1%	11,4%	11,0%	10,7%	10,5%	10,3%	10,2%	10,0%
7	13,0%	11,9%	11,3%	10,5%	10,1%	9,9%	9,7%	9,6%	9,5%	9,3%
10	11,6%	10,7%	10,0%	9,3%	9,0%	8,9%	8,7%	8,6%	8,5%	8,4%

Figure 6.1 – ATM swaption Black volatilities at 13/02/2001

We have implemented the different calibration methodologies, presented in chapter 4, to this new market data. Still, we met some problems with the G2++ model, in the sense that we still had to impose some restrictive constraints on the parameters. The market consistency of the resulting model was of course more satisfactory (because of the low market volatilities used in this chapter) than the one derived in chapter 4, even though the calibrated parameters were influenced by the applied constraints. On the contrary, the calibration of the standard LMM was much satisfactory, that is to say the parameters directly calibrated (without any additional constraint) could be used without generating aberrant or unstable results. We will thus only consider the LMM in this section.

6.2 LMM calibration to older market data

Firstly, by jointly calibrating the LMM to the caps and swaptions markets, we have obtained the following parameters:

Table 6.2 – LMM parameters jointly calibrated to the caps and swaptions markets

a	b	c	d	β	ρ_∞
0.0666	0.0837	0.403	0.0638	0.0842	-0.552

When calibrating the LMM to the market swaptions Black volatilities (considering squared relative errors), the resulting calibrated parameters are:

Table 6.3 – LMM parameters calibrated to the market swaptions volatilities

a	b	c	d	β	ρ_∞
0.133	0.0011	0.186	0.0519	3.127	0.531

whereas when calibrating to the market swaptions prices (considering squared relative errors):

Table 6.4 – LMM parameters calibrated to the market swaptions prices

a	b	c	d	β	ρ_∞
0.133	0.0012	0.185	0.0517	3.21	0.538

The resulting matrices of relative errors are presented in appendix A.5. Note that in the three cases (except for the fit to the market swaption volatilities when jointly calibrating the model to the caps and swaptions markets), the fit is very good since the model parameters were left free. Note that the two calibrations to the swaptions market lead to the same parameters. However, the caps market imply different volatility parameters than the swaption one, even if the fitting to the caps market using the parameters calibrated to the swaptions market is quite good.

6.3 Impact of the calibration on ESG quantities

We will now consider the same ESG quantities as in chapter 5 with the new calibrated parameters, to study the impact of the derivative used in the LMM calibration process.

6.3.1 Impact on the one-year rate

Firstly, we consider the evolution of the one year rate in figure 6.2. The global evolutions are the same, especially for small maturities. Still, from 2025 to 2039, the one year rates derived from the joint calibration always take higher values than the ones implied by the swaptions market only. However, for higher maturities, the latter continue to increase while the ones derived from both the caps and swaptions markets take a more concave shape.

Note that the curves derived from the market prices or volatilities of swaptions seem identical. On the contrary, unlike chapter 5, we can notice a certain difference between one year rates implied by different derivatives (cap or swaption). This difference will be also present in every ESG quantity considered in the following. This confirms our doubt expressed in chapter 5, in which the different one year rate curves of a same model were almost identical, because of the additional constraints put on the parameters in the calibration processes.

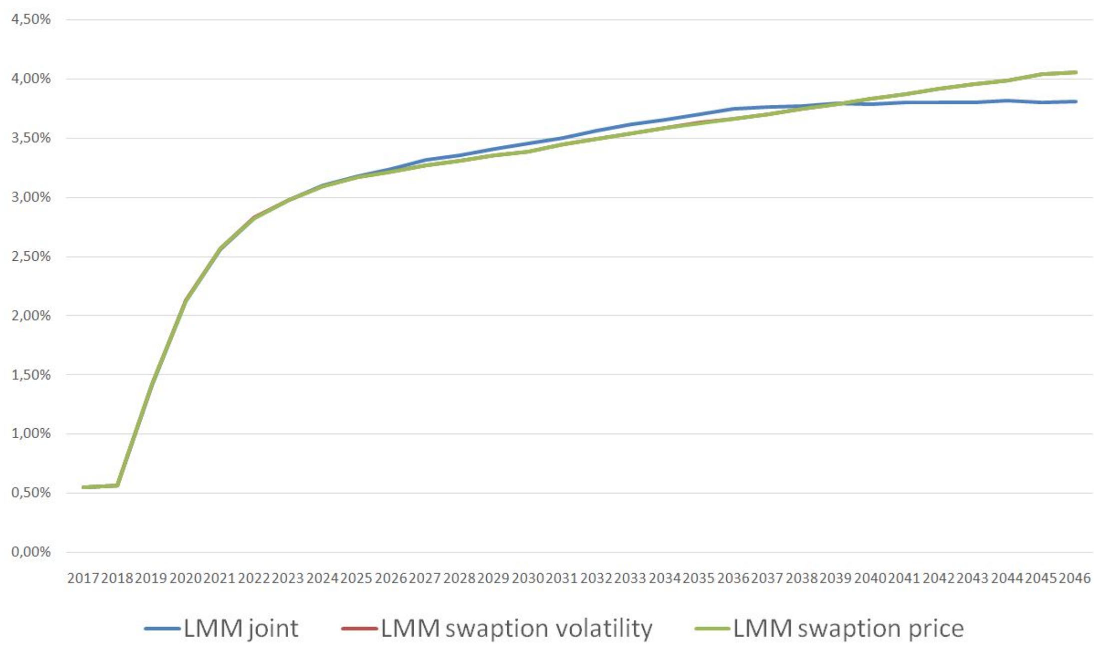


Figure 6.2 – Evolution of the one-year rate $t \to R(t, t+1)$ derived from the LMM calibrations to older market data

6.3.2 Impact on share returns

In figure 6.3, share returns curves take approximately the same shape as one year rates in figure 6.2, even though they do not reproduce the higher values of the one year rates derived from the swaptions market for long maturities, since share returns implied by the joint LMM calibration are always higher.

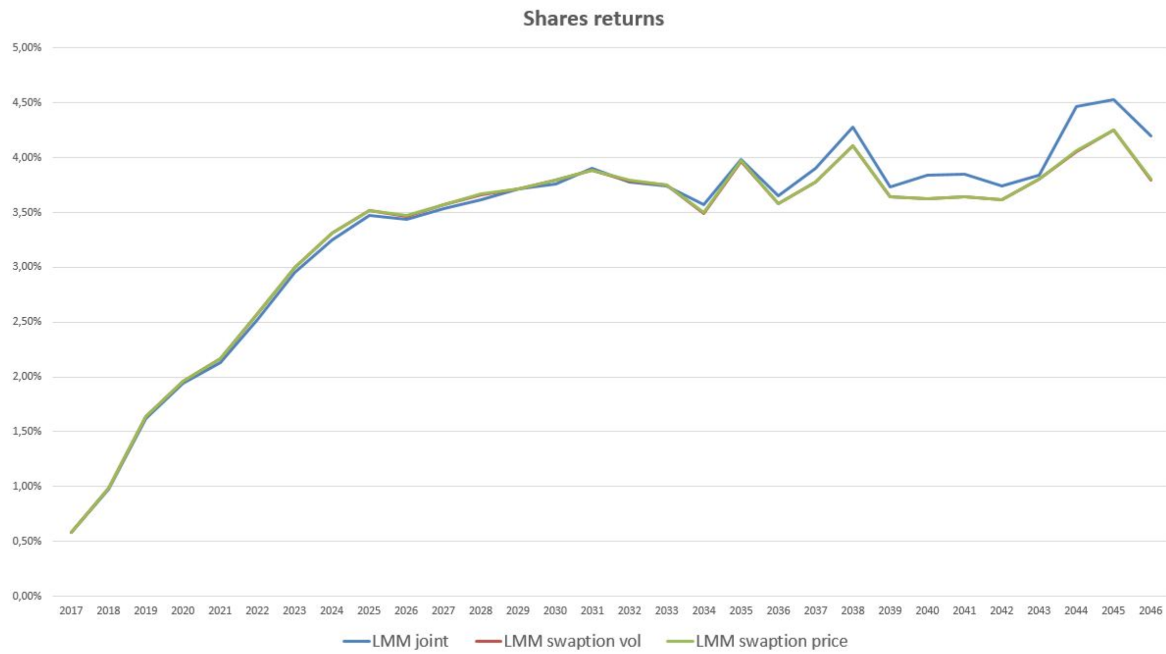
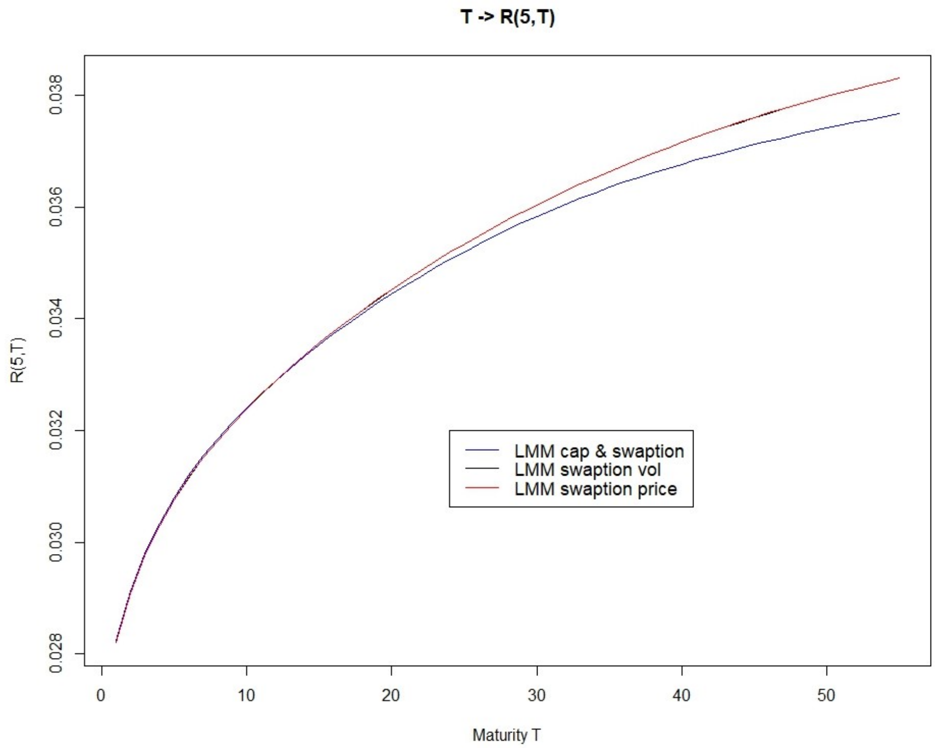


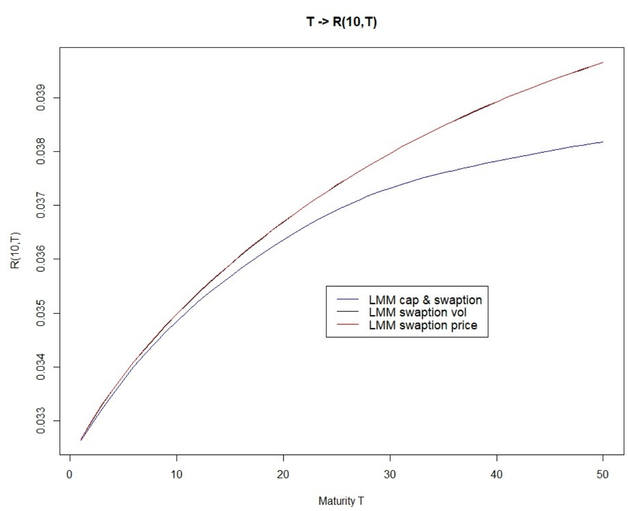
Figure 6.3 – Evolution of the mean of share returns derived from the different LMM calibrations to older market data

6.3.3 Impact on spot rate curves

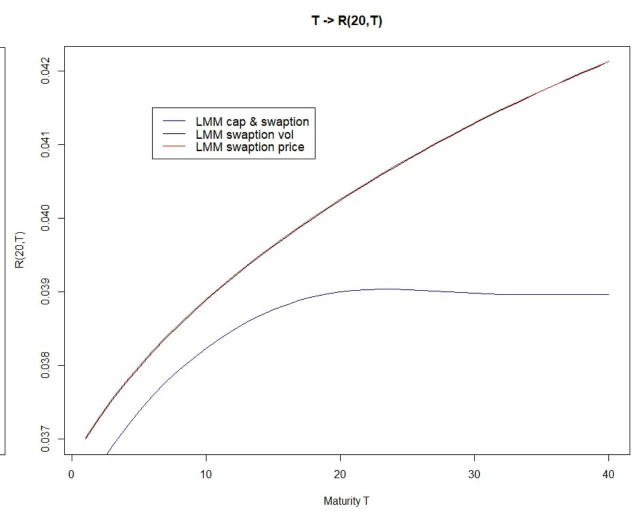
We have then plotted the spot rate curves in 5, 10 and 20 years implied by the different calibrations. Just like in chapter 5, the more distant from now the spot rate curves are, the more different they become. For small maturities, the spot rate curves are almost identical. Still, the spot rate curves implied by the swaptions market always take higher values than the one jointly derived from the caps and swaptions markets, especially for high maturities. Particularly, the spot rate curves in 20 years take completely different shapes. Just like for the other ESG quantities considered, using the market price or volatility of swaptions to calibrate the model does not seem to imply any difference.



(a) $t = 5$



(b) $t = 10$



(c) $t = 20$

Figure 6.4 – Spot rate curve $T \to R(t,T)$ in t years using the different LMM parameters calibrated to older market data

6.3.4 Impact on zero-coupon prices

We have plotted the distributions of the zero-coupon price of maturity 10 years in 5 years in figure 6.5. The ZC prices are concentrated around the same value. We can only notice a higher variance for the ZC prices jointly derived from the caps and swaptions markets, probably due to the higher value of the long-term volatility of the forward rates (i.e. the parameter d).

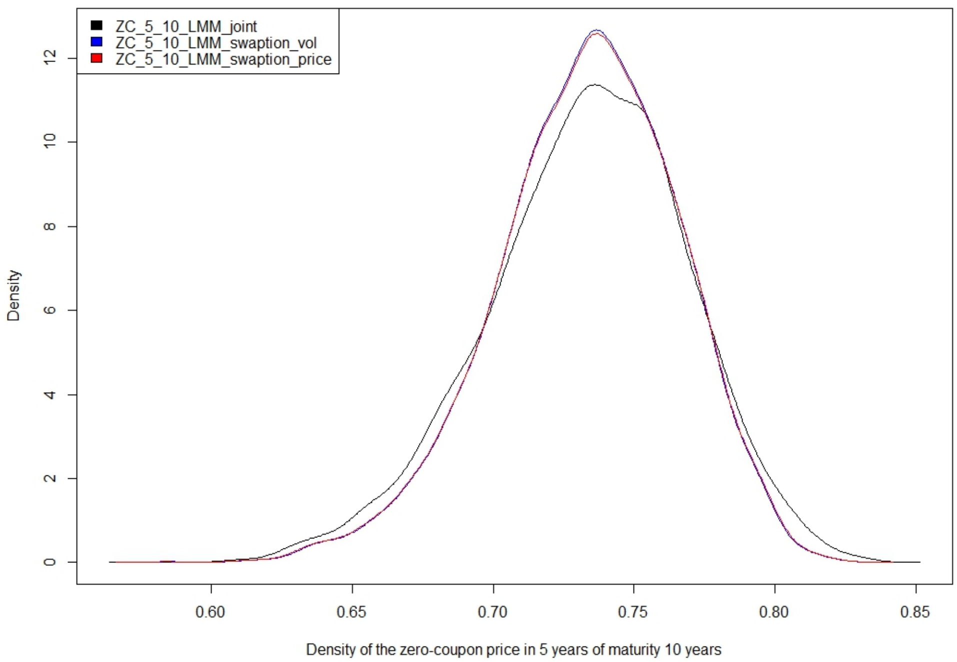


Figure 6.5 – Density of the zero-coupon price of 10 years maturity in 5 years

6.4 Impact on the Best Estimate

The resulting Best Estimates are presented in table 6.5. We do not present the technical provisions densities since they do not give any valuable information. Despite the differences noticed previously between the generated economic scenarios, the resulting BE's are quite close, even though they are derived from calibrations to different derivatives.

Table 6.5 – Resulting Best Estimates using the different calibrations of the LMM to older market data

Model	Calibration	Best Estimate
LMM	Cap & Swaption	78 304 700
	Swaption volatility	78 732 231
	Swaption price	78 758 674

To conclude this first part, we have seen that the Best Estimate was sensitive to the interest rate model and the derivative used in the calibration process, even if the differences observed were not high. Therefore, we recommend to document the choice of the interest rate model and the derivative used in the calibration process because these choices should be subjects of reflection for the actuarial function when developing an ESG.
Still, we must be cautious about the results obtained in this part for several reasons. Firstly, since the Best Estimate evaluation is sensitive to the asset and liability mechanisms implemented, we would maybe obtain different results if we used another ALM tool. Secondly, in the current economic environment characterized by high implied volatilities and very low interest rates, it is difficult to draw a solid conclusion about the impact of the derivative used in the calibration process. As we have seen in chapter 5, by calibrating the interest rate model to recent market data, the derivative used did not seem to have any meaningful impact on the generated scenarios, using either the G2++ or the LMM, because of the restrictive constraints put on the calibrated parameters. Thus, these interest rate models may not be well-adapted to reproduce high market volatilities without generating aberrant and unstable economic scenarios. Yet, we have considered quite basic formulations of these models and if we implemented more sophisticated extensions of the LMM (more accurate parametrization of the volatility and correlation structures, displaced diffusion, stochastic volatility) or the G2++ (constant piecewise volatility, functional instantaneous correlation), we would probably better fit the market. Despite this clear limit of the study, we were able to notice some impacts of the interest rate model on the generated scenarios and the Best Estimate. Nonetheless, if we use older market data to calibrate the interest rate model, even though we have only considered the LMM in chapter 6, we can note that the derivative used in the calibration process has a definite impact on the generated scenarios. Still, when considering the resultant Best Estimates, the impacts were quite low.

Part II

Testing Risk Neutral Economic Scenarios

Introduction

When developing an ESG, it is essential to check the reliability of the generated scenarios by verifying that the theoretical properties of the underlying models and the regulatory requirements for ESG are respected. In the case of a risk-neutral ESG, one especially tests the market consistency of the scenarios and the martingale property of the discounted simulated processes.

To begin with, a risk neutral ESG should generate market consistent economic scenarios, i.e. the prices and implied volatilities derived from the scenarios should be consistent with the ones observed on the market since the models have been calibrated to market data quoted on the evaluation date. An easy way to test the market consistency of a model is to reprice the derivatives used in the calibration process based on Monte Carlo simulations. This time, we don't use the theoretical formulas obtained from the interest rate model, but the universal discounted payoff formulas. Precisely, to price cap, we use formula 4.1 and for swaption pricing, we use formula 4.2. In addition, this allows us to check the calibration process by comparing the prices or implied volatilities computed from the calibration process with the ones from Monte Carlo repricing.

However, as mentioned in the first part, because of the current high market volatilities, we had to apply some additional constraints to the interest rate model calibration processes in order not to generate aberrant results, which can considerably weaken the market consistency of the resultant economic scenarios. This limit of the resulting ESG has been confirmed by the Monte Carlo repricing because the generated scenarios do not manage to reproduce the high market volatilities observed on the market.

Still, we present in figure 6.6 the caps implied volatilities obtained from the G2++ calibrated to the caps market volatilities, and in 6.7 the implied volatilities of swaptions with tenor 5 years derived from the G2++ calibrated to market swaptions volatilities. Clearly, the ESG does not manage to reproduce high market volatilities, especially for long maturities (either for caps or swaptions).

Remark 6.1. *We note quite large differences (about 50%) between the caps implied volatilities derived from Monte Carlo repricing and the ones computed from the G2++ closed-form formula. This can be explained by the presence of the standard normal CDF*[1] Φ *in the G2++ closed form formula for cap pricing. Indeed, when the absolute value of* $\Sigma\left(t, T_{i-1}, T_i\right)$ *takes high values, the* $\Phi(...)$ *term is either 0 or 1, which leads to under-estimate the cap price in the G2++ calibration process. This gives us another reason to consider swaption in the G2++ calibration process rather than cap.*

[1] Cumulative Distribution Function

Maturity	Market volatility	Market price	Model implied volatility	Model price - Monte Carlo	Model price - G2++ Formula
9	152,4%	0,054	125,8%	0,052	0,023
10	115,2%	0,063	91,5%	0,058	0,027
12	85,8%	0,081	66,3%	0,071	0,033
15	70,3%	0,109	53,7%	0,094	0,042
20	63,3%	0,168	43,8%	0,135	0,055
25	64,2%	0,249	35,2%	0,172	0,071
30	67,1%	0,348	26,9%	0,200	0,094

Figure 6.6 – Comparison between the caps prices and implied volatilities derived from the G2++ calibrated to caps market volatilities and Monte Carlo repricing (and the G2++ closed formula for cap pricing), and the ones quoted on the market

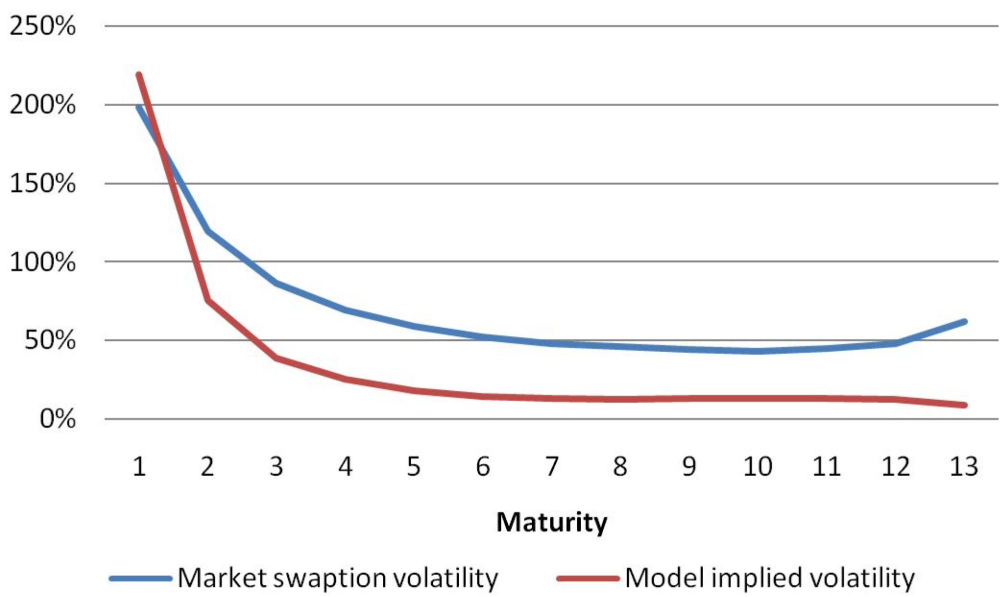


Figure 6.7 – Comparison between the market volatilities and the G2++ implied volatilities (computed with Monte Carlo repricing) for swaptions with tenor 5 years

Secondly, when considering a risk neutral ESG, a widely used approach to test the risk neutrality of the resulting scenarios is the martingale test, which consists in testing that discounted asset prices are martingales under the risk neutral probability. One should be demanding with the martingale test results since we know that discounted asset prices are supposed to be martingales in the risk neutral framework. Therefore, the martingale test does not aim to know if an unknown random process is a martingale but rather check that is indeed the case.

The martingale test usually applied in a risk neutral ESG consists in testing that the mean of discounted asset prices is equal to the time zero asset price (in other words, the mean of the discounted asset prices is constant). The main advantage of this approach lies in its simplicity. Still, in this second part, we want to study how statistical tests for martingale hypothesis can be used to test a risk neutral ESG from the point of view of the actuarial function in the context of a risk-neutral ESG development or audit, and to see what could be their advantages or drawbacks compared to the martingale test commonly applied.

The second part is organized as follows. In a first chapter, we present the mathematical framework of the martingale hypothesis and introduce the main existing statistical tests for martingale hypothesis. Then, in a second chapter, we apply these statistical tests to the risk neutral ESG of GALEA & Associés, in order to study whether they can represent a reliable source of information for the actuarial function about the quality and consistency of the generated scenarios, and in particular to see if they can detect implementation errors in an ESG.

Chapter 7

Statistical tests for Martingale Hypothesis

7.1 Introduction

Martingale testing is a crucial step in many economic and financial studies. It is, amongst others, related to the fundamental theorem of asset pricing which tells that if the market is in equilibrium and arbitrage-free, then discounted asset prices are martingales under the risk neutral probability. The martingale hypothesis implies that the best forecast (in the sense of least mean squared error) of future values of a time series, given the currently available information, is the current value of the time series. This theme of lack of predictability is also fundamental in econometrics (see for example Hall (1978) who explains that changes in consumption between two consecutive periods should be martingales). Thus, testing that a stochastic process is a martingale allows to evaluate a model adequacy and its assumptions that are fundamental in financial and econometric theories.

A voluminous literature has been dedicated to the subject. We will therefore not try to be exhaustive, but rather explain the main martingale properties that can be tested and present the principal classes of tests that have been proposed. We will then apply in the next chapter these statistical martingale tests to some ESG quantities that theoretically form martingales. To our knowledge, we have not found many applications of these statistical tests within a risk neutral ESG, while statistical tests (generally speaking) are widely used when considering real world ESG. However, one can easily find numerous studies that apply martingale hypothesis tests to financial data (especially exchange rates). For example, as concerns testing the martingale difference hypothesis in exchange rates, we can note Belaire Franch & Opong (2005) who applied variance ratio test, Escanciano & Lobato (2009*b*) who used Box-Pierce autocorrelation tests and Kuan & Lee (2004) who applied spectral shape tests. More recently, we note Charles et al. (2011) who compared the power of different martingale hypothesis tests, and Phillips & Jin (2014) who developed new martingale tests, based on Kolmogorov-Smirnov test and the Cramér-von Mises goodness-of-fit test. Nonetheless, the conclusions of these articles are quite mixed and sometimes confusing from each other. See Escanciano & Lobato (2009*b*) (p.4) and Phillips & Jin (2014) (p.548) for more examples of articles which applied statistical martingale tests to exchange rates returns.

7.2 Mathematical formalization

As mentioned previously, the martingale hypothesis is closely linked to the idea of lack of predictability. In the literature, this concept is often referred to as the random walk hypothesis. Actually, one can distinguish three types of random walks (see Campbell et al. (1997) for more details), each one corresponding to a specific dependence structure of time series. The first random walk corresponds to serial uncorrelatedness. In this case, as explained by Kuan (2008), "no linear function of the lagged variables can account for the behavior of the current variable". The random walk 2 is characterized by mean-independent increments, which corresponds to the concept of martingale or alternatively, the martingale difference sequence hypothesis (see 7.3 for a precise definition). In this case, the behavior of the current variable can not be characterized by any function (linear or non-linear) of the lagged variables. The random walk 3 corresponds to independent increments. If a time series is serially independent, there exists no relationship between the past and current variables. When considering financial data, the random walk 3 is not relevant for different reasons, especially the fact that the variance of current asset returns conditional on lagged asset returns is not constant.
These three different dependence structures are closely related to each other since, as we will see later on, serial independence implies the martingale difference sequence hypothesis, which itself implies serial uncorrelatedness.

Formally, consider a sequence of integrable random variables $\{X_t\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\{\mathcal{F}_t\}$ an increasing family of sub-σ-algebras of the filtration $\mathcal{F}$.

Definition 7.1. $\{X_t\}$ *is a* $(\mathcal{F}_t)$-***martingale*** *if:*

1. $\mathbb{E}(|X_t|) < +\infty$ *, i.e* $X_t \in L^1(\Omega)$, $\forall\ t \geq 0$,
2. $\mathbb{E}(X_t|\mathcal{F}_s) = X_s$, $\forall s \leq t$, *almost surely.*

Definition 7.2. $\{X_t\}$ *is said to be a martingale if its relative filtration is its natural one, i.e. the* σ*-field generated by* $\{X_t, X_{t-1}, X_{t-2}, ...\}$.

Proposition 5. *If* $\{X_t\}$ *is a martingale, then* $\mathbb{E}(X_t) = \mathbb{E}(X_0)$, $\forall t$.

When looking at asset prices, the martingale hypothesis means that the best predictor of future asset price is today's. Then, since they are unpredictable, asset returns are said to be a martingale difference sequence. In practice, it is easier to consider asset returns because asset prices are not stationary. Therefore, instead of testing that asset prices are a martingale, one rather tests that asset returns form a martingale difference sequence. Formally, consider the first differences $Y_t = X_t - X_{t-1}$.

Definition 7.3. $\{Y_t\}$ *is a* ***martingale difference sequence (MDS)*** *relative to* $\{\mathcal{F}_t\}$ *if, and only if:*

1. $\{Y_t\}$ *is* $\{\mathcal{F}_t\}$*-measurable,*
2. $\mathbb{E}\left(Y_t \mid \mathcal{F}_{t-1}\right) = 0$, $\forall t$.

Remark 7.1. *If* $\{Y_t\}$ *is a martingale difference sequence, then its cumulative sums form a martingale.*

Remark 7.2. *Since the MDS property only depends on the mean behavior and not on higher moments, a conditionally heteroskedastic sequence can be a MDS.*

The MDS property depends significantly on the associated filtration $\mathcal{F}_t$. Nevertheless, when testing whether $\{Y_t\}$ is a MDS, we are in particular interested in the filtration $\{\sigma(Y_t, Y_{t-1}, ...)\}$. Therefore, in the following, we will only consider this filtration for simplicity, even if this case may be very restrictive.

More generally, most of the tests of the martingale hypothesis focus on testing the Martingale Difference Hypothesis. Compared to a MDS, the formulation is slightly modified as it allows for an unknown and non-zero unconditional mean μ for Y_t. Let $I_t = \{Y_t, Y_{t-1}, ...\}$ be the information set at time t and $\mathcal{F}_t$ the σ-field generated by I_t.

Definition 7.4. *The **Martingale Difference Hypothesis (MDH)** holds for a real-valued stationary time series $\{Y_t\}$ if:*

$$\mathbb{E}\left[Y_t \mid I_{t-1}\right] = \mu, \quad \mu \in \mathbb{R}, \quad (a.s.).$$

The MDH means that past and current information do not help to predict future values of a MDS since the best predictor (in the sense of least mean squared error) of these future values, given the past and current information set, is the unconditional expectation. In other words, $\{Y_t\}$ is not predictable with respect to the filtration $\mathcal{F}_t$. Formally, the conditional expectations $\mathbb{E}(Y_t \mid Y_{t-1}, Y_{t-2}, ...)$ are the same as the unconditional one $\mathbb{E}(Y_t) = \mu$. In the literature, the MDH is also referred to as conditional mean independence.

To test the MDH, most tests are based on the following fundamental equivalence:

$$\boxed{\mathbb{E}\left[Y_t \mid I_{t-1}\right] = \mu, \quad \mu \in \mathbb{R} \quad \Longleftrightarrow \quad \mathbb{E}\left[(Y_t - \mu)\, w\left(I_{t-1}\right)\right] = 0,} \tag{7.1}$$

for any $\mathcal{F}_{t-1}$-measurable function w (such that the moment exists).

Suppose that $\{Y_t\}$ is a MDS and w is a linear function. Then, since $\mathbb{E}\left[(Y_t - \mu)\, w\left(I_{t-1}\right)\right] = 0$, Y_t is serially uncorrelated with $Y_{t-1}, ..., Y_1$. Hence, the MDS hypothesis implies serial uncorrelatedness, but the converse is not true. Secondly, suppose that $\{Y_t\}$ is serially independent with mean zero, then $\mathbb{E}\left[Y_t w\left(I_{t-1}\right)\right] = \mathbb{E}\left[Y_t\right] \mathbb{E}\left[w\left(I_{t-1}\right)\right] = 0$ for any measurable function w, which means that $\{Y_t\}$ is necessarily a MDS. Therefore, serial independence implies the MDS hypothesis, but the converse is not true. See Kuan (2008) (p.10) for counter-examples.

The equivalence 7.1 formalizes the fundamental property of MDS previously given, that is to say the fact that the behavior of Y_t is linearly unpredictable given any function (linear or non-linear) of the past $w\left(I_{t-1}\right)$. From this equivalence, two challenging issues arise. First, the fact that the associated filtration is the past information set I_t means that it includes the infinite past of the considered time series Y_t. Secondly, the equivalence includes an infinite number of functions w.

Given that it is practically impossible to test the equivalence 7.1 against all measurable functions w, MDH tests restrict themselves to specific classes of functions. Actually, we can classify MDH tests according to the class of functions considered in the test. Firstly, the first historical MDH tests considered w to be a linear function, which directly lead to some well-known autocorrelation-based tests. However, these tests were restricted to linear measure of dependence, ignoring possible non-linear dependence between Y_t and I_{t-1}. Therefore, non-linear tests considering non-linear functions w have also been proposed and still represent today a research subject that receives a lot of attention. More recent tests focused for example on avoiding the use of bootstrap procedure to estimate the asymptotic null distribution or tried to directly test the martingale hypothesis, rather than the MDH. See Escanciano & Lobato (2009*b*) for a detailed and clear review of the existing statistical tests for MDH.

7.3 Linear tests for the Martingale Difference Hypothesis

As seen previously, serial uncorrelatedness is a necessary, but not sufficient, condition for the MDH to hold. Ideally, all autocovariances $\gamma_j = Cov\left(Y_t, Y_{t-j}\right)$, or alternatively autocorrelations $\rho_j = \gamma_j/\gamma_0$, should be tested to be null. Still, the most famous MDH tests only test a finite number of autocorrelations to be zero.

7.3.1 Box-Pierce Portmanteau Q test

The Box-Pierce Portmanteau Q test from Box & Pierce (1970) is the most popular serial uncorrelatedness test in the time domain. It focuses on testing the first m autocorrelations but ignores ρ_k for large k, meaning that it only considers the linear dependence up to lag m. For a given number m,

$$H_0: \quad \rho_1 = ... = \rho_m = 0.$$

From sample data $\{Y_t\}_{t=1}^n$, the sample autocovariance $\hat{\gamma}_j$ can consistently estimate γ_j:

$$\hat{\gamma}_j = \frac{1}{n-j} \sum_{t=j+1}^{n} \left(Y_t - \bar{Y}\right)\left(Y_{t-j} - \bar{Y}\right),$$

with $\bar{Y}$ the sample mean. Thus, the j^{th} sample autocorrelation $\hat{\rho}_j = \hat{\gamma}_j/\hat{\gamma}_0$ is the estimator of ρ_j. The Q statistic of the Box-Pierce Portmanteau test is then:

$$Q_m = n\sum_{j=1}^{m} \hat{\rho}_j^2 \quad \xrightarrow{D} \chi^2(m) \quad \text{under the null hypothesis.}$$

In practice, the modified $\tilde{Q}$ test of Ljung & Box (1978), defined by:

$$\tilde{Q}_m = n(n+2)\sum_{j=1}^{m} \frac{\hat{\rho}_j^2}{n-j} \quad \xrightarrow{D} \chi^2(m) \quad \text{under the null hypothesis,}$$

has to be preferred because of its correction term which allows the Ljung and Box test to have a better finite-sample performance, even though the two tests are asymptotically equivalent. Other modifications of the Q test have been proposed in the literature. For example, Lobato et al. (2001) proposed a modified Q-type test more robust to conditional heteroskedasticity because of a test statistic standardized by a consistent estimator of the asymptotic variance. The robustified Portmanteau statistic takes the form:

$$Q_m^* = n\sum_{j=1}^{m} \tilde{\rho}_j^2,$$

where $\tilde{\rho}_j^2 = \hat{\gamma}_j^2/\hat{\tau}_j$, and:

$$\hat{\tau}_j = \frac{1}{n-j}\sum_{t=j+1}^{m}\left(Y_t - \bar{Y}\right)^2\left(Y_{t-j} - \bar{Y}\right)^2.$$

This test does not require serial independence hypothesis for $\{Y_t\}$ to obtain the χ^2 limiting distribution, contrary to Box-Pierce and Ljung-Box tests.
Furthermore, Escanciano & Lobato (2009*a*) proposed an automatic Portmanteau test where the number m of autocorrelations employed is not arbitrarily chosen but takes an optimal value determined by a data-dependent procedure. The proposed test statistic is defined by:

$$AQ = Q^*_{\tilde{m}},$$

where:

$$\begin{aligned}\tilde{m} &= \min\left\{m : 1 \le m \le d; L_m \ge L_h, h = 1, 2, ..., d\right\}, \quad d \quad \text{is a fixed upper bound,}\\ L_m &= Q^*_m - \pi\left(m, n, q\right).\end{aligned}$$

The penalty term $\pi\left(m, n, q\right)$ is a balance between AIC[1] and BIC[2] and takes the form:

$$\pi\left(m, n, q\right) = \begin{cases} m\ln(n) & \text{if} \max\limits_{1\le j\le d}\sqrt{n}\left|\tilde{\rho}_j\right| \le \sqrt{q\ln(n)},\\ 2m & \text{otherwise,}\end{cases}$$

with q some fixed positive number (fixed to 2.4 by Escanciano & Lobato (2009*a*)).

Nonetheless, the Q-type tests may not be well-adapted to test financial time series. Indeed, the limiting χ^2 null distribution is valid if the data $\{Y_t\}$ is either *iid* or has at least a finite $(4+k)^{th}$ moment for some $k > 0$, which is a condition not satisfied by many financial time series.

7.3.2 Variance Ratio test

The Variance Ratio (VR) test, explored in detail by Cochrane (1988), Lo & MacKinlay (1989) and Poterba & Summers (1987), is another well-known test employed to test the MDH, which has undergone many robustified versions. See Charles & Darne (2009) for a detailed review of the existing VR tests. The variance ratio of m-period return is defined as:

$$VR(m) = \frac{Var\left(Y_t + Y_{t-1} + ... + Y_{t-m+1}\right)}{m.Var\left(Y_t\right)} = 1 + 2\sum_{j=1}^{m-1}\left(1 - \frac{j}{m}\right)\rho_j.$$

The VR test is based on the intuition that if returns are uncorrelated over time, then:

$$Var\left(Y_t + Y_{t-1} + ... + Y_{t-m+1}\right) = m.Var\left(Y_t\right), \quad \text{i.e.} \quad VR(m) = 1.$$

The related statistic is:

$$VR_m = 1 + 2\sum_{j=1}^{m-1}\left(1 - \frac{j}{m}\right)\hat{\rho}_j,$$

[1] Akaike information criterion
[2] Bayesian information criterion

which should be close to 1 under the null hypothesis. Under serial independence,

$$\sqrt{n.m}\,(VR_m - 1) \xrightarrow{D} \mathcal{N}(0, 2(m-1)).$$

The main drawback of this test is that it is not consistent against all alternatives, because of the possible compensations between autocorrelations with different signs. Moreover, in order to compute variance of long term returns, the VR test is customary performed using overlapping data[1] as suggested by Lo & MacKinlay (1989) to improve the test power. However, it is then more difficult to estimate the exact distribution of the test statistic. Therefore, the VR tests are often asymptotic tests since the sampling distribution is approximated by its asymptotic distribution instead of its exact one. Nevertheless, the normal distribution may badly approximate the sampling distribution, especially if the sample size does not justify asymptotic approximation, because the sample null distribution can be biased and right-skewed. Several modifications of the VR test have been proposed to circumvent this problem.

Chen & Deo (2006) proposed a power transformation of the VR statistic that provides a better approximation than the normal distribution (if m is not too large) in finite samples and solves the right skewness problem. The distribution of their test statistic is also robust to conditional heteroskedasticity. They proposed a VR statistic based on the periodogram (n is the sample size):

$$VR_{CD}(m) = \frac{1}{n-m}\frac{4\pi}{\hat{\sigma}^2}\sum_{j=1}^{(n-1)/2} W_m(\lambda_j)\, I_{\Delta y}(\lambda_j),$$

with:

$$\begin{aligned}
I_{\Delta y}(\lambda_j) &= (2\pi n)^{-1}\left|d_y(\lambda_j)\right|^2,\\
d_y &= \sum_{t=1}^{n-1}\left[Y_t - Y_{t-1} - \hat{\mu}\right]e^{-i\lambda_j t},\\
W_m(\lambda_j) &= m^{-1}\left[\frac{\sin(m\lambda_j/2)}{\sin(\lambda_j/2)}\right]^2,\\
\lambda_j &= 2\pi j/n, \quad j = 1, ..., n.
\end{aligned}$$

Given that Lo-MacKinlay tests are biased and right skewed, Wright (2000) proposed another approach based on the use of signs and ranks. This led to rank and sign tests that have exact sampling distribution, and therefore do not require asymptotic approximation. These rank and sign tests can also be more powerful than the classic VR test against a wide class of alternatives displaying serial correlation. Wright (2000) first proposed two rank statistics R_1 and R_2:

$$R_i(m) = \phi(m)^{-1/2}\left(\frac{(n.m)^{-1}\sum_{t=m}^{n}\left(r_{i,t} + ... + r_{i,t-m+1}\right)^2}{n^{-1}\sum_{t=m}^{n} r_{i,t}^2} - 1\right), \quad i = 1, 2,$$

with:

$$\begin{aligned}
r_{1,t} &= \frac{r(Y_t) - (n+1)/2}{\sqrt{(n-1)(n+1)/12}},\\
r_{2,t} &= \Phi^{-1}\left(\frac{r(Y)}{n+1}\right),\\
\phi(m) &= \frac{2(2m-1)(m-1)}{3m.n}.
\end{aligned}$$

[1]overlapping data means that two samples can not be considered to be independent.

where Φ^{-1} is the inverse of the standard normal CDF and $r(Y)$ is the rank of Y_t among $\{Y_1, ..., Y_n\}$. The two test statistics based on the signs of first differences are:

$$S_1(m) = \phi(m)^{-1/2} \left(\frac{(n.m)^{-1} \sum_{t=m}^{n} (s_t + ... + s_{t-m+1})^2}{n^{-1} \sum_{t=m}^{n} s_t^2(t)} - 1 \right),$$

$$S_2(m) = \phi(m)^{-1/2} \left(\frac{(n.m)^{-1} \sum_{t=m}^{n} \left(s_t(\bar{Y}) + ... + s_{t-m+1}(\bar{Y}) \right)^2}{n^{-1} \sum_{t=m}^{n} s_t^2(\bar{Y})} - 1 \right),$$

where $s_t = 2u(Y_t, 0)$, $s_t\left(\bar{Y}\right) = 2u\left(Y_t, \bar{Y}\right)$, and $u(Y_t, q) = 0.5 * 1_{\{Y_t > q\}} - 0.5 * 1_{\{Y_t \leq q\}}$.

Furthermore, in practice, the statistic VR_m is commonly tested for several m values. Since we should ideally test all values of m, several procedures have been proposed either to estimate the optimal m value or to perform a joint test based on multiple comparisons of VRs over different time horizons, called multiple VR test in contrast to conventional individual one. In this case, the joint null hypothesis becomes $H_0 : VR(m_i) = 1, \forall i = 1, ..., p$.

Chow & Denning (1993) proposed a joint VR test that considers a set of individual VR tests and rejects the null hypothesis if one of the ratios is significantly different from one. Based on the intuition that the test decision can be obtained from the maximum value of the set of individual VR statistics, Chow & Denning (1993) proposed the following test statistic for iid time series:

$$MV_1 = \sqrt{n} \max_{1 \leq i \leq p} |M_1(m_i)|,$$

where M_1 is the M_1 statistic from Lo & MacKinlay (1989):

$$M_1(m) = \frac{VR(Y, m) - 1}{\sqrt{\phi(m)}},$$

$$VR(Y, m) = \frac{(n.m)^{-1} \sum_{t=m}^{n} \left(Y_t + ... + Y_{t-m+1} - m\bar{Y} \right)^2}{n^{-1} \sum_{t=1}^{n} \left(Y_t - \bar{Y} \right)^2}.$$

For uncorrelated series with possible heteroskedasticity, Chow & Denning (1993) proposed a second test statistic:

$$MV_2 = \sqrt{n} \max_{1 \leq i \leq p} |M_2(m_i)|,$$

where M_2 is the M_2 statistic from Lo & MacKinlay (1989):

$$M_2(m) = \frac{VR(Y, m) - 1}{\sqrt{\phi^*(m)}},$$

$$\phi^*(m) = \sum_{j=1}^{m-1} \left[\frac{2(m-j)}{m} \right]^2 \delta(j),$$

$$\delta(j) = \frac{\sum_{t=j+1}^{n} \left(Y_t - \bar{Y} \right)^2 \left(Y_{t-j} - \bar{Y} \right)^2}{\left\{ \sum_{t=1}^{n} \left(Y_t - \bar{Y} \right)^2 \right\}^2}.$$

Based on a similar approach, Belaire-Franch & Contreras (2004) proposed joint versions of Wright's rank and sign tests, based on statistics that take the maximum value of a set of one of the individual rank and sign tests proposed by Wright (2000). Belaire-Franch & Contreras (2004) proved that the joint tests based on ranks are more powerful than the ones based on signs.

Chen & Deo (2006) proposed another multiple VR test based on the power transformed statistic VR_{CD} previously given. They considered the following Wald statistic which follows a $\chi^2(m)$ distribution:

$$QP(m) = \left(\mathbf{V}_{p,\beta}(m) - \mu_\beta\right)' \Sigma_\beta^{-1} \left(\mathbf{V}_{p,\beta}(m) - \mu_\beta\right),$$

where $\mathbf{V}_{p,\beta}(m) = \left[VR_{CD}^{\beta}(2), ..., VR_{CD}^{\beta}(m)\right]$, μ_β and Σ_β are respectively measures of the expectation and covariance matrix of $\mathbf{V}_{p,\beta}$. Chen & Deo (2006) showed that their joint VR test was more powerful than the individual one.

Rather than considering asymptotic approximation, some VR tests based on bootstrap approach to estimate the sampling distribution have also been introduced by researchers. Kim (2006) applied wild bootstrap to Lo & MacKinlay (1989) and Chow & Denning (1993) tests to approximate the sampling distribution of the corresponding test statistics. Moreover, Kim (2009) extended with a bootstrap procedure the work of Choi (1999), who proposed an automatic variance ratio test where m takes its optimal value based on a fully data-dependent method. Kim (2009) proved that the use of wild bootstrap can considerably improve the properties of Choi's test in small samples.

7.3.3 Spectral based test

The two families of tests presented previously test only a finite number of autocorrelations and therefore, may miss some dependence structure due to omitted lags. To test all autocorrelations, a convenient approach, proposed by Durlauf (1991), consists in considering the frequency domain since it allows to use the spectral density function, which summarizes the information contained in all autocorrelations. We are then able to asymptotically test that all autocorrelations are null:

$$H_0: \quad \rho_1 = \rho_2 = ... = 0.$$

The spectral density $f(\omega)$ is defined by:

$$f(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{+\infty} \gamma_j e^{i\omega j}, \quad \omega \in [-\pi, \pi].$$

The null hypothesis implies that $\gamma_k = 0, \forall k \neq 0$. Thus, under the null hypothesis, the spectral density $f(\omega)$ reduces to the constant $\gamma_0/2\pi, \forall\omega \in [-\pi, \pi]$, or equivalently, the spectral distribution, i.e. the integral of the spectral density, is linear in w. To analyze the spectral shape, the Durlauf test consists in analyzing the convergence of random functions that can estimate the spectral density (or spectral distribution). The standard approach considers the periodogram $I_T(\omega)$, which estimates the spectral density:

$$I_T(\omega) = \frac{1}{2\pi} \sum_{j=-(T-1)}^{T-1} \hat{\gamma}_j e^{i\omega j}.$$

The difference between $I_T(\omega)$ and $\gamma_0/2\pi$, which should be close to 0 under the null hypothesis, could appear to be a natural statistic to consider. Still, the periodogram does not converge pointwise to a constant because of the "inconsistency of the individual frequency estimates"(Durlauf (1991), p.360). Durlauf (1991) then proposed an alternative approach that estimates the deviations of the spectral distribution function from its theoretical shape based on the periodogram normalized by the sample variance:

$$U_T(t) = \frac{\sqrt{2T}}{\pi} \sum_{j=1}^{m(T)} \hat{\rho}_j \frac{\sin(j\pi t)}{j}, \quad t \in [0,1],$$

where $m(T)$ is less than T but grows with T at a slower rate. Under the null hypothesis,

$$U_T(t) \xrightarrow{D} U(t) \quad t \in [0,1],$$

with $U(t)$ the Brownian bridge on $t \in [0,1]$. Based on $U_T(t)$, Durlauf (1991) proposed several test statistics, such as:

1. Cramér-von Mises statistic: $\int_0^1 U_T(t)^2 dt \xrightarrow{D} \int_0^1 U(t)^2 dt$ under the null hypothesis,
2. Kolmogorov-Smirnov statistic: $\sup |U_T(t)| \xrightarrow{D} \sup |U(t)|$ under the null hypothesis.

However, these tests may be over-sized under conditional heteroskedasticity because of thicker right-tails of the limiting distribution and thus, reject too often the null hypothesis. Robustified versions of Durlauf's test have been proposed, based for example on the use of other estimator of the spectral density than the periodogram. Still, tests in the frequency domain often require strong moment assumptions that restrict the dependence structure.

7.4 Non-linear tests for the Martingale Difference Hypothesis

The three categories of linear tests presented previously can be powerful in testing for the necessary, but not sufficient, condition of serial uncorrelatedness, but even their robustified versions are not consistent against non-martingale difference alternatives with zero autocorrelations. Since non-linear dependence can be present in financial data, tests for the MDH have been proposed to test non-linear martingale process.

A first approach, called the smoothing or local approach, consists in directly estimating the conditional expectation $\mathbb{E}\left[Y_t \mid Y_{t-1}, ..., Y_{t-m}\right]$ for some finite m. However, this type of test is known to be very bias when m is large, even for large sample size. See Hart (1997) for a review of local approach with $m = 1$.
Demestrescu (2009) proposed to use panel unit root test to test the MDH in stock prices since testing in panels can increase the power of a test. These tests are robust to conditional heteroskedasticity, can test time series that features infinite kurtosis and do not require strong moment conditions. Nonetheless, panel unit root test leads to consider possible cross-dependence between the units of the panel, and the power of these tests may be reduced against certain properties implied by the MDH, such as autocorrelation, compared to other existing tests. Moreover, the unit root hypothesis is more general than the martingale hypothesis.

7.4.1 Integrated tests

Recall that the equivalence 7.1 should hold true for all measurable function w. Still, it is in practice not necessary to consider a very large class of functions w to conduct a consistent MDH test, because testing the orthogonality condition over a convenient parametric class of functions, satisfying certain properties, is sufficient. A parametric class still contains an infinite number of functions, but allows to considerably reduce the complexity of the problem. In the literature, this method is referred to as the integrated approach or the global one. Unlike linear MDH tests, integrated tests are consistent for testing the MDH and have power against a much larger class of non-martingale difference alternatives (correlated or not). Moreover, they often require weaker moment condition and do not rely on the assumption of conditional homoskedasticity. However, their main weakness is that their limiting null distribution can not be identified given that they are data dependent and thus, critical values can often not be tabulated. Therefore, most of these tests require the use of bootstrap procedure to estimate the asymptotic null distribution.

Integrated tests in finite dimension

If we first consider integrated test in finite dimension, we now have:

$$w\left(I_{t-1}\right) = w_0\left(\tilde{Y}_{t,m}, x\right),$$

where $\tilde{Y}_{t,m} = (Y_{t-1}, ..., Y_{t-m})$ for some finite m and w_0 is a known function indexed by a nuisance parameter x. The global issue of these tests is the specification of the function w_0. Globally, two functions have been proposed by different authors for w_0: the exponential function and the indicator one

Firstly, Kuan & Lee (2004) developed a class of MDH tests based on the following equivalence proved by Bierens (1982):

$$\mathbb{E}\left[Y \mid \sigma\left(\mathbf{X}\right)\right] = 0 \Longleftrightarrow \mathbb{E}\left[Y e^{i\omega'\mathbf{X}}\right] = 0, \quad \forall \omega \in \mathbb{R}^k,$$

with Y an integrable random variable and $X \in \mathbb{R}^k$. Consequently, Kuan and Lee considered the exponential function $w_0\left(\tilde{Y}_{t,m}, \omega\right) = \mathbb{E}\left[Y_t \exp\left(i\omega'\tilde{Y}_{t,m}\right)\right]$, which led them to focus on the correlations between Y_t and exponential functions of its past. Indeed, the MDH can then be expressed as:

$$h_m(\omega) = \mathbb{E}\left[Y_t exp\left(i\omega'\tilde{Y}_{t,m}\right)\right] = 0, \quad \forall \omega \in \mathbb{R}^k. \tag{7.2}$$

Kuan and Lee summarized the above condition with a weighting function g of ω, and integrate out of ω to obtain the following implication of 7.2:

$$\int_{\mathbb{R}^{k+}} h_m(\omega) g(\omega) dw = 0.$$

The choice of the proper function g is one of the main issues of this test. Then, they considered a norm on the sample mean $\bar{\psi}_{j,g}$ which led them to the test statistic (n is the sample size):

$$KL_{g,k} = \frac{n-m}{\hat{\sigma}^2_{c,g}\hat{\sigma}^2_{s,g} - \hat{\sigma}^2_{cs,g}} \left[\hat{\sigma}^2_{s,g}\bar{\psi}^2_{c,g} + \hat{\sigma}^2_{c,g}\bar{\psi}^2_{s,g} - 2\hat{\sigma}_{cs,g}\bar{\psi}_{c,g}\bar{\psi}_{s,g}\right],$$

where, for $j = c, s$:

$$\begin{aligned}
\bar{\psi}^2_{j,g} &= \frac{1}{n-m}\sum_{t=m+1}^{n} \psi_{j,g}\left(Y_t, \tilde{Y}_{t,m}\right), \\
\hat{\sigma}^2_{j,g} &= \frac{1}{n-m}\sum_{t=m+1}^{n} \psi_{j,g}\left(Y_t, \tilde{Y}^2_{t,m}\right), \\
\hat{\sigma}_{cs,g} &= \frac{1}{n-m}\sum_{t=m+1}^{n} \psi_{c,g}\left(Y_t, \tilde{Y}_{t,m}\right)\psi_{s,g}\left(Y_t, \tilde{Y}_{t,m}\right), \\
\psi_{c,g}\left(Y_t, \tilde{Y}_{t,m}\right) &= Y_t \int_{\mathbb{R}^k} \cos\left(\omega'\tilde{Y}_{t,m}\right) g(\omega) d\omega, \\
\psi_{s,g}\left(Y_t, \tilde{Y}_{t,m}\right) &= Y_t \int_{\mathbb{R}^k} \sin\left(\omega'\tilde{Y}_{t,m}\right) g(\omega) d\omega.
\end{aligned}$$

Kuan and Lee derived closed-form formula for $\psi_{j,g}, j = c, s$, using a multivariate exponential density function g (see Kuan & Lee (2004) p.9). In the ESG tests, we have considered $m = 3$ and g as the multivariate exponential distribution with parameter $\tilde{\beta} = \left(n^{-1}\sum_{t=1}^{n} Y_t^2\right)^{-1/2}$. Contrary to most integrated test, Kuan and Lee' test has a standard limiting distribution (a $\chi^2(2)$ distribution). Furthermore, it is quite adapted to test the MDH in financial time series because it is robust to conditional heteroskedasticity and does not require strong moment conditions (two properties often featured by financial time series). However, Kuan and Lee do not consider a whole family of functions w, but only an arbitrary one. Indeed, unlike some other MDH tests, Kuan and Lee first integrate the nuisance parameter ω in the exponential function and then consider a norm. Hence, they only check for a necessary condition of the MDH because they only test the serial uncorrelatedness between Y_t and a unique function w_0. As a result, their test has power only against the alternatives that are correlated with the function w_0.

Secondly, Dominguez & Lobato (2003) considered the indicator function $w_0\left(\tilde{Y}_{t,m}, x\right) = 1\left(\tilde{Y}_{t,m} \leq x\right)$, $x \in \mathbb{R}$, and proposed two tests based on Cramér-von Mises (CvM) and Kolmogorov-Smirnov (KS) statistics:

$$CvM_{n,m} = \frac{1}{\hat{\sigma}^2 n^2} \sum_{j=1}^{n} \left[\sum_{t=1}^{n} \left(Y_t - \bar{Y}\right) 1\left(\tilde{Y}_{j,m} \leq \tilde{Y}_{j,m}\right) \right]^2,$$

$$KS_{n,m} = \max_{1 \leq i \leq n} \left| \frac{1}{\hat{\sigma}\sqrt{n}} \sum_{t=1}^{n} \left(Y_t - \bar{Y}\right) 1\left(\tilde{Y}_{j,m} \leq \tilde{Y}_{i,m}\right) \right|.$$

Nevertheless, the asymptotic distributions of these tests statistics depend on the data. Dominguez and Lobato proposed a modified wild bootstrap procedure to tabulate p-values. In addition, Escanciano (2007) proposed another class of MDH tests (CvM tests) also based on the use of indicator functions for w_0.

Integrated tests in infinite dimension

Secondly, integrated tests in infinite dimension have also been introduced in order not to miss some dependence structure due to omitted lags. Dominguez & Lobato (2003) suggested to use as a test statistic a weighted average of all the statistics established for a fixed number of lags, but they did not go deeper into this proposition (in particular about the weight of the different statistics).
The main approach for integrated tests in infinite dimension is based on a generalization of the spectral density methodology. Hong & Lee (2005) proposed a bootstrap test based on a generalized spectral density approach. However, as explained by Escanciano & Lobato (2009*b*), this test requires three choices: a kernel, a bandwidth parameter and an integrating measure. To circumvent this problem, Escanciano & Velasco (2006*a*) introduced another methodology, based on a generalized spectral distribution function, which requires less restrictive hypotheses. Nevertheless, the asymptotic null distribution of the proposed test still depends on the data, so a bootstrap procedure is required.
The null hypothesis, implied by the MDH, of the generalized spectral approach is:

$$H_0: \quad \gamma_{j,w}(x) = 0, \quad \forall j \geq 1, \forall x \in \mathbb{R},$$

with $\gamma_{j,w}(x) = \mathbb{E}\left[(Y_t - \mu)\, w_0\left(Y_{t-j}, x\right)\right]$. The generalized spectral density is defined as the Fourier transform of the functions $\gamma_{j,w}(x)$:

$$f_w\left(\varpi, x\right) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_{j,w}(x) e^{-ij\varpi}, \quad \forall \varpi \in \left[-\pi, \pi\right],$$

and contains the same information about H_0 as the whole sequence $\{\gamma_{j,w}(x)\}_{j=0}^{\infty}$. Still, Escanciano & Velasco (2006*a*) proposed to consider the generalized spectral distribution function:

$$H_w\left(\lambda, x\right) = 2\int_0^{\lambda\pi} f_w\left(\varpi, x\right) d\varpi = \gamma_{0,w}(x)\lambda + 2\sum_{j=1}^{\infty} \gamma_{j,w}(x) \frac{\sin\left(j\pi\lambda\right)}{j\pi}, \quad \lambda \in \left[0, 1\right], x \in \mathbb{R},$$

whose sample analogue is:

$$\hat{H}_w\left(\lambda, x\right) = \hat{\gamma}_{0,w}(x)\lambda + 2\sum_{j=1}^{n-1} \sqrt{1 - \frac{j}{n}}\hat{\gamma}_{j,w}(x) \frac{\sin\left(j\pi\lambda\right)}{j\pi}.$$

Rather than considering $\hat{H}_w(\lambda, x)$, Escanciano and Velasco use the discrepancy process between $\hat{H}_w(\lambda, x)$ and $\hat{H}_{0,w}(\lambda, x) = \lambda\gamma_0(x)$:

$$S_{n,w}(\lambda, x) = \sqrt{\frac{n}{2}}\left[\hat{H}_w(\lambda, x) - \hat{H}_{0,w}(\lambda, x)\right] = \sum_{j=1}^{n-1}\sqrt{n-j}\,\hat{\gamma}_{j,w}(x)\frac{\sqrt{2}\sin(j\pi\lambda)}{j\pi},$$

which is a generalization of the process used in Durlauf (1991). Escanciano and Velasco then considered the Cramér-von Mises norm $D^2_{n,w}$ to evaluate the distance of $S_{n,w}(\lambda, x)$ to zero, which presents the advantage of being free of any smoothing parameter or kernel, and can lead to tests with good power properties:

$$D^2_{n,w} = \int_{\mathbb{R}}\int_0^1 |S_{n,w}(\lambda, x)|^2 W(dx)d\lambda = \sum_{j=1}^{n-1}(n-j)\frac{1}{(j\pi)^2}\int_{\mathbb{R}}|\hat{\gamma}_{j,w}(x)|^2 W(dx),$$

where W is a weighting function. Escanciano & Velasco (2006*a*) considered $w_0(Y_{t-j}, x) = \exp(ixY_{t-j})$ and the weighting function W to be the standard normal cdf, which led to the test statistic:

$$D^2_{n,exp} = \hat{\sigma}^{-2}\sum_{j=1}^{n-1}(n-j)\frac{1}{(j\pi)^2}\sum_{t=j+1}^{n}\sum_{s=j+1}^{n}\left(Y_t - \bar{Y}_{n-j}\right)\left(Y_s - \bar{Y}_{n-j}\right)\exp\left(-0.5\left(Y_{t-j} - Y_{s-j}\right)^2\right),$$

while Escanciano & Velasco (2006*b*) used $w_0(Y_{t-j}, x) = 1(Y_{t-j} \leq w)$ and the empirical cdf as the weighting function W which led to the statistic:

$$D^2_{n,ind} = \hat{\sigma}^{-2}\sum_{j=1}^{n-1}\frac{n-j}{n(j\pi)^3}\sum_{t=1}^{n}\hat{\gamma}_{j,ind}(X_t).$$

7.5 Explicit tests for the Martingale Hypothesis

Explicit tests for the martingale hypothesis have also been proposed, in contrast to the tests for the MDH previously presented. Hong (1999) introduced one of the pioneer approaches to test the martingale hypothesis, based on a generalized spectral test. However, this test relies on the hypothesis of strict stationarity and requires the choice of smoothing parameters.

One of the most well-known martingale hypothesis tests is Park and Whang' one in Park & Whang (2005), which is rather powerful against non-martingales. Park and Whang proposed two consistent tests: a generalized Kolmogorov Smirnov type test and a Cramér-von Mises type test. Under the hypothesis that the process is first order Markovian in mean, Park and Whang' test features limiting distributions that are nicely characterized and do not require resampling procedure, so that critical values can easily be tabulated (see Park & Whang (2005) p.9 for the asymptotic critical values). They restricted their study to the case in which $\mathbb{E}[Y_t \mid \mathcal{F}_{t-1}] = \mathbb{E}[Y_t \mid Y_{t-1}]$, and used the equivalence proved by Billingsley (1995) ($\Delta Y_t = Y_t - Y_{t-1}$):

$$\mathbb{E}[\Delta Y_t \mid Y_{t-1}] = 0 \Longleftrightarrow \mathbb{E}\left[\Delta Y_t 1_{\{Y_{t-1}\leq x\}}\right] = 0, \quad \forall x \in \mathbb{R}.$$

This led them to consider the following process as the basis of their statistics:

$$Q_n(x) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\Delta Y_t 1_{\{Y_{t-1}\leq x\}}.$$

Park and Whang proposed two statistics from $Q_n(x)$. The first one is a Kolmogoriv-Smirnov type statistic:

$$S_n = \sup_{x \in \mathbb{R}} |Q_n(x)|,$$

while the second one is a Cramér-von Mises type statistic defined as:

$$T_n = \int Q_n^2(x) \mu_n(dx),$$

with μ_n a measure. Park and Whang proposed to define μ_n as the empirical distribution of (Y_{t-1}), so that T_n becomes:

$$T_n = \frac{1}{n} \sum_{t=1}^{n} Q_n^2 \left(Y_{t-1}\right).$$

Note that these two statistics depend on the levels of the considered time series $\{Y_t\}$. This is the major difference between martingale hypothesis tests and MDH tests since the first rely on the levels of the times series, while the latter MDH rely on the first differences. As a consequence, the limiting distributions derived from martingale hypothesis tests are very different from those of MDH tests. However, Park and Whang' tests are inconsistent against explosive processes, such as $AR(1)$. Furthermore, they assumed the drift of the process to be zero, that is to say they only test for a pure martingale process.

Phillips & Jin (2014) confirmed Park and Whang' test weaknesses and tried to overcome them, especially the low power against explosive alternatives. In addition, they took into account the possibility of a non-zero drift. This represents a quite desirable feature because a small (sometimes negligible) drift can be present in many financial time series, without representing the dominant component of the process. The formulation of the small drift retained by Phillips & Jin (2014) $\mu = \mu_0 n^{-\gamma}$ is thus quite appealing when testing the martingale hypothesis in financial time series. The asymptotic distributions of these tests remain easy to compute and do not require bootstrap procedure to tabulate critical values, even when the drift is not null. Nevertheless, these tests are restricted to the case of univariate first-order Markovian process following a martingale and require the specification of some parameters.

Chapter 8

Testing the Martingale Hypothesis in a risk-neutral ESG

8.1 Introduction

After presenting the classic martingale test commonly used, we will apply the statistical tests for the martingale hypothesis introduced in the last chapter to the risk neutral ESG of GALEA & Associés.

The most common martingale test applied in a risk neutral ESG, referred to as the **standard martingale test** in the following, consists in testing that the mean of discounted asset prices (derived from Monte Carlo simulation) is equal to the time zero asset price (in other words, the mean of the discounted asset prices is constant), which is a necessary, but not sufficient, condition for a martingale process. Despite its simplicity, this method requires a high number of simulations for the Monte Carlo method to perform relevant results.

On the contrary, statistical tests can be applied to each simulation, allowing to know if the corresponding scenario is theoretically relevant or should be rejected. Consequently, statistical martingale tests may represent an attractive approach to perform ESG simulations, while trying to keep only the theoretically relevant scenarios, by rejecting the ones that fail suitable statistical tests for martingale hypothesis. Still, rejecting scenarios may weaken the market consistency of the model because they may represent a reality of the market. We will therefore also try to consider these statistical tests from the point of view of the actuarial function to study how they can be used in the context of a risk-neutral ESG development or audit.

8.2 Standard martingale test in an ESG

8.2.1 Martingale test for deflators

The first test, commonly referred to as a martingale test, consists in checking that the deflator used to compute the present value of future asset prices is equal to the zero coupon prices given by the EIOPA zero coupon curve., i.e. the asset prices are discounted at the risk free rate.

Definition 8.1. *The deflator $D(t)$ is the discount factor applied to the future cash flows at time t and is equal to the product of annual discount factors based on the risk-free one year rate. Formally:*

$$D(t) = \prod_{i=0}^{t-1} P\left(i, i+1\right).$$

In a risk neutral ESG, one should have: $\mathbb{E}^{\mathbb{Q}}\left[D(t)\right] = P_{Market}\left(0, t\right)$, where $P_{Market}\left(0, t\right)$ denotes the zero coupon price with maturity t implied by the EIOPA zero coupon curve.
We present in figure 8.1 the deflators obtained with the LMM and G2++, calibrated to the market swaptions Black volatilities quoted on 31/12/2016. The test is very satisfactory for the G2++ since the deflators are very close to the market zero coupon prices. However, as concerns the LMM, we can notice a higher difference between the estimated deflators and market data.
Even though this test is referred to as a martingale test, it can clearly not be conducted by applying a statistical test for martingale hypothesis.

Figure 8.1 – Martingale test for the deflators derived from the G2++ and LMM calibrated to the market swaptions volatilities quoted on 31/12/2016

8.2.2 Application of the standard martingale test to zero-coupon

In a risk neutral framework with no arbitrage opportunities, the expected discounted future cash flows is equal to the price of the cash flow steam. Thus, a second martingale test consists in checking that the discounted price of a zero coupon bought at time t whose maturity is $T > t$ is equal to the time zero price of the zero coupon whose maturity is T. Formally,

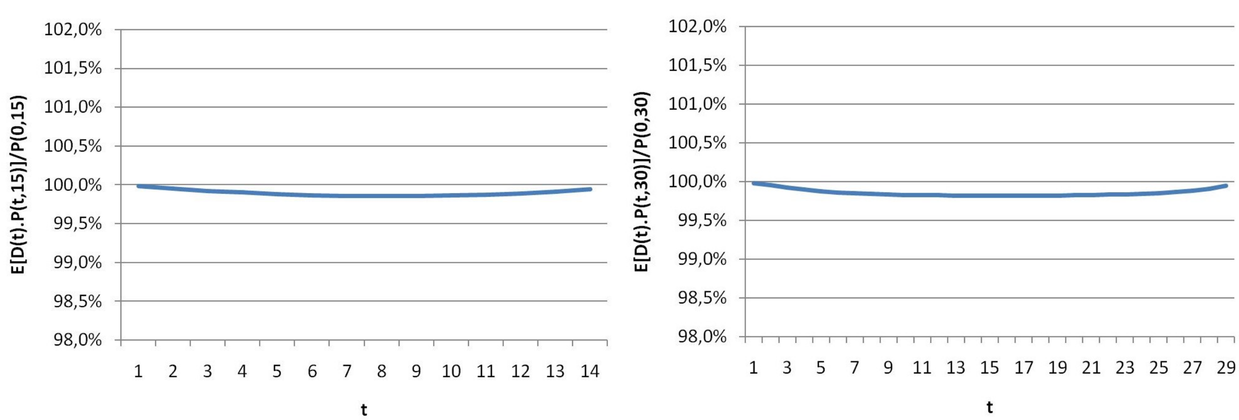
$$\mathbb{E}^{\mathbb{Q}}\left[D(t)P\left(t,T\right)\right] = P\left(0,T\right), \quad \forall t \leq T.$$

This test can be conducted with two different approaches:

1. consider the function $t \to D(t)P\left(t,T\right)$ for a fixed T and check if its mean over time is equal to $P\left(0,T\right)$,
2. consider the function $T \to D(t)P\left(t,T\right)$ for a fixed t and check if the mean evolution over time is close to the evolution of the given market zero-coupon prices.

We present the results of both approaches obtained with the G2++ and LMM calibrated to the market swaptions volatilities quoted on 31/12/2016. As concerns the first approach (see figures 8.2 and 8.3), we have considered the two cases $T = 15$ and $T = 30$ years. For the second one (see figure 8.4), we present the results for $t = 1$ and $t = 5$ years.

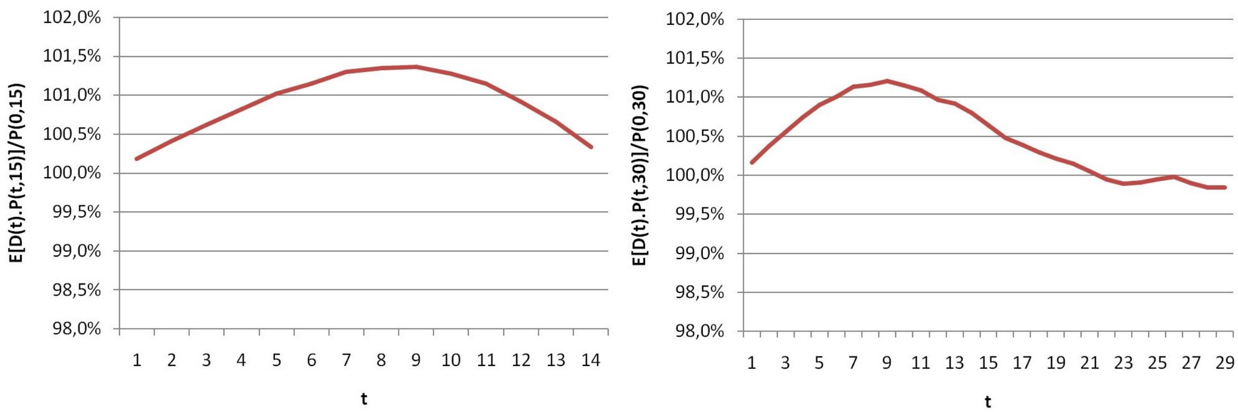
The results of these tests are satisfactory with both models. We can still notice more fluctuations of the ratio in the first method when considering the LMM. As concerns the second approach, obviously, the higher t is, the more differences appear between the discounted zero coupon price and the time zero market prices, which lead to underestimate zero coupon prices with both models.



(a) $T = 15$ years

(b) $T = 30$ years

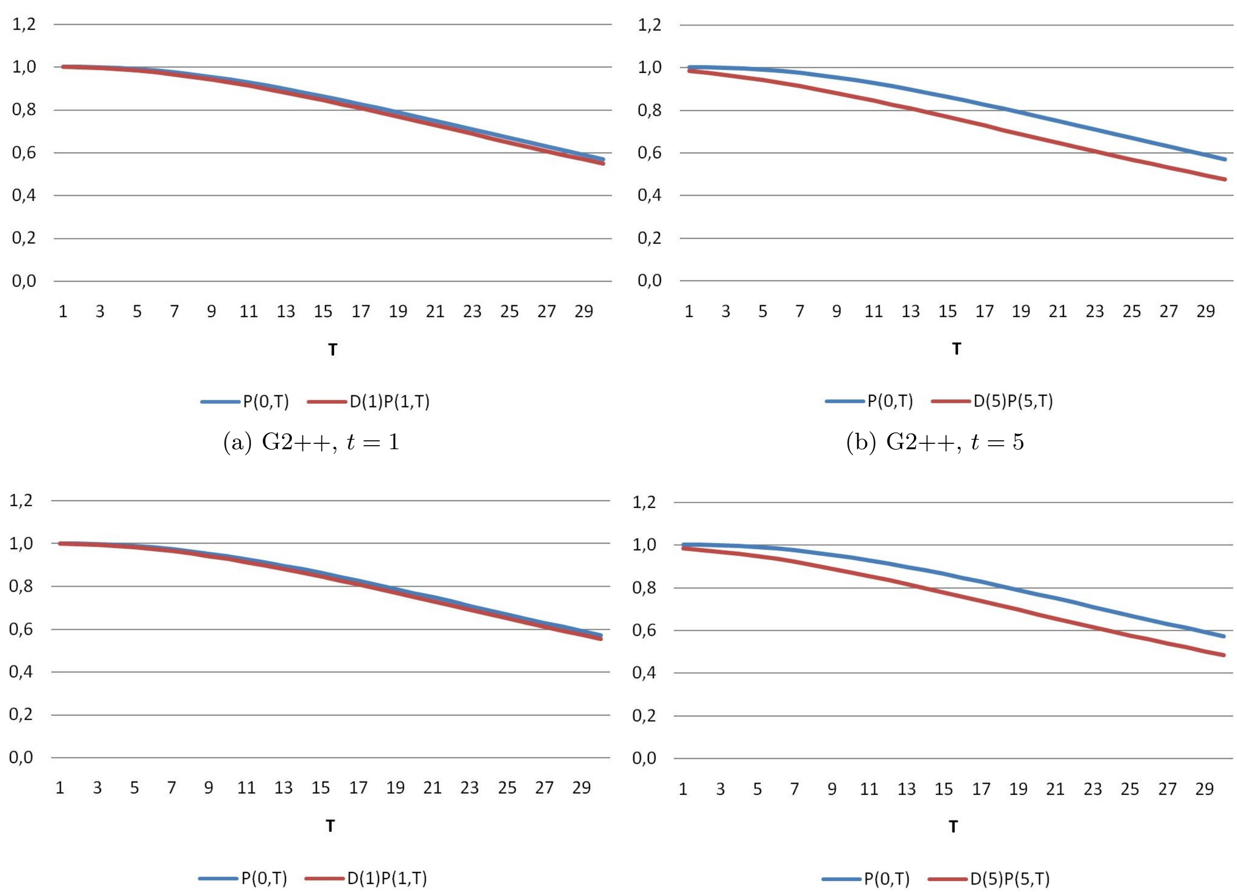
Figure 8.2 – Evolution over time t of $\mathbb{E}^{\mathbb{Q}}\left[D(t)P\left(t,T\right)\right]/P\left(0,T\right)$ derived from the G2++ calibrated to the market swaptions volatilities quoted on 31/12/2016



(a) $T = 15$ years

(b) $T = 30$ years

Figure 8.3 – Evolution over time t of $\mathbb{E}^{\mathbb{Q}}\left[D(t)P\left(t,T\right)\right]/P\left(0,T\right)$ derived from the standard LMM calibrated to the market swaptions volatilities quoted on 31/12/2016



(a) G2++, $t = 1$

(b) G2++, $t = 5$

(c) LMM, $t = 1$

(d) LMM, $t = 5$

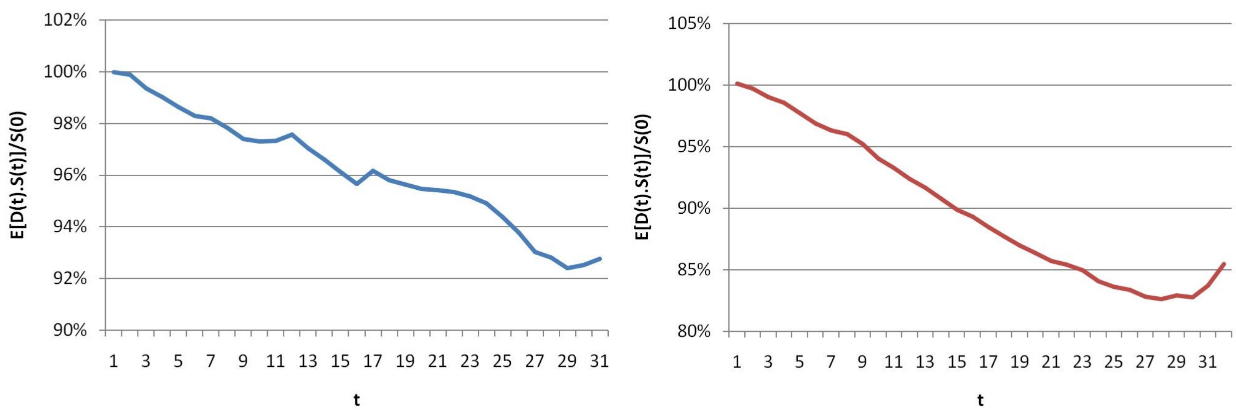
Figure 8.4 – Evolution over time T of $\mathbb{E}^{\mathbb{Q}}\left[D(t)P\left(t,T\right)\right]$ and $P_{Market}\left(0,T\right)$ derived from the G2++ and LMM calibrated to the market swaptions volatilities quoted on 31/12/2016

8.2.3 Application of the standard martingale test to share prices

Martingale tests can be applied to any other kind of asset prices. In practice, one compares the average discounted future asset price to the time zero price of the asset. In the following, we will only consider shares and not equity, since the approach would be identical.
A specific treatment is necessary when a dividend paying asset is tested. In this case, the martingale test can be based on the discounted asset price, plus the expected value of the discounted dividend cash flows over the horizon. In our case, we have considered that the dividend rate was null for the martingale test.
Formally, if $S(t)$ denotes the shares market value at time t, we test:

$$\mathbb{E}^{\mathbb{Q}}\left[D(t)S\left(t\right)\right] = S\left(0\right).$$

Once again, the test results (presented in figure 8.5) are far more satisfactory with the G2++ than the LMM, where the discounted shares values tend to take values further away from the time zero price.



(a) G2++ (b) LMM

Figure 8.5 – Evolution over time t of $\mathbb{E}^{\mathbb{Q}}\left[D(t)S(t)\right]/S\left(0\right)$ derived from the G2++ and LMM calibrated to the market swaptions volatilities quoted on 31/12/2016

8.3 Statistical martingale test within a risk neutral ESG

Despite the simplicity of the standard martingale test, it may be difficult to use it to reject the economic scenarios that fail the martingale criterion. A possible approach could consist in simulating the ESG variables for all simulations for a given maturity, excluding the scenarios in which the ESG variables take values too far away from their theoretical ones, based on a tolerance threshold previously defined, and then repeating this step for every projection step.
Based on a similar approach, statistical tests for martingale hypothesis could be applied but in this case, the decision rule (i.e. rejecting or not an economic scenario) would be more properly defined by a statistical criterion, rather than an arbitrary tolerance threshold. They can also be used by the actuarial function not to reject scenarios (since it may weaken the market consistency of the resultant model) but to have another reliable source of information about the quality and consistency of the generated scenarios with their theoretical properties, especially when testing or auditing a risk neutral ESG.

Therefore, we now propose to use the statistical tests for martingale hypothesis introduced in the previous chapter to test if discounted zero coupon prices and shares prices generated by the ESG form martingales. A major convenience of these tests is that they allow to test each scenario, i.e. each simulation, and not just the mean of the resulting quantities.

8.3.1 Presentation of the statistical martingale tests used

In the following, we will consider 15 different statistical martingale tests. Most of them are directly available through the R package *vrtest*[1]. We have implemented on our own Kuan and Lee' test and Park and Whang' one. A brief description of each test is given below (more information can be found in the previous chapter). Obviously, many more statistical tests among the voluminous literature dedicated to the subject could be applied. Still, the retained tests allow us to cover the different families of statistical tests for the martingale hypothesis. When a lag value had to be chosen, we selected a lag equal to 10, as suggested by Hyndman & Athanasopoulos (2014), since we supposed that there was no seasonality in our data (especially because of the low number of data points).

- Box Pierce Portmanteau type tests:
  1. Classic Box Pierce Portmanteau test (R function *Box.test*);
  2. Classic Ljung Box test (R function *Box.test*);
  3. Automatic Portmanteau Test of Escanciano & Lobato (2009*a*), denoted Auto.Q in the results tables (R function *Auto.Q*);
- Variance Ratio (VR) type tests:
  4. The extended version from Kim (2009), based on wild bootstrap, of the automatic VR test of Choi (1999), denoted AutoBoot.test (R function *AutoBoot.test*);
  5. A second wild bootstrap VR test, from Kim (2006), based on the VR test of Chow & Denning (1993), denoted Boot.test (R function *Boot.test*);

[1] https://cran.r-project.org/web/packages/vrtest/vrtest.pdf

6. The multiple VR test of Chen & Deo (2006), denoted Chen.Deo (R function *Chen.Deo*);
7. Chow-Denning multiple VR Tests from Chow & Denning (1993), denoted Chow.Denning (R function *Chow.Denning*, we have used the statistic noted “CD2”in the package for uncorrelated series with possible heteroskedasticity);
8. Joint version of Wright’s rank and sign tests from Belaire-Franch & Contreras (2004), denoted Joint.Wright (R function *Joint.Wright*);

- Spectral tests:

9. Spectral test of Durlauf, based on the Anderson-Darling statistic, denoted Spec. test - AD (R function *Spec.shape*);
10. Spectral test of Durlauf, based on the Cramér-von Mises statistic, denoted Spec. test - CvM (R function *Spec.shape*);
11. Generalized spectral test from Escanciano & Velasco (2006*a*), denoted Gen.Spec.Test (R function *Gen.Spec.Test*);

- Tests for MDH or Martingale Hypothesis:

12. Dominguez-Lobato test for MDH, from Dominguez & Lobato (2003), based on the Cramér-von Mises statistic, denoted DL.test Cramer (R function *DL.test*);
13. Dominguez-Lobato test for MDH, based on the Kolmogorov-Smirnov statistic, denoted DL.test Kolmog. (R function *DL.test*);
14. Kuan and Lee’ test, from Kuan & Lee (2004);
15. Park and Whang’ test, from Park & Whang (2005).

8.3.2 Application of statistical martingale tests in a risk neutral ESG

For each ESG variable, we have N trajectories of the variable over the horizon T, with N the number of simulations, i.e. a sample data $\{Y_t^i\}_{t=1,\dots,T}^{i=1,\dots,N}$ with Y the considered ESG variable. We have applied the statistical martingale tests to each trajectory $\{Y_t^i\}_{t=1,\dots,T}$ of the considered ESG variable, contrary to the standard martingale test which only tests the mean of the variable at each time step t. For each economic scenario, the statistical test indicates whether the null hypothesis H_0 (in each test, H_0 is either the martingale hypothesis, the MDH or a necessary condition, typically serial uncorrelatedness) is rejected or not, by applying the corresponding test statistic to the time series $\{Y_t^i\}_{t=1,\dots,T}$ (or its first differences when considering a MDH test).
Hence, we have chosen to present the results through the rejection probability of H_0, obtained by generating 1500 scenarios with different calibrations of the G2++ and LMM to market data quoted on 31/12/2016. In other words, we present the percentage of economic scenarios for which the null hypothesis was rejected by the considered statistical test. This will allow us to see better which tests can give relevant results when applying to a risk neutral ESG. In the following, we consider a significance level $\alpha = 5\%$.

As explained previously, the standard martingale test in a risk neutral ESG tests the constancy of the mean of discounted asset prices, which is a necessary but not sufficient condition for a martingale process. Thus, the statistical tests of serial uncorrelatedness (i.e. a necessary condition for a martingale difference sequence) do not suffer from a specific drawback compared to the standard martingale test. On the other hand, statistical tests that directly test the MDH or martingale hypothesis are much more appealing than the standard martingale test.

However, statistical tests suffer from a major drawback in our case: the low number of data points. Indeed, since we consider in our study a horizon of 30 years for a projection discretized annually, we consider time series that only count at most 30 data points (remember that we now consider each simulation to perform test and not all the simulations), which is a quite low number of data points to perform a statistical test, especially compared to the articles mentioned earlier that test the martingale difference hypothesis in financial data. This lack of data points may make it more difficult to derive relevant results, and leads us to consider our results with caution. Of course, a shorter time step would increase the number of data points but would also increase the computation time.

As a consequence, bootstrap approach (necessary to estimate the null distribution in some tests) may not be appropriate for testing risk neutral scenarios. Indeed, the low number of initial data points makes it difficult to estimate the distribution of the sample. Thus, many bootstrap iterations have to be performed in order to obtain results that can be reasonably considered to be relevant. Still, this approach may considerably increase the computation time when applied to practical situations (for example, performing 500 bootstrap replicates for each simulation of an ESG that generates 3000 scenarios).

Furthermore, another drawback of statistical tests is that, even if they allow to test each simulation in a more robust way than the standard martingale test, they require the simulation of the entire trajectory of the considered variable before performing the test, in order to have a sufficiently high number of data points. Unlike the standard martingale test, statistical tests can not be performed at each simulation step and therefore, the simulation of the entire trajectory over a sufficiently long horizon is necessary before determining whether the null hypothesis is rejected or not according to the considered test.

To apply these statistical tests, we have considered the same ESG variables as in the previous section, i.e. the time series $\{D(t)P(t,15)\}^i_{t=0,\ldots,15}$ (see figure 8.6), $\{D(t)P(t,30)\}^i_{t=0,\ldots,30}$ (see figure 8.7) and $\{D(t)S(t)\}^i_{t=0,\ldots,30}$ (see figure 8.8), with $i=1,\ldots,1500$, derived from the different calibrations of the LMM and G2++ to market data quoted on 31/12/2016. The denotations of the tests in the results tables are given above in the description of the tests. Moreover, in the results tables, "G2++ cap"means that we consider the G2++ calibrated to the caps market, "LMM joint"means that we consider the LMM jointly calibrated to the caps and swaptions markets, etc.

We can notice a large disparity between the rejection probabilities implied by the different tests. Some differences can be theoretically justified by the nature of the tests. For example, serial uncorrelatedness tests in infinite dimension, i.e. the spectral tests of Durlauf, tend to imply higher rejection probabilities than serial uncorrelatedness tests in finite dimension, such as VR tests, because the first ones consider all the autocorrelations, unlike the latter that omit some lags and may miss some dependence structure.
Nevertheless, some results are a bit surprising. In particular, tests for serial uncorrelatedness (a necessary condition for a martingale difference sequence) have higher rejection probabilities than the tests for MDH or martingale hypothesis, even though we would expect the opposite since the first ones may not be powerful against non martingale processes alternatives with zero autocorrelations.

As hinted before, some tests that required bootstrap approach are disappointing, especially the Dominguez-Lobato test and the Generalized spectral test, in the sense that they imply low rejection probabilities and do not seem to be suitable here. It is even more disappointing given that these are direct tests for the MDH (and not for a necessary condition for it). Furthermore, some tests requiring bootstrap (the VR tests AutoBoot.test and Boot.test) lead to results that seem to be relevant. However, these tests require a high computation time to be performed because of the numerous bootstrap iterations.

Kuan & Lee' and Park & Whang' tests are theoretically the most appealing tests in our study since they directly test the MDH and martingale hypothesis respectively and do not require the use of bootstrap. However, Park & Whang' test seems not to be reliable because it either always or never rejects the null hypothesis (using either the Kolmogorov-Smirnov type statistic or the Cramér-von Mises type statistic proposed by Park & Whang (2005)).

We would first recommend the use of Kuan & Lee' test because it is the only test of the MDH that seems to be reliable here in the sense that it has rejection probabilities consistent with other tests. The spectral tests of Durlauf can also be suitable as statistical tests in infinite dimension, even if they only test the necessary condition of serial uncorrelatedness and seem to possibly over reject the null hypothesis (compared to other tests). We could also add Chen and Deo' VR test and the joint version of Wright's rank and sign tests, two variance ratio tests that seem to give relevant results. Note that none of these tests requires the use of bootstrap.

Family	Test	Rejection Probabilities - G2++			Rejection Probabilities - LMM		
		G2++ cap	G2++ swap vol	G2++ swap price	LMM joint	LMM swap vol	LMM swap price
Box Pierce Port. tests	Box Pierce test	0,3%	0,3%	0,3%	0,3%	0,2%	0,2%
	Ljung-Box test	7,1%	7,1%	6,7%	7,0%	7,1%	7,1%
	Auto.Q	7,1%	6,8%	6,5%	8,4%	8,3%	8,3%
Variance Ratio tests	AutoBoot.test	4,3%	4,3%	4,3%	10,3%	10,3%	10,3%
	Boot.test	8,1%	8,5%	7,7%	13,3%	13,3%	13,2%
	Chen.Deo	2,2%	2,3%	2,1%	0,9%	0,8%	0,8%
	Chow.Denning	2,7%	2,6%	2,2%	1,7%	1,7%	1,7%
	Joint.Wright	6,9%	7,6%	6,0%	3,9%	4,0%	4,0%
Spectral tests	Spec. test - AD	9,8%	9,5%	9,1%	13,3%	13,2%	13,2%
	Spec. test - CvM	7,1%	7,1%	6,7%	9,0%	8,9%	8,9%
	Gen.Spec.Test	0,3%	0,3%	0,3%	5,2%	5,1%	5,1%
Other MDH or MH tests	DL.test Cramer	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
	DL.test Kolmog.	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
	Kuan & Lee	1,0%	1,1%	0,9%	0,9%	0,9%	0,9%
	Park & Whang	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Figure 8.6 – Rejection probabilities obtained considering the time series $\{D(t)P(t, 15)\}$

Family	Test	Rejection Probabilities - G2++			Rejection Probabilities - LMM		
		G2++ cap	G2++ swap vol	G2++ swap price	LMM joint	LMM swap vol	LMM swap price
Box Pierce Port. tests	Box Pierce test	5,9%	6,1%	5,2%	4,9%	4,3%	4,3%
	Ljung-Box test	15,9%	15,5%	15,3%	13,5%	12,3%	12,3%
	Auto.Q	16,5%	17,2%	14,5%	7,9%	7,8%	7,9%
Variance Ratio tests	AutoBoot.test	9,1%	9,1%	7,5%	5,3%	4,8%	4,8%
	Boot.test	11,5%	11,4%	10,0%	5,9%	6,3%	6,3%
	Chen.Deo	13,8%	13,1%	11,6%	3,9%	3,9%	3,9%
	Chow.Denning	4,7%	4,9%	4,0%	1,1%	0,9%	0,9%
	Joint.Wright	10,3%	10,4%	9,1%	4,9%	4,8%	4,8%
Spectral tests	Spec. test - AD	10,0%	10,3%	9,2%	7,5%	6,7%	6,7%
	Spec. test - CvM	10,9%	10,2%	9,5%	7,9%	7,3%	7,3%
	Gen.Spec.Test	0,0%	0,0%	0,0%	3,1%	3,5%	3,5%
Other MDH or MH tests	DL.test Cramer	0,0%	0,0%	0,0%	0,9%	0,9%	0,8%
	DL.test Kolmog.	0,0%	0,0%	0,0%	0,2%	0,1%	0,2%
	Kuan & Lee	5,0%	4,7%	4,1%	2,0%	2,3%	2,3%
	Park & Whang	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Figure 8.7 – Rejection probabilities obtained considering the time series $\{D(t)P(t, 30)\}$

Family	Test	Rejection Probabilities - G2++			Rejection Probabilities - LMM		
		G2++ cap	G2++ swap vol	G2++ swap price	LMM joint	LMM swap vol	LMM swap price
Box Pierce Port. tests	Box Pierce test	3,8%	3,6%	3,7%	4,1%	4,1%	4,1%
	Ljung-Box test	10,2%	10,1%	9,9%	11,3%	11,3%	11,3%
	Auto.Q	7,1%	6,9%	6,9%	8,3%	8,1%	8,1%
Variance Ratio tests	AutoBoot.test	6,9%	6,3%	6,9%	7,9%	7,9%	7,9%
	Boot.test	6,3%	6,5%	6,3%	7,4%	7,7%	7,7%
	Chen.Deo	4,5%	4,5%	4,8%	4,2%	4,5%	4,5%
	Chow.Denning	2,1%	2,1%	2,1%	2,3%	2,3%	2,3%
	Joint.Wright	6,0%	5,9%	5,7%	5,7%	5,6%	5,6%
Spectral tests	Spec. test - AD	6,8%	6,7%	6,9%	7,9%	7,7%	7,7%
	Spec. test - CvM	7,5%	7,4%	8,1%	7,7%	7,6%	7,6%
	Gen.Spec.Test	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Other MDH or MH tests	DL.test Cramer	0,8%	0,7%	0,8%	0,5%	0,5%	0,7%
	DL.test Kolmog.	0,3%	0,3%	0,3%	0,0%	0,0%	0,0%
	Kuan & Lee	4,0%	4,0%	4,2%	3,9%	4,0%	4,0%
	Park & Whang	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Figure 8.8 – Rejection probabilities obtained considering the time series $\{D(t)S(t)\}$

8.3.3 Impact of the tested ESG variable on the rejected scenarios

An interesting issue is knowing if the statistical tests reject the null hypothesis in the same economic scenarios whether we consider the discounted shares value, the discounted zero-coupon prices of maturity 15 years, etc. The following table presents the percentage of scenarios rejected by a test considering a specific ESG variable (the lines) that are also rejected by the same test considering another ESG variable (the columns). We have considered the four tests recommended in the previous paragraph: Kuan & Lee' test, the spectral tests of Durlauf and the joint version of Wright's rank and sign tests. The table in figure 8.9 was obtained considering the G2++ calibrated to the market swaptions volatilities. For example, the second cell of the first table means that 25% of the scenarios rejected by Kuan & Lee' test considering the discounted ZC prices of maturity 15 years are also rejected by the same test considering the discounted ZC prices of maturity 30 years.

We easily notice that the ESG variable to be tested has an important impact on which scenarios fail the statistical test. We can still note a certain consistency between the tests based on the discounted zero coupon price with maturity 15 or 30 years. However, there is much more disparity between the rejected scenarios when considering discounted shares prices and the ones rejected when testing discounted zero-coupon prices. This raises a major issue for the actuarial function, about which ESG variable should be considered to apply martingale statistical tests since it may considerably influence the test results and thus, the conclusion drawn from them.

Kuan & Lee Test		... also rejected considering discounted ...		
		ZC with mat. 15	ZC with mat. 30	Shares prices
Percentage of scenarios rejected considering discounted ...	ZC with mat. 15	100,0%	25,0%	0,0%
	ZC with mat. 30	5,6%	100,0%	1,4%
	Shares prices	0,0%	1,7%	100,0%

Joint Wright Test		... also rejected considering discounted ...		
		ZC with mat. 15	ZC with mat. 30	Shares prices
Percentage of scenarios rejected considering discounted ...	ZC with mat. 15	100,0%	52,6%	7,2%
	ZC with mat. 30	31,3%	100,0%	5,5%
	Shares prices	7,9%	9,9%	100,0%

Spectral test - AD		... also rejected considering discounted ...		
		ZC with mat. 15	ZC with mat. 30	Shares prices
Percentage of scenarios rejected considering discounted ...	ZC with mat. 15	100,0%	38,6%	8,3%
	ZC with mat. 30	36,6%	100,0%	6,5%
	Shares prices	11,9%	9,9%	100,0%

Spectral test - CvM		... also rejected considering discounted ...		
		ZC with mat. 15	ZC with mat. 30	Shares prices
Percentage of scenarios rejected considering discounted ...	ZC with mat. 15	100,0%	41,7%	5,8%
	ZC with mat. 30	27,0%	100,0%	5,0%
	Shares prices	5,4%	7,2%	100,0%

Figure 8.9 – Impact of the tested ESG variable on the rejected scenarios, using four different statistical martingale tests

8.3.4 Statistical martingale tests and explosive economic scenarios

We have seen in the first part that the LMM can generate explosive rates, which lead to explosive scenarios. Thus, we wanted to know if statistical tests for martingale hypothesis were able to detect these extreme scenarios, which would allow us to reject them. We know consider the LMM calibrated to the market swaptions volatilities. After computing the Best Estimate for each simulation, we have identified 17 extreme scenarios, based on the criterion that the value of the associated BE was superior to the 99% quantile of the resulting BE's. However, statistical martingale tests do not seem to be very helpful here, because, as can be seen in figure 8.10 which presents the number of extreme scenarios for which the null hypothesis was rejected by the statistical tests, they can only detect a few of them.

Family	Test	Number of extreme scenarios rejected considering discounted ...		
		ZC with mat. 15	ZC with mat. 30	Shares prices
Box Pierce Port. Test	Ljung-Box test	1	1	2
Variance Ratio tests	Boot.test	1	2	4
	Chen.Deo	0	2	2
	Chow.Denning	0	1	1
	Joint.Wright	0	1	1
Spectral tests	Spec.shape - AD	1	2	1
	Spec.shape - CvM	0	2	1
MH test	Kuan & Lee	0	0	2

Figure 8.10 – Number of extreme economic scenarios in which the martingale hypothesis was rejected by different statistical tests, among the 17 extreme scenarios generated by the LMM

8.3.5 Impact of the statistical tests on the standard martingale test

Moreover, we also wanted to see the impact of the statistical tests on the standard martingale test (which tests the constancy of the mean) commonly applied in a risk neutral ESG, i.e. see the consequences of rejecting the scenarios that fail a statistical test on the standard martingale test. Once again, we have considered the LMM calibrated to the market swaptions volatilities quoted on 31/12/2016, and tested the discounted zero coupon prices with maturity 30 years (in figure 8.11) and the discounted shares prices (in figure 8.12). We have applied the joint version of Wright's rank and sign tests, the spectral test of Durlauf based on the Cramér-von Mises statistic and Kuan & Lee' test. Nonetheless, the interpretation of the results is difficult since the means of the considered ESG variables do not seem to take closer values to their theoretical ones, even after rejecting the scenarios that fail the statistical test used. In other words, the standard martingale test does not give more satisfactory results after having rejected the scenarios that fail the statistical martingale test.

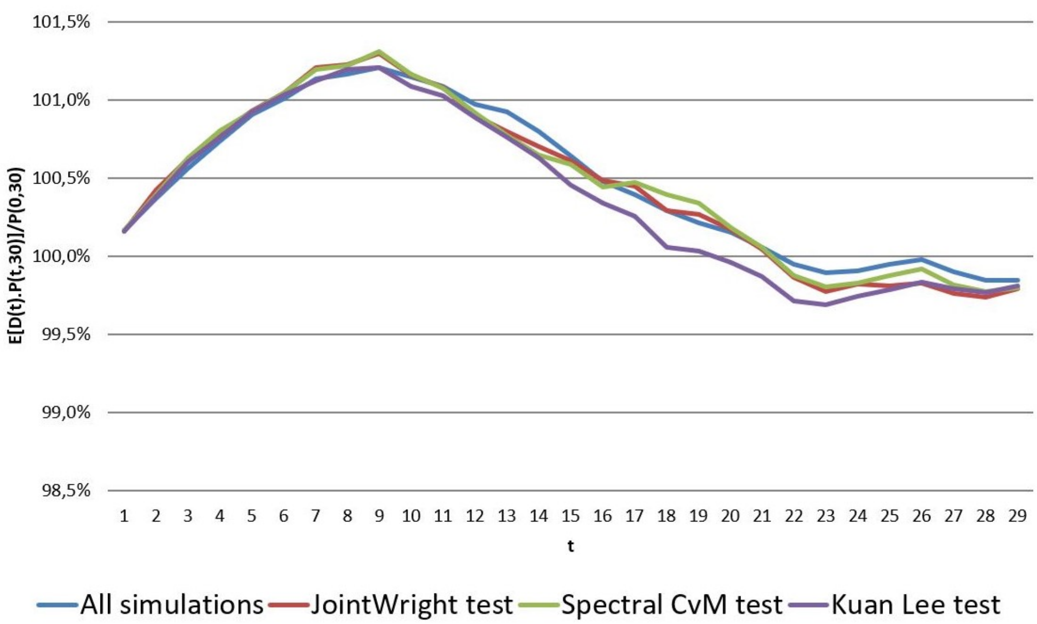


Figure 8.11 – Evolution over time T of $\mathbb{E}^{\mathbb{Q}}\left[D(t)P\left(t,30\right)\right]/P\left(0,30\right)$ derived from the LMM calibrated to the market swaptions volatilities quoted on 31/12/2016, considering all the simulations or only the ones that were not rejected by the considered statistical test

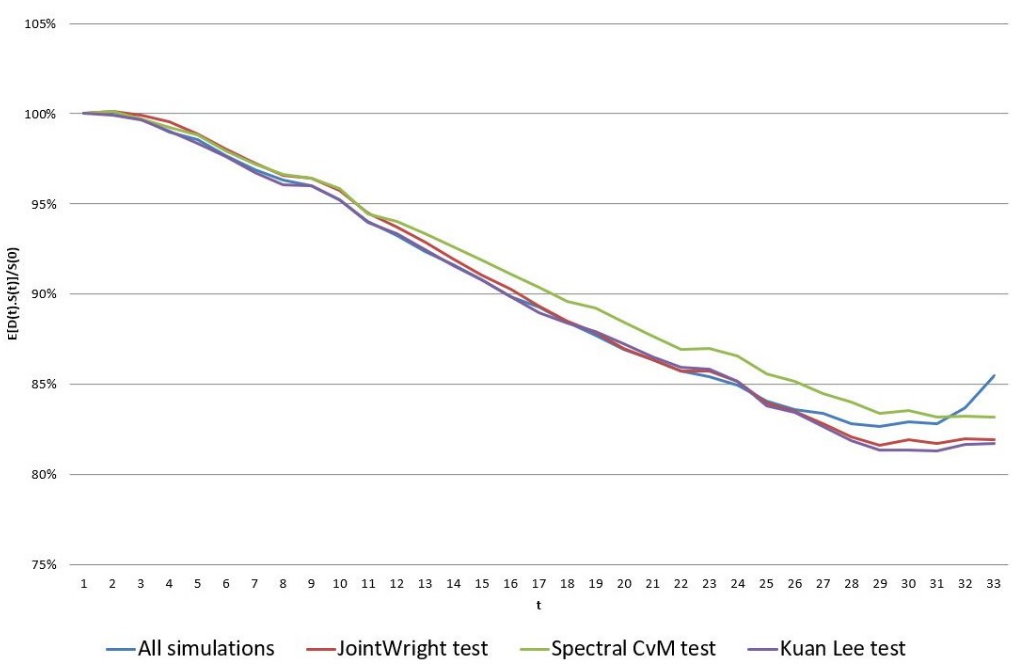


Figure 8.12 – Evolution over time T of $\mathbb{E}^{\mathbb{Q}}\left[D(t)S\left(t\right)\right]/S\left(0\right)$ derived from the LMM calibrated to the market swaptions volatilities quoted on 31/12/2016, considering all the simulations or only the ones that were not rejected by the considered statistical test

8.3.6 Detecting implementation errors with statistical martingale tests

Afterwards, we wanted to know if statistical martingale tests could detect implementation error (typically discretization error) of a model in a risk neutral ESG. To do so, we have voluntarily included some mistakes in the implementation of different models of the ESG and tested the resulting scenarios to see the impact of these errors on the null hypothesis rejection. Especially, we wanted to determine what is the best method (supposing there exists one) for the actuarial function to detect implementation errors between the standard martingale test and the statistical ones.

Implementation error in the G2++ discretization

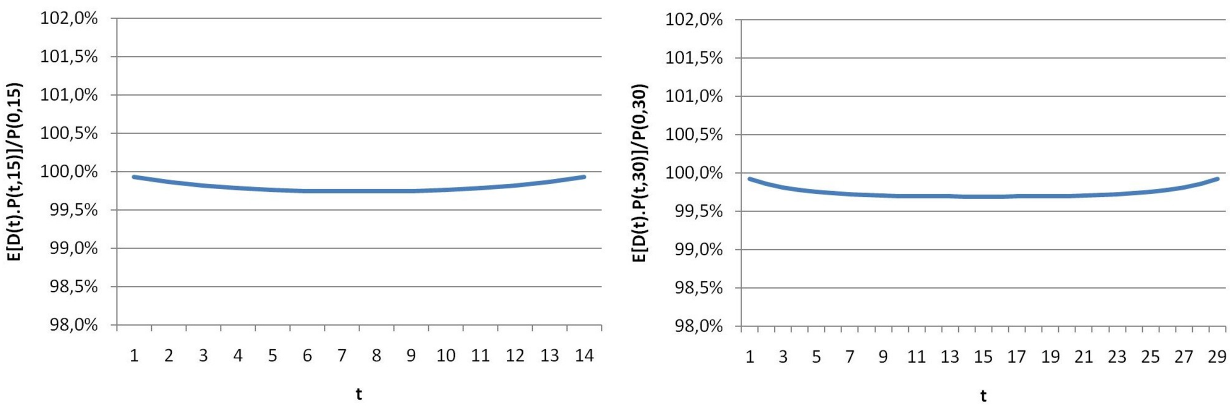
Firstly, we have included a mistake in the discretization of the G2++ model, calibrated to the market swaptions volatilities. The correct discretization of the two processes x_t and y_t composing the short rate model in the G2++ is given below in the first equation, while the incorrect discretization implemented is the second one. This error might seem aberrant in practice, however keep in mind that we precisely want to estimate the capacity of statistical tests to detect this kind of mistake, in particular compared to the standard martingale test.

$$
\begin{aligned}
x_{t+1} &= x_t e^{-a} + \epsilon_{x,t}\sigma\sqrt{\frac{1-e^{-2a}}{2a}} \qquad \text{: exact G2++ discretization,} \\
x_{t+1} &= x_t e^{-a} + \epsilon_{x,t}\left(\sigma\sqrt{\frac{1-e^{-2a}}{2a}}\right)^2 \qquad \text{: implemented G2++ discretization.}
\end{aligned}
$$

We first present in figure 8.13 the results obtained with the standard martingale test, through the evolution of the mean of discounted zero coupon prices of maturity 15 and 30 years. Note that the results could seem satisfactory to an unaware user because we do not notice a major difference between the resulting graphs and the ones previously presented in 8.2, despite the discretization error, even though the resulting ratios are further from 100% than previously.

On the contrary, the statistical martingale tests applied to the same scenarios (see figure 8.14) highlight the implementation error in the model discretization, through high rejection probabilities of the null hypothesis in the generated scenarios (particularly with tests based on bootstrap), such that an unaware user could almost be sure there is an error in the G2++ implementation. However, we must qualify this assertion because the rejection probabilities of some tests (in particular Kuan & Lee and Box Pierce test) remain close to the ones obtained previously without the implementation error, which means that an unaware user would probably not be able to detect the discretization error in the G2++ if he only considered these tests. This emphasizes the importance of the choice of the statistical martingale test used and the necessity of applying several tests.

In addition, statistical martingale tests did not detect every implementation error that we had voluntarily included in the ESG. For example, after including small mistakes in the discretization of the G2++ or in the G2++ zero-coupon price closed-form formula, statistical tests did not reject more scenarios than previously and thus, did not really detect the implementation errors (even if we noticed a small increase in the rejection probabilities when considering discounted zero coupon prices with long maturity), which were not detected by the standard martingale test neither. Moreover, we also put some errors that could be called aberrant in the discretization of the LMM or of the Black, Scholes and Merton dynamics, but, except for enormous mistakes that led to aberrant scenarios, neither the statistical martingale tests nor the standard one detected these implementation errors. Finally, statistical martingale tests may not represent an ideal method able to highlight any kind of mistake that could be present in a risk neutral ESG implementation, but as hinted previously, they might be potentially more robust than the standard martingale test to detect implementation errors.



(a) $T = 15$ (b) $T = 30$

Figure 8.13 – Evolution over time T of $\mathbb{E}^{\mathbb{Q}}\left[D(t)P\left(t,T\right)\right]/P\left(0,T\right)$ derived from the G2++ voluntarily implemented with an error in its discretization

Family	Test	Rejection Probabilities considering discounted	
		ZC with mat. 15	ZC with mat. 30
Box Pierce Port. tests	Box Pierce test	2,3%	6,3%
	Ljung-Box test	71,7%	17,7%
	Auto.Q	58,5%	61,3%
VR tests	AutoBoot.test	84,5%	77,5%
	Boot.test	94,2%	95,3%
	Joint.Wright	2,9%	13,9%
Spectral test	Spec. test - AD	30,1%	12,6%
	Spec. test - CvM	17,0%	17,3%
	Gen.Spec.Test	43,1%	60,2%
MDH test	Kuan & Lee	2,6%	4,9%

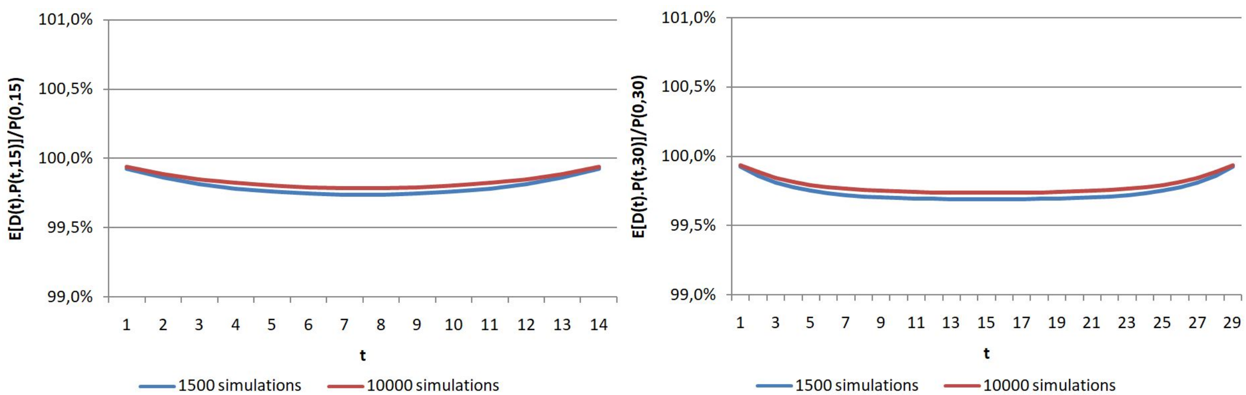
Figure 8.14 – Rejection probabilities obtained considering the time series $\{D(t)P\left(t,15\right)\}$ and $\{D(t)P\left(t,30\right)\}$ in the scenarios generated with a discretization error in the G2++ implementation

Impact of the number of simulations

It might seem interesting to study the impact of the number of simulations performed on the test results obtained when trying to detect implementation error in an ESG. We now consider the same situation as previously, i.e. the G2++ is implemented with the discretization error given above. While the previous test results were obtained by running 1500 simulations, we now present the results of the same tests but this time, obtained by running 10 000 simulations.

We first present in figure 8.15 the results obtained with the standard martingale test. We notice that the results seem to be more satisfactory than previously since the resulting ratios are closer to 100% than in figure 8.13. Hence, the higher number of simulations makes it even more difficult to detect the discretization error with the standard martingale test.

As concerns the statistical martingale tests in figure 8.16, the results are globally stable, we can only observe a small increase of the rejection probabilities of the null hypothesis. Therefore, the statistical martingale tests results do not seem to be very sensitive to the number of simulations performed.



(a) $T = 15$ (b) $T = 30$

Figure 8.15 – Evolution over time T of $\mathbb{E}^{\mathbb{Q}}\left[D(t)P\left(t,T\right)\right]/P\left(0,T\right)$ derived from the G2++ voluntarily implemented with an error in its discretization, after performing 10 000 simulations

Family	Test	Rejection Probabilities considering discounted	
		ZC with mat. 15	ZC with mat. 30
Box Pierce Port. tests	Box Pierce test	2,5%	6,6%
	Ljung-Box test	72,5%	17,7%
	Auto.Q	58,5%	61,8%
VR tests	AutoBoot.test	84,9%	77,9%
	Boot.test	93,5%	95,3%
	Joint.Wright	2,5%	14,2%
Spectral test	Spec. test - AD	29,4%	13,1%
	Spec. test - CvM	16,5%	17,9%
	Gen.Spec.Test	42,8%	60,5%
MDH test	Kuan & Lee	2,3%	5,0%

Figure 8.16 – Rejection probabilities obtained considering the time series $\{D(t)P\left(t,15\right)\}$ and $\{D(t)P\left(t,30\right)\}$ in the scenarios generated with a discretization error in the G2++ implementation, after performing 10 000 simulations

8.3.7 Impact of the interest rate model discretization on martingale tests

It could be interesting to study the impact of the discretization scheme and step of the interest rate model on the martingale tests, given that, on the one hand, the ESG simulations are more precise with a smaller discretization step and on the other hand, it allows to have more data points available for the statistical tests, which can then give more relevant results.

Smaller discretization step for the G2++

First of all, we consider the G2++ model discretized semi-annually (i.e. $\Delta t = 1/2$), calibrated to the market swaptions volatilities. We have applied the standard martingale test and the statistical ones to the discounted zero coupon prices with maturity 15 and 30 years. Compared to the results obtained with an annual discretization, the rejection probabilities of all statistical tests in figure 8.17 have clearly increased. This is not so surprising because the increase in the number of data points (due to the decrease of the discretization step) allows statistical tests to be more precise and to probably better detect the scenarios that do not respect the martingale property (at least according to those tests). The standard martingale test does not give much information here about the impact of the discretization step since the results were very close to those obtained previously.

Family	Test	Rejection Probabilities considering...	
		ZC price with mat. 15	ZC price with mat. 30
Box Pierce Port. tests	Box Pierce test	5,8%	14,1%
	Ljung-Box test	16,2%	21,3%
	Auto.Q	15,6%	30,5%
Variance Ratio tests	Boot.test	27,7%	27,7%
	Chen.Deo	10,9%	32,1%
	Chow.Denning	4,5%	12,8%
	Joint.Wright	11,1%	30,7%
Spectral tests	Spec. test - AD	10,7%	21,9%
	Spec. test - CvM	11,5%	23,0%
	Gen.Spec.Test	0,0%	0,0%
MDH test	Kuan & Lee	4,9%	15,6%

Figure 8.17 – Rejection probabilities obtained considering the time series $\{D(t)P(t,15)\}$ and $\{D(t)P(t,30)\}$ in the scenarios generated by the G2++ with a semi-annual discretization

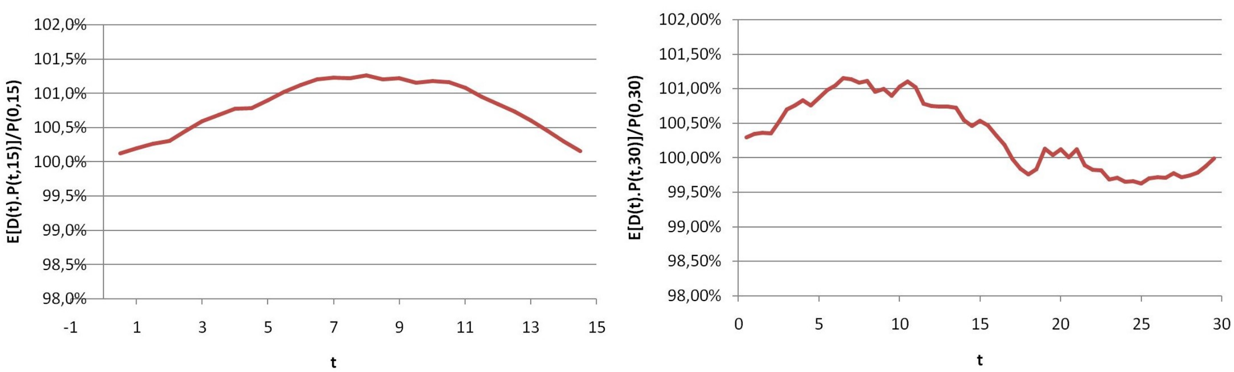
Smaller discretization step for the LMM

Secondly, we have applied the same approach with the LMM, discretized semi-annually and calibrated to the market swaptions volatilities. Studying the impact of the discretization step of the LMM on the martingale criterion is very interesting since, as explained by Brigo & Mercurio (2006) (p.266), discretizing the continuous time dynamics of the LMM does not lead to arbitrage free LMM in discrete time. Nevertheless, the violation of the arbitrage free condition can be considered as negligible if the discretization step is small enough. It is therefore interesting to see if decreasing the discretization step improves the results obtained with the different martingale tests.

However, contrary to what we expected, we notice an increase in the rejection probabilities of the statistical tests in figure 8.18, just like with the G2++. The standard martingale test in figure 8.19 gives similar results to the ones previously obtained with an annual discretization (figure 8.3).

Family	Test	Rejection Probabilities considering...	
		ZC price with mat. 15	ZC price with mat. 30
Box Pierce Port. tests	Box Pierce test	6,0%	11,4%
	Ljung-Box test	14,1%	17,5%
	Auto.Q	11,8%	9,6%
Variance Ratio tests	Boot.test	12,1%	8,2%
	Chen.Deo	3,3%	4,6%
	Chow.Denning	2,9%	3,4%
	Joint.Wright	9,0%	8,6%
Spectral tests	Spec. test - AD	10,7%	8,6%
	Spec. test - CvM	11,1%	9,6%
	Gen.Spec.Test	8,3%	4,8%
MDH test	Kuan & Lee	8,6%	5,1%

Figure 8.18 – Rejection probabilities obtained considering the time series $\{D(t)P(t,15)\}$ and $\{D(t)P(t,30)\}$ in the scenarios generated by the LMM with a semi-annual discretization



(a) $T = 15$ (b) $T = 30$

Figure 8.19 – Evolution over time T of $\mathbb{E}^{\mathbb{Q}}[D(t)P(t,T)]/P(0,T)$ derived from the LMM discretized semi-annually

Euler discretization scheme for the G2++

Furthermore, we also wanted to study the impact of the discretization scheme on the martingale tests. We have only considered the G2++ model here. There exists an exact discretization of the G2++, which is the one implemented in the ESG. Nevertheless, we have also implemented the G2++ with Euler discretization scheme and an annual step, and then applied martingale tests to the resulting zero coupon prices to study the impact of the discretization scheme (see Kloeden & Platen (1995) or Planchet & Therond (2005) for more details on discretization schemes).

Definition 8.2. *Consider a process X_t whose dynamics is given by:*

$$dX_t = \mu(X_t,t)\,dt + \sigma(X_t,t)\,dB_t, \quad X_0 = x,$$

with B a Brownian motion. The discretization $\left(\tilde{X}_{k\delta}\right)_{k\in[1;T/\delta]}$ of the process X_t based on Euler scheme is given by (with δ the discretization step):

$$\tilde{X}_{t+\delta} = \tilde{X}_t + \mu\left(\tilde{X}_t, t\right)\delta + \sigma\left(\tilde{X}_t, t\right)\sqrt{\delta}\epsilon.$$

Hence, the Euler discretization for the two processes x_t and y_t composing the G2++ model (discretized annually) is given by:

$$x_{t+1} = x_t - a.x_t + \sigma.\epsilon_{x,t}.$$

Since Euler discretization is an approximate discretization scheme, we expect less satisfactory results than with the exact discretization. This intuition was confirmed by the statistical tests (see 8.20) because the rejection probabilities are higher with Euler discretization scheme than with the exact one (figures 8.6 and 8.7). It would be more difficult to conclude with the standard martingale test, which does not give much relevant information.

Family	Test	Rejection Probabilities considering...	
		ZC price with mat. 15	ZC price with mat. 30
Box Pierce Port. tests	Box Pierce test	0,5%	6,9%
	Ljung-Box test	8,2%	18,7%
	Auto.Q	7,2%	20,3%
Variance Ratio tests	Boot.test	9,9%	15,0%
	Chen.Deo	2,1%	16,5%
	Chow.Denning	2,9%	6,1%
	Joint.Wright	7,7%	14,2%
Spectral tests	Spec. test - AD	10,2%	12,8%
	Spec. test - CvM	8,1%	14,1%
	Gen.Spec.Test	0,3%	0,0%
MDH test	Kuan & Lee	1,1%	7,4%

Figure 8.20 – Rejection probabilities obtained considering the time series $\{D(t)P(t,15)\}$ and $\{D(t)P(t,30)\}$ in the scenarios generated by the G2++ discretized with Euler scheme

8.3.8 Opposition between market consistency and the martingale property ?

We have seen in the first part that, because of current high market volatilities, we have to apply some constraints to the ESG calibration process in order not to generate aberrant results, which can considerably weaken the market consistency of the resulting economic scenarios. Yet, we might intuitively think that the more volatile a model is, the more difficult it is in practice to respect the martingale property. Therefore, one may think that there exists an opposition between the two main validation criteria for a risk neutral ESG (i.e. the market consistency and the martingale property).

In this section, we use the theoretical calibrations of the G2++ and LMM to the market swaptions volatilities obtained in the first part (i.e. the calibrations derived from market data without any additional constraint). Thus, the ESG now generates scenarios that are far more market consistent than previously. We will study the impact on the martingale tests to see if there exists an operational contradiction between the two validation criteria.

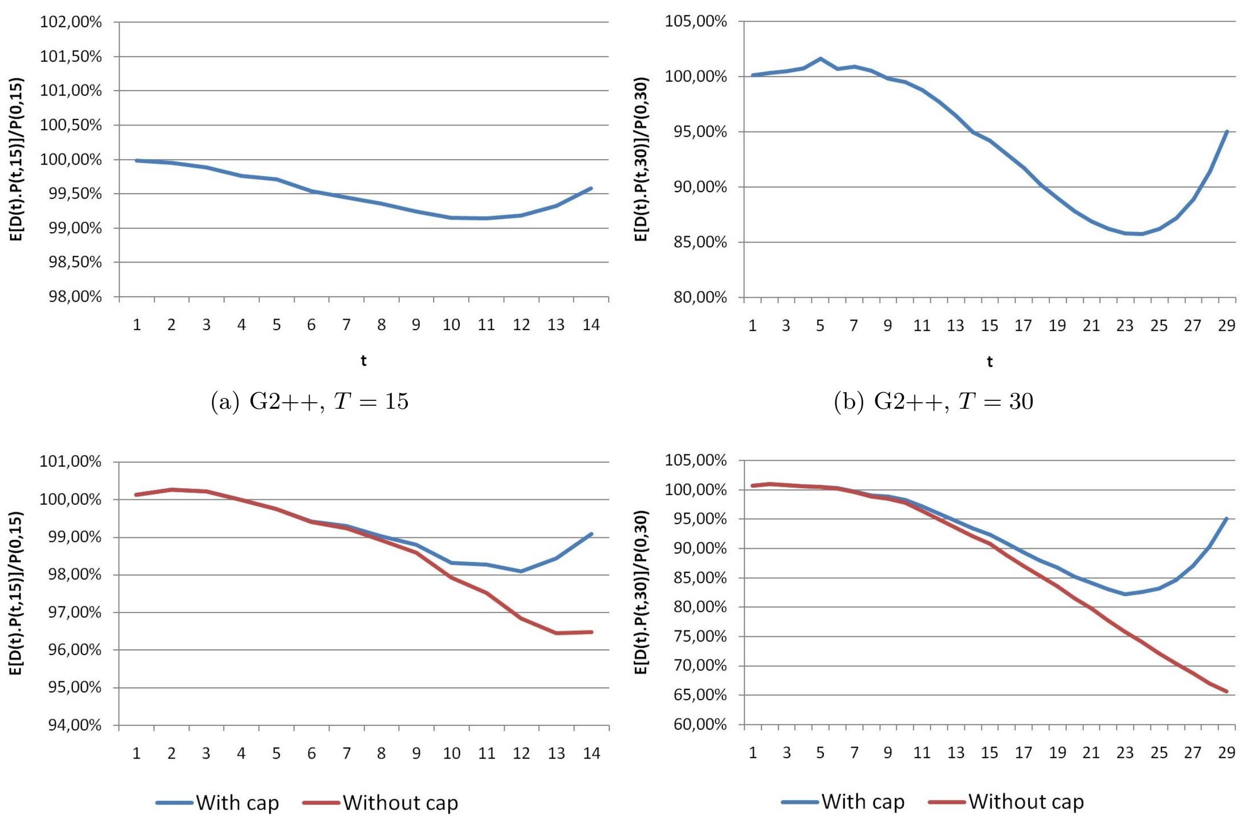
Furthermore, it is common to put a cap on the forward rates generated by the LMM because of its capacity of generating explosive rates. However, the use of a cap weakens the market consistency of the generated scenarios since we truncate some scenarios that could represent a market reality and thus, create a bias in the simulation. On the other hand, one may think that explosive rates weaken the martingale criterion. Therefore, we have also applied martingale tests to economic scenarios generated by the LMM without applying a cap on the forward rates.

We notice a high disparity in the rejection probabilities of the statistical tests in figure 8.21, and that the tests with the highest rejection probabilities are not the same as previously. Still, we easily note a global increase in the rejection probabilities, especially with Kuan and Lee' test which is the test with the highest rejection probability and the most impacted one by the use of the theoretical calibrations. Moreover, the standard martingale test (see figure 8.22) is also clearly impacted by the theoretical calibrations since the discounted zero coupon prices are further away from their theoretical value than previously.
As concerns the impact of the cap applied on the forward rates generated by the LMM, the statistical tests results are quite confusing because, when the cap is not used, some of the tests have lower rejection probabilities while some others have higher ones. Nonetheless, the standard martingale test (figure 8.22) clearly gives less satisfactory results without the application of a cap, since the mean of the discounted zero coupon prices take values much further away from their time zero value than when we applied a cap on the forward rates.

When considering these results, we may think that the objective of market consistency can be contradictory to the martingale criterion, because the generated scenarios are, on the one hand, more market-consistent than previously, but on the other hand, the martingale tests results are less satisfactory. However, the additional constraints applied in the calibration process and the cap put on the forward rates generated by the LMM weaken the market consistency, as already mentioned, but the bias that they create also theoretically weakens the martingale property. As can be seen in figure 8.22, they may improve the convergence speed, but the bias weakens both the market consistency and the martingale property from a theoretical point of view. Hence, we might think that there is a contradiction between the two validation criteria of a risk neutral ESG if we only consider the operational level (see figure 8.22), but this does not prove that there exists a theoretical opposition between the market consistency criterion and the martingale property.

Family	Test	Rejection Probabilities G2++ ZC with mat. 15	G2++ ZC with mat. 30	LMM with cap ZC with mat. 15	LMM with cap ZC with mat. 30	LMM without cap ZC with mat. 15	LMM without cap ZC with mat. 30
Box Pierce Port. tests	Box Pierce test	0,1%	6,0%	0,1%	5,2%	0,3%	6,1%
	Ljung-Box test	5,0%	13,3%	5,0%	9,3%	7,4%	11,4%
	Auto.Q	6,3%	14,7%	6,3%	13,7%	12,7%	13,5%
Variance Ratio tests	Boot.test	10,7%	19,6%	19,0%	21,9%	18,5%	17,2%
	Chen.Deo	1,9%	3,3%	0,9%	3,9%	0,8%	2,9%
	Chow.Denning	4,2%	3,7%	4,5%	3,6%	4,3%	4,7%
	Joint.Wright	4,9%	13,1%	4,9%	21,1%	5,0%	30,5%
Spectral test	Spec. test - AD	13,5%	14,5%	15,5%	17,1%	16,8%	17,1%
	Spec. test - CvM	8,7%	15,7%	10,0%	18,2%	11,2%	17,5%
	Gen.Spec.Test	3,5%	6,6%	3,8%	6,4%	3,3%	3,3%
MDH test	Kuan & Lee	0,6%	21,6%	9,4%	32,8%	9,5%	31,3%

Figure 8.21 – Rejection probabilities obtained considering the time series $\{D(t)P(t,15)\}$ and $\{D(t)P(t,30)\}$ in the scenarios generated by the G2++ and LMM, with and without a cap applied on the forward rates, with the theoretical calibrations to the market swaptions volatilities



(a) G2++, $T = 15$

(b) G2++, $T = 30$

(c) LMM, $T = 15$

(d) LMM, $T = 30$

Figure 8.22 – Evolution over time T of $\mathbb{E}^{\mathbb{Q}}[D(t)P(t,T)]/P(0,T)$ derived from the G2++ and the LMM, with and without a cap applied on the generated forward rates, with the theoretical calibration to the market swaptions volatilities

8.3.9 Recommendations

We will now give some recommendations about the statistical martingale tests to apply in a risk neutral ESG, based on our study. First of all, as already explained, the most appealing test are theoretically the ones for the MDH or martingale hypothesis. Still, we have seen that Park and Whang' test, Dominguez and Lobato' test and the generalized spectral test are quite disappointing because their resulting rejection probabilities were very low compared to other tests. The only test for the MDH that we would recommend is Kuan and Lee' one, which gave us several interesting results in the different studied problems.

To test the necessary condition of serial uncorrelatedness, we think that Durlauf spectral tests are the most reliable tests since they can test all autocorrelations to be null, unlike Box-Pierce type and VR type tests, even if they may over-reject the null hypothesis, as explained in the previous chapter. Box Pierce Portmanteau type tests (especially Ljung Box test and Auto.Q test) seem to be the statistical tests that reject the most the null hypothesis among the different tests used, but still may be useful when checking that there is no implementation error in a risk-neutral ESG. As concerns variance ratio tests, which can give results with high disparity based on which VR test is used, the most relevant ones would be Chen-Deo' and Joint Wright' tests. When trying to detect implementation errors, tests based on bootstrap were the tests with the highest rejection probabilities. Ljung Box test, Auto.Q and spectral tests, even the generalized spectral test, were also able to detect implementation errors.

We would rather not recommend tests based on bootstrap approach because of the high necessary computation time. Still, some of them can give reliable results, such as AutoBoot.test and Boot.step. Nevertheless, we did not go deep into the bootstrap methodology retained. Indeed, the R functions allow to use different bootstrap methods, but we have always retained the standard approach proposed by the test. Thus, studying which bootstrap approach is the most appropriate for testing the martingale hypothesis in a risk neutral ESG might represent an interesting issue to invest in the future.

To conclude the second part, statistical martingale tests definitely represent a reliable source of information for the actuarial function when testing a risk-neutral ESG, especially because they can potentially better detect implementation errors than the standard martingale test. Still, we do not pretend that we have found an ideal approach through statistical martingale test to check the martingale criterion in a risk neutral ESG, since we recommend to use statistical martingale tests and the standard martingale test in a complementary manner. In addition, many more statistical tests among the voluminous literature dedicated to the subject could be tested. Nevertheless, we can once more emphasize the importance of the actuarial function which, in the context of statistical martingale tests, would have to arbitrate many problems and justify as many choices, such as:

- The statistical tests retained: how to justify that the tests used are the most reliable and do not over or low reject the null hypothesis. As concerns tests based on bootstrap approach, one should also be able to justify the bootstrap methodology retained;
- The choice of the ESG variable to be tested: we have seen that the considered ESG variable may considerably influence the test results and thus the conclusion drawn from them;
- The actuary should also keep in mind that the statistical test results can be greatly influenced by the discretization scheme and step, and obviously by the projection horizon.

Conclusion

In the first part of this study devoted to the study of the impact of the interest rate model of the ESG and of its calibration on the generated economic scenarios and the results of the ALM model, we have seen that the Best Estimate was sensitive to the choice of the interest rate model and the derivative used in the calibration process (even if the differences observed were not high) and that therefore, these choices should be documented and subjects of reflection for the actuarial function. More precisely, we first observed a definite impact of the interest rate model on several ESG quantities (yield curve shape, zero-coupon prices, etc.). However, despite these differences in the economic scenarios, we calculated close Best Estimate values with the ALM model. It can still be noted that the Best Estimates obtained with the standard Libor Market Model were higher because of the presence of extreme scenarios.

The derivative used to calibrate the interest rate model seems to have a lower impact on both the generated scenarios and the Best Estimate. It is then difficult to draw a solid conclusion about the consequences of the choice of the derivative considered in the calibration process, for several reasons. Firstly, since we have considered a risk-neutral ESG, we have only used market data quoted on 31/12/2016, and thus, if we used data quoted even shortly after, we would maybe obtained different results in terms of the impact of the derivative (for example, the highest BE value would not result from the same derivative). Secondly, and this is the main limit of the first part, since the current market implied volatilities are very high, we have seen that the interest rate models considered in this study were inadequate to reproduce these high market volatilities without generating aberrant and unstable scenarios. Consequently, we had to constrain the model parameters in the calibration process, which may have led the ESG to produce similar scenarios and thus complicate the study of the impact of the derivative used to calibrate the model. In addition, because of these operational constraints, the market consistency criterion was not very well respected by the generated scenarios.

A chapter devoted to the calibration of the ESG to an imaginary economic environment characterized by much lower market volatilities has shown that without additional constraints, the derivative used to calibrate the interest rate model could lead to generate economic scenarios with marked differences (particularly the yield curve shapes). Nonetheless, as previously, the resulting Best Estimate values were close.

To extend this first part, we could implement more recent and complex interest rate models that would be able to reproduce high market volatilities without generating unstable scenarios and therefore, better respect the market consistency criterion. We would then be more confident about the results obtained on the impact of the derivative considered in the calibration process.

The second part focused on the tests that can be used to audit a risk-neutral ESG, and in particular aimed at presenting the main statistical martingale tests that can be found in the literature and applying them to a risk neutral ESG. These tests can be theoretically very different and lead to confusing results in terms of rejection of the martingale hypothesis, some being more adapted to certain situations than others. Still, we have shown that they can be a reliable source of information for the actuarial function when testing a risk-neutral ESG. The principal points highlighted were the importance of the ESG quantity considered to apply those statistical tests, their sensitivity to the discretization scheme and step and their capacity of potentially better detecting implementation errors in the ESG than the standard martingale test (which only tests the constancy of the mean). We have concluded this part by proposing some recommendations for use and underlining the responsibilities of the actuarial function if one wants to employ statistical martingale tests within a risk-neutral ESG, especially when choosing and justifying the tests to use and the ESG quantity to consider to apply them.

Possible extensions of the second part could be to implement other statistical martingale tests (we have mainly considered the tests directly available through the R package *vrtest*[1], but we have not used all of them) and to apply them to other market situations than the 31/12/2016 to see if we would make the same observations as in this study.

[1]https://cran.r-project.org/web/packages/vrtest/vrtest.pdf

Bibliography

Andersen, L. & Piterbarg, V. (2010), *Interest Rate Modeling Volume III: Products and Risk Management*, Atlantic Financial Press.

Belaire-Franch, G. & Contreras, D. (2004), *Ranks and signs-based multiple variance ratio tests*, Working paper, University of Valencia.

Belaire Franch, J. & Opong, K. (2005), 'Some evidence of random walk behaviour of euro exchange rates using ranks and signs', *Journal of Banking and Finance* **29**, 1631–1643.

Bierens, H. (1982), 'Consistent model specification tests', *Journal of Econometrics* **20**, 102–134.

Billingsley, P. (1995), *Probability and Measure, 3rd ed*, John Wiley: New York.

Black, F., Derman, E. & Toy, W. (1990), 'A one-factor model of interest rates and its application to treasury bond options', *Financial Analysts Journal* **46**, 33–39.

Black, F. & Karasinski, P. (1991), 'Bond and option pricing when short rates are lognormal', *Financial Analysts Journal* **47**, 52–59.

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *The Journal of Political Economy* **81**, 637–654.

Box, G. & Pierce, D. (1970), 'Distribution of residual autocorrelations in autoregressive integrated moving average time series models', *Journal of the American Statistical Association* **65**, 1509–1526.

Brace, A. (1996), *Dual Swap and Swaption Formulae in Forward Models*, School of Mathematics, UNSW, Australia.

Brace, A. (1997), *Rank-2 Swaption Formulae*, UNSW preprint.

Brace, A., Gatarek, D. & Musiela, M. (1997), 'The market model of interest rate dynamics', *Mathematical Finance* **7**, 127–147.

Brennan, M. & Schwartz, E. (1982), 'An equilibrium model of bond pricing and a test of market efficiency', *Journal of Financial and Quantitative Analysis* **17**, 1–30.

Brigo, D. & Mercurio, F. (2006), *Interest Rate Models - Theory and Practice With Smile, Inflation and Credit*, Springer.

Campbell, J., Lo, A. & MacKinlay, A. (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Charles, A. & Darne, O. (2009), ‘Variance ratio tests of random walk: An overview’, *Journal of Economic Surveys* **23**, 503–527.

Charles, A., Darne, O. & Kim, J. (2011), ‘Small sample properties of alternative tests for martingale difference hypothesis’, *Economics Letters* **110**, 151–154.

Chen, W. & Deo, R. (2006), ‘The variance ratio statistic at large horizons’, *Econometric Theory* **22**, 206–234.

Choi, I. (1999), ‘Testing the random walk hypothesis for real exchange rates’, *Journal of Applied Econometrics* **14**, 293–308.

Chow, K. & Denning, K. (1993), ‘A simple multiple variance ratio rest’, *Journal of Econometrics* **58**, 385–401.

Cochrane, J. (1988), ‘How big is the random walk in gnp?’, *Journal of Political Economy* **96**, 893–920.

Collin Dufresne, P. & Goldstein, R. (2002), ‘Pricing swaptions within an affine framework’, *The Journal of Derivatives* **10**, 9–26.

Cox, J., Ingersoll, J. & Ross, S. (1985), ‘A theory of the term structure of interest rates’, *Econometrica* **53**, 385–407.

Demestrescu, M. (2009), ‘Panel unit root testing and the martingale difference hypothesis for german stocks’, *Economics Bulletin* **29**, 1749–1759.

Di Francesco, M. (2012), *A General Gaussian Interest Rate Model Consistent with the Current Term Structure*, International Scholarly Research Network.

Dominguez, M. & Lobato, I. (2003), ‘Testing the martingale difference hypothesis’, *Econometric Review* **22**, 371–377.

Durlauf, S. (1991), ‘Spectral based testing of the martingale hypothesis’, *Journal of Econometrics* **50**, 355–376.

Dybvig, P. (1988), ‘Increases in risk aversion and portfolio choice in a complete market’, *Yale University, Cowles Foundation Discussion Papers* **859**.

Dybvig, P., Ingersoll, J. & Ross, S. (1996), ‘Long forward and zero-coupon rates can never fall’, *J. Business* **60**, 1–25.

El Karoui, N., Loisel, S., Prigent, J.-L. & Vedani, J. (2015), Market inconsistencies of the market-consistent European life insurance economic valuations: pitfalls and practical solutions. working paper or preprint.
URL: *https://hal.archives-ouvertes.fr/hal-01242023*

Escanciano, J. (2007), ‘Model checks using residual marked empirical processes’, *Statistica Sinica* **17**, 115–138.

Escanciano, J. & Lobato, I. (2009*a*), ‘Au automatic portmanteau test for serial correlation’, *Journal of Econometrics* **151**, 140–149.

Escanciano, J. & Lobato, I. (2009*b*), ‘Testing the martingale hypothesis’, *Palgrave Hand-Book of Econometrics, eds. K. Patterson and T. C. Mills* **29**, 972–1003.

Escanciano, J. & Velasco, C. (2006*a*), 'Generalized spectral tests for the martingale difference hypothesis', *Journal of Econometrics* **134**, 151–185.

Escanciano, J. & Velasco, C. (2006*b*), 'Testing the martingale difference hypothesis using integrated regression functions', *Computational Statistics and Data analysis* **51**, 2278–2294.

Ferranto, M. (2015), Calibration And Simulation Of The Gaussian Two-Additive-Factor Interest Rate Model, PhD thesis, Universita' Degli Studi Di Padova.

Gilli, M. & Schumann, E. (2010), *A note on 'good starting values' in numerical optimization*, COMISEF Working Papers Series.

Glasserman, P. & Zhao, X. (2000), 'Arbitrage-free discretization of lognormal forward libor and swap rate models', *Finance and Stochastic* **4**, 35–68.

Grubisic, I. (2002), Interest Rate Theory The BGM Model, PhD thesis, Leiden University.

Hall, R. (1978), 'Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence', *Journal of Political Economy* **89**, 971–987.

Hart, J. (1997), *Nonparametric Smoothing and Lack-of-Fit Tests*, Springer Verlag, New-York.

Heath, D., Jarrow, D. & Morton, A. (1992*a*), 'Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation', *Econometrica* **60**, 77–105.

Heath, D., Jarrow, R. & Morton, A. (1990), 'Bond pricing and the term structure of interest rates : A discrete time approximation', *Journal of Financial and Quantitative Analysis* **25**, 419–440.

Heath, D., Jarrow, R. & Morton, A. (1992*b*), 'Bond pricing and the term structure of interest rates : A new methodology for contingent claims valuation', *Econometrica* **60**, 77–105.

Ho, T. & Lee, S. (1986), 'Term structure movements and pricing interest rate contingent claims', *Journal of Finance* **41**, 1011–1029.

Hogan, M. & Weintraub, K. (1993), *The Log-Normal Interest Rate Model and Eurodollar Futures*, Working Paper, Citibank, New York.

Hong, Y. (1999), 'Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach', *Journal of the American Statistical Association* **84**, 1201–1220.

Hong, Y. & Lee, Y. (2005), 'Generalized spectral tests for conditional mean models in time series with conditional heteroskedasticity of unknown form', *Review of Economic Studies* **43**, 449–541.

Hull, J. & White, A. (1990), 'Pricing interest-rate-derivative securities', *The Review of Financial Studies* **3**, 573–592.

Hull, J. & White, A. (1999), 'Forward rate volatilities, swap rate volatilities, and the implementation of the libor market model', *The Journal of Fixed Income* **10**.

Hyndman, R. & Athanasopoulos, G. (2014), *Forecasting: Principles and Practice*, University of Western Australia.

Jamshidian, F. (1988), *The One-Factor Gaussian Interest Rate Model: Theory and Implementation*, Working paper, Financial Strategies Group, Merrill Lynch Capital Markets.

Jamshidian, F. (1997), 'Libor and swap market models and measures', *Finance and Stochastics* **1**, 293–330.

Jamshidian, F. & Zhu, Y. (1997), 'Scenario simulation: Theory and methodology', *Finance and Stochastic* **1**, 43–67.

Joshi, M. & Rebonato, R. (2003), 'A stochastic-volatility, displaced-diffusion extension of the libor market model', *Quantitative Finance* **3**, 458–469.

Joshi, M. S. (2008), *The Concepts and Practice of Mathematical Finance*, Cambridge University Press.

Kim, J. (2006), 'Wild bootsraping variance ratio tests', *Economic Letters* **92**, 3–401.

Kim, J. (2009), 'Automatic variance ratio test under conditional heteroskedascity', *Finance research letters* **6**, 179–185.

Kloeden, P. & Platen, E. (1995), *Numerical Solution of stochastic differential equations, Second Edition*, Springer-Verlag.

Kuan, C. (2008), *Lecture on time series diagnostic tests*, Institute of Economics, Academia Sinica.

Kuan, C. & Lee, W. (2004), 'A new test of the martingale difference hypothesis', *Studies in Nonlinear Dynamics and Econometrics* **8**, 1–24.

Lesniewski, A. (2008), 6. LIBOR market model. Lecture NYU.
URL: *https://www.math.nyu.edu/∼alberts/spring07/Lecture6.pdf*

Ljung, G. & Box, G. (1978), 'On a measure of lack of fit in time series models', *Biometrika* **65**, 297–303.

Lo, A. & MacKinlay, A. (1989), 'The size and power of the variance ratio test in finite samples : a monte carlo investigation', *Journal of Econometrics* **40**, 203–238.

Lobato, Ignacio, Nankervis, J. & Savin, N. (2001), 'Testing for autocorrelation using a modified box-pierce q test', *International Economic Review* **42**, 187–205.

Lutz, M. (2011), Libor Market Models with Stochastic Volatility and CMS Spread Option Pricing, PhD thesis, Ulm University.

Miltersen, K., Sandmann, K. & Sondermann, D. (1997), 'Closed form solutions for term structure derivatives with log-normal interest rates', *The Journal of Finance* **52**, 409–430.

Morton, A. (1988), *Arbitrage and Martingales*, Technical Report 821, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University.

Park, J. & Whang, Y. (2005), 'Testing for the martingale hypothesis', *Studies in Nonlinear Dynamics and Econometrics* **9**, 1–30.

Phillips, P. & Jin, J. (2014), 'Testing the martingale hypothesis', *Cowles Foundation Paper* **1443**, 537–554.

Planchet, F. & Therond, P. (2005), 'Simulation de trajectoires de processus continus', *Belgian Actuarial Bulletin* **5**, 1–13.

Planchet, F., Therond, P. & Aymeric, K. (2009), *Scénarios économiques en assurance : Modélisation et simulation*, Economica.

Poterba, J. & Summers, L. (1987), 'Mean reversion in stock prices: Evidence and implications', *Journal of Financial Economics* **12**, 27–59.

Qureshi, T. (2013), Mémoire - Scénarios Economiques dans un contexte Solvabilite II, PhD thesis, Université Paris Dauphine.

Rebonato, R. (1998), *Interest Rate Option Models. Second Edition*, Wiley Series in Financial Engineering.

Rebonato, R. (1999), 'On the simultaneous calibration of multifactor lognormal interest rate models to black volatilities and to the correlation matrix', *The Journal of Computational Finance* **2**.

Rebonato, R. (2002), *Modern Pricing of Interest Rate Derivatives: The LIBOR Market Model and Beyond*, Princeton University Press.

Richard, Q. (1978), 'An arbitrage model of the term structure of interest rates', *Journal of Financial Economics* **6**, 33–57.

Sandmann, K. & Sondermann, D. (1994), *On the Stability of Log-Normal Interest Rate Models*, Discussion Paper No. B-263, Department of Statistics, University of Bonn.

Schatz, D. (2011), Robust Calibration of the Libor Market Model and Pricing of Derivative Products, PhD thesis, Universitat Ulm Institut fur Finanzmathematik.

Schoenmakers, J. & Coffey, B. (2000), 'Stable implied calibration of a multi-factor libor model via a semi-parametric correlation structure', *Weierstrab-Institut fur Angewandte Analysis und Stochastik* **611**.

Schrager, D. & Pelsser, A. (2005), 'Pricing swaptions and coupon bond options in affine term structure models', *Mathematical Finance* **16**, 673–694.

Singleton, K. & Umantsev, L. (2002), 'Pricing coupon-bond options and swaptions in affine term structure models', *Mathematical Finance* **16**, 673–694.

Tanimura, H. (2006), 'An efficient calibration method for the multifactor libor market model and its application to the japanese market', *International Journal of Theoretical and Applied Finance* **9**, 1123–1139.

Vasicek, O. (1977), 'An equilibrium characterization of the term structure', *Journal of Financial Economics* **5**, 177–188.

West, G. (2010), *Interest Rate Derivatives: Lecture notes*, Financial Modeling Agency.

Wright, J. (2000), 'Alternative variance-ratio tests using ranks and signs', *Journal of Business and Economic Statistics* **18**, 1–9.

Appendix A

Appendices

A.1 Structural schemes of the principal integrated models

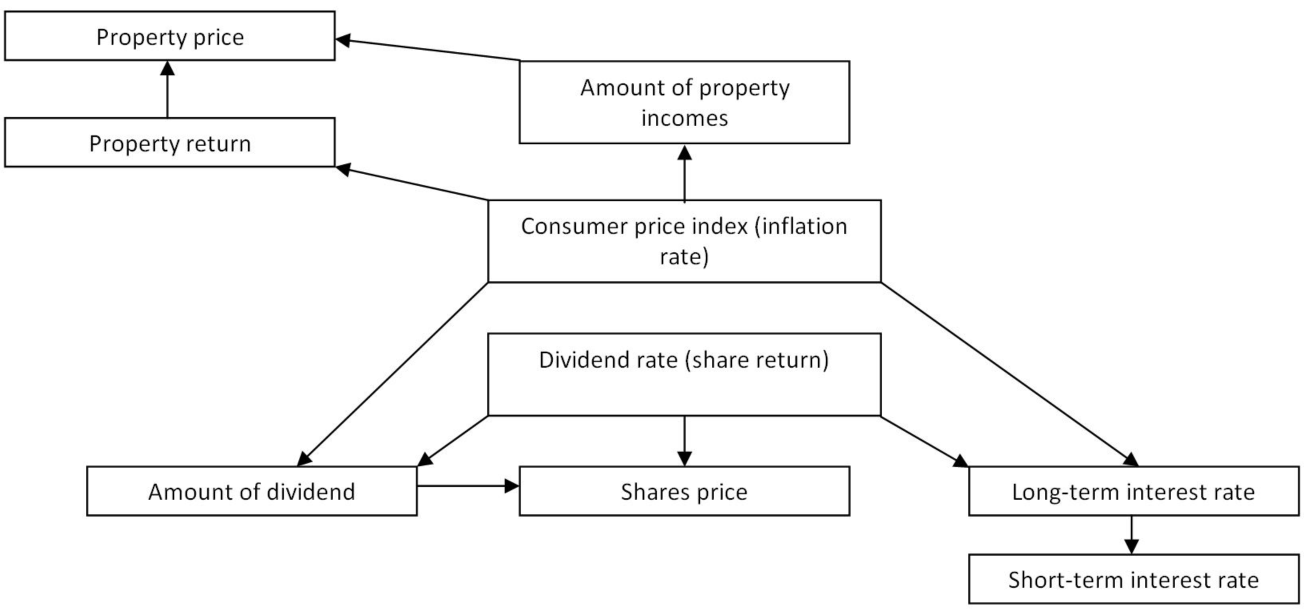


Figure A.1 – Structure of Wilkie model

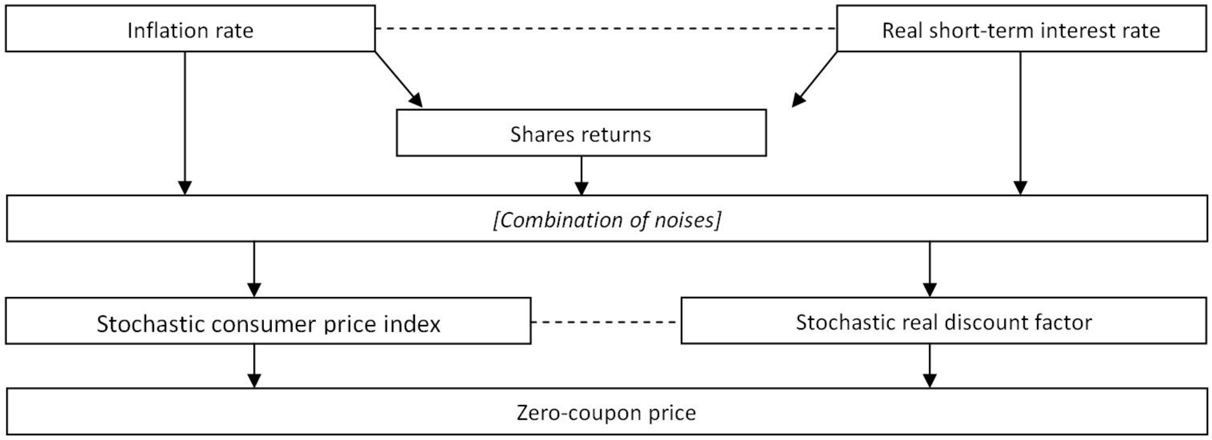


Figure A.2 – Structure of Brennan and Xia model

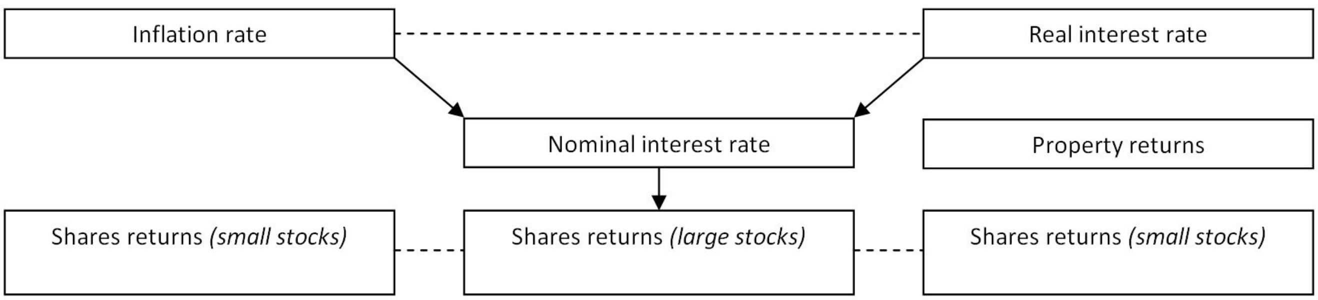


Figure A.3 – Structure of Ahlgrim model

A.2 Analytical integral of the “abcd”formula for instantaneous volatility of the LMM

If the instantaneous volatility of the forward rates in the LMM is expressed as follows:

$$\sigma_i(t) := [a + b\,(T_{i-1} - t)]\, e^{-c(T_{i-1}-t)} + d$$

then, if $s \leq T_{i-1}$:

$$\int_s^{T_{i-1}} \sigma_i(u)\sigma_i(u)du = I\,(T_{i-1}; T_{i-1}, T_{j-1}) - I\,(s; T_{i-1}, T_{j-1})$$

with:

$$I\,(t, T_n, T_m) = \frac{ad}{c}\left(e^{c(t-T_n)} + e^{c(t-T_m)}\right) + d^2 t - \frac{bd}{c^2}\left(e^{c(t-T_n)}\left[c\,(t-T_n)-1\right] + e^{c(t-T_m)}\left[c\,(t-T_m)-1\right]\right)$$
$$+\frac{e^{c(2t-T_n-T_m)}}{4c^3}\left(2a^2c^2 + 2abc\,(1 + c\,(T_n + T_m - 2t)) + b^2\left[1 + 2c^2\,(t-T_n)\,(t-T_m) + c\,(T_n + T_m - 2t)\right]\right)$$

A.3 Caplet volatilities estimated from the caps market with the second method presented in section 4.2.1

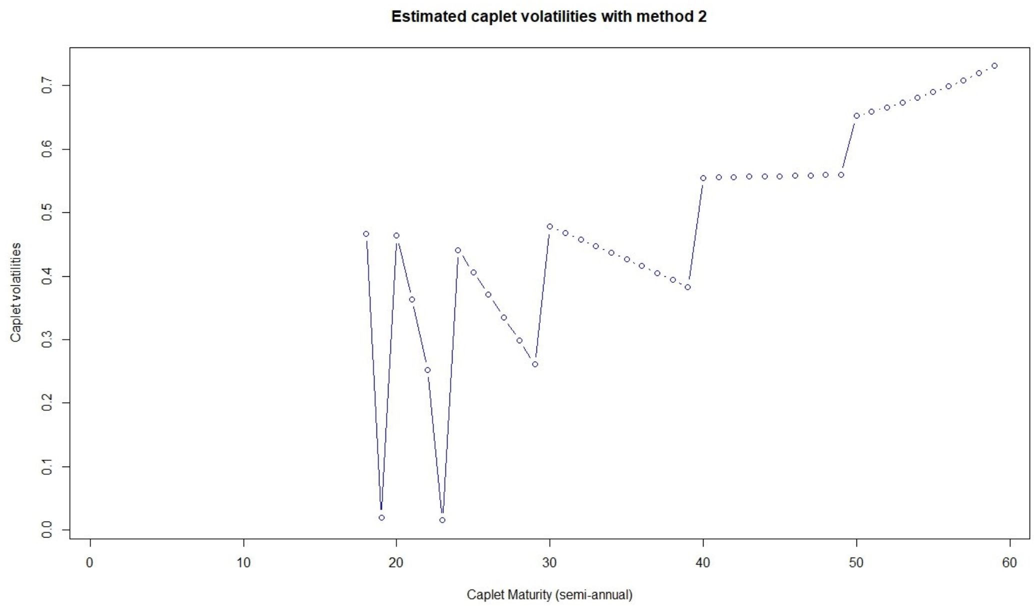


Figure A.4 – Estimated caplet volatilities from market caps volatilities with method 2

A.4 Evolution of the one-year rate $t \to R(t,1)$ using G2++ calibrated parameters

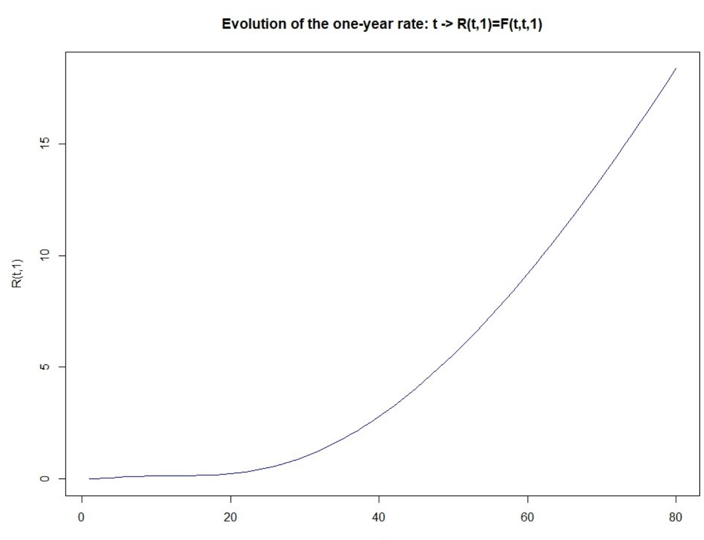


Figure A.5 – Evolution of the one-year rate $t \to R(t,1)$ using G2++ parameters calibrated to the caps market in table 4.16

A.5 Matrices of relative errors, resulting from the LMM calibration to older market data

Table A.1 – Relative errors between caps prices derived from the LMM jointly calibrated to older cap and market swaption volatilities and the market caps prices

Maturity	1	2	3	4	5	7	10	15	20
Relative error	-6.08%	-1.66%	-0.74%	-0.35%	-0.055%	0.118%	0.023%	-0.03%	0.008%

Mat/Tenor	1	2	3	4	5	6	7	8	9	10
1	-3,8%	5,2%	11,9%	16,1%	14,3%	9,8%	3,4%	-1,6%	-7,8%	-12,5%
2	1,6%	8,3%	12,6%	14,2%	12,2%	6,2%	0,5%	-4,7%	-9,4%	-14,4%
3	3,1%	9,5%	13,1%	14,7%	11,2%	6,7%	0,7%	-3,9%	-8,9%	-13,7%
4	7,0%	13,0%	16,0%	15,4%	12,5%	6,8%	2,4%	-2,7%	-7,5%	-12,1%
5	9,6%	15,9%	16,5%	16,5%	13,3%	9,2%	4,2%	-0,6%	-6,2%	-10,7%
7	9,9%	15,9%	16,4%	18,7%	16,5%	11,9%	7,3%	1,5%	-4,0%	-8,4%
10	11,2%	16,4%	19,0%	21,4%	18,3%	12,5%	8,1%	2,8%	-1,8%	-5,8%

Figure A.6 – Relative errors between the swaptions volatilities derived from the LMM jointly calibrated to older caps and swaptions market volatilities and the market swaptions volatilities

Mat/Tenor	1	2	3	4	5	6	7	8	9	10
1	1,3%	-1,4%	-0,9%	1,2%	0,8%	-0,5%	-2,7%	-3,5%	-5,7%	-6,6%
2	-1,1%	-1,5%	-0,3%	1,4%	1,6%	-0,7%	-2,5%	-4,0%	-5,0%	-6,7%
3	-4,0%	-2,2%	0,0%	2,6%	1,9%	0,8%	-1,6%	-2,7%	-4,5%	-5,9%
4	-2,8%	0,0%	2,4%	3,3%	3,0%	0,5%	-0,7%	-2,6%	-4,1%	-5,3%
5	-1,7%	2,0%	2,6%	3,9%	3,1%	1,8%	-0,2%	-1,8%	-4,1%	-5,0%
7	-2,2%	1,6%	2,1%	5,2%	4,9%	3,0%	1,5%	-0,7%	-2,6%	-3,1%
10	-0,7%	2,7%	5,2%	8,6%	8,2%	5,8%	4,9%	3,2%	1,7%	0,5%

Figure A.7 – Relative errors between the swaptions volatilities derived from the LMM calibrated to older swaptions market volatilities and the market swaptions volatilities

Mat/Tenor	1	2	3	4	5	6	7	8	9	10
1	1,4%	-1,2%	-0,7%	1,4%	1,0%	-0,3%	-2,5%	-3,3%	-5,5%	-6,4%
2	-0,9%	-1,3%	-0,1%	1,5%	1,8%	-0,5%	-2,3%	-3,8%	-4,8%	-6,4%
3	-3,8%	-2,0%	0,2%	2,8%	2,1%	1,0%	-1,4%	-2,5%	-4,2%	-5,6%
4	-2,6%	0,2%	2,6%	3,5%	3,2%	0,7%	-0,5%	-2,3%	-3,8%	-5,0%
5	-1,5%	2,1%	2,8%	4,1%	3,3%	2,0%	0,0%	-1,6%	-3,8%	-4,8%
7	-2,0%	1,8%	2,3%	5,3%	5,1%	3,2%	1,7%	-0,5%	-2,3%	-2,9%
10	-0,5%	2,9%	5,3%	8,8%	8,3%	5,9%	5,1%	3,4%	1,9%	0,6%

Figure A.8 – Relative errors between the swaptions prices derived from the LMM calibrated to older swaptions market prices and the market swaptions prices

List of Figures

List of Tables