FINANCIAL MATHEMATICS TEAM CHALLENGE

A collection of the five reports from the 2015 Financial Mathematics Team Challenge.

ACQuFRR

AFRICAN COLLABORATION FOR QUANTITATIVE FINANCE AND RISK RESEARCH

Preamble

The first Financial Mathematics Team Challenge (FMTC) took place in the June-July Winter Break at the University of Cape Town in 2014. Since it was the first time we had attempted it, we had little idea of what to expect. Fortunately, the event proved to be much more successful than we had hoped for, so we decided to forge on with a second edition in 2015. Our vision had always been to see if we could create an annual event.

The purpose of the FMTC is for South African postgraduate students in Financial and Insurance Mathematics to have the opportunity to focus (ostensibly without distraction) on a topical, industry-relevant research project, while simultaneously developing links with international students and academics in the field. An allied aim we have is to bring a variety of international researchers to South Africa to give them a glimpse of the dynamic environment that is developing at UCT in the African Institute of Financial Markets & Risk Management. One of the goals of the FMTC is for students to learn to work in diverse teams and to be exposed to a healthy dose of fair competition.

The Second Financial Mathematics Team Challenge was held from the 2nd to the 14th of July 2015. The challenge brought together five teams of Masters and PhD students from Switzerland, South Africa and the UK to pursue intensive research in Financial Mathematics. Each team worked on a distinct research problem during the twelve days. Professional and academic experts from France, Switzerland, South Africa, and the UK individually mentored the teams; fostering teamwork and providing guidance. Once again, the students applied themselves with incredible dedication and exemplary vigour.

This years research included topical projects on *expected shortfall in a multi-currency framework, the accuracy of the Rebonato formula for swap rates and swaption volatilities in single and multi-curve models, linear commodity models with unspanned stochastic volatility, Basel III Tier 2 capital pricing models, and multivariate risk measures for margin computations*. These were either proposed directly by our industry partners or chosen from areas of current relevance to the finance industry. In order to prepare the teams, guidance and preliminary reading was given to them a month before the meeting in Cape Town. During the final two days of the challenge, the teams presented their conclusions and solutions in extended seminar talks. The team whose research findings were adjudged to be the best was awarded a floating trophy. Each team wrote a report containing a critical analysis of their research problem and the results that they obtained. This volume contains these five reports, and will be available to future FMTC participants. It may also be of use and inspiration to Masters and PhD students in Financial and Insurance Mathematics. The second Financial Mathematics Team Challenge added to the success of the first, and we are already planning its third version.

David Taylor, *University of Cape Town*
Andrea Macrina, *University College London & University of Cape Town*

Contents

---

[1]Winning team of the second Financial Mathematics Team Challenge

A study on expected shortfall in a multi-currency environment

Team 1

ALEX BACKWELL, University of Cape Town
CHRISTOPHER ROBERTS, University of Cape Town
GAUKHAR SHAIMERDENOVA, University College London
FERGUS WEGENER, University of Cape Town

Supervisor:
CAMILO GARCIA TRILLOS, University College London and OpenGamma Ltd

Contents

1 Introduction

This paper is concerned with risk management, and addresses two themes within this topic. The first is the risk measure *expected shortfall*; the second is a financial environment with multiple currencies.

Expected shortfall is one of a number of prevalent risk measures. While it is not as prevalent as the pre-eminent value-at-risk measure, it has some theoretical advantages over its rival. These – described below in detail – have increased interest in expected shortfall from both academics and practitioners, and is therefore certainly good fodder for research in risk management.

Risk management in the presence of multiple currencies is also of increasing relevance. Risky entities have, generally speaking, become more globalised and international. Regulators need to supervise these entities, and are motivated to provide regulations in a more global and universal manner. We investigate risk measurement in a multiple currency environment by specifying models. These models' parameters are estimated from equity and exchange rate historical time-series. The parameters are then manipulated and the effects measured and interpreted. We show how the presence of multiple currencies has implications for risk measurement, and systematically study these implications.

The rest of the paper is organised as follows. Section 2 reviews some literature apposite to our focus. Here we make some key definitions and establish a foundation to continue. Section 3 pertains to modelling; we describe the models we use and their estimation and implementation details. Section 4 addresses the matter of *aggregation currency*. We show that choice of currency in which one measures expected shortfall makes a difference to the calculation, and we study the sign and size of this discrepancy. Section 5 introduces our second sub-problem; that of *capital composition*: following some papers in the literature review, we suppose that risk-free assets of different currencies can be added to risky positions to make them *acceptable* (a notion defined in Section 2), and examine the effects of this composition. Section 6 concludes.

2 Literature Review

This being a study on expected shortfall and its application in a multi-currency framework, it is vital that a clear description of this and related risk measures is given. The risk measure value-at-risk (VaR) has been prevalent in industry since the mid-1990s. It was recommended by both the Basel Committee on Banking Supervision in Europe and the Securities and Exchange Commission in the US for the first time in 1995 (Jorion, 1996). It has been maintained in spite of much criticism, largely because of its conceptual simplicity, as well as it being relatively straightforward to compute and apply (Yamai and Yoshiba, 2002). Acerbi and Tasche (2002)

define the VaR of a financial position X with

$$\text{VaR}_\alpha(X) = -\inf\{x \in \mathbb{R} : \mathbb{P}[X \le x] > \alpha\},$$

where α is the chosen level of significance. Note that X does not represent losses, as is occasionally the convention, but rather the final profit-and-loss random variable, and we are accordingly interested in its left-hand tail. The minus sign on the definition allows the VaR figure to be in terms of losses, and a larger VaR to represent larger risk. Note also that the left tail would be found by setting, say, $\alpha = 0.05$, rather than the occasionally used $\alpha = 0.95$.

Expected shortfall, the risk measure of focus in this paper, is given its technical definition in terms of VaR by setting

$$\mathbb{ES}_\alpha[X] = -\mathbb{E}[X | X < -\text{VaR}_\alpha(X)].$$

The above definition is useful as it allows easy comparison with VaR. It highlights the fundamental difference between the two measures – VaR gives the minimum loss that is expected given that the worst quantile occurs, whereas expected shortfall gives an average of the losses that would occur over and above this level.

It is obvious that summarising the risk of a position or risk into a single number is a very useful thing to do, if the summary is an intelligent and appropriate one. This would greatly aid risk managers and regulators, for example, who require heuristics and rules to apply to complex and dynamic environments. However, as is well known, it is impossible to fully summarise the risk of a position with a single measure, and this leads into a discussion of the shortcomings of each measure. The above definitions show the first of the basic problems with VaR – given that large losses do occur, the extent or size of these losses are not quantified. In the light of recent financial crises, this problem holds a lot of weight. The second prominent issue with VaR is that it is not *sub-additive* (Artzner et al., 1997, 1999). This problem can be expressed mathematically as follows, in that the following does not necessarily hold true:

$$\text{VaR}_\alpha(X + Y) \le \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y).$$

The classic example of VaR failing sub-additivity involves two independent risks that yield losses only 4% of the time, and zero otherwise. The $\text{VaR}_{0.05}$ of one of the these individually is clearly zero, and, while the distribution of the combined position is not immediately obvious, it is intuitively clear that a loss, and therefore a positive VaR, will occur at the 5%-quantile. In words, according to VaR, the risk of a sum of positions is not necessarily less than the sum of the individual risks, which contradicts the idea of diversification. As a result, the use of VaR in risk management may not encourage diversification of risk, and in some cases may motivate against it (Acerbi and Tasche, 2002) (Embrechts et al., 2015). Generally speaking, expected shortfall does account for the magnitude of tail-risk, and, in

particular, obeys sub-additivity. In fact, it meets the more stringent condition of *coherence*. The idea of a risk measure being coherent was introduced by Artzner et al. (1997, 1999). The four properties that a risk measure must satisfy in order for it to be coherent are presented below. In what follows, the mapping $\rho : V \to \mathbb{R}$ is a risk measure, with V being a space of random variables representing financial positions, with $X, Y \in V$:

1. Sub-additivity: $X, Y, X + Y \in V \Rightarrow \rho(X + Y) \leq \rho(X) + \rho(Y)$

2. Positive Homogeneity: $X \in V, a > 0, aX \in V \Rightarrow \rho(aX) = a\rho(X)$

3. Monotonicity: $X \in V, X \geq 0 \Rightarrow \rho(X) \leq 0$

4. Translation Invariance: $X \in V, a \in \mathbb{R} \Rightarrow \rho(X + a) = \rho(X) - a$.

Expected shortfall satisfies these properties, including sub-additivity, and is thus coherent. A proof of the sub-additivity of expected shortfall is available in the Appendix of Acerbi and Tasche (2002).

In addition to the positive/negative and α versus $(1 - \alpha)$ convention choices, there are definitional issues surrounding expected shortfall. This point is addressed thoroughly by Acerbi and Tasche (2002), who delineate a number of related notions such as conditional VaR, worst conditional expectation, and tail conditional expectation. These largely depend on whether, and in what combination, the inequalities in our above definitions are strict or not. Acerbi and Tasche (2002) develop a robust and general definition of expected shortfall, given by

$$\mathbb{ES}_\alpha[X] = -\alpha^{-1}(\mathbb{E}[X\mathbb{I}_{\{X \leq x_\alpha\}}] + x_\alpha(\alpha - \mathbb{P}[X \leq x_\alpha])),$$

where x_α is the α-quantile of X (in fact, the lower quantile, which they carefully define). This definitional issue is important when there are discontinuities in the underlying loss distribution, and the robust definition is necessary to guarantee coherence in general. In this paper we consider continuous distributions, though, and do not require great sensitivity to this issue.

Also prominent in the literature, Föllmer and Schied (2002) extend the idea of risk measure coherence to risk measure convexity. The main idea here is that the risk of a position may change in a non-linear fashion as the size of the position changes. Föllmer and Schied (2002) present the situation where the conditions of sub-additivity and positive homogeneity are relaxed to a weaker property of convexity, defined, for $\lambda \in [0, 1]$, by

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y).$$

Convexity is related to sub-additivity and is readily interpreted in terms of diversification, but the positive homogeneity property is relaxed. This might be appropriate and necessary when explicitly modelling liquidity risk, where positions are not

assumed to scale in a simplistic way. VaR is not a convex measure of risk, whereas expected shortfall, being coherent, is.

The idea of 'acceptance sets' - a set of acceptable financial positions - is central to the mathematical literature on risk measures. We draw here from the seminal paper of Artzner et al. (1999). For a particular circumstance, we can imagine a subset of all possible random variables (representing positions) that are considered acceptable (because, perhaps, a business has policy that defines acceptability, or alternatively, a regulator might simply define what is acceptable), and we call this set $\mathcal{A}$. If one starts with such a set in mind, it can induce a risk measure ρ by setting

$$\rho_{\mathcal{A}}(X) = \inf\{m|m + X \in \mathcal{A}\}.$$

The interpretation here is that m is a a deterministic amount needed to shift the risk into the acceptance set. Certain conditions on $\mathcal{A}$ ensure that the infimum exists. Conversely, one can induce an acceptance set from a particular risk measure by defining

$$\mathcal{A}_{\rho} = \{X|\rho(X) \leq 0\}.$$

This interpretation – acceptability being synonymous with the risk measure not exceeding zero – is essential to Section 5 and will be expanded there. Because of translation invariance (with coherent risk measures in mind), making a position acceptable can simply involve adding an amount of a risk-free asset (a deterministic amount) until the risk measure is equal to zero.

The correspondence between risk measure and acceptance set is very important in the mathematical literature, as authors such as Artzner et al. (1999) will make assumptions or prove results on one side of the correspondence and explore the implications on the other. We do not rely heavily on the very formal mathematical framework, and so the above summary is sufficient for our more practical purposes.

If a particular zero-coupon bond is assumed to be truly risk-free, the effect of including this pay-off in a position is simply a deterministic shift of the distribution. But if there is more than one currency in a model, deterministic amounts can be paid out in each currency, and not be deterministic when denominated by another currency. The question then arises as to whether adding risk-free assets in several currencies might lead to greater efficiency in capital management. Another question that arises in this context is the effect of measuring the risk in terms of different currencies (i.e., allowing different currencies to denominate the position). It turns out that expected shortfall varies depending on the currency used to denominate the position (the aggregation currency), and, in fact, Artzner et al. (2009) show that this incompatibility exists for all coherent risk measures. VaR, on the other hand, has been shown to be currency-invariant, in that the acceptability of a financial position does not depend on the aggregation currency (Koch-Medina and Loubet, 2014).

While we are not aware of literature studying the composition of risk-free assets of different currencies, Koch-Medina and Loubet (2014) have studied the issue of aggregation currency. They present a one-period, dual-currency theoretical examination of the currency or exchange rate risk that arises in the situation where a financial position is made up of foreign and domestic assets. They question the contribution of currency risk to the total risk of a portfolio. In order to do this they separate currency risk into translation and structural risk, where translation risk arises purely from the need to translate assets or liabilities of a position into one currency for risk aggregation, and structural risk is the general uncertainty of where the exchange rate will lie. They present a theoretical framework for capturing structural risk. Their highly theoretical study will be well accompanied by our much more practically-oriented one.

Before developing a framework to study these two issues of aggregation currency and capital composition, we end the section with some general remarks on expected shortfall and risk measurement, based on our review of the literature.

Research regarding the robustness of risk measures is becoming increasingly well-developed. Soon after its introduction as an industry standard, Jorion (1996) heeded the risks associated with VaR estimates and proposed a methodology to analyse estimation error and improve accuracy in the estimation of VaR. As the suitability of VaR has come under question, the method of reaching a VaR figure has been addressed in detail. Embrechts et al. (2013) presented a thorough examination of theoretical bounds for the estimation of VaR when the dependence structure between various sources of risk are unknown. This research has since been extended to other risk measures, most notably expected shortfall (Embrechts et al., 2015). The idea of robustness for a risk measure is crucial for its usefulness in regulation and business practice, while aggregation-robustness specifically relates to a risk measure's insensitivity to the dependence structure of the underlying risk factors (Embrechts et al., 2015). Under their own definition of aggregation-robustness, expected shortfall was found to display a narrower spread of uncertainty than VaR in the face of model uncertainty. However, it has been noted that under different definitions of robustness, there are contrasting views in the literature as to which of VaR and expected shortfall are more robust (Embrechts et al., 2014).

The Basel Committee on Banking Supervision (BCBS, 2012) specified a move towards expected shortfall as the risk measure used in practice. The operational challenges of the move were acknowledged but believed to be outweighed by the need to better account for risk of extreme negative cases (tail-risk). In an academic response to the change in regulatory recommendations (Embrechts et al., 2014), the unfavourability of the back-testing process for expected shortfall compared with that for VaR is cited as a crucial challenge. See Acerbi and Székely (2014) for a discussion on back-testing expected shortfall in contrast to VaR. Due to this practical disadvantage, as well as the currency (in)variance that is addressed directly in this paper, expected shortfall's superiority over VaR as a risk measure remains unclear,

despite its theoretical advantages. The practical implementation of expected short-fall as a risk measure is addressed in detail in this research, and hence what follows will provide insight into the ongoing evaluation of its performance.

3 Modelling

We focus on a parametric/distributional approach to estimating risk measures. See Coleman and Litterman (2012) for a thorough treatise on the different methods.

We therefore need to introduce the models we use in our estimates of expected shortfall, as defined in the previous section. After outlining the models and providing their discretisation schemes, we will address their estimation, and finally give a few details about the expected shortfall computation.

3.1 The models

The Constant Elasticity Variance (CEV) model is a stochastic volatility diffusion model, which was introduced by Cox and Ross (1976) and is characterised by the following SDE:

$$dS_t = \mu S_t dt + \sigma S_t^{\alpha} dW_t,$$

where μ, σ, α are constant parameters, and $\{W_t, t \geq 0\}$ is a standard Wiener process.

The model presents the following simple relationship between local volatility and stock prices:

$$\sigma(S, t) = \sigma S^{\alpha-1}.$$

Here, $(\alpha - 1)$ is the so-called elasticity of return variance with respect to price. In the case of $0 < \alpha < 1$, there is an inverse relationship between volatility and price. This is the *leverage effect*; the tendency for volatility to increase when asset price falls. Conversely, when $\alpha > 1$, we will observe an inverse leverage effect, where volatility of a stock rises as its price rises. Because of this, and also because of negative bias in volatility skewness, the $\alpha > 1$ case is not given much interest.

Several empirical investigations approved the fact that the variance of stock returns and stock prices have a strong inverse association (Schroder, 1989). In particular, Beckers (1980) and Christie (1982) studied the CEV option pricing model, where they found that variance elasticities are generally negative and concluded that CEV model could describe market price behaviour much better than the Black-Scholes model (introduced below).

It is a well-known fact in practice that the probability density function is usually defined by higher kurtosis (known as being leptokurtic) and by a heavy decay of tails (Cont, 2001). Therefore, the case $0 < \alpha < 1$ is considered more realistic, since it can produce a fatter left tail.

An Euler method can easily be specified for the CEV model. It consists of the following algorithm: $\hat{S}_{t_0} = S_{t_0}$ for $i = 0, ..., n - 1$:

$$\hat{S}_{t_{i+1}} = \hat{S}_{t_i} + \mu \hat{S}_{t_i} \Delta t + \sigma \hat{S}_{t_i}^{\alpha} \varepsilon_i \sqrt{\Delta t},$$

where ε_i are i.i.d. standard Gaussian random numbers, and $\Delta t = t_{i+1} - t_i$. However, in our problem we will use the Student t-distribution (instead of the Gaussian), which is an example of a distribution with fat-tails, where the parameter degrees of freedom allows control over heaviness of the tail (at the limit of infinite degrees of freedom, the distribution converges to the standard Gaussian). We choose this in order to capture the stylised fact of heavy-tail (Cont, 2001). Then our discretised scheme is simply adjusted thus:

$$\hat{S}_{t_{i+1}} = \hat{S}_{t_i} + \mu \hat{S}_{t_i} \Delta t + \sigma \hat{S}_{t_i}^{\alpha} T_i \sqrt{\Delta t},$$

where T_i are t-distributed random variables.

In summary, it can be said that our CEV model captures the stylised facts of the leverage effect and heavy-tails. Note that we are not concerned with some stylised facts – Cont (2001) shows that there tends to be gain/loss asymmetries in return data, but we are not concerned about the gain side of the distribution, and can lower the importance we place on this stylised fact.

The Black-Scholes-Merton (BSM) model is a special case of the CEV model with $\alpha = 1$, that is, when stock prices follow Geometric Brownian Motion (GBM):

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

and the local volatility is constant:

$$\sigma(S, t) = \sigma.$$

The BSM model was developed originally by Black and Scholes (1973) and Merton (1976). In the next sections, the BSM model (also applied to the exchange rate) is referred to as our *basic model*. The CEV model, and the mean-reverting log-normal and Heston models (yet to be introduced) are more complex models that will be used thereafter in an aim to more realistically capture the stock- and exchange rate-dynamics.

In what follows, a mean-reverting log-normal diffusion model, which evolves according to the following SDE, is presented:

$$dX_t = a(b - X_t)dt + \sigma X_t dW_t,$$

where $b > 0$ is a long run mean, $a > 0$ is a reversion speed, σ is the volatility coefficient and $\{W_t\}_{t \geq 0}$ is a standard Wiener process.

The GARCH(1,1) model, which is the discrete version of the above model, is popular for derivatives pricing and widely used in general modelling of the financial market (Zhao, 2009). In particular, it was investigated and suggested for modelling the foreign exchange market (Erdemlioglu et al., 2013).

The Euler-Maruyama discretisation scheme for the mean-reverting log-normal model is given by the following:

$$\hat{X}_{t_{i+1}} = a(b - \hat{X}_{t_i})\Delta t + \sigma \hat{X}_{t_i} \varepsilon_i \sqrt{\Delta t},$$

where ε_i are i.i.d. Gaussian random numbers.

The Heston model is a stochastic volatility model describing a joint process between stock price and volatility. The general Heston model assumes that the asset price is described by the following SDE:

$$dS_t = \mu S_t dt + \sqrt{\nu_t} S_t dW_t^S,$$

where μ is the drift of the stock process and instantaneous squared volatility ν_t is defined by the CIR process:

$$d\nu_t = k(\theta - \nu_t)dt + \sigma \sqrt{\nu_t} dW_t^\nu,$$

where dW_t^S, dW_t^ν come from correlated Brownian Motions with correlation coefficient $\rho \in [-1, 1]$; $k > 0$, $\theta > 0$ is a mean reversion rate and level respectively; $\sigma > 0$ is a volatility-of-volatility (Heston, 1993).

If the parameters satisfy the following Feller condition, then the mean-reverting square-root dynamics for the volatility will remain strictly positive:

$$2k\theta < \sigma^2.$$

The Heston model captures the stylised facts of volatility clustering and the leverage effect. Moreover, volatility is mean-reverting.

The Euler discretisation scheme for the Heston model is defined by the following algorithm:

$$\hat{S}_{t_{i+1}} = \hat{S}_{t_i} + \mu \hat{S}_{t_i} \Delta t + \sqrt{\hat{\nu}_{t_i}} \hat{S}_{t_i} \varepsilon_i^s \sqrt{\Delta t}$$

$$\hat{\nu}_{t_{i+1}} = \hat{\nu}_{t_i} + k(\theta - \hat{\nu}_{t_i})\Delta t + \sigma \sqrt{\hat{\nu}_{t_i}} \varepsilon_i^\nu \sqrt{\Delta t}$$

where ε_i^S is a standard normal random variable that has correlation ρ with ε_i^ν. Even if the Feller condition is met, one may need to adopt a *truncation* or *reflection* scheme in the discrete implementation to preclude negative values going into the square-root volatility.

3.2 Estimation

The models require parameter estimates. Historical estimation of parameters is an extremely large topic, and there are many methods from which one can chose.

Firstly note that historical estimation is quite different from calibration to prevailing market data. The former, which is what we attempt here, takes averages, in a certain sense, over a period of history, which is assumed to have some stationary properties. The real-world measure is necessarily involved. Calibration to prevailing prices does not involve a period of history, but instead determines the value of risk-neutral parameters assuming the model is correct. Some of the real-world and risk-neutral parameters are common, so the two approaches can sometimes be combined (the circumstances of the modelling entity will dictate whether this approach is taken). Here, however, we focus on the historical approach, and do not have any derivative price information as an input.

We estimate the BSM models in the standard and straightforward way; we take daily log returns of the relevant series and estimate their mean and volatility, and then convert these to parameter estimates using the classic solution to the Geometric Brownian Motion stochastic differential equation. Correlations are straightforward to estimate; the standard formula is applied to the daily log returns. We provide the estimates in the next sub-section.

The CEV and Heston models are more difficult to estimate. Both models are used primarily for derivative pricing and hedging, and therefore the literature addresses their calibration more heavily. The primary method to estimate the Heston model, in the absence of derivative price information, is Markov Chain Monte Carlo (Cape et al., 2015). Implementing such an approach has proven to be beyond the scope of this paper, and we were unable to develop an alternative method. We will at least mention some of the experiments we would like to have performed using the Heston model, had we estimated the model suitably.

The CEV model, having fewer parameters, can be estimated with a simple method-of-moments approach. As described above, including the degrees of freedom of the t-distributed increments, there are four parameters to estimate. We can therefore equate the first four moments by manipulating the four parameters. The moments are not known in closed-form, so they need to be estimated by Monte Carlo. The sampling errors in each computation that a numerical solver employs are a challenge to any optimisation algorithm, and it requires many different initialisations to ensure avoidance of local minima. As seen below, we end up achieving, if not the true global minimum, a very close fit to the first four moments. We used a sum of squared percentage differences as the objective function to be minimised. The mean-reverting currency model only has three parameters, so a close fit to four moments is not possible, but using the same objective function, and a number of different initialisations, a reasonable fit is attained (these are shown below).

11

3.3 Computation

There are straightforward numerical estimates for both VaR (simply the standard quantile estimator), as well as, expected shortfall (which then simply involves averaging the sample points in the empirical tail). See Nadarajah et al. (2014) for an outline of some of the numerical details here in a variety of parametric environments.

In the case of a BSM-modelled stock, the VaR and expected shortfall are known in closed-form. We can test our Monte Carlo coding algorithm by comparing our estimates with their target. Both measures appear to converge at roughly the same rate. Their two different natures (expected shortfall has some averaging, while VaR is simply based on order statistics) make this difficult to discern a priori. Note also that errors from the discretisation scheme are also present, because the closed-form values are of course based on the full continuous model.
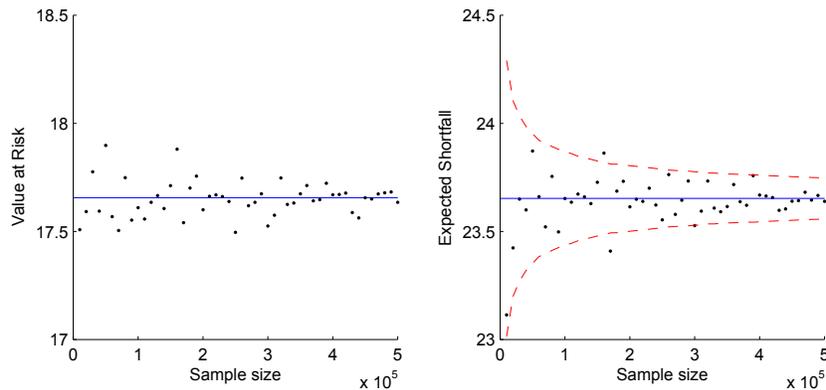
Figure 1: Monte Carlo convergence test

Besides applying antithetic samples (all of the random innovation terms are symmetric around zero and this is therefore easy to do), no other numerical techniques are necessary to ensure that the Monte Carlo is fast and accurate enough to make the estimation feasible.

Our fundamental data includes three time-series, with daily entries over a period of just over four years. Two of these are stock prices, which we refer to as LDNS1 and NYS1, which are denominated in different currencies. We refer to the currency of the second as domestic, and use this as our primary denomination basis. The third series, ER, is the exchange rate: the amount of domestic currency needed to purchase one unit of foreign. We treat the series in an abstract fashion and do not consider the underlying economics in an explicit way.

Our estimated parameters are displayed in the Tables 1, 2 and 3 below.

Table 4 shows the moments of the CEV estimation procedure. With four degrees of freedom, we achieve a very close fit between the empirical first four raw

12

Table 1: Basic (BSM) model parameter estimates

	μ	σ	$\rho(\text{LDN1}, \cdot)$	$\rho(\text{NY1}, \cdot)$	$\rho(\text{ER}, \cdot)$
LDN1	0.1424	0.1676	1	0.0422	0.0268
NY1	0.1299	0.1865	0.0422	1	-0.0125
ER	0.0084	0.0743	0.0268	-0.0125	1

Table 2: CEV model parameter estimates

	μ	σ	α	ν
LDN1	0.1351	0.1398	0.9101	3.7988
NY1	0.1362	0.0899	0.9410	6.5102

moments and the (Monte-Carlo-estimated) ones implied by the model. The fit for the currency model is of course not as close, but we show something of a reasonable fit.

We are almost in a position to study our two sub-problems. In the next section, we refer to random variables X, E and Y. X is the random variable relating to the profit-and-loss of an initial investment of 50 units in the NY1 stock, which we call domestic, for a period of one year (the initial investment is subtracted so that we are dealing with profit-and-loss rather than a gross final position). E is the distribution of final exchange rates - the cost of a unit of foreign currency, so that X and E have the same basis. Y represents profit-and-loss of the foreign stock, denominated in foreign currency. The initial investment is also 50 units *in domestic currency terms*, so that the total position costs 100 domestic units to enter, and the random variable of the final position, $X + YE$, represents the profit-and-loss on this investment in domestic terms.

Our realisations of X, E and Y are generated by 50 Euler steps over the fixed, one-year horizon. This relatively small number of steps allows the computation to be feasible, and is not a concern because we are not especially interested in close convergence to the continuous-time models (and this exact scheme was used in the estimation procedure).

4 Aggregation Currency

In this section, we study the issue of choice of aggregation (or numeraire) currency when calculating expected shortfall. As mentioned in the introduction, expected shortfall can depend on this choice; in other words, expected shortfall calculated in the domestic currency, thus

$$\mathbf{ES}_\alpha = \mathbb{ES}_\alpha[X + YE], \tag{1}$$

13

Table 3: Full ER model parameter estimates

	a	b	σ
ER	2.5273	1.5386	0.1224

Table 4: Matched moments

	LDN1		NY1		ER	
	Empirical	Estimated	Empirical	Estimated	Empirical	Estimated
Moment 1	0.1309	0.1301	0.1342	0.1323	0.8784	1.0102
Moment 2	0.0264	0.0266	0.0239	0.0249	2.8729	3.0094
Moment 3	0.0060	0.0059	0.0050	0.0052	0.0335	0.0170
Moment 4	0.0016	0.0016	0.0012	0.0012	0.0241	0.0272

is not necessarily equal to the expected shortfall *aggregated* in the foreign currency, but still expressed in domestic terms, namely

$$\mathbf{ES}^*_\alpha = \mathbb{ES}_\alpha[X/E + Y].E_0, \tag{2}$$

where we have followed the conventions of the previous section. In particular, our risk measures are calculated at the one-year horizon; that is, only the variation at that point in time, and nothing more frequent, is considered. Note that $\mathbf{ES}^*_\alpha$ involves the conversion back to the domestic currency (multiplication by E_0), so that the two metrics are comparable in principle (in the next section, expected shortfall will explicitly be considered as the amount of capital to be added to the position now – hence the E_0 – to make it acceptable). They turn out not to coincide in general however, and we study this discrepancy, firstly in the context of our elementary BSM-type models, and then in the context of our full modelling framework. Specifically, we attempt to identify and isolate particular effects driving the aggregation discrepancy.

4.1 Basic models and principles

We begin by presenting, in Table 5, expected shortfall calculations using our BSM models, at a few levels of α.

Table 5: BSM calculations

α	0.01	0.02	0.05	0.1
$\mathbf{ES}_\alpha$	20.105	17.479	13.366	9.799
$\mathbf{ES}^*_\alpha$	20.296	17.755	13.615	10.064

Notice firstly that the figures are of the order of magnitude one would roughly expect, considering an initial investment of 100 – given that we are in the worst percentile (α=0.01) of cases, say, we *expect to lose* 20 units. The low correlation between the different risks ensures that this is not especially high. Secondly, the discrepancy between the two aggregation currencies is not negligibly small. This phenomenon, induced by straightforward and standard modelling, is of different significance to different kinds of participants, and we elaborate on these perspectives towards the end of the section. We first attempt to understand the principles at work in this discrepancy.

Assuming the exchange rate to be constant turns out to be a natural starting point for the investigation, because expected shortfall does not depend on the aggregation currency in this case. Constant currency can be recovered from our model by setting μ_E and σ_E to zero, and indeed the aggregation discrepancy can be seen to vanish in Table 6 (there is no Monte Carlo sampling variation visible as we have used common random numbers). Here, $\mathbf{ES}^*_\alpha$ involves an initial currency conversion and then the exact opposite conversion at the end of the period, yielding the exact same result as $\mathbf{ES}_\alpha$.

Table 6: Constant currency

α	0.01	0.02	0.05	0.1
$\mathbf{ES}_\alpha$	19.175	16.347	12.579	9.080
$\mathbf{ES}^*_\alpha$	19.175	16.347	12.579	9.080

If, however, the exchange kept deterministic ($\sigma_E = 0$) but allowed to vary with time ($\mu_E \neq 0$), an aggregation discrepancy is introduced. Our BSM-type currency model involves a simple drift μ_E, which we manipulate in Table 7.

Table 7: Drifting currency

μ_E	-0.1	0	0.0084	0.1
$\mathbf{ES}_{0.05}$	17.819	13.715	13.461	9.297
$\mathbf{ES}^*_{0.05}$	8.985	13.287	13.693	17.264
$\mathbf{ES}_{0.05} - \mathbf{ES}^*_{0.05}$	8.834	0.429	-0.231	-7.966

Notice that our estimated value, $\mu_E = 0.0084$, recovers our estimate for $\mathbf{ES}_{0.05}$ in Table 5 above. The interpretation of this *trend effect* is straightforward; if you (and your model calculating expected shortfall) expect a currency to strengthen, as in the right-hand side of Table 7, you are better off measuring your risks in terms of this currency – $\mathbf{ES}_{0.05}$ decreases as we move right in the table, reflecting less risk.

To the end of analysing the full expressions (1) and (2), we first look at currency and position interactions (that is, the random variables X/E and YE and their

expected shortfalls), and then turn to the addition of the two positions. To make the comparisons fair, we set $E_0 = 1$ (which eases the exposition) and make the currency process driftless (to remove the trend effect). Table 8 compares $\mathbb{ES}_\alpha[XE]$ and $\mathbb{ES}_\alpha[X/E]$ for various levels of correlations between the Brownian motions driving the stock and the currency.

Table 8: Currency correlation

$\rho_{X,E}$	-1	-0.5	0	0.5	1
$\mathbb{ES}_{0.05}[X]$	23.628	23.567	23.717	23.744	23.646
$\mathbb{ES}_{0.05}[X/E]$	34.072	30.301	25.553	19.696	11.124
$\mathbb{ES}_{0.05}[XE]$	11.648	20.147	26.121	30.656	34.311

When X and E are well-correlated, the quotient X/E tends to have less variation (as movements in the numerator tend to be matched by movements in the denominator) and thus a lower expected shortfall. The intuitive analogue pattern emerges when comparing X and XE. This *currency-correlation effect* is one component of the aggregation discrepancy. While there are other effects, a positive correlation between X and E will decrease the risk in X/E and therefore decrease **ES**$_\alpha^*$.

Table 9 shows the effect of, all other things equal, converting the same position to and from a particular currency (when the rate is unit-initialised and driftless to make the comparison fair).

Table 9: Convexity

σ_E	0	0.1	0.2	0.3
$\mathbb{ES}_{0.05}[XE]$	23.636	27.777	37.289	47.951
$\mathbb{ES}_{0.05}[X/E]$	23.609	27.056	34.725	43.037
Difference	0.027	0.721	2.564	4.914

Although one might naïvely think that multiplying and dividing by exchange rates that center around one would have the same effect, this ignores the well-known Jensen's inequality. The position part of the hyperbolic function is positively convex, which can be seen to increase expectation and thereby lower the measured risk. The effect is more pronounced for high variability in the currency.

Before pulling all the effects together, we examine the *diversification effect* that arises when measuring a joint position; in other words, we comment on how $\mathbb{ES}_\alpha[X+Y]$ relates to $\mathbb{ES}_\alpha[X] + \mathbb{ES}_\alpha[Y]$ (taking currency out of the picture, in an attempt to isolate effects). As outlined in the introduction, we require our risk measure to respect sub-additivity, which is to say that the risk of the positions must be less than or equal to the sum of the individual risks. Equality arises when the risks offer no

diversification at all, for instance if you add two identical risks:

$$\mathbb{ES}_\alpha[X + X] = \mathbb{ES}_\alpha[2X] \tag{3}$$

$$= 2\mathbb{ES}_\alpha[X] \tag{4}$$

$$= \mathbb{ES}_\alpha[X] + \mathbb{ES}_\alpha[X], \tag{5}$$

where the positive homogeneity property is used to move from (3) to (4). Using expected shortfall as a risk measure, and assuming a model and calculation framework, the diversification between two risks can be measured by how much lower the combined risk is compared to the sum of two individual ones. This is easily demonstrated with the BSM models – Table 10 looks at this difference for many levels of correlation.

Table 10: Diversification effect

$\rho_{X,Y}$	-1	-0.5	0	0.5	1
$\mathbb{ES}_{0.05}[X]$	23.697	23.588	23.619	23.644	23.695
$\mathbb{ES}_{0.05}[Y]$	23.630	23.625	23.684	23.615	23.673
$\mathbb{ES}_{0.05}[X + Y]$	-19.079	13.328	27.940	38.745	47.123
$\mathbb{ES}_{0.05}[X + Y]$ $-\mathbb{ES}_{0.05}[X] - \mathbb{ES}_{0.05}[Y]$	66.405	33.885	19.363	8.513	0.245

While the dependence structure between the two stocks is very easily controlled in this context (it is measured by the single coefficient $\rho_{X,Y}$), there are modelling and practical situations where things are not so clear. For instance, we wanted to explore two stocks from the Heston model, where, say, the two volatility processes are well-correlated. This measure of diversification – the amount of expected shortfall reduction in the composition – is an interesting one. It is dependent on the risk measure used, and the underlying modelling, but is a potentially informative way of summarising the relationship between the two risks.

As an example, let us again consider our $\mathbf{ES}^*_{0.05}$ calculation using the BSM-type model in Table 5. With the heuristic effects above, we can reconcile the result that $\mathbf{ES}^*_{0.05} > \mathbf{ES}_{0.05}$. Firstly, μ_E is positive; exposure to this risk is, all other things equal, beneficial to your position and will decrease the measure (in this case $\mathbf{ES}_{0.05}$ is clearly more exposed to E). Reinforcing this, the one estimated negative correlation induces a currency-correlation effect (because $\rho_{XE} < 0$) and a diversification effect that both increase the discrepancy of $\mathbf{ES}^*_{0.05}$ over $\mathbf{ES}_{0.05}$.

We might then conceive another set of parameters that would cause the opposite sign of the discrepancy, that is, $\mathbf{ES}^*_{0.05} < \mathbf{ES}_{0.05}$.

The first adjustment we make is to remove currency drift so as to negate the trend effect; this results in a small reversal of the ordering of the two expected shortfalls (Table 11). Instead of this, we might reverse the sign of the correlation between the stock and the exchange rate, as well as amplify it. This results in a

Table 11: Reversal example

	original	$\mu_E = 0$	$\rho_{XE} = 0.2$	$\mu_E = 0, \rho_{XE} = 0.2$
$\mathbf{ES}_{0.05}$	13.307	13.720	14.325	14.641
$\mathbf{ES}^*_{0.05}$	13.664	13.344	12.673	12.434
$\mathbf{ES}_{0.05} - \mathbf{ES}^*_{0.05}$	0.357	-0.376	-1.652	-2.207

much larger reversal of the difference in expected shortfall. The two effects above, when combined, result in an even greater reversal of the difference.

4.2   Full models

We now consider the CEV model of stock processes, with a mean-reverting log normal model of exchange rates. As before, we present expected shortfall calculations in Table 12 for a few values of α.

Table 12: CEV calculations

α	0.01	0.02	0.05	0.1
$\mathbf{ES}_\alpha$	5.858	3.835	1.245	-1.025
$\mathbf{ES}^*_\alpha$	5.848	4.031	1.448	-0.856

These values might seem surprisingly low (negative, even, for $\alpha = 0.1$), when compared to the BSM values, since we would expect the CEV model to produce fatter tails in the portfolio value distribution, and hence to raise the expected shortfall. This discrepancy might be explained, however, by comparing the (raw) empirical moments of the one year log return series of the two stocks (LDNS1: 0.0016, NYS1: 0.0012), to those of the BSM (LDNS1: 0.0054, NYS1: 0.0065). It appears that the actual return series exhibit narrower tails than the BSM. This effect is not present in the CEV model, since the parameters were chosen so as to match the empirical moments of the return series.

The introduction of mean reversion in the mean-reverting log normal currency model augments the convexity effect mentioned previously. In particular, holding all other parameters fixed, we vary exchange rate volatility and mean reversion rate. We see in Table 13 that as the mean reversion rate increases, the convexity effect becomes less pronounced (sampling errors notwithstanding) across different currency volatility effects. This is because a higher rate of mean reversion reduces, in effect, the overall volatility of the process, and thus reduces the convexity effect.

As mentioned, we would have liked to have experimented with the Heston model in particular, but time limitations unfortunately precluded this.

Table 13: Mean-reverting currency

a	σ_E				
		0	0.1	0.2	0.3
0.5	$\mathbb{ES}_{0.05}[XE]$	4.621	10.820	22.262	33.747
	$\mathbb{ES}_{0.05}[X/E]$	4.735	10.298	20.835	30.764
	Difference	-0.114	0.522	1.427	2.982
2.53 (estimated)	$\mathbb{ES}_{0.05}[XE]$	4.436	6.490	11.754	18.038
	$\mathbb{ES}_{0.05}[X/E]$	4.675	6.556	11.262	17.318
	Difference	-0.239	-0.066	0.492	0.721
10	$\mathbb{ES}_{0.05}[XE]$	4.700	5.175	6.878	9.582
	$\mathbb{ES}_{0.05}[X/E]$	4.269	5.004	6.631	9.218
	Difference	0.431	0.171	0.247	0.364

4.3   Perspectives

While we have pointed out some of the principles that affect the currency aggregation gap, there are a few different perspectives from which one can view the discrepancy.

A regulator might be interested in whether, and to what extent, changing the aggregation currency would have an effect on the entities they supervise. The above offer some heuristics to answer the question. For instance, if they are considering a very volatile exchange rate, they will know that the aggregation discrepancy will be large (as it is zero in the constant-currency case), and giving businesses the option to elect one of the currencies might decrease their apparent risk measures.

From the point of view of an entity, their goal might be purely cynical, in wanting to take as much risk as the fine print of the regulations will allow. As above, the basic heuristics will probably allow you to identify the favourable currency, and some basic quantitative modelling will allow you to estimate its magnitude. On this view, the option to elect your measurement currency can be a valuable one.

An entity will also, of course, be concerned about its own risk management. What should their reaction be to, for instance, $\mathbf{ES}_\alpha^*$ increasing but $\mathbf{ES}_\alpha$ remaining the same? We could easily think of an combination of factors, using the above results, that would cause this. It is not completely obvious. An actuarially-based answer depends on whether the liabilities (in the general sense) are considered to be denominated in domestic currency; in this case, then $\mathbf{ES}_\alpha$ is extremely important. An international entity might consider its liabilities to be denominated in many currencies, and might therefore be concerned with both $\mathbf{ES}_\alpha^*$ and $\mathbf{ES}_\alpha$.

5 Capital Composition

This section addresses the issue of risk capital in the context of a multiple currency environment.

We have seen that the expected shortfall measurement itself can vary with the choice of aggregation currency. Once an aggregation choice is made and an expected shortfall figure calculated, recall from the introduction the important and often-used interpretation: capital, generally in the form of a risk-free *reference asset*, must be added to the position until expected shortfall is zero – that is, until the position is *acceptable*.

Adding capital in this way has the simple effect of shifting the risk measure (so one simply translates – recall the axiom of translation invariance – until one reaches acceptability), but *only when the reference asset corresponds to the aggregation currency*. Suppose we are measuring the expected shortfall of a risk Z (expressed in the aggregation currency), and, pre-empting the need to make the position acceptable, we add an amount a in domestic currency and an amount b/E_0 in foreign currency. We then have

$$\mathbb{ES}_\alpha \left[Z + a + \frac{bE}{E_0} \right].$$

The foreign amount b/E_0 is set in this way so that its initial cost in domestic terms is b. This random variable E is necessary to convert this risk-free investment back for domestic aggregation. As mentioned, the effect of the domestic capital is predictable; we have

$$\mathbb{ES}_\alpha \left[Z + a + \frac{bE}{E_0} \right] = \mathbb{ES}_\alpha \left[Z + \frac{bE}{E_0} \right] - a, \tag{6}$$

but the effect of the foreign capital is not trivial; it involves, from the point of view of domestic aggregation, adding E-risk to the position, which will interact with the Z-risk in some way.

We assume that no interest is earned on the reference investments. This is more or less the case in the two economies which provide our data, but can easily be accommodated – the a on either side of (6) would need to be related by an accumulation factor.

A regulator may very well allow, or consider allowing, an entity to post capital in more than one currency, and we will expand on the possible perspectives on this optionality later in the section. Before that, we analyse mathematically this problem of capital composition, with the particular goal of optimising the composition.

Before performing calculations, we differentiate between three approaches that one may take to this acceptability and optimisation problem. Firstly, one may consider a fixed ratio in one's capital composition. Achieving acceptance by this *fixed-*

ratio approach amounts to solving

$$\min\{a + ka | \mathbb{ES}_\alpha \left[Z + a + \frac{(ka)E}{E_0} \right] = 0, a \in \mathbb{R}^+\},$$

where k is the ratio of foreign to domestic capital, perhaps specified by the regulator. Alternatively, one may take a *fixed-capital* approach, where the amount in the foreign reference asset, say, is fixed, and the the domestic capital is increased until sufficient; that is,

$$\min\{a + b | \mathbb{ES}_\alpha \left[Z + a + \frac{bE}{E_0} \right] = 0, a \in \mathbb{R}^+, b \text{ given}\}.$$

This might be appropriate if, for instance, a business happens to have a certain amount of foreign capital available, but raising any more would incur liquidity costs. Finally, one may take a *global* approach,

$$\min\{a + b | \mathbb{ES}_\alpha \left[Z + a + \frac{bE}{E_0} \right] = 0, a, b \in \mathbb{R}^+\},$$

where the ratio between foreign and domestic capital becomes flexible in the optimisation.

5.1 Basic models and principles

As a first approximation we consider the expected shortfall on our portfolio for various combinations of local and foreign capital. Figure 2 shows how adding enough capital will bring the position below the zero plane into the acceptable region.

We then consider the total capital required (in domestic currency) to make our portfolio acceptable, for various ratios of capital holdings. As can be seen in Figure 3 the minimum capital requirement is obtained when fully invested in the local reference asset.

If we allow the correlation between the foreign stock and exchange rate to vary, as in Figure 4, we see that a minimum capital value can be obtained by investing fully in the foreign reference asset when the foreign stock is very negatively correlated with the exchange rate. The opposite holds true when a very high positive correlation exists.

Varying the volatility of the exchange rate (Figure 5), keeping all else fixed, simply increases the curvature of these lines. It does not, however, change the optimal capital allocation. Similarly, we can see in Figure 6 that increasing currency drift does not alter the optimum capital allocation, but does raise the overall level of capital required.

We now consider the joint effect of the capital allocation ratio and correlation between the foreign stock and exchange rate. Figure 7 shows that for all capital
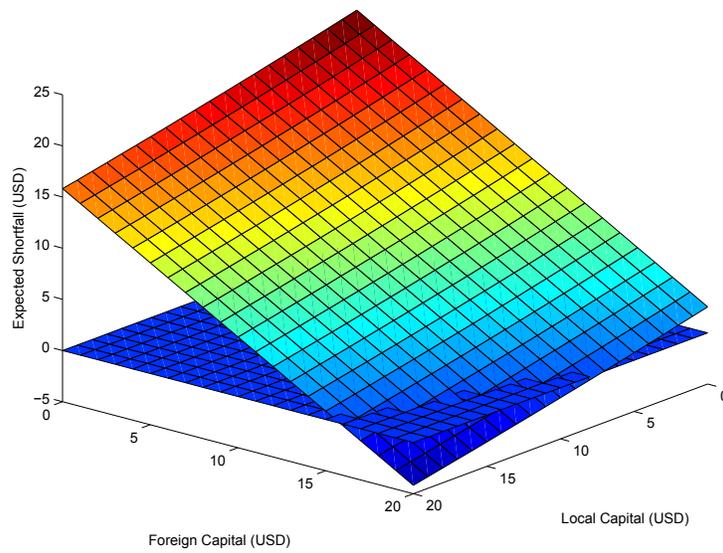
Figure 2: Expected shortfall for different capital holdings
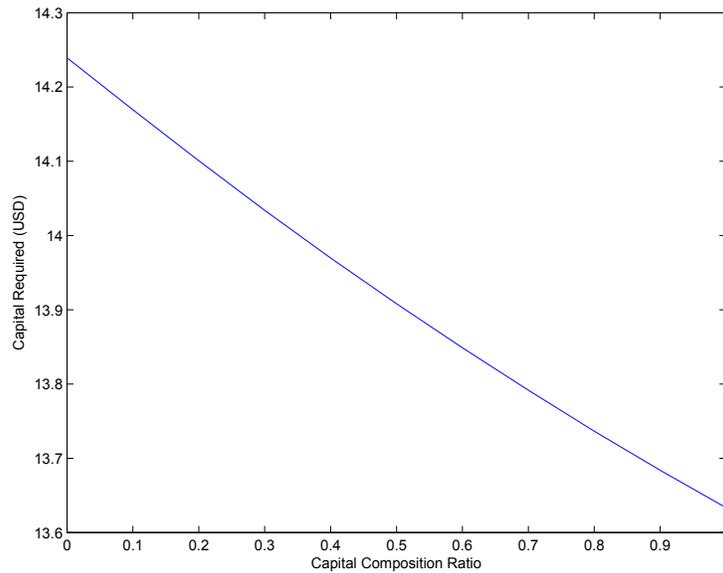
Figure 3: Capital required for different ratios

compositions, a smaller capital requirement can be obtained for a foreign stock which is more negatively correlated with the exchange rate. For very negative correlations, an investment fully in the foreign reference asset is preferable, whereas

22
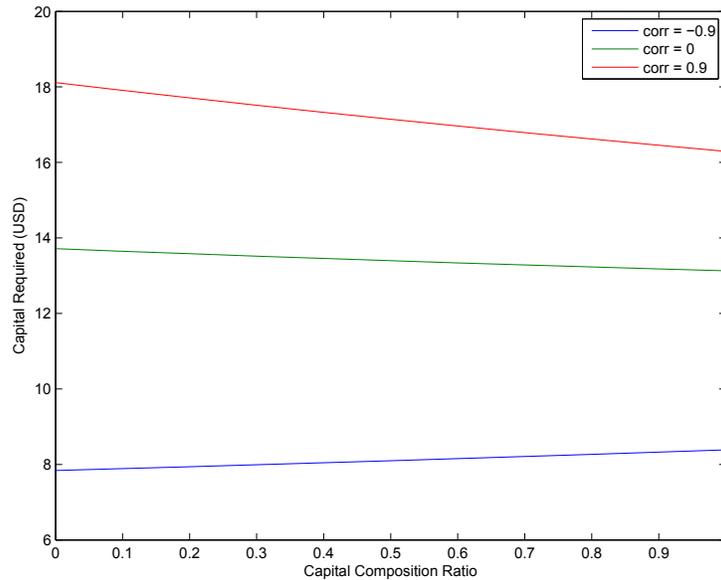
Figure 4: Varying ρ_{YE}

for very large positive correlations the opposite holds true. If, however, we increase the exchange rate volatility to 0.3, then an investment fully in the foreign asset is never preferable (Figure 8). This added volatility also increases the curvature of the surface.

Holding correlation fixed, and allowing the stock allocation to vary (with shorting permitted), we see that a minimum capital amount can be achieved by holding an equal weighting of the two stocks (see Figure 9). For portfolios which have a greater long holding in the local stock, foreign capital provides a minimum amount. The opposite is true for portfolios which have a greater long holding in the foreign stock. Figure 10 shows capital requirements, but at a higher exchange rate volatility. In particular, it shows that the riskiest position, in our example, is one in which the investor is invested fully in the foreign reference asset, as well as being short one unit of the foreign stock (since this position has the greatest exposure to exchange rate risk).

Finally, in Figure 11, we consider the minimum capital requirement (allowing the capital allocation ratio to vary freely; that is, to take the optimum at each point), that can be achieved for different stock allocations and correlations between the foreign stock and exchange rate. This shows that the minimum capital requirement, for a fixed stock allocation, falls as the correlation between the foreign stock and the exchange rate decreases. As correlation falls, the stock allocation minimising shifts more towards the foreign stock.

Since the minimum capital requirement occurs for a portfolio invested fully in
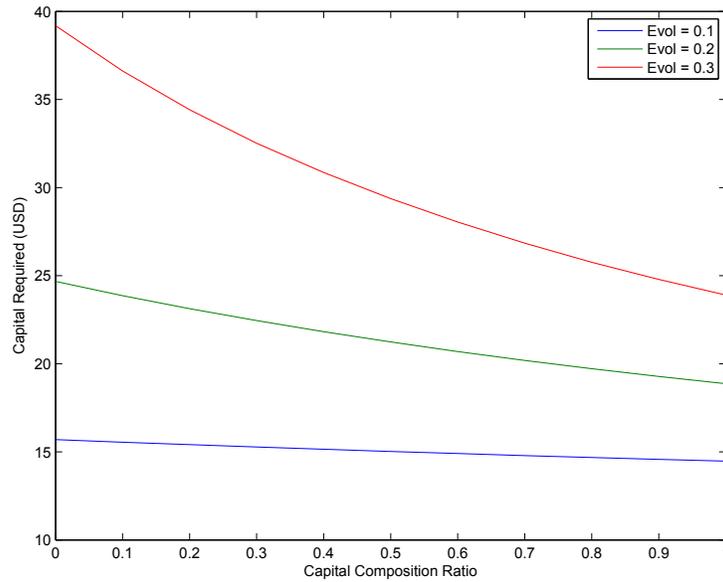
Figure 5: Varying σ_E
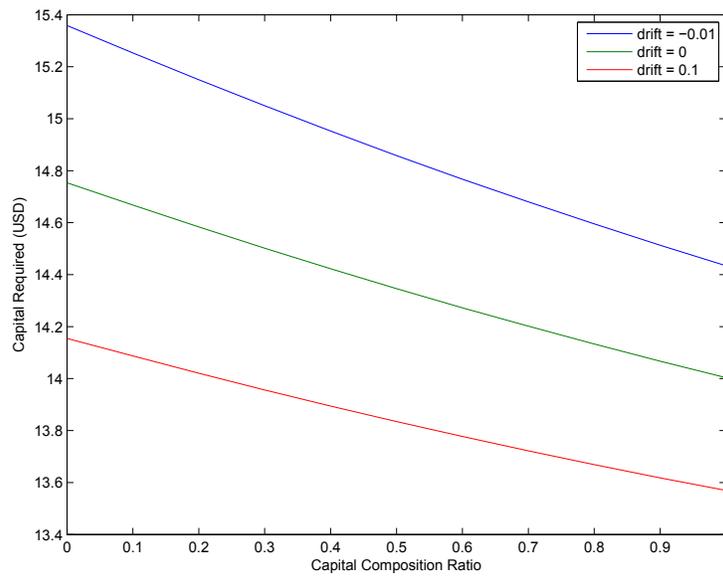
Figure 6: Varying μ_E

either the local or foreign reference asset, we consider the difference between the minimum capital surface, and the capital requirement when fully invested in the local reference asset (Figure 12). This shows, roughly, the crossover boundary at
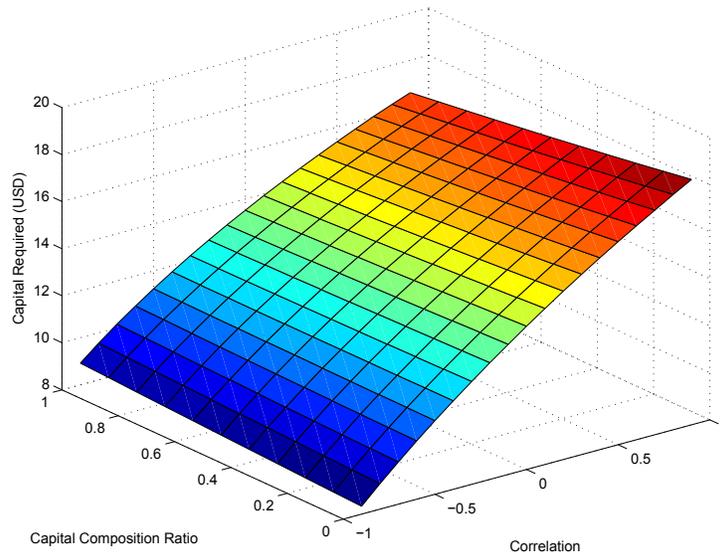
Figure 7: Varying composition and ρ_{YE}
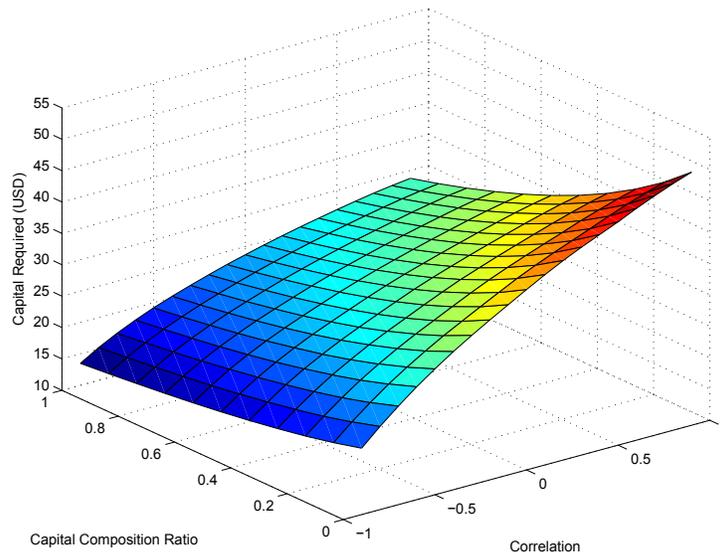
Figure 8: Varying composition and ρ_{YE}

which an investor would switch between being fully invested in the local reference asset, and fully invested in the foreign one. It also shows the *value of the option* of being allowed to post capital in the foreign currency in addition to the local one.
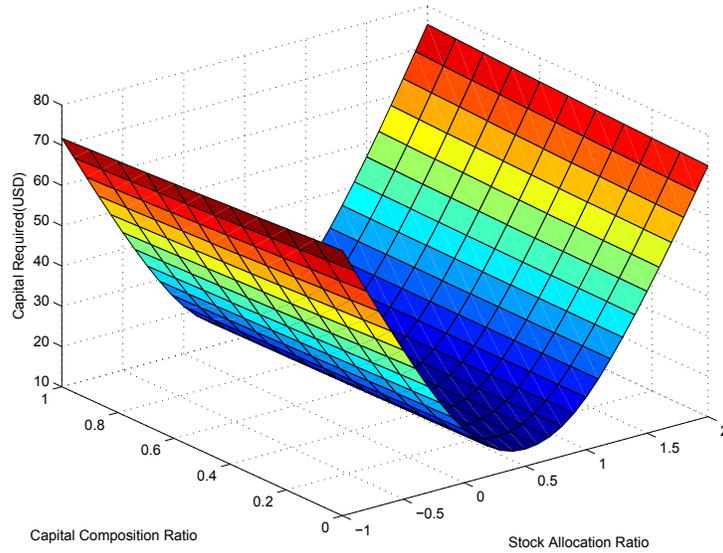
Figure 9: Portfolio and composition joint effects
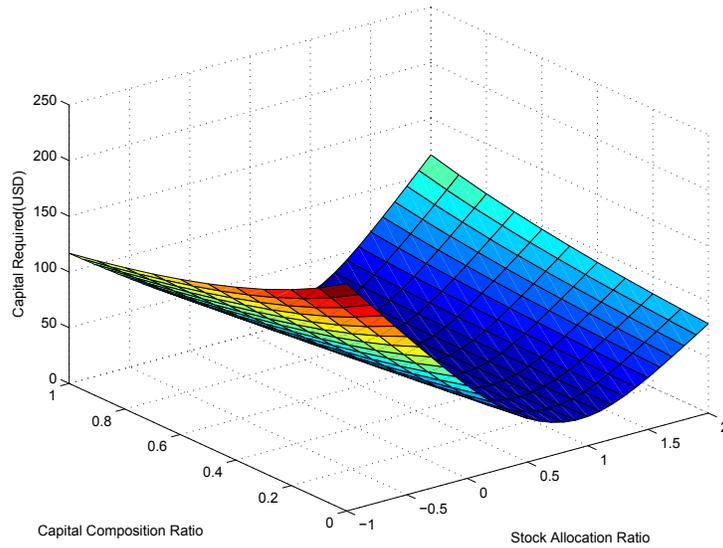
Figure 10: Portfolio and composition joint effects

5.2 Full models

We attempt, in a similar fashion to the BSM, to find an optimal capital allocation for the CEV model with mean reverting currency. Using our estimated parameters we get Figure 13, indicating that it would be optimal to invest fully in the local
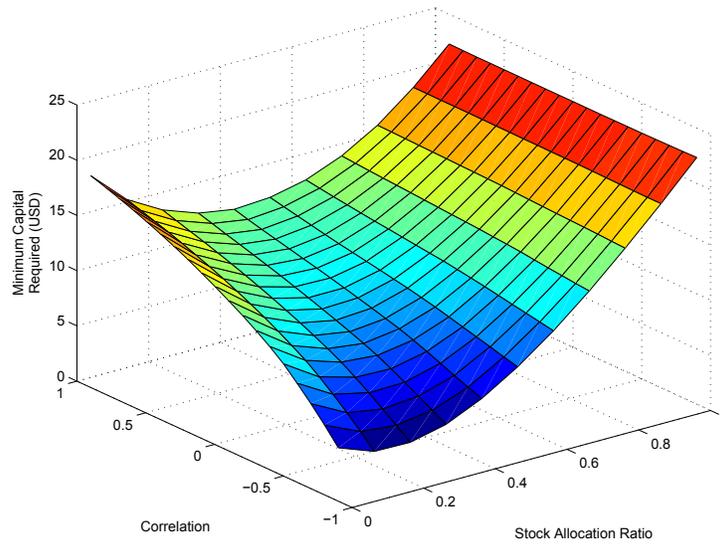
Figure 11: Optima and for different correlations and allocations
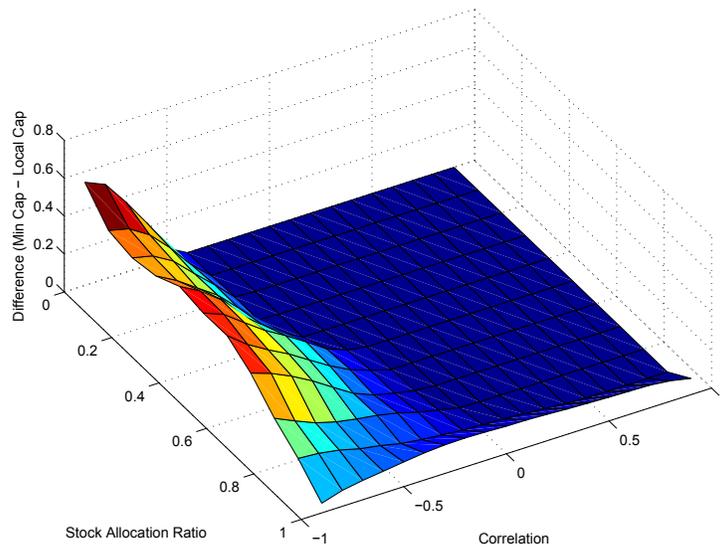
Figure 12: Difference between optimal and fully local capital

reference asset. Figure 14 plots the same, but for a currency volatility of 0.3. All this effects is the minimum currency value, and the curvature of the line. It does not, however, influence the optimal capital allocation. The most striking feature is how low the capital requirements are, which is explained by precisely the same reasons cited in the discussion around Table 12.
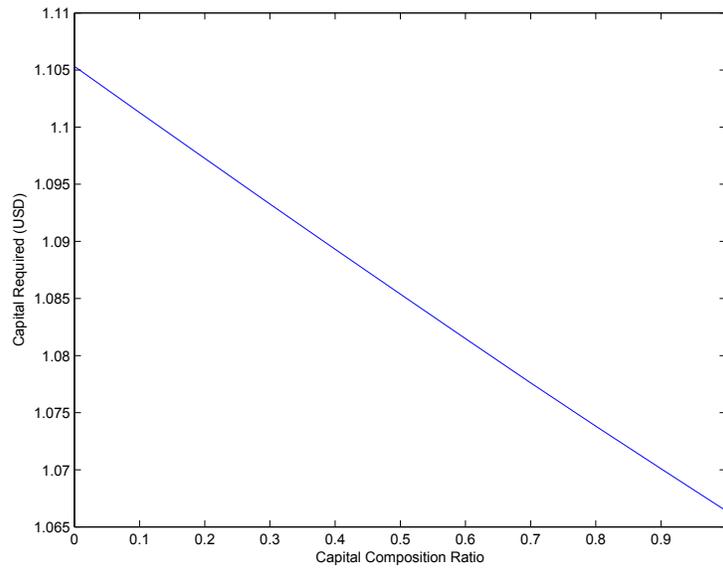
Figure 13: CEV capital compositions
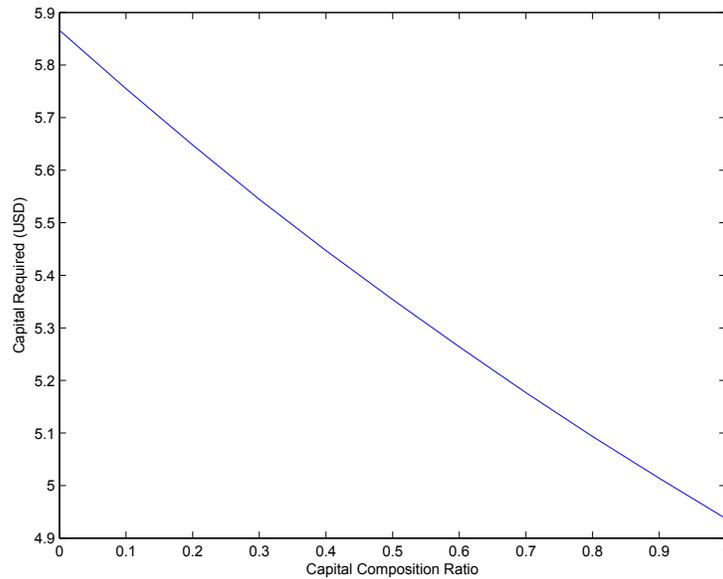
Figure 14: CEV capital compositions

5.3 Perspectives

As with the aggregation currency, we can conceive a variety of perspectives on this problem. The regulator might be considering allowing entities the multiple

currency option, and the results in this section will provide some heuristics as how to think about this problem.

An entity might be curious to get a sense of the value of this option, and determine whether formally optimising would be materially beneficial.

6 Conclusion

We have undertaken a literature review on the general topic of risk management and risk measurement. We have focussed on expected shortfall and therefore on coherent risk measures. The literature survey also included the few papers dealing with multiple currencies in this context.

We went on to establish a modelling framework with which we could perform computations of expected shortfall and thereby study the issues of aggregation currency and capital composition.

With regard to aggregation currency, we calculated the aggregation discrepancy on a position in the two shares with both our simple and full models. We then attempted to isolate the effects at work and retrospectively understand why the discrepancy was of the sign that it in fact was. While we were unable to complete the analysis we envisaged with our complex models, the rules we established could be useful heuristics to different parties considering this matter.

With regard to capital composition, we established a framework in which one can consider providing risk capital in more than one currency and attempt to optimise the composition. We confirmed some intuitive factors at work in this optimisation and performed calculations under a variety of models and conditions. Again, we were not able to add to the basic results to our satisfaction. We found the optimisation problem, under a wide variety of circumstances, has a *boundary solution*: that is, one is very likely to want to provide capital in full in one or the other currency, if one is allowed to do so.

We end by outlining a few research questions that one could address in a continuation of this work.

Further research questions

Because of the complex nature of the underlying problem, and the time limitations of the project, it is quite easy to outline a few problems that could be the basis for future work. These are listed below.

- Would the results be enhanced by qualitatively different data (e.g. access to historical derivative prices to reflect on volatility states, or series with more pronounced correlation structures)? More sophisticated estimation methods would be required for this, taking cognisance of the fact that derivative prices are often considered under the risk-neutral measure only, whereas risk calculations are under the real-world measure.

29

- What insights and results can be obtained when the Heston model is estimated and used in the aggregation setting of Section 4? It is the ability of the this model to capture richer correlation structures (e.g. between volatility processes) that makes this potentially interesting.

- Would anything important be added to the analysis if more than two assets were involved in the position? Are there any interesting asymptotic results or features that hold (for example, the occurrence of boundary solutions to the capital composition problem) when the number of assets becomes large?

- Can the measure of diversification defined in Section 4 – the amount of expected shortfall reduction in the composition of positions – be interrogated and studied further in a useful way? While the measure is model and horizon dependent, it has the *a priori* benefit of focussing directly on the loss-tail (capturing all of the dependencies at work on the extreme side of the final distribution).

- What would the impact of more sophisticated models be in the context of Section 5? Does our result of boundary solutions hold under these models? The motivation for this point is that, due to time limitations, we could not extend our analysis much beyond our basic models.

Bibliography

Acerbi, C., Székely, B., 2014. Back-testing expected shortfall. Risk, 76.

Acerbi, C., Tasche, D., 2002. On the coherence of expected shortfall. Journal of Banking & Finance 26 (7), 1487–1503.

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1997. Thinking coherently. Risk 10, 68–71.

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent measures of risk. Mathematical Finance 9 (3), 203–228.

Artzner, P., Delbaen, F., Koch-Medina, P., 2009. Risk measures and efficient use of capital. Astin Bulletin 39 (01), 101–116.

BCBS, 2012. Fundamental review of the trading book. Basel Committee on Banking Supervision.Basel: Bank for International Settlements.

Beckers, S., 1980. The constant elasticity of variance model and its implications for option pricing. The Journal of Finance 35 (3), 661–673.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. The Journal of Political Economy, 637–654.

Cape, J., Dearden, W., Gamber, W., Liebner, J., Lu, Q., Nguyen, M. L., 2015. Estimating Heston and Bates models parameters using Markov chain monte carlo simulation. Journal of Statistical Computation and Simulation 85 (11), 2295–2314.

Christie, A. A., 1982. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. Journal of Financial Economics 10 (4), 407–432.

Coleman, T., Litterman, B., 2012. Quantitative Risk Management: A Practical Guide to Financial Risk. Wiley Finance. Wiley.

Cont, R., 2001. Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance 1 (2), 223–236.

Cox, J. C., Ross, S. A., 1976. The valuation of options for alternative stochastic processes. Journal of Financial Economics 3 (1), 145–166.

Embrechts, P., Puccetti, G., Rüschendorf, L., 2013. Model uncertainty and var aggregation. Journal of Banking & Finance 37 (8), 2750–2764.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R., Beleraj, A., 2014. An academic response to Basel 3.5. Risks 2 (1), 25–48.

Embrechts, P., Wang, B., Wang, R., 2015. Aggregation-robustness and model uncertainty of regulatory risk measures. Finance Stochastics, Forthcoming.

Erdemlioglu, D., Laurent, S., Neely, C. J., 2013. Econometric modeling of exchange rate volatility and jumps. Handbook of Research Methods and Applications in Empirical Finance, 373.

Föllmer, H., Schied, A., 2002. Convex measures of risk and trading constraints. Finance and Stochastics 6 (4), 429–447.

Heston, S. L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of financial studies 6 (2), 327–343.

Jorion, P., 1996. Measuring the risk in value at risk. Financial Analysts Journal 52 (6), pp. 47–56.

Koch-Medina, P., Loubet, E., 2014. Currency risk in capital adequacy. Available at SSRN 2402594.

Merton, R. C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3 (1), 125–144.

Nadarajah, S., Zhang, B., Chan, S., 2014. Estimation methods for expected shortfall. Quantitative Finance 14 (2), 271–291.

Schroder, M., 1989. Computing the constant elasticity of variance option pricing formula. Journal of Finance, 211–219.

Yamai, Y., Yoshiba, T., 2002. On the validity of value-at-risk: comparative analyses with expected shortfall. Monetary and Economic Studies 20 (1), 57–85.

Zhao, B., 2009. Inhomogeneous geometric Brownian motion. Available at SSRN 1429449.

Margin optimisation for central counterparty applications

Team 2

SARA SVALUTO-FERRO, Swiss Federal Institute of Technology
MELUSI MAVUSO, University of Cape Town
CHRIS MCPETRIE, University of Cape Town
GRAHAM ZIERVOGEL, University of Cape Town

Supervisor:
STÉPHANE CRÉPEY, Université d'Evry

Contents

1   Introduction

This report is concerned with the optimisation of the initial margin (im) and default fund contribution (df) demanded by a central clearing counterparty (CCP) in the clearing of derivatives. Also discussed is the formulation of optimal liquidation strategies of the positions of defaulting clearing members. Before exploring the technicalities of the models used, it is important to understand the context in which this investigation was motivated.

The financial crisis exposed systemic fragilities in the global financial system: the interdependence between large financial institutions was shown to be susceptible to catastrophic contagion effects. In the aftermath of the crisis, a number of ongoing regulatory initiatives have been undertaken to reduce the overall counterparty risk in the system. One such initiative, and perhaps the most significant, is the shift to central clearing of OTC derivatives. Regulations pushing for the move to central clearing are the Dodd Frank Act in the US and the European Market Infrastructure Regulation (EMIR) in Europe. Depending on the product and the classification of the derivatives, estimates suggest that $50 - 80\%$ of the OTC derivative market will eventually transact through a CCP (InteDelta, 2013).

Two major risk mitigation features of CCPs are the requirements to post an initial margin with daily maintenance coming from the variation margin and contribute to the pooled default fund. This report is concerned with optimising the allocation of these values, from the point of view of competing CCPs. CCPs, in competition with one another, are incentivised to minimise their margin demands - to attract clients while maintaining a level of margins that protects the system from default by any number and combination of counterparties. This leads to an optimisation problem: how best to minimise margins while protecting the integrity of the system.

The concept of initial margin is a long-held feature of financial markets. It is most pertinent in the context of exchange traded derivatives and cash equity whereby members of the exchange post initial margins and settle the daily variation margin. In the context of OTC derivatives, the concept of initial margin is less well developed although it is similar to the independent amount under a credit support annex, in the terminology of the International Swaps and Derivatives Association (ISDA) (InteDelta, 2013).

CCPs are free to choose their own margining methodologies. Many CCPs are moving towards a 5-day worst case loss model based on five years of historic data. Other CCPs use Value at Risk (V@R) methodology (InteDelta, 2013). We propose a margining methodology based on a modified, multivariate extension of the ex-

pected shortfall risk allocation approach, of which loss functions are a central part. This methodology operates at the level of sets of interconnected components and accounts for correlation between clearing members, something which aggregated univariate risk measures fail to do. For this reason, the margins specified by our models are generally higher than those suggested by aggregated univariate risk measures, but better protect against systemic risk.

Figure 1 illustrates how CCPs place themselves as a counterparty to transactions between participating firms. Firms that clear directly with a CCP are known as clearing members. In the case of default of one or more clearing member's, the CCP may reduce the disruptive effects of defaults. The CCP may enable the structured replacement of the defaulting members position by, for example, auctioning its portfolio to other clearing members (InteDelta, 2013). We explore the strategies that the CCP may take to liquidate the defaulting member's portfolio as an extension to the central problem of margin optimisation.
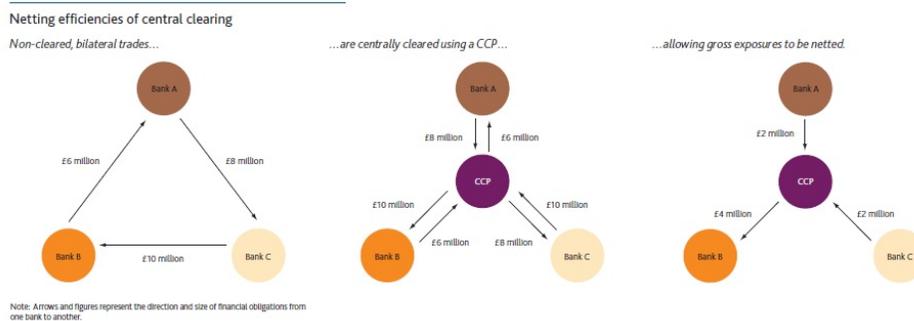
Figure 1: The netting effect of central clearing.

When a clearing member defaults, there is a hierarchy of collateral available to the CCP to protect itself and the other clearing members. Firstly, the defaulting member's initial margin can be used to cover any incurred losses or obligation. Ideally, this initial margin, maintained daily through the variation margin, should be sufficient to cover all losses faced by the CCP. This has been the experience of CCPs based in the UK following defaults such as those of Lehman Brothers and MF Global. If the initial margin does not adequately cover the amount owed, the CCP can then access the defaulting clearing members default fund contribution. Before accessing the default fund contributions of the other clearing members, the CCP may contribute some of its own equity towards resolving the incurred losses. This motivates the CCP to optimise the allocation of each clearing member's initial margin and default fund contribution. This is the focus of this investigation. Thereafter, if losses are still not adequately covered, the CCP may access the default fund

contribution of the remaining non-defaulting members, mutualising losses across them. Thereafter, CCPs may call on surviving members to make further contributions; sometimes termed rights of assessment. If the extent to which CCPs can call on surviving members to make further contributions is limited, the last remaining loss-absorbing resource available to the CCP is its remaining equity. If losses are in excess of this remaining equity, the CCP would itself become insolvent. This hierarchy, often termed the default waterfall, is captured in Figure 2 (Rehlon and Nixon, 2013). In our modelling we focus on the initial margin and the default fund as loss absorbing mechanisms.

Figure 2: The hierarchy of collateral available to the CCP in the event of a clearing member defaulting.

The report is structured as follows. We first establish the theoretical framework in which our models are constructed. We thereafter discuss the loss functions and distributions used. We then provide a detailed explanation of the calibration and margining scheme procedure developed. We then move onto the issue of optimising the liquidation strategy of the constituent positions of the defaulting clearing member's portfolio, with a discussion of the liquidity constraint and integrated approaches we use. We conclude with a discussion of our results and proposals as well as a brief discussion of potential future extensions.

2 Theoretical framework

Let $X = (X_1, \ldots, X_d) \in L^0$ be a random vector representing the respective loss profiles of the CCP's constituent clearing members, i.e. negative values of X_k represent profits whereas positive values represent losses. We aim to minimise the sum of the total margins $m_1, \ldots, m_k$ levied on each clearing member while still maintaining an acceptable level of systemic risk. In order to solve this problem, we build on the theoretical framework proposed in Armenti et al. (2015).

We start with a loss function l over $\mathbb{R}^d$, used for measuring the expected loss $E[l(X)]$ of a financial loss profile X.

Definition 2.1. A function $l : \mathbb{R}^d \to (-\infty, \infty]$ is called a *loss function* if:
(A1) l increasing, that is $l(x) \geq l(y)$ if $x_k \geq y_k$ for all $k = 1 \ldots, d$;
(A2) l is convex, lower semi-continuous and satisfies $l(x_0) < \infty$ for some $x_0 > 0$;
(A3) $l(0) = 0$ and $l(x) \geq \sum x_k$ on $\mathbb{R}^d$.

A risk neutral assessment of the losses would correspond to $E[\sum X_k] = \sum E[X_k]$. Hence, **(A3)** expresses a form of risk aversion, whereby the loss function weights high losses more heavily than a risk neutral evaluation. **(A1)** and **(A2)** express the respective normative facts about risk that "the more losses, the riskier" and "diversification should not increase risk".

The corresponding *acceptance set* is given by

$$A(X) = \left\{ m \in \mathbb{R}^d : E[l(X - m)] \leq c \right\}.$$

We say that a monetary allocation $m \in \mathbb{R}^d$ is *acceptable* at the loss level $c \geq 0$ if $E[l(X - m)] \leq c$, i.e. if $m \in A(X)$.

We are now ready to state the main problems explored in this report.

Definition 2.2. An *optimal monetary risk allocation* is an acceptable allocation m^* such that

$$\sum m_k^* = \inf_{m \in A(X)} \sum m_k \qquad \text{(primal problem)}.$$

The dual formulation of the proposed optimisation problem is given by

$$\begin{cases} 0 = 1 - \lambda E[\partial_i l(X - m)] & \text{for } i = 1, \ldots, d, \\ c = E[l(X - m)], \end{cases} \qquad \text{(dual problem)}. \qquad (1)$$

2.1 Distributions

In our modelling, we make use of the Multivariate Gaussian distribution, as well as the Multivariate Affine Generalised Hyperbolic (MAGH) distribution to model the vector X of losses and profits. Indeed, our framework is as general as possible so as to allow for extensions employing other distributions. What follows is a brief motivation for and discussion of the distributions used in this report.

The Multivariate Gaussian distribution is a widely used modelling component in mathematical finance, largely due to its tractability. However, there is wealth of literature criticising its applicability to the modelling of financial asset-returns, most of which focuses on its inability to account for the empirically observed 'fat-tails' of financial returns. While aware of its well-documented limitations, we include the Gaussian distribution as a preliminary modelling distribution. We specify the distribution as follows for a d-dimensional loss and profit vector X:

$$X \sim \mathrm{N}_d(\mu, \Sigma)$$

where μ is a d-dimensional mean vector

$$\mu = (\mu_1, \mu_2, \ldots, \mu_d)^T := (\mathbb{E}[X_1], \mathbb{E}[X_2], \ldots, \mathbb{E}[X_d])^T$$

and $\Sigma = (\sigma_{ij})_{ij}$ is a $d \times d$ covariance matrix

$$\sigma_{ij} := [\mathrm{Cov}[X_i, X_j]], \quad i = 1, 2, \ldots, d; \, j = 1, 2, \ldots, d.$$

The joint moment generating function of the random vector X is given by

$$M_X(t) = \exp\left(\mu^T t + \frac{1}{2} t^T \Sigma t\right),$$

while each component has moment generating function

$$M_{X_i}(t) = \exp\left(\mu_i t + \frac{1}{2} \sigma_{ii}^2 t^2\right).$$

As an alternative to the Multivariate Gaussian distribution, we propose to use the Multivariate Affine Generalised Hyperbolic distribution to model the vector X of losses. This distribution has been favoured because of its ability to capture many of the stylised facts of the distribution asset returns.

A random vector X having the MAGH distribution with location vector μ and scaling matrix Σ is constructed as

$$X = L^T Y + \mu,$$

where $Y_1, Y_2, \ldots, Y_d$ are independent random variables, each with density given by

$$f(y_i) = \frac{(\gamma_i/\delta_i)^{\lambda_i}}{\sqrt{2\pi} K_{\lambda_i}(\delta_i \gamma_i)} e^{\beta_i(y_i - \mu_i)} \times \frac{K_{\lambda - 1/2}(\alpha_i \sqrt{\delta_i^2 + (y_i - \mu)^2})}{(\sqrt{\delta^2 + (y_i - \mu_i)^2}/\alpha_i)^{1/2 - \lambda_i}}.$$

Here K_λ is the hyperbolic Bessel function of the second kind. Each Y_i has the following moment generating function

$$M_{Y_i}(t) = \frac{e^{\mu_i t} \gamma_i^{\lambda_i}}{(\sqrt{\alpha_i^2 - (\beta_i + t)^2})^{\lambda_i}} \times \frac{K_{\lambda_i}(\delta_i \sqrt{\alpha_i^2 - (\beta_i + t)^2})}{K_{\lambda_i}(\delta_i \gamma_i)}.$$

The joint moment generating function of X is given by

$$M_X(t) = e^{t^T \mu} M_Y(t^T L)$$

where

$$M_Y(t) = \prod_{i=1}^{d} M_{Y_i}(t_i).$$

As a comparison', we show in Equation 3 the acceptance sets for the Normal and MGH distributions for increasing values of c (here blue corresponds to low c and red corresponds to high c). These distributions have been manipulated so that the moment generating functions when evaluated at (2,0) and (0,2) are the same.

2.2 Loss functions

We now attempt to solve the proposed optimisation problem analytically by making simplifying assumptions. We assume X has a bivariate Gaussian distribution and our loss function $l(x_1, x_2)$ takes the following, relatively tractable, form:

$$X \sim \mathcal{N}_2 \left(0, \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix} \right), \quad l(x_1, x_2) = x_1^+ + x_2^+ + x_1 \vee x_2.$$

It can be easily verified that $l(x, y)$ satisfies **(A1-3)** and is a permissible loss function. We start characterising the corresponding acceptance set at level c.

$$\begin{aligned} A(X) &= \big\{ m \mid E[(X_1 - m_1)^+ + (X_2 - m_2)^+ + (X_1 - m_1) \vee (X_2 - m_2)] \le c \big\} \\ &= \big\{ m \mid E[(X_1 - m_1)^+] + E[(X_2 - m_2)^+] \\ &\qquad + E[(X_1 - m_1 + 0 \vee (X_2 - m_2 - (X_1 - m_1))] \le c \big\} \\ &= \big\{ m \mid E[(X_1 - m_1)^+] + E[(X_2 - m_2)^+] - m_1 \\ &\qquad + E[((X_2 - X_1) - (m_2 - m_1))^+] \le c \big\}. \end{aligned}$$

8
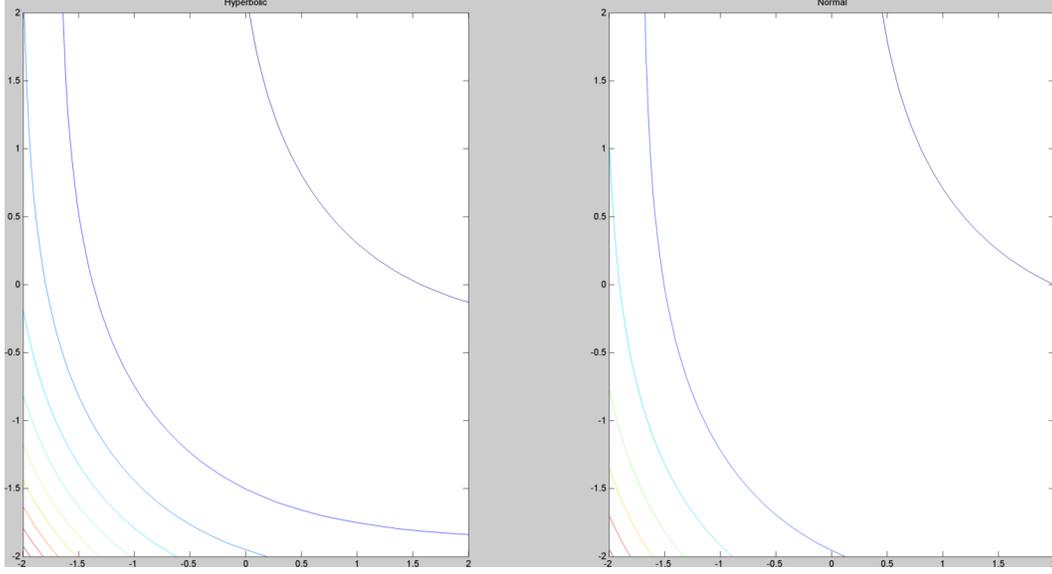
Figure 3: Comparison of normal and hyperbolic acceptance sets.

Making use of Bachelier's pricing model for vanilla call options we can simplify the expressions using the following identities. Let $\bar{\sigma} := \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$, then

$$E[(X_1 - m_1)^+] = -m_1\Phi(\frac{-m_1}{\sigma_1}) + \sigma_1\phi(\frac{-m_1}{\sigma_1}),$$

$$E[(X_2 - m_2)^+] = -m_2\Phi(\frac{-m_2}{\sigma_2}) + \sigma_2\phi(\frac{-m_2}{\sigma_2}),$$

$$E[((X_2 - X_1) - (m_2 - m_1))^+] = -(m_2 - m_1)\Phi(\frac{-(m_2 - m_1)}{\bar{\sigma}})$$

$$+ \bar{\sigma}\phi(\frac{-(m_2 - m_1)}{\bar{\sigma}}).$$

Attempting to solve the dual problem (Equation 1) from these expectations we have:

$$\frac{\partial}{\partial m_1}E[l(X - m)] = -\Phi(\frac{-m_1}{\sigma_1}) + \Phi(\frac{-(m_2 - m_1)}{\bar{\sigma}}) - 1,$$

$$\frac{\partial}{\partial m_2}E[l(X - m)] = -\Phi(\frac{-m_2}{\sigma_2}) + \Phi(\frac{-(m_2 - m_1)}{\bar{\sigma}}) - 1.$$

9

This leads to simple implicit expressions for the optimal choices of m_1 and m_2:

$$m_2 = -\Phi^{-1}\left(\frac{\lambda - 1}{\lambda} + \Phi(\frac{-m_1}{\sigma_1})\right)\bar{\sigma} + m_1,$$

$$m_1 = -\Phi^{-1}\left(\frac{\lambda - 1}{\lambda} + \Phi(\frac{-m_2}{\sigma_2})\right)\bar{\sigma} + m_2.$$

Clearly attempts to solve for the optimal risk allocation analytically are impractical. However, if we make the simplifying assumption that $\sigma_1 = \sigma_2$ (which is not too unreasonable given an appropriate ρ), we find that this leads to $m_1 = m_2$. This allows us to simplify as follows, with $\sigma_1 = \sigma_2 = \sigma^*$:

$$m_1^* = m_2^* = -\Phi^{-1}(-\frac{1}{2} + \frac{1}{\lambda})\sigma^*.$$

A pragmatic approach to calibrating the model can now be adopted by choosing λ such that an appropriate value for c is realised in our limiting condition $E[l(X - m)] = c$.

A logical comparison to check whether this simple Gaussian framework is an appropriate model for the risk allocation is to compare results with the easily computed V@R and CV@R statistics. For choices of $c = 0.2$, $\sigma_1 = \sigma_2 = 0.9$ and 99% value at risk statistics we obtain the following acceptance sets, as shown in figure 4 below.
Clearly this model is far too lenient and permits combinations of m_1 and m_2 that are far outside of the acceptable value at risk ranges. As a result, we motivate the following loss function which will allow us to calibrate the model so that the convergent tails of the acceptance sets align with the calculated value at risk statistics:

$$l(x_1, x_2) = \alpha x_1^+ + \beta x_2^+ + \gamma x_1^+ x_2^+.$$

We elaborate on this argument in the chapter on calibration and margin scheme proposal.

Is also interesting to note that the acceptance set reduces in size if the correlation between the clearing members is lower. This is inconsistent with our general model of acceptable margins, since higher correlation between clearing members should increase systemic risk. This indicates a need for care when choosing loss functions, as some may prove to be inconsistent with the general model.

We propose now another interesting loss faction, given by a d-dimensional multivariate extension to the exponential risk function proposed by Armenti et al. (2015)

$$l(\mathrm{x}) = \frac{1}{1 + \alpha}\left(\frac{1}{d}\sum_{i=1}^{d} e^{dx_i} + \alpha\exp\left(\sum_{i=1}^{d} x_i\right)\right) - 1.$$

10
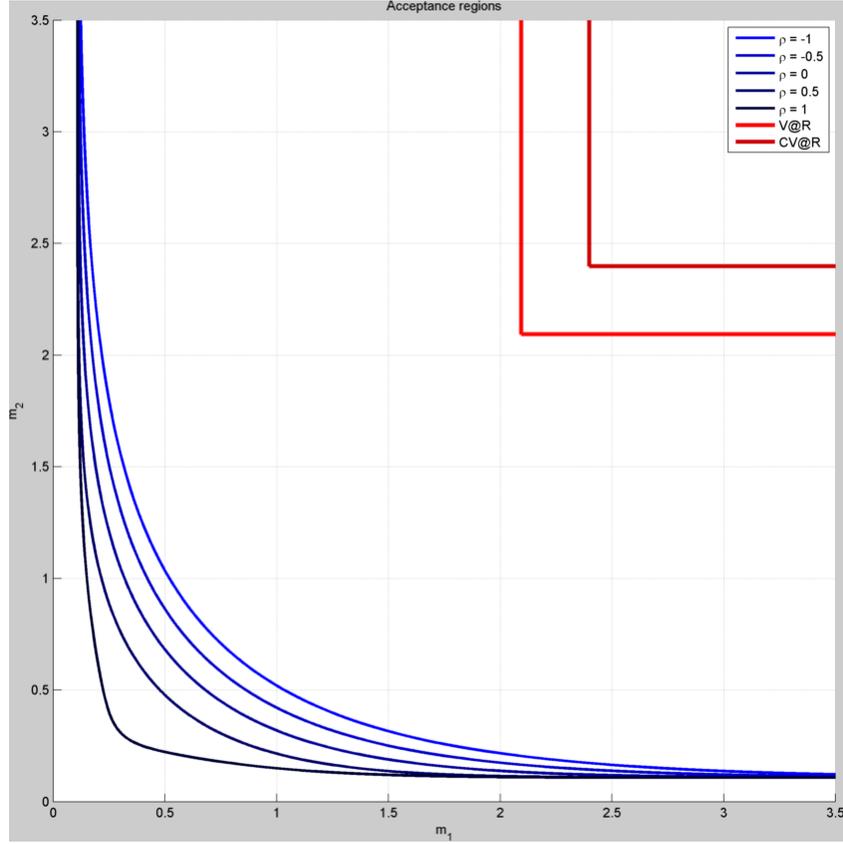
Figure 4: Acceptance sets for pre-specified choices of c, σ_1 and σ_2 and 99% value at risk statistics.

The high tractability of this function permits us to compute optimal monetary risk allocation analytically (for a general distribution and loss profile) where the moment generating function is known.

To find the optimal margins m^*, we first need to compute $\mathbb{E}\left[l(X - m)\right]$ and $\mathbb{E}\left[\frac{\partial}{\partial m_i}l(X - m)\right]$ for each $i = 1, 2, \ldots, d$. We first present some preliminary calculations

$$\mathbb{E}\left[\sum_{i=1}^{d} e^{dX_i}\right] = \sum_{i=1}^{d} M_{X_i}(d), \quad \mathbb{E}\left[e^{\sum_{i=1}^{d} X_i}\right] = M_X(\mathbf{1}).$$

This allows us to compute the necessary expectation of the loss function as follows:

$$\mathbb{E}\left[l(X - m)\right] = \frac{1}{1 + \alpha}\left(\frac{1}{d}\sum_{i=1}^{d} M_{X_i}(d)e^{-dm_i} + \alpha M_X(\mathbf{1})e^{-\sum_{i=1}^{d} m_i}\right) - 1,$$

$$\frac{\partial}{\partial m_i} l(x - m) = -\frac{1}{1+\alpha}\left(e^{d(x_i - m_i)} + \alpha \exp\left(\sum_{i=1}^{d}(x_i - m_i)\right)\right).$$

We use these results to express the dual problem (1) in the following manner:

$$\frac{1}{\lambda} = \frac{1}{1+\alpha}\left(M_{X_i}(d)e^{-dm_i} + \alpha M_X \exp\left(-\sum_{i=1}^{d}m_i\right)\right),$$

$$\tilde{c} := c + 1 = \frac{1}{1+\alpha}\left(\frac{1}{d}\sum_{i=1}^{d}M_{X_i}(d)e^{-dm_i} + \alpha M_X(\mathbf{1})e^{-\sum_{i=1}^{d}m_i}\right).$$

Solving these equations yields the following results:

$$m_i^* = \frac{\ln\left(M_{X_i}(d)\right)}{d} - \frac{1}{d}\ln\left(\tilde{c}(1+\alpha)\right) + \frac{1}{d}\ln\left(1 + \alpha M_X(\mathbf{1})\prod_{i=1}^{d}[M_{X_i}(d)]^{-\frac{1}{d}}\right),$$

$$\sum_{i=1}^{d}m_i^* = \frac{1}{d}\sum_{i=1}^{d}\ln\left(M_{X_i}(d)\right) - \ln\left(\tilde{c}(1+\alpha)\right) + \ln\left(1 + \alpha M_X(\mathbf{1})\prod_{i=1}^{d}[M_{X_i}(d)]^{-\frac{1}{d}}\right).$$

We now use these results in an example. Consider the same multivariate d-dimensional loss function, and consider a vector X which has a d-dimensional multivariate Gaussian distribution:

$$X \sim \mathrm{N}_d(\mu, \Sigma) \quad X_i \sim \mathrm{N}(\mu_i, \sigma_i^2).$$

Making use of our earlier results, we find that:

$$m_i^* = \mu_i + \frac{1}{2}d\sigma_i^2 + \frac{1}{d}\ln\left(1 + \alpha\exp\left[\frac{1}{2}\hat{\sigma}^2 - \frac{1}{2}d\sum_{i=1}^{d}\sigma_i^2\right]\right) - \frac{1}{d}\ln(\tilde{c}(1+\alpha)),$$

$$\sum_{i=1}^{d}m_i^* = \sum_{i=1}^{d}\mu_i + \frac{1}{2}d\sum_{i=1}^{d}\sigma_i^2 + \ln\left(1 + \alpha\exp\left[\frac{1}{2}\hat{\sigma}^2 - \frac{1}{2}d\sum_{i=1}^{d}\sigma_i^2\right]\right) - \ln(\tilde{c}(1+\alpha)),$$

where $\hat{\sigma}^2 = \mathbf{1}^T\Sigma\mathbf{1}$.

3 Calibration procedure and margin scheme proposal

The ability to choose different loss functions and distributions makes the expected shortfall framework extremely versatile. However, this versatility also implies that an effective calibration approach is required to make the model effective. In this section we discuss methods for specifying the loss function and determining the optimum threshold level c. This calibration approach leads naturally on to a margin scheme in which we can not only define an optimum monetary allocation, m^*, but also two layers of margins: the initial margin im^* and the default fund contribution $df^* = m^* - im^*$.

As an illustrative example we make use of the following loss function:

$$l(x_1, x_2) = cw_1 x_1^+ + cw_2 x_2^+ + w x_1^+ x_2^+. \tag{2}$$

Clearly this function can be easily generalised for dimension $d > 2$ and the various coefficients can be chosen such that the axioms in Definition 2.1 are satisfied. Further, the presence of an idiosyncratic contribution from each risk factor as well as a pooled contribution from both risk factors suggests plausibility. Additional quadratic terms can be introduced to improve smoothness (Feinstein et al., 2015).

The proposed strategy is to calibrate each w_k to a target value $\widehat{im_k}$ corresponding to some individual measure of risk, such as value at risk or conditional value at risk. If we assume that one of the members posts an extremely large margin, we are effectively eliminating the risk contribution of that member and reducing the loss function to a function of only one risk factor:

$$l_k(x_k) = cw_k x_k^+.$$

Thus, setting the optimal m_k^* in this asymptotic scenario to the relevant value at risk statistic we can solve $\mathbb{E}[w_k(X_k - m_k^*)^+] = 1$ for the appropriate weighting w_k. Finally, by setting $c = \frac{1}{\min(w_1, w_2)}$ to satisfy the risk aversion requirement for our loss function, we are now able to plot acceptance sets as a function of w. In particular, w will lie in a specific range $[0, \hat{w}]$ where $\hat{w}$ is the largest possible weighting for the dependence structure which does not break the convexity of the loss function. The results of this calibration approach are shown in Figure 5 below.

This calibration has a number of pleasing properties - most importantly that we can make a clear proposal for the determination of initial margins and default fund contributions. Namely, the asymptotic values (here set to value at risk) represent the initial margins and the difference between these values and the corresponding point chosen on the acceptance set curve represent the default fund contributions. As the acceptance curves deviate from the asymptotic values more significantly for

larger values of w, we can choose w according to the size of default fund contributions we desire. In addition, the methodology can be used for any distribution of X and can be extended to any loss function $l(x)$ such that $l(0, 0, \ldots, 0, x_k, 0, \ldots, 0, 0)$ reduces to $l_k(x_k)$ (Armenti et al., 2015).
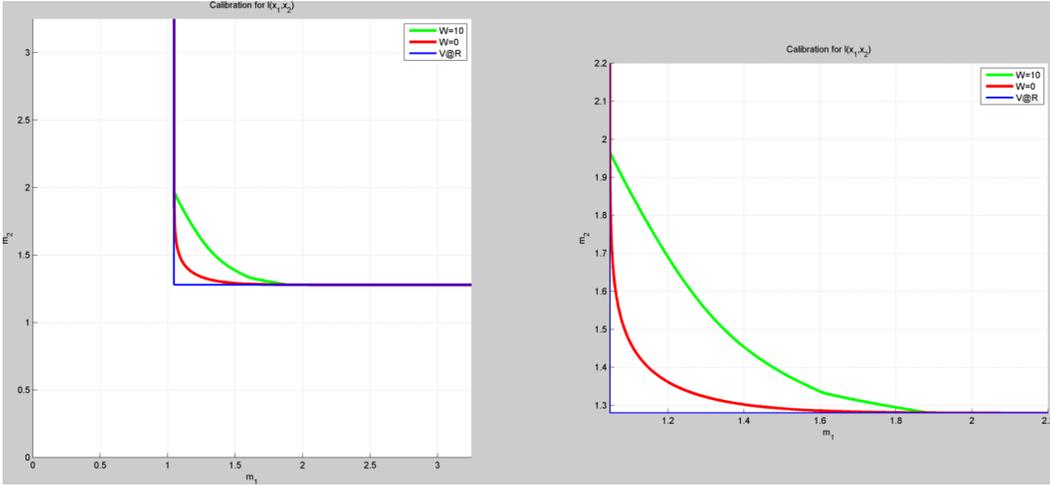
Figure 5: The initial calibration approach.

An obvious criticism of this approach is that the weightings applied to each risk factor result in the losses arising from different system components being treated unfairly. For example, simply swapping x_1 and x_2 will result in different weights applying to the risk factors and will consequently yield different output - clearly this is undesirable. Essentially, this implies that the loss function cannot be canonical and varies depending on the properties of its inputs. In addition, using the target idiosyncratic risk measures (in this case value at risk) directly as initial margins potentially places too much emphasis on this input data.

3.1 Pooled weighting approach

We propose a simple adjustment to the approach which both adjusts the weightings so as to allow for the risk factors x_k to determine losses. This also shifts the focus away from the $\widehat{im_k}$, treating them rather as a sensible first guess for initial margins. The adjustment is simply to calculate an average weighting based on the w_1 and w_2 calculated above, $\bar{w} = \frac{w_1 + w_2}{2}$, and then set the threshold level to $c = \frac{1}{\bar{w}}$. The resulting acceptance set hence solves the following equation:

$$\mathbb{E}[X_1^+ + X_2^+ + wX_1^+ X_2^+] \leq \frac{1}{\bar{w}}$$

14

and the results are shown in figure 6.
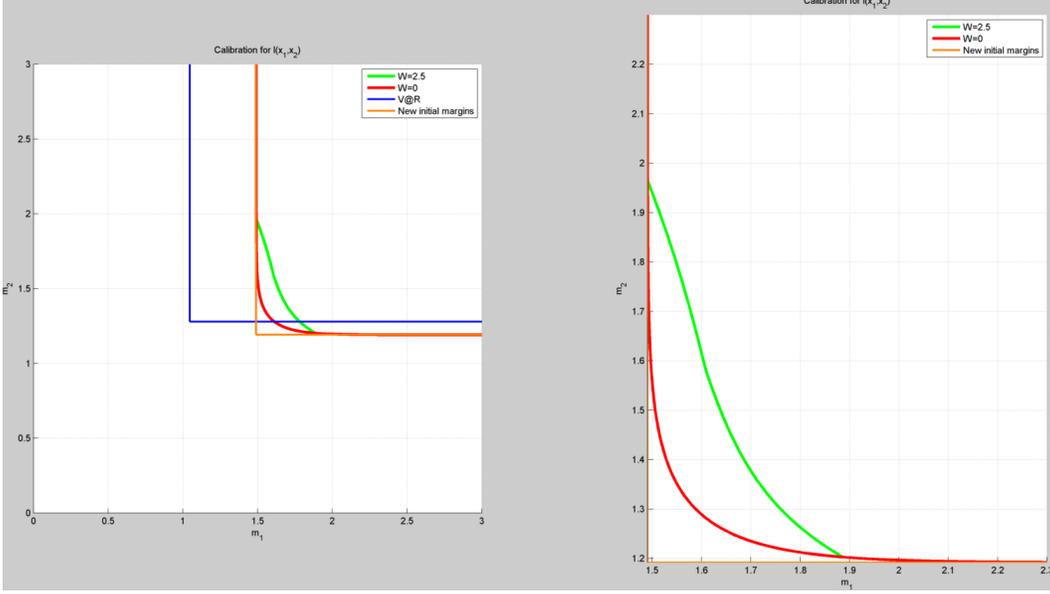
Figure 6: Pooled weighting approach.

3.2 Optimal threshold approach

An additional improvement to the approach which can be considered is to remove the dependence on the averaging of weights and simply solve for the optimum threshold level c directly. This makes intuitive sense as we would like to have that c is the primary factor in determining the level of margins required. The simplest method for doing so is to minimise the distance between our optimal margins m_k^* and the initial estimates $\widehat{im_k}$. Thus, for the simple bivariate example we use, we simply solve for c as follows:

$$\min_c \left[\left(m_1(c) - \widehat{im_1} \right)^2 + \left(m_2(c) - \widehat{im_2} \right)^2 \right]. \tag{3}$$

An extension to higher dimensions is obvious. Then setting $\bar{w} = \frac{1}{c}$ we again have a simple equation from which acceptance sets can be constructed. We note in the results shown in Figure 6 that, while this approach yields initial margins which are closer to the input value at risk values, the size of the default fund contribution and the range of acceptable w does not change significantly from the results in Figure

15

7.
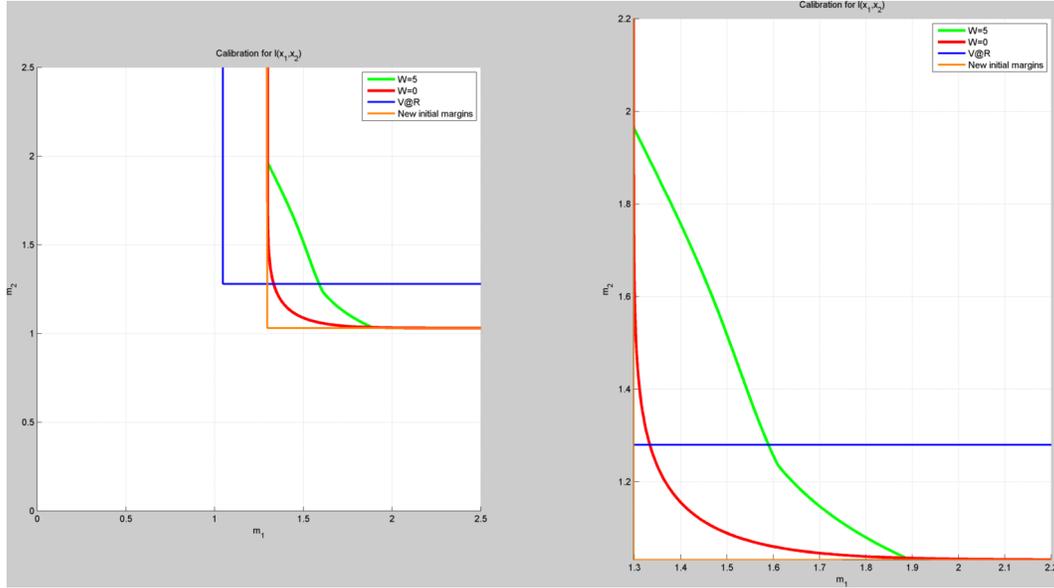
Figure 7: Optimum threshold approach.

3.3 Default fund contribution approach

We now propose an alternative approach which, in contrast to previous efforts, focusses on the determination of the default fund contribution df^* as opposed to the initial margins im^*. In order to do so, we define a new vector X^* such that $X^* = X - \widehat{im}$. Thus, rather than setting the margins m^* to our target values $\widehat{im}$ or minimising the distance between m^* and $\widehat{im}$ as in previous approaches, we solve the primal problem of $\mathbb{E}[l(X^* - m)] \leq c$ where our vector of risk factors X is simply reduced by the appropriate target values $\widehat{im}$. As a result, our model focusses exclusively on the size of the default fund contribution. Naturally this places more stringent limits on the potential values w we may apply, but gives us the freedom to use varying threshold values. Note the unique aspect of this approach as shown in Figure 8 where the calculated acceptance sets lie (obviously) well below the value at risk acceptance range. Here, arguments can be made for the idiosyncratic components of the acceptance sets to be added to value at risk to determine initial margin values versus simply leaving initial margins set at the target values.
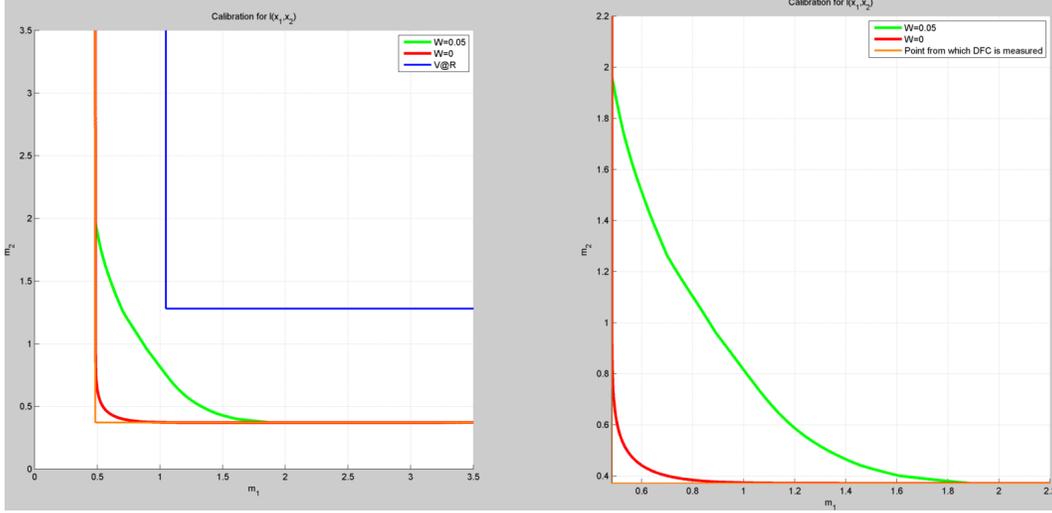
Figure 8: DFC focussed approach.

3.4 Multi-weighted approach

Our final and preferred approach for both calibration and margin determination is now discussed. This approach is particularly rich in that it has the symmetry of earlier approaches whilst also being calibrated well against the target values. For this section we expand our example loss function to include quadratic terms as follows:

$$l(x_1, x_2) = w_1 x_1^+ + w_2 (x_1^+)^2 + w_1 x_2^+ + w_2 (x_2^+)^2 + w x_1^+ x_2^+.$$

Now we have two parameters (w_1 and w_2) to calibrate, which gives us some more freedom to manipulate the model. These can be determined by solving the system of equations:

$$\begin{cases} w_1 \mathbb{E}[(X_1 - \widehat{im_1})^+] + w_2 \mathbb{E}[((X_1 - \widehat{im_1})^+)^2] = 1, \\ w_1 \mathbb{E}[(X_2 - \widehat{im_2})^+] + w_2 \mathbb{E}[((X_2 - \widehat{im_2})^+)^2] = 1. \end{cases}$$

We then choose to set our threshold level to give $c = \frac{1}{w_1}$, but arguments can be made for $c = \frac{1}{w_2}$ as well. As shown clearly in Figure 9 this approach is not only pinned closely to our target values $\widehat{im}$, but also allows us great flexibility in terms of default fund contribution size. The range of w for which convexity is maintained is significantly larger than in previous approaches and this is the primary reason why we strongly advocate this approach for calibration and determination of margins.

17
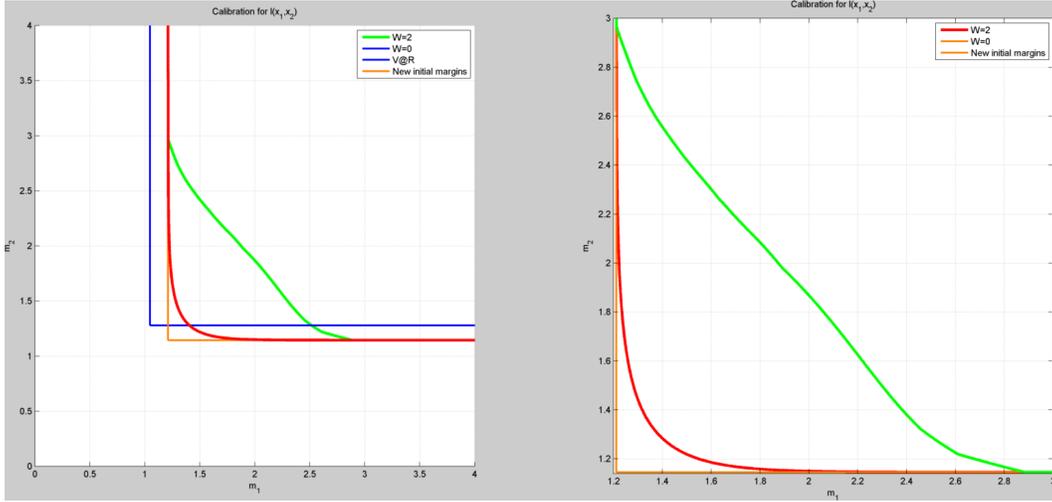
Figure 9: Final approach.

3.5 Table of results

Tables 1 and 2 are constructed using the following parameters:

$$\sigma_1 = 0.55, \quad \sigma_2 = 0.45, \quad \rho = 0.25, \quad \mu_1 = \mu_2 = 0.$$

Unfortunately, the multi-weighted approach required us to use non-zero means $\mu_1 = \mu_2 = 1$, preventing these results from being directly comparable with the earlier approaches. However, we hypothesise that for parameters chosen consistently our conclusions will still hold.

18

Table 1: Tables of results for selected parameters.

Approach	im_1	im_2	$w = 0$		$w = w_{max}$	
			df_1	df_2	df_1	df_2
Basic calibration	1.05	1.28	0.12	0.10	0.31	0.27
Pooled weighting	1.50	1.20	0.09	0.10	0.18	0.20
Optimum threshold	1.29	1.01	0.11	0.15	0.24	0.36
Default fund contribution	1.61	1.48	0.09	0.12	0.42	0.50
Multi-weighted	1.20	1.14	0.18	0.22	0.68	0.73

Table 2: Tables of results for selected parameters continued.

Approach	$w = 0$		$w = w_{max}$	
	$im_1 + df_1$	$im_2 + df_2$	$im_1 + df_1$	$im_2 + df_2$
Basic calibration	1.17	1.38	1.36	1.55
Pooled weighting	1.59	1.30	1.68	1.40
Optimum threshold	1.40	1.16	1.53	1.37
Default fund contribution	1.70	1.60	2.03	1.98
Multi-weighted	1.38	1.36	1.88	1.87

A number of important interpretations can be made from tables 1 and 2. Firstly, the basic calibration clearly places too much emphasis on the value at risk and, as a result, produces margins which are in contrast to all other approaches. Secondly, the default fund contribution approach leads to significantly larger initial margins as a result of the additional idiosyncratic elements added to V@R. Finally, the multi-weighted approach (which we previously claimed to be the approach we find most appealing) provides significantly greater freedom with respective to the default fund contribution. The range of values of w for which l maintains convexity is significantly larger than for the other approaches and this is the the primary criterion which makes this approach preferable.

4 Optimal liquidation strategies

We now consider the problem of calculating the initial margin of a single member of the CCP. As in the previous sections, we want to choose this margin to be as low as possible while keeping the level of systemic risk, as measured using loss functions, acceptable. We now extend the model to include a multiple-day liquidation horizon.

4.1 Problem specification

Let $X = \{X_t : t = 1, 2, 3, \ldots\}$ be a stochastic process representing the profit and loss (P&L) of the portfolio of the member. From now onwards, we will refer to the member as X. We assume that X has positions of size $Q = (Q_1, Q_2, \ldots, Q_n)$ in n assets whose mark-to-markets (MTM) are

$$\{(\mathrm{MTM}^1(t), \mathrm{MTM}^2(t), \ldots, (\mathrm{MTM}^n(t))) : t = 0, 1, 2, \ldots\}.$$

Finally we define the P&L of each position by Y^i, i.e.

$$Y_t^i = \mathrm{MTM}^i(t) - \mathrm{MTM}^i(0) \quad t \geq 1.$$

With this notation, X's total P&L is given by

$$X_t = \sum_{i=1}^{n} Q_i Y_t^i$$

for $t = 1, 2, \ldots$. We want to choose the initial margins $m = (m_1, m_2, \ldots, m_n)$ charged for each contract to minimise the total margin charged to X, while keeping the risk at an acceptable level. To this end, we assume that X has defaulted on at least one contract, and the CCP has to liquidate the remaining position X by a certain time T.

At first it seems better for the CCP to liquidate all the positions on the first day in order to eliminate any future risks. However, this may not be possible due to liquidity constraints and price movements caused by large transactions. Also, some price trends may motivate for a non-trivial liquidation strategy.

We will denote a liquidation strategy by $q = \{q_t^i : t = 1, 2, \ldots, T, i = 1, 2, \ldots, n\}$. Here q_t^i represents the proportion of asset i which is liquidated (sold) on day t. This means that $q_t^i Q_i$ units of asset i are liquidated at time t; we therefore have

$$\sum_{t=1}^{T} q_t^i = 1 \quad \text{for every } i = 1, 2, \ldots, n.$$

20

Let $\mathcal{C}$ denote the set of all admissible liquidation strategies. For now we will assume that there are no external trading constraints on the strategies and therefore take $\mathcal{C}$ to be the simplex

$$\mathcal{C} = \{q = (q_t^i) : q_t^i \in [0,1] \text{ and } \sum_{t=1}^{T} q_t^i = 1 \text{ for } i = 1, 2, \ldots, n\}.$$

Given a strategy $q \in \mathcal{C}$ and for each $t = 1, 2, \ldots, T$ we define the realised profit $R_t^i(q)$ by

$$R_t^i(q) := \sum_{s=1}^{t} Q_i q_s^i Y_s^i \text{ for } i = 1, 2, \ldots, T,$$

and the unrealised profit by

$$U_t^i(q) := Q_i \left(1 - \sum_{s=1}^{t} q_s^i\right) Y_t^i = Q_i \left(\sum_{s=t+1}^{T} q_s^i\right) Y_t^i.$$

The total profit at time t is defined by $P_t^i(q) := R_t^i(q) + U_t^i(q)$ for every position i. We note that $P_t^i(q)$ can be written as

$$P_t^i(q) = \sum_{s=1}^{t} Q_i q_s^i Y_s^i + Q_i \left(\sum_{s=t+1}^{T} q_s^i\right) Y_t^i = \sum_{s=1}^{T} Q_i q_s^i Y_{t \wedge s}^i.$$

For risk management purposes, it will be useful to work with the loss process $L(q) = \{L_t^i(q) : i = 1, 2, \ldots, n \text{ and } t = 1, 2, \ldots, T\}$ defined by $L_t^i(q) = -P_t^i(q)$. So from now onwards, we will deal with the process $L = L(q) = (L^1(q), \ldots, L^n(q))$.

In the literature, different functionals of L have been used to measure the riskiness of a strategy q. Some include the following random vectors

1. The worst loss:

$$\max_{1 \leq t \leq T} L_t(q) = (\max_{1 \leq t \leq T} L_t^1(q), \ldots, \max_{1 \leq t \leq T} L_t^n(q)).$$

2. The terminal loss:
$$L_T(q) = (L_T^1(q), \ldots, L_T^n(q)).$$

3. The average loss:
$$L_{\text{av}} := \left(\sum_{t=1}^{T} L_t^1(q), \ldots, L_t^n(q)\right).$$

For more information on these risk measures, see Avellaneda and Cont (2013). For presentation purposes, we will choose the terminal loss $L_T(q)$ as a risk measure. The main reason for this choice is simplicity, since for each i we have that

$$L_T^i(q) = \sum_{t=1}^{T} Q_i q_t^i Y_t^i,$$

implying that L_T^i is simply a linear combination of the Y^i's. This will later be useful for the models we use.

We now assume a loss function l. Fix a strategy $q \in \mathcal{C}$ and let $m = (m_1, \ldots, m_n)$ be a vector of margins charged for each contract (for $i = 1, 2, \ldots, n$). Again choosing a risk level $c > 0$, we say that m is acceptable at level c for the strategy q if and only if

$$\mathbb{E}\left(l(L_T(q) - m)\right) \leq c.$$

For each $q \in$, define $A_q(X)$ to be the set of all margins m that are acceptable for the strategy q, i.e.

$$A_q(X) := \{m \in \mathbb{R}^n : \mathbb{E}\left(l(L_T(q) - m)\right) \leq c\}.$$

Our aim is to find a strategy $q^* \in \mathcal{C}$ and margins $m^* \in A_{q^*}(X)$ that minimise the total margins

$$\sum_{i=1}^{n} m_i.$$

This can be interpreted as finding $q^* \in \mathcal{C}$ and $m^* \in A_{q^*}(X)$ such that

$$\sum_{i=1}^{n} m_i^* = \inf_{q \in \mathcal{C}} \left(\inf_{m \in A_q} \left(\sum_{i=1}^{n} m_i \right) \right).$$

4.2 Liquidity constraint

In the previous subsection we discussed a method for optimising the liquidation strategy of the portfolio of a defaulting clearing member, without considering a possible price impact on the strategy.

Our first attempt to model this effect is to include a daily market liquidity constraint associated with the instruments constituting the defaulting member's positions (see Avellaneda and Cont (2013)).

Definition 4.1 (Daily market-liquidity constraint)**.** The daily liquidity constraint represents an upper bound on the number of contracts that can be traded, long or short, without impacting the price.

Let $\ell_1, \ldots, \ell_n$ be the daily liquidity constraint corresponding to each position. Recall that in our model the traded volume of the i-th position at time t is represented by the quantity $Q_i q_t^i$. We then define the set $\widetilde{\mathcal{C}} \subseteq \mathcal{C}$ of all admissible liquidation strategies which satisfy the daily liquidity constraint. A liquidation strategy in this set satisfies

$$\sum_{t=1}^{T} q_t^i = 1 \qquad \text{(positions are liquidated before time } T\text{),}$$

$$0 \leq q_t^i \leq \ell_i/Q_i \qquad \text{(liquidity constraints are satisfied),}$$

for all $i = 1, \ldots, n$. The optimisation problem reduces to finding $q^* \in \widetilde{\mathcal{C}}$ and $m^* \in A_{q^*}$ such that

$$\sum_{i=1}^{n} m_i^* = \inf_{q \in \widetilde{\mathcal{C}}} \left(\inf_{m \in A_q} \left(\sum_{i=1}^{n} m_i \right) \right).$$

A further potential approach to modelling a liquidity constraint is to directly model the price impact of a liquidation strategy on the MTM of the defaulting clearing member's constituent positions . In this sense Bertsimas et al (see Bertsimas and Lo (1998)) present a model whereby the dynamics of the price include two distinct components: an arithmetic random walk modelling the price in the absence of a trade, and the impact of the trade.

4.3 Example

We provide an example to illustrate the proposed methodology.

Assume that the portfolio of the member is composed of only two positions, the P& L processes thereof being denoted by $Y^i = \{Y_t^i : t = 1, \ldots, T\}$, for $i = 1, 2$. Suppose that $Y^i = \sigma_i W^i$ for all i, where W^1, W^2 is a pair of Brownian Motions satisfying $\text{Cov}(W^1, W^2) = \rho\sigma_1\sigma_2$. This structure allows us to easily describe the distribution of the terminal loss $L_T(q) = (\sum_{t=1}^{T} Q_1 q_t^1 Y_t^1, \ldots, \sum_{t=1}^{T} Q_n q_t^n Y_t^n)$. Setting $q^i := (q_1^i, \ldots, q_T^i)'$ and $(A_T)_{ij} = i \wedge j$ we indeed obtain that

$$L_T(q) \sim \mathcal{N}(0, \Sigma),$$

where

$$\Sigma_{ii} = \sigma_i^2 Q_i^2 (q^i)' A_T q^i, \ i = 1, 2 \quad \text{and} \quad \Sigma_{21} = \Sigma_{12} = \rho\sigma_1\sigma_2 Q_1 Q_2 (q^1)' A_T q^2.$$

Consider the loss function proposed in Chapter 3 given by

$$l(x) = \frac{1}{1+\alpha} \left(\frac{1}{2} e^{2x_1} + \frac{1}{2} e^{2x_2} + \alpha e^{x_1+x_2} \right) - 1,$$

and recall that, in this framework, the sum of the corresponding optimal allocation is given by:

$$R(\sigma_1, \sigma_2, \rho, q) := R(L_T(q)) := \Sigma_{11} + \Sigma_{22} + \ln\left(1 + \alpha e^{\Sigma_{12} - \frac{1}{2}(\Sigma_{11} + \Sigma_{22})}\right) - \ln((c+1)(1+\alpha)).$$

Figures 10, 11, and 12 illustrate the differences between the optimal strategy, determined by the methodology above, and more naïve ones, such as "liquidate as soon as possible", "liquidate as late as possible" or "liquidate the same amount every day". We also compare optimal strategies assuming different daily liquidity constraints. The parameters for figures 10, 11, and 12 are:

$$\sigma_1 = 0.1, \quad \sigma_2 = 0.5, \quad Q_1 = 20, \quad Q_2 = 20.$$
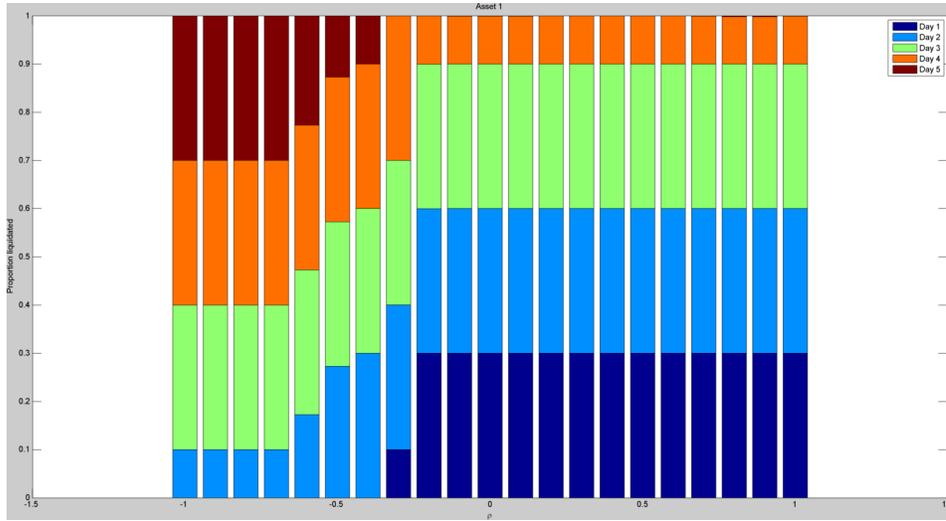
Figure 10: Optimal close out strategy assuming a daily market constraint of the 30% of the initial quantity.

Figures 10, 11, and 12 illustrate that within our framework, the optimal liquidity strategy coincides with "liquidate as soon as possible" for large enough ρ. What is of interest is that this is not the case when the two underlying Brownian Motions are strongly negatively correlated. In this case, the optimal liquidation strategy as determined by our model is strictly better than "liquidate as soon as possible". It is also important to note that a more restrictive liquidity constraint improves the performance of a more staggered liquidation strategy.

Figure 13 confirms the sub-optimality of the "liquidate as soon as possible" liquidation strategy for ρ close to -1. It also illustrates how $\sum m_i^*$ increases with the correlation.
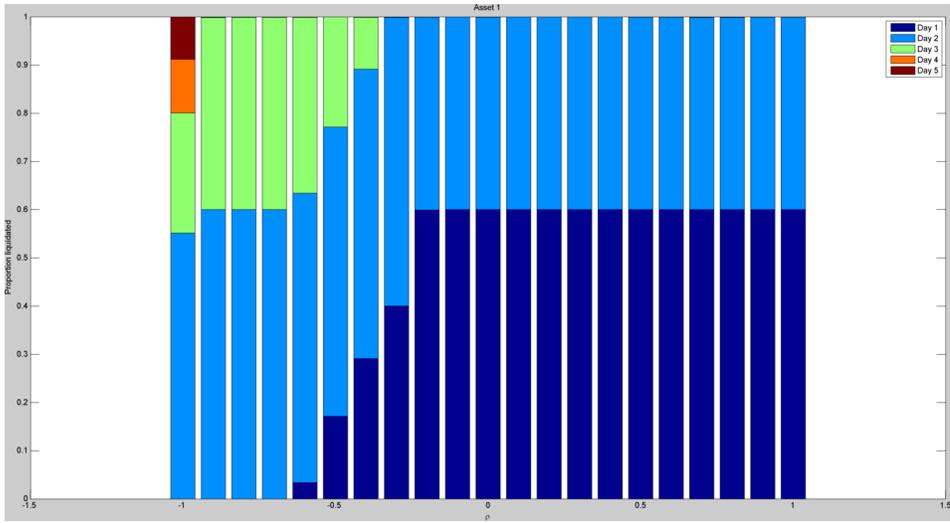
Figure 11: Optimal close out strategy assuming a daily market constraint of the 60% of the initial quantity.
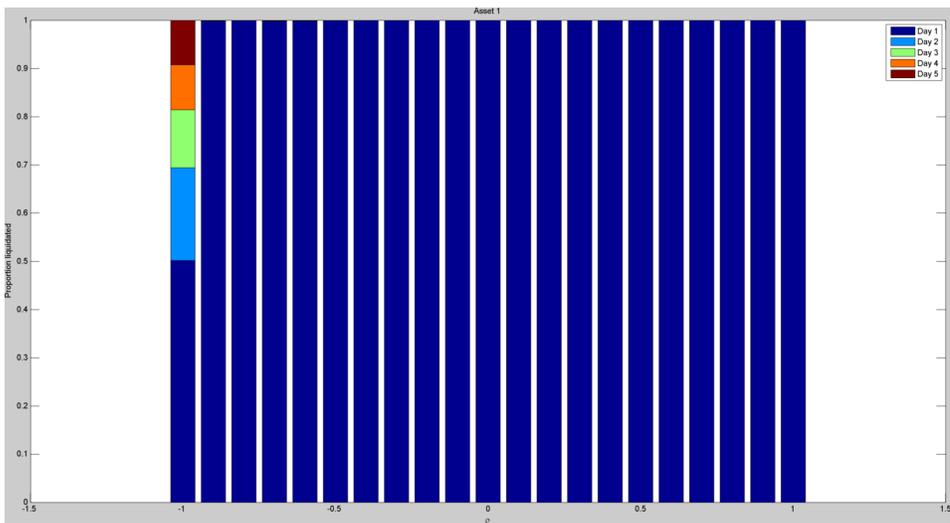
Figure 12: Optimal close out strategy assuming no daily market constraint.
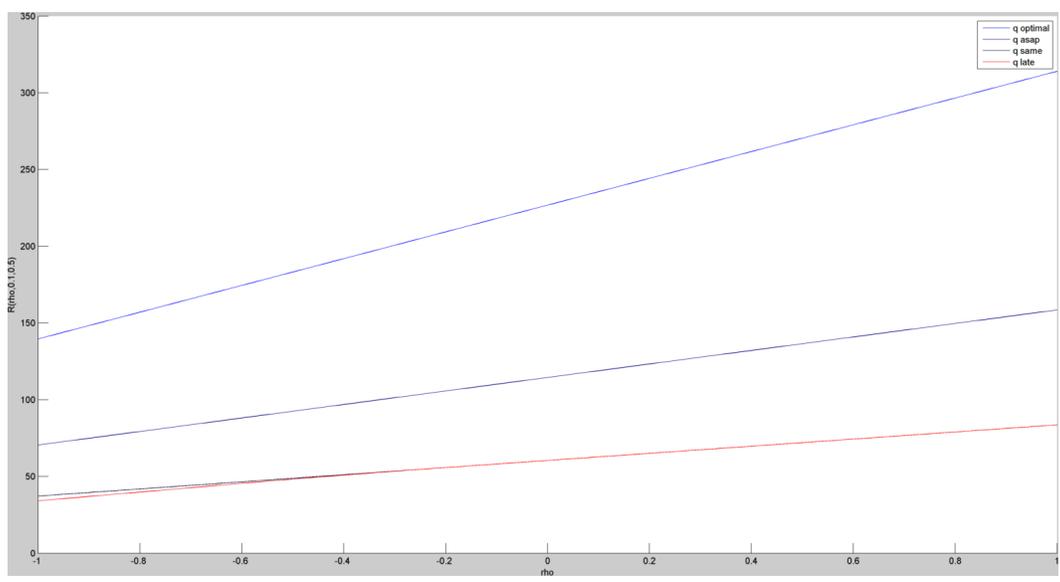
Figure 13: Sum of optimal allocation following the strategies "liquidate as late as possible", "liquidate the same amount every day", "liquidate as soon as possible", and the optimal one.

5 Conclusion and discussion

In this report, we have outlined strategies for optimising margins in a clearing counterparty context, whilst also including a discussion on optimal liquidation strategies which take into account liquidity constraints. The flexibility of our model - specifically with regard to the possible choices for the underlying distributions and loss functions - allows for both an analytical approach and a numerical approach. In particular, a number of methodologies for calibrating the model are proposed each of which flows naturally into a potential scheme for allocating margins.

Our investigation can be extended to the issue of pro-cyclicality. CCPs may adjust initial margin demands in response to changes in market conditions. While perhaps pertinent to their own risk management, these margin changes could have a destabilising effect on clearing members if sufficiently large. A CCP may, for example, increase initial margin requirements in response to increased price volatility. This could occur if initial margin requirements were established under less volatile market conditions, necessitating a large increase in margin requirements when volatility increases. This increased obligation may force clearing members to liquidate their positions, or seek other sources of funding to meet margin calls. This can further impact price volatility at a time when markets may already be illiquid and credit controls are tight. It has been suggested that a better strategy is for margins to remain at higher levels in good times, even if this puts them above the regulatory minimum requirements (Rehlon and Nixon, 2013).
Suggestions for improvement on the model in order to avoid this pro-cyclical effect, include stochastic modelling of the optimal threshold level c as a function of volatility of the clearing members as well as putting bounds on the amount of margins that the CCP can demand.

Open questions

Hereunder, we propose a list of research questions which have come up during the work presented int his report.

- The calibration procedure is limited to a class of loss functions with a specific set of properties (described in the relevant chapter). Future research should look into finding an analogous method for other forms of loss functions.

- The liquidity constraint proposed is modeled by an upper bound on the number of contracts that can be traded per day without impacting the price. One then considers as admissible only the strategies where traded volume per day remains below this threshold. This assumption is very restrictive and one should consider a price impact proportional to both the traded quantity and the price, without restricting the set of admissible strategies.

- The optimal liquidation strategy proposed focuses on a single clearing member only. Can it be improved by considering the simultaneous liquidation of an arbitrary number of members?

- As mentioned in the conclusion, our model does not consider the pro-cyclicality issue. Is it reasonable to improve it in this sense by putting bounds on the amount of margins that the CCP can demand? From the CCP's point of view, it may be more convenient to model the optimal threshold level c as a stochastic process. In particular, is it reasonable to model it as function of the volatility of its members?

Bibliography

Armenti, Y., Crepey, S., Drapeau, S., Papapantoleon, A., 2015. Multivariate shortfall risk allocation. Working Paper.

Avellaneda, M., Cont, R., 2013. Close-out risk evaluation (core): a new risk management approach for central counterparties. Available at SSRN 2247493.

Bertsimas, D., Lo, A. W., 1998. Optimal control of execution costs. Journal of Financial Markets 1 (1), 1–50.

Feinstein, Z., Rudloff, B., Weber, S., 2015. Measures of systemic risk. arXiv preprint arXiv:1502.07961.

InteDelta, 2013. Initial margin: A commentary on issues for centrally cleared and non-cleared business. White Paper.

Rehlon, A., Nixon, D., 2013. Central counterparties: What are they, why do they matter, and how does the bank supervise them? Bank of England Quarterly Bulletin, Q2.

Linear Commodity Models with Unspanned Stochastic Volatility

Team 3

THANOS ALEVIZOS, University College London
MICHAEL KATTEREGA, University of Cape Town
DIVANISHA PILLAY, University of Cape Town
RALPH RUDD, University of Cape Town

Supervisor:
MARTIN LARSSON, ETH Zürich

Contents

1 Introduction

Energy and commodity markets have grown in complexity and sophistication over the past decade, with a range of derivatives on futures being liquidly traded. Trading in such contracts has become an integral part of operations. When developing models for these derivatives, spot price, the term structure of futures prices, as well as derivatives prices should ideally be treated within a single framework in a consistent manner. In addition, a sound understanding of the dynamics of volatility in commodity markets is key in pricing, hedging and risk-managing commodity options.

The extent to which volatility is spanned refers to the extent to which volatility risk can be hedged by trading in the underlying commodities themselves or their corresponding futures, forwards or swap contracts (Trolle and Schwartz, 2009). In the presence of unspanned stochastic volatility, options cannot be completely hedged and risk-managed by trading in only the underlying instruments. Trolle and Schwartz (2009) develop a framework, based on Heath, Jarrow, and Merton (1992), which incorporates unspanned stochastic volatility in the pricing of commodity derivatives. They estimate the model on an extensive data set for New York Mercantile Exchange (NYMEX) crude oil derivatives; and find that two volatility factors, which are largely unspanned by futures contracts, are required to fit options on futures. This makes it difficult to find a model that provides an accurate representation of the temporal dynamics observed in time series of prices.

This paper develops a model for commodity spot, futures, and option prices, with a view towards accommodating USV. It also details estimation and implementation procedures for the model. The modeling approach draws on the linear-rational framework developed by Filipovic et al. (2014) for the term structure of interest rates. Filipovic et al. (2014) introduce a new class of term structure models, the linear rational. The linear rational model is highly tractable and easily incorporates unspanned volatility factors affecting the volatility of bond prices. Filipovic et al. (2014) specify a multivariate factor process with linear drift and a state price density, which is a linear function of the current state. This specification means that bond prices and the short rate become linear-rational functions (ratios of linear functions) of the current state. An important feature of the framework they develop is that the martingale component of the factor process does not affect the term structure. This means that factors which affect the prices of interest rate derivatives without affecting bond prices can be easily included. They demonstrate that the state vector can be partitioned into factors that affect the term structure, factors that affect interest rate volatility but not the term structure, and factors that affect neither term structure nor volatility but may have an indirect effect on the distribution of future bond prices. This particular model is called the Linear-Rational Square-Root (LRSQ) model.

This paper derives pricing formulas for future and option prices through adap-

tation of the LRSQ model to the commodities framework. $\mathbb{P}$-dynamics are then obtained by specifying an appropriate market price of risk. Restriction of the model parameters so as to enforce unspanned stochastic volatility is investigated. Model price computation and trajectory simulation is performed using the developed model. In addition, quasi-maximum likelihood estimation using the Kalman filter is investigated.

2 Model

2.1 Affine Diffusion Processes

The linear-rational diffusion model has dynamics of the form

$$dX_t = \kappa(\theta - X_t)dt + \sigma(X_t)dW_t, \tag{1}$$

for some d-dimensional Brownian Motion W_t and some dispersion function $\sigma(x)$. What makes this process affine is that the drift and diffusion matrix, $a(x) = \sigma(x)\sigma(x)^T$, are affine in X_t.

   The aim in this section is to construct a large class of Linear Square-Root models LSQ(m, n) with m term factors and n unspanned stochastic volatility factors with $m \geq n$ and $m + n = d$. The LSQ model is introduced in the following section.

2.2 The Linear Square-Root Model

The linear commodity spot model is specified by a multivariate factor process X_t together with a linear (affine) spot price of the form

$$S_t = u^T X_t = u_1 X_{1t} + \cdots + u_d X_{dt}. \tag{2}$$

The factor process dynamics under the pricing measure $\mathbb{Q}$ is specified by

$$dX_t = \kappa(\theta - X_t)dt + dM_t^{\mathbb{Q}}, \tag{3}$$

where $\kappa \in \mathbb{R}^{d \times d}$ is the mean-reversion matrix, $\theta \in \mathbb{R}^d$ contains the levels of mean reversion, $M_t^{\mathbb{Q}}$ is a $\mathbb{Q}$-martingale. This report investigates the square-root factor process specified as

$$dX_t = \kappa(\theta - X_t)dt + \begin{pmatrix} \sigma_1\sqrt{X_{1t}} & & \\ & \ddots & \\ & & \sigma_d\sqrt{X_{dt}} \end{pmatrix} dW_t^{\mathbb{Q}}, \tag{4}$$

where $\sigma_i > 0$ $(i = 1, \cdots, d)$, and $W_t^{\mathbb{Q}} = (W_{1t}^{\mathbb{Q}}, \cdots, W_{dt}^{\mathbb{Q}})$ is a $d-$dimensional Brownian motion under $\mathbb{Q}$. The solution to (4) exists and is unique if and only if κ has

non-positive off-diagonal elements and $\kappa\theta \in \mathbb{R}_+$. Consequently, the solution satisfies $X_t \in \mathbb{R}_+$ for all $t \geq 0$ as shown in Filipovic (2009). In addition, any positive u_i in (2) can be normalised by coordinate-wise scaling to $u_i = 1$ without loss of generality. It is important to normalise u in order for the model to be identifiable.

Under this specification, it is possible to derive a number of useful properties:

- Futures prices are linear (affine) in the current state X_t, i.e.,

$$\mathbb{E}_{\mathbb{Q}}[S_T|\mathcal{F}_t] = A(T-t) + B(T-t)^\top X_t$$

  for some deterministic functions $A(\tau)$ and $B(\tau)$ with values in $\mathbb{R}$ and $\mathbb{R}^d$ respectively (Filipovic et al., 2014).

- Assuming a deterministic risk-free rate r, the price of a European call option on a futures contracts is given by

$$C(t, T, T', K) = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}[(F(T,T') - K)^+|\mathcal{F}_t],$$

  where T is the option expiry, T' the maturity date of the underlying futures contract, and K is the strike price. The conditional expectation takes the form

$$\mathbb{E}_Q[(a + b^T X_T)^+|X_t]$$

  for some deterministic values $a \in \mathbb{R}$ and $b \in \mathbb{R}^d$. Under the square-root factor specification, such expectations can be computed efficiently using transform methods (see Section 3).

2.3 Market Price of Risk and $\mathbb{P}$-Dynamics

In order to take advantage of the temporal information contained in price observations over time, these observations should not be considered as a sample from their risk-neutral ($\mathbb{Q}$) distribution, but from their distribution under the historical measure $\mathbb{P}$. A $\mathbb{R}^d$-valued market price of risk process of the following form was considered

$$\lambda_t = (\lambda_1\sqrt{X_{1t}}, \ldots, \lambda_d\sqrt{X_{dt}})^T, \tag{5}$$

and the historical measure $\mathbb{P}$ specified by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\int_0^t \lambda_s^T dW_s^{\mathbb{Q}} - \frac{1}{2}\int_0^t \|\lambda_s\|^2 ds\right). \tag{6}$$

Then $dW_t^{\mathbb{P}} = dW_t^{\mathbb{Q}} - \lambda_t dt$ is a Brownian motion under $\mathbb{P}$ which gives the $\mathbb{P}$-dynamics of the factor process as

$$dX_t = (\kappa\theta + (\mathrm{Diag}(\sigma \circ \lambda_t) - \kappa)X_t)dt + \mathrm{Diag}(\sigma \circ \sqrt{X_t})dW_t^{\mathbb{P}}. \tag{7}$$

Figure 1: Attainable shapes of the futures curve
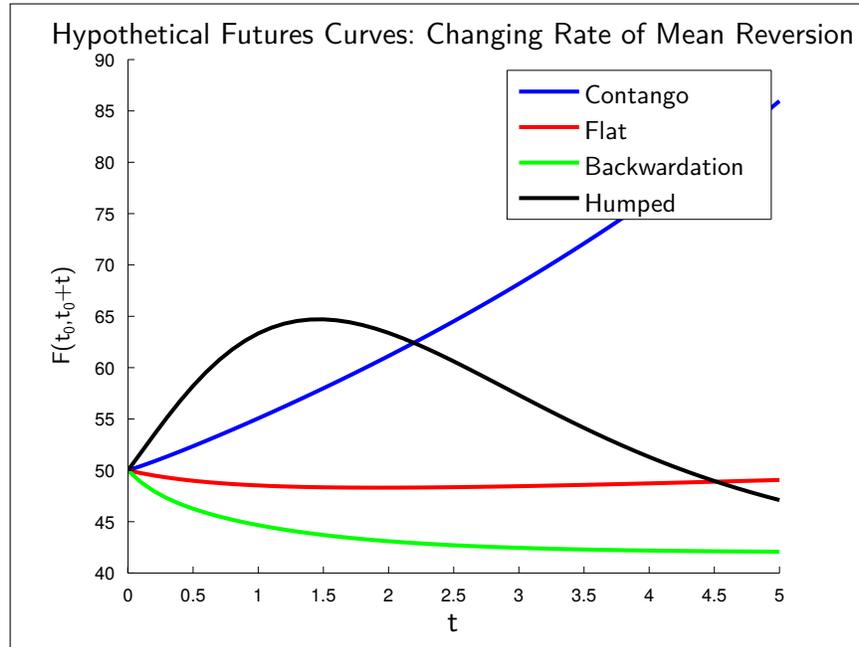
This specification ensures that the factor process remains an LSQ process under the change of measure, making estimation and implementation procedures which are widely reviewed in the literature applicable to this model.

2.4 Suitability to the Commodity Market

The behaviour of the commodity market and commodity term structure are characterised by four "stylised facts". The ability of the LSQ model to capture these effects was investigated. In particular, a five-factor LSQ model was used.

2.4.1 Backwardation and Contango

The commodity futures curve is in contango when it is upward-sloping and in backwardation when it is downward-sloping. Contango corresponds to a situation where the futures price of a commodity is above the expected future spot price of the commodity, and backwardation a situation where the futures price is below the expected future spot price. A humped futures curve is also common. The rate of mean reversion parameter, κ, was successfully adjusted in order to produce the different futures curve shapes observed in the market, see Figure 1.

2.4.2 Mean Reversion

Spot commodity markets show strong evidence of mean-reverting behaviour (Nielsen and Schwartz, 2004), largely believed to be the result of the dynamics of supply and demand interactions. Essentially, prices rise when shortages occur and this tends to raise the level of investments which will increase supply and decrease prices again (see Back and Prokopczuk (2013)). Whilst there are some studies which question the presence of mean reversion in some commodities (see Barkoulas et al. (1997)), it is generally accepted that mean reversion is an important feature of the commodities market (Pindyck, 2001).

The spot price process is simply a linear combination of mean-reverting factors. Therefore, under the LSQ specification, the spot commodity market displays mean reversion.

2.4.3 Seasonality

Seasonality is a common feature in certain commodities (see Geman (2009)). It is driven by supply side factors (seasonal production cycles, such as agricultural commodities) or demand side factors (such as an increase in demand in the USA for natural gas during the winter months). Beyond the seasonality in the price level and convenience yield, there is evidence of seasonality in the volatility corresponding to that exhibited in the spot.

In the LSQ model, deterministic seasonality could easily be incorporated through the inclusion of a deterministic time-dependent adjustment, $\Pi(t)$, as follows

$$S_t = \Pi(t)u^T X_t.$$

2.4.4 The Samuelson Effect

(Geman, 2009) describes the Samuelson Effect as the observation that, all else being equal, the volatility of futures prices tends to increase as the time to maturity decreases. It is believed that this is due to the increased sensitivity of the futures price to current information as it nears its time to maturity, as was originally proposed in Samuelson (1965).

Figure 2 was obtained by once again varying the rate of mean reversion parameter, κ. It can be seen that the five-factor LSQ specification does not always capture a decreasing volatility term structure. This opens the door to further research on whether further parameter restrictions are necessary in order to more consistently capture this decreasing volatility.

2.5 Unspanned Factors

To discuss the potential presence of unspanned factors, it is necessary to derive an expression for the futures term structure. The futures price is a $\mathbb{Q}$-expectation of

Figure 2: The Samuelson Effect illustrated
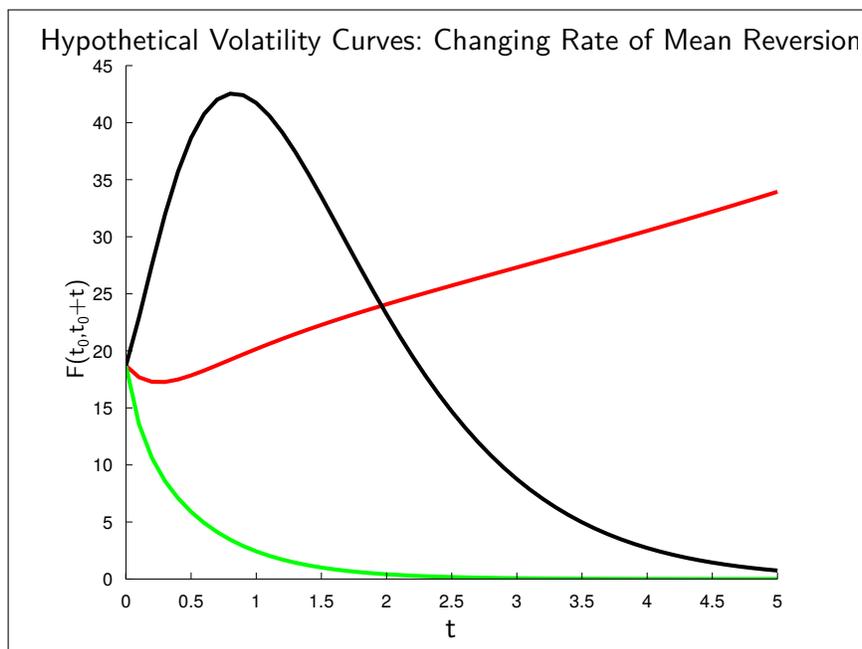
the spot price,

$$F(t, T) = E_{\mathbb{Q}}[S_T | \mathcal{F}_t].$$

The first moment of the LSQ model is explicitly derived in Section 4.2 and obeys the general form for first moments of affine processes, see Fisher and Gilles (1996). Thus,

$$F(t, T) = u^\top(\theta - e^{-\kappa(T-t)}\theta) + u^\top e^{-\kappa(T-t)} X_t \qquad (8)$$
$$= G(T - t, X_t), \qquad (9)$$

with

$$G(\tau, x) = u^\top(\theta - e^{-\kappa\tau}\theta) + u^\top e^{-\kappa\tau} x \qquad (10)$$

being a convenient formulation.

The intent is to describe a set, $\mathcal{U}$, of directions, $\xi \in \mathbb{R}^d$, such that the futures term structure is unchanged with respect to movements of X_t along ξ. All the results presented in this section are equivalent to those derived in Filipovic et al. (2014) Section 2.2, where the role of the interest rate term structure has been replaced by the futures term structure. This allows the specification of the state price density to be neglected and eventually leads to a different set of parameter constraints.

8

Definition 2.1 (Function Kernel). *Consider a differentiable function f on some space E, then the kernel of f is defined by*

$$\ker f = \left\{ \xi \in \mathbb{R}^d : \nabla f(x)^T \xi = 0 \text{ for all } x \in E \right\}.$$

Taking $f(x)$ as $G(\tau, x)$ in the above definition would imply that $G(\tau, x)$ is insensitive to movements of x along ξ. That is, the location of X_t along the direction ξ cannot be recovered by only considering information at time t of the futures price. As in Filipovic et al. (2014), the term structure kernel will now denote the set of all such directions.

Definition 2.2 (Term Structure Kernel). *The term structure kernel, denoted by $\mathcal{U}$, is given by*

$$\mathcal{U} = \bigcap_{\tau \geq 0} \ker G(\tau, \cdot) \tag{11}$$

Finally, Theorem 2.1 yields an expression for $\mathcal{U}$ in terms of the model parameters.

Theorem 2.1. *The term structure kernel, $\mathcal{U}$, is the largest subspace of $\ker u^\top$ that is invariant under κ. Equivalently,*

$$\mathcal{U} = span \left\{ u, \kappa^\top u, \cdots, \kappa^{(d-1)\top} u \right\}^\perp. \tag{12}$$

In the case where κ is diagonalizable, Proposition 2.2 provides the sufficient conditions on κ and u to ensure that there are no unspanned factors.

Proposition 2.2. *Assume κ is diagonizable with real eigenvalues, i.e. $\kappa = S^{-1}\Lambda S$ where S is invertible and Λ is diagonal and real. Then the term structure kernel is trivial, $\mathcal{U} = \{0\}$, if and only if all eigenvalues of κ are distinct and all components of $S^{-\top}u$ are non zero.*

The idea now is to construct an invertible linear transformation S on $\mathbb{R}^d$ for the state space such that the unspanned directions ξ correspond to the last components of the transformed state vector. The transformed process $\hat{X}_t = SX_t$ satisfies the linear drift dynamics

$$d\hat{X}_t = \hat{\kappa}(\hat{\theta} - \hat{X}_t)dt + dM_t^\mathbb{Q}, \tag{13}$$

where $\hat{\kappa} = S\kappa S^{-1}$, $\hat{\theta} = S\theta$, $\hat{M}_t = SM_t$ and in the LSRQ specification, $dM_t^Q = \text{Diag}(\sigma \circ \sqrt{X_t})dW_t^\mathbb{Q}$. Note that the transformed process is observationally equivalent to the original process.

If the linear transformation S can be constructed in such a way as to map the term structure kernel onto the last n components of $\mathbb{R}^d = \mathbb{R}^m \times \mathbb{R}^n$, it can be written as

$$S(\mathcal{U}) = \{0\} \times \mathbb{R}^n, \tag{14}$$

9

where $n = \dim \mathcal{U}$ and $m = d - n$. It can then be shown that the application of S to X_t allows the transformed process to be decomposed as $\hat{X}_t = (Z_t, U_t)$, where Z_t impacts the futures term structure, but U_t does not.

Theorem 2.3. *Let $m, n \geq 0$ be integers with $m + n = d$ and define $\hat{u} = S^{-\top}u$. Then Equation 14 holds if and only if the transformed model parameters satisfy:*

1. *$\hat{u} = (\hat{u}_Z, 0) \in \mathbb{R}^m \times \mathbb{R}^n$;*

2. *$\hat{\kappa}$ has block lower triangular structure, $\hat{\kappa} = \begin{pmatrix} \hat{\kappa}_{ZZ} & 0 \\ \hat{\kappa}_{UZ} & \hat{\kappa}_{UU} \end{pmatrix} \in \mathbb{R}^{(m \times n) \times (m+n)}$;*

3. *The upper left block $\hat{\kappa}_{ZZ}$ of $\hat{\kappa}$ satisfies span $\left\{ \hat{u}_Z, \hat{\kappa}_{ZZ}^\top \hat{u}_Z, \cdots, \hat{\kappa}_{ZZ}^{(m-1)\top} \hat{u}_{ZZ} \right\} = \mathbb{R}^m$.*

Proof. The proof for the equivalent theorem for linear-rational models for the term structure of interest rates can be found in Filipovic et al. (2014). Note that here the role of the state-price density coefficient is played by u. $\qquad \square$

Now suppose Equation 14 holds and write $Sx = (z, u) \in \mathbb{R}^m \times \mathbb{R}^n$ and $\hat{\theta} = (\hat{\theta}_Z, \hat{\theta}_U)$. Through simple substitution it is clear that

$$\hat{G}(\tau, z) = G(\tau, x) = \hat{u}_Z^\top [\hat{\theta}_Z + e^{-\hat{\kappa}_{ZZ}(\tau)}(z - \hat{\theta}_Z)] \tag{15}$$

does not depend on u. Thus $F(t, T) = \hat{G}(\tau, Z_t)$ which clearly illustrates that the components of U_t are *unspanned factors*. The realizations of U_t do not influence the futures term structure and cannot be recovered from the current futures curve. On the other hand, Z_t directly determines the term structure of futures and can be recovered from a realization of the futures curve. The components of Z_t will be denoted as *term structure factors*.

2.6 Unspanned Stochastic Volatility

The discussion on unspanned factors can now be specialized to those factors that give rise to unspanned stochastic volatility. This implies the presence of unspanned factors, which do not influence the term structure, but which still influence the volatility of the term structure.

Using Itō's Lemma, the volatility dynamics of any future can be derived. Let $F(t, T) = g(t, x; T)$ where

$$g(t, x; T) = u^\top [\theta - e^{\kappa(T-t)\theta}] + u^\top e^{-\kappa(T-t)}x.$$

The necessary partial derivatives are given by

$$\frac{\partial g(t, x; T)}{\partial t} = -\kappa e^{-\kappa(T-t)}u^\top \theta + \kappa e^{-\kappa(T-t)}u^\top x, \tag{16}$$

10

and

$$\frac{\partial g(t, x; T)}{\partial x} = u^\top e^{-\kappa(T-t)}. \tag{17}$$

The application of Itō's Lemma now yields

$$dg(t, x; T) = [-\kappa e^{-\kappa(T-t)} u^\top \theta + \kappa^{-\kappa(T-t)} u^\top X_t] dt +$$
$$+ u^\top e^{-\kappa(T-t)} \text{Diag}(\sigma \circ \sqrt{X_t}) dW_t^{\mathbb{Q}}. \tag{18}$$

Thus the volatility vector of $F(t, T)$ is given by

$$\nu(t, T) = u^\top e^{-\kappa(T-t)} \text{Diag}(\sigma \circ \sqrt{X_t}). \tag{19}$$

And finally, the squared volatility at time t of the future with maturity T is given by $\|\nu(t, T)\|^2 = V(T - t, X_t)$ with

$$V(\tau, x) = u^\top e^{-\kappa(\tau)} a(x) u e^{-\kappa^\top(\tau)}, \tag{20}$$

and

$$a(x) = \text{Diag}(\sigma^2 \circ x). \tag{21}$$

This allows for the definition of the *volatility kernel*.

Definition 2.3. *The volatility kernel, denoted by $\mathcal{W}$, is defined as*

$$\mathcal{W} = \bigcap_{\tau \geq 0} \ker V(\tau, \cdot) \tag{22}$$

It should be intuitively clear that the model will exhibit unspanned stochastic volatility if there are elements of the term structure kernel that do not lie in the volatility kernel. In a similar fashion to Section 2.5, the desire is to decompose the unspanned factors such that $U_t = (V_t, W_t)$ where movenents of W_t affect neither the term structure of futures or their volatiliies, whereas V_t will not affect the term structure of futures but will impact the volatilities. Thus V_t will be the *USV factors*, whereas W_t are known as *residual factors*.

This means that S must be an invertible linear transformation that satisfies Equation 14 and possesses the additional property that

$$S(\mathcal{U} \cap \mathcal{W}) = \{0\} \times \{0\} \times \mathbb{R}^q, \tag{23}$$

where $q = \dim \mathcal{U} \cap \mathcal{W}$ and $p + q = n = \dim \mathcal{U}$.

Although it is not immediately clear how this additional required property influences the parameter restrictrions required by Theorem 12, Filipovic et al. (2014) provides a canonical representation for linear-rational square-root models for the term structure with unspanned stochastic volatility and this representation is easily translated into the current context.

2.6.1 The Linear Square-Root Model

In this section, u is taken as the unit vector with no loss of generality and the mean-reversion matrix κ will be denoted in block form as

$$\kappa = \begin{pmatrix} \kappa_{II} & \kappa_{IJ} \\ \kappa_{JI} & \kappa_{JJ} \end{pmatrix}, \tag{24}$$

where κ_{IJ} denotes the submatrix whose rows are indexed by I and columns by J.

Definition 2.4 (LSR). *Fix nonnegative integers $m \geq n$ with $m + n = d$, representing the desired number of term structure and USV factors, respectively. The LSQ(m,n) specification is obtained by choosing $\kappa_{II} \in \mathbb{R}^{m \times m}$ with non-positive off-diagonal elements and such that*

$$span \left\{ u, \kappa_{II}^{\top} u, \cdots, \kappa_{II}^{(m-1)\top} u \right\} = \mathbb{R}^m. \tag{25}$$

The mean reversion matrix is defined by

$$\kappa = \begin{pmatrix} \kappa_{II} & \kappa_{II} A - A A^{\top} \kappa_{II} A \\ 0 & A^{\top} \kappa_{II} A \end{pmatrix}, \tag{26}$$

with $A \in \mathbb{R}^{m \times n}$ given by

$$A = \begin{pmatrix} \mathbb{I}_n \\ 0 \end{pmatrix}. \tag{27}$$

The level of mean reversion is taken to be a vector $\theta \in \mathbb{R}^d$ with $\kappa\theta \in \mathbb{R}^d_+$, and the volatility parameters are taken to be nonnegative, $\sigma_1, \cdots, \sigma_d \geq 0$.

Finally, Theorem 2.4 below provides the important result for this section: the further parameter restrictions required to ensure the desired number of USV factors in the LSQ(m,n) specification. Again, for a proof, refer to Filipovic et al. (2014).

Theorem 2.4. *The LSQ(m,n) specification exhibits m term structure factors and n unspanned factors. Assume that $u^{\top}\theta \neq 0$ and κ_{II} is invertible. Then the number of USV factors equals the number of indices $1 \leq i \leq n$ such that $\sigma_i \neq \sigma_{m+i}$. If $\sigma_i \neq \sigma_{m+i}$ for all $1 \leq i \leq n$ then every unspanned factor is a USV factor.*

An explicit algorithm for the construction of the appropriate linear transformation matrix S is outlined in Filipovic et al. (2014) Section 4. It is applied below to illustrate the LSQ(1,1) model. The effect on the futures volatility of moving along an unspanned direction is displayed for the LSQ(1,1) and LSQ(3,2) models in Figures 3 and 4 respectively. The effect on the European call price is illustrated in

Example 2.1. The LSQ$(1, 1)$ implies $m = n = 1$.
The second column of S^{-1} must form a basis for $\mathcal{U}$ to ensure that Equation 14 holds and the first column must be selected in such a way that the columns remain linearly independent.

Figure 3: LSQ(1,1) illustrated
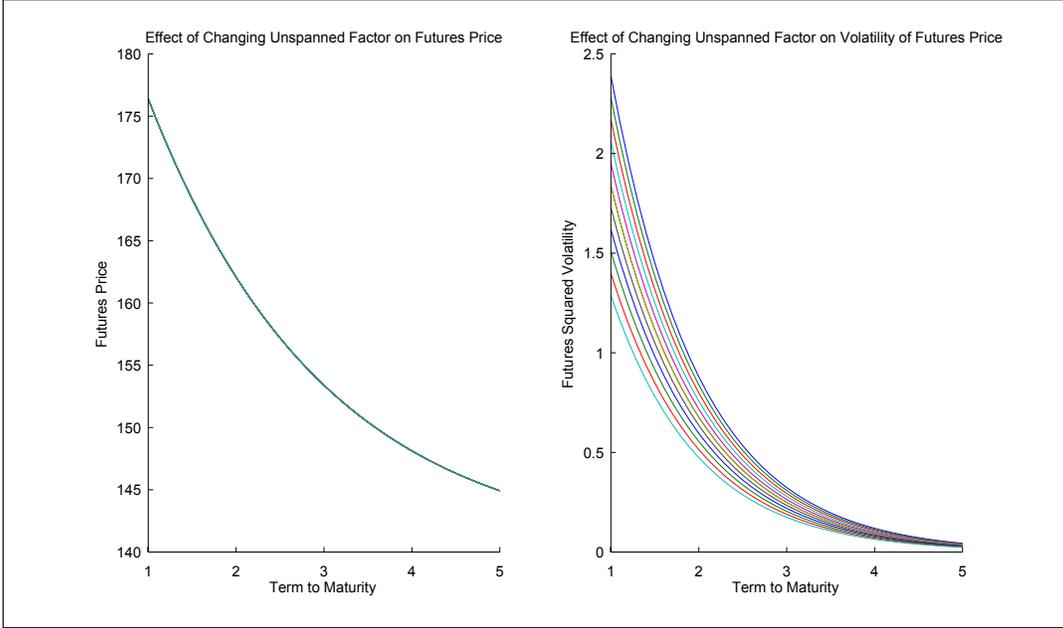
Observe that $\ker u^\top$ is spanned by vectors of the form $-e_i + e_j$, for $i < j$, where e_i denotes the i-th standard basis vector in $\mathbb{R}^d$. Choose $-e_1 + e_2$ as a basis for $\mathcal{U}$ and the first column of S^{-1} as e_1 such that

$$S^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{28}$$

The corresponding transformed process $\hat{X}_t = SX_t$ follows

$$SX_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} X_{1t} + X_{2t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} Z_t \\ U_t \end{pmatrix}, \tag{29}$$

where $Z_t = X_{1t} + X_{2t}$ is the term structure factor and $U_t = X_{2t}$ is the unspanned factor.

This transforms the volatility matrix as

$$S\mathrm{Diag}(\sigma \circ \sqrt{X_t}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \sigma_1\sqrt{X_{1t}} & 0 \\ 0 & \sigma_1\sqrt{X_{2t}} \end{pmatrix} = \begin{pmatrix} \sigma_1\sqrt{z-v} & \sigma_1\sqrt{v} \\ 0 & \sigma_1\sqrt{v} \end{pmatrix}, \tag{30}$$

and, similarly, the transform for κ is given by

$$\hat{\kappa} = S\kappa S^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{22} \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{22} \end{pmatrix}. \tag{31}$$

13

Figure 4: LSQ(3,2) illustrated
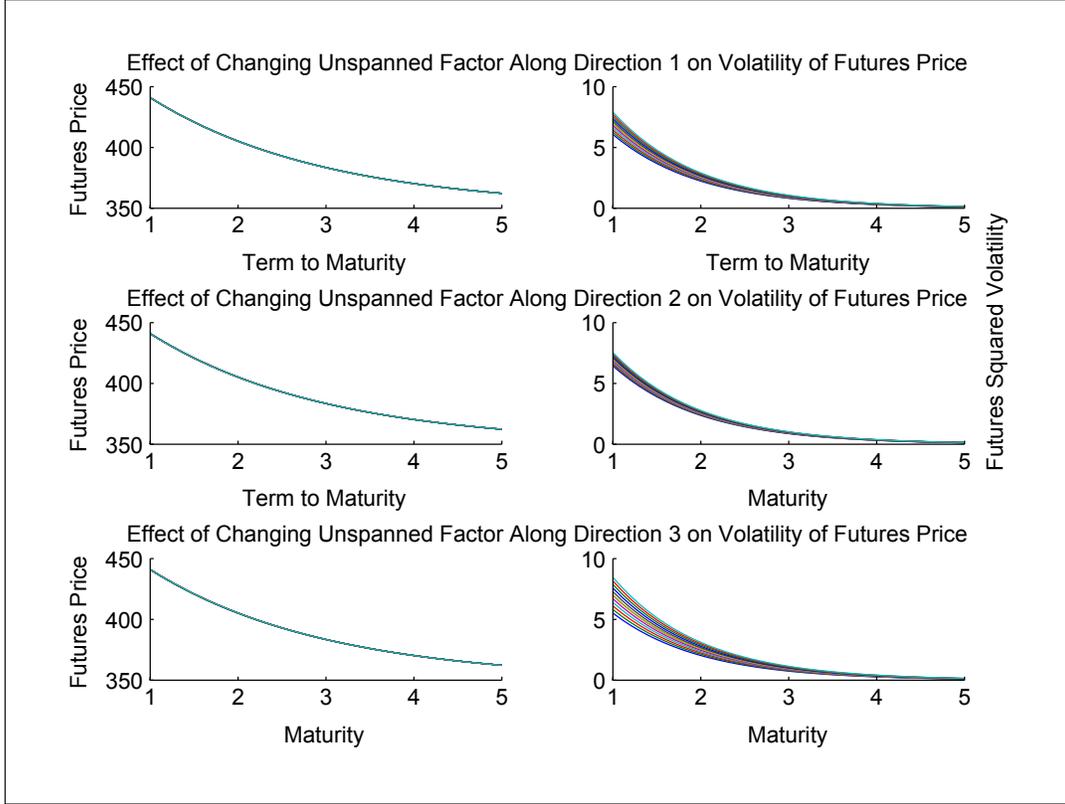
According to Theorem 2.4, V_t is a USV factor if $\sigma_1 \neq \sigma_2$, $\kappa_{11} \neq 0$, and $\mathbf{u}^\top \theta \neq 0$.

3 Pricing

3.1 Pricing Options on Futures

In this section the formula for a European call on a future is expounded for the LSR model.

Proposition 3.1. *Consider an European call option with maturity T and strike price K written on a futures contract with expiry T'. Given a deterministic risk-free rate r, the price of the option, denoted $C(t, T, T', K)$, is given by*

$$C(t, T, T', K) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(a + b^\top X_T)^+ | X_t]. \tag{32}$$

14

Figure 5: The effect on the European call price of moving along an unspanned direction in the LSQ(1,1) model
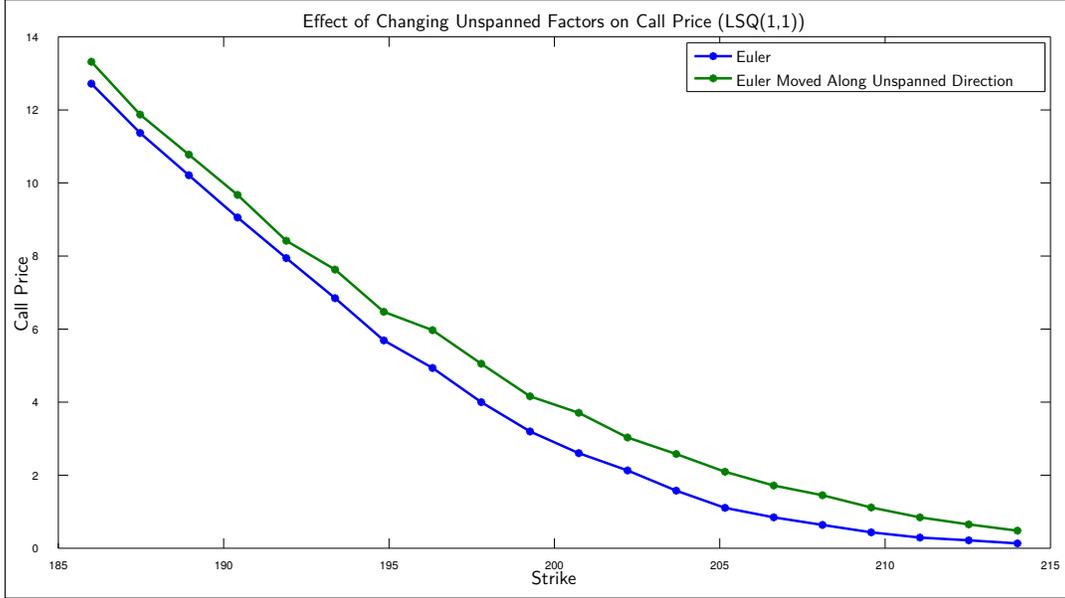
where

$$b^\top = u^\top e^{-\kappa(T'-T)}, \qquad (33)$$

and

$$a = u^\top[\theta - e^{-\kappa(T'-T)}\theta] - K. \qquad (34)$$

The expectation in Equation 32 requires the computation of a multidimensional integral. This could be done numerically if the conditional distribution of X_T were known. However, it is more efficient to use Fourier transform methods. The required result is summarized in the next theorem without proof.

Theorem 3.2. *Define*

$$\hat{q}(z) = E[e^{z(a+b^\top X_T)}|\mathcal{F}_t] \qquad (35)$$

for every $z \in \mathbb{C}$ such that the conditional expectation is well-defined. Choose any $\mu > 0$ such that $\hat{q}(\mu) < \infty$. Then

$$C(t, T, T', K) = \frac{1}{\pi}e^{-r(T-t)}\int_0^\infty Re\left[\frac{\hat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2}\right]d\lambda. \qquad (36)$$

15

Since X_t is a multifactor square-root process, $\hat{q}(z)$ can be computed using the exponential-affine transform formula (Duffie and Pan, 2000). It reproduced below for completeness.

Lemma 3.3. *For any $0 \leq t \leq T$, $u \in \mathbb{C}$ and $v \in \mathbb{C}^d$ such that $E[|e^{v^\top X_T}|] < \infty$,*

$$E[e^{u+v^\top X_T}|\mathcal{F}_t] = e^{\Phi(T-t)+\Psi(T-t)^\top X_t}, \tag{37}$$

where $\Phi : \mathbb{R}_+ \to \mathbb{C}$, $\Psi : \mathbb{R}_+ \to \mathbb{C}^d$ solve the system of Ricatti equations

$$\Phi'(\tau) = \theta^\top \kappa^\top \Psi(\tau) \tag{38}$$

$$\Psi_i'(\tau) = -\kappa_i^\top \Psi(\tau) + \frac{1}{2}\sigma_i^2 \Psi_i(\tau)^2, \qquad i = 1, \cdots, d \tag{39}$$

with initial conditions $\Phi(0) = u$, $\Psi(0) = v$ and where κ_i is the i-th column of the matrix κ.

3.2 Implementation of the Call Option Price

Pricing an European call option in the LSR model is complicated by the fact that the integral in Equation 36 requires a solution to an ODE system at each evaluation point. This slows pricing considerably and can lead to situations where parameter estimation becomes untractable. This section briefly introduces two techniques to aid speeding up the computation.

3.2.1 Gauss-Legendre Quadrature

A n-point Gaussian quadrature rule is a technique for finding points x_i and weights w_i such that an approximation of the form

$$\int_{-1}^{1} f(x)dx = \sum_{i=1}^{n} w_i f(x_i) \tag{40}$$

is exact for polynomials of degree less than or equal to $2n - 1$. It can be shown that the evaluation points x_i are the roots of a polynomial belonging to the class of orthogonal polynomials. Gauss-Legendre quadrature is the simplest case; where the associated polynomials are the Legendre polynomials. A standard implementation of Guass-Legendre quadrature was used in MATLAB to accelarate the evaluation of the integral in Equation 36. The effect on the computation time is illustrated in Figure 6.

3.2.2 Control Variate

Figure 7 illustrates the evaluation of the integrand for the European call price using Gauss-Legendre quadrature. A potential computational advantage could be

Figure 6: The effect of Guass-Legendre quadrature on option pricing
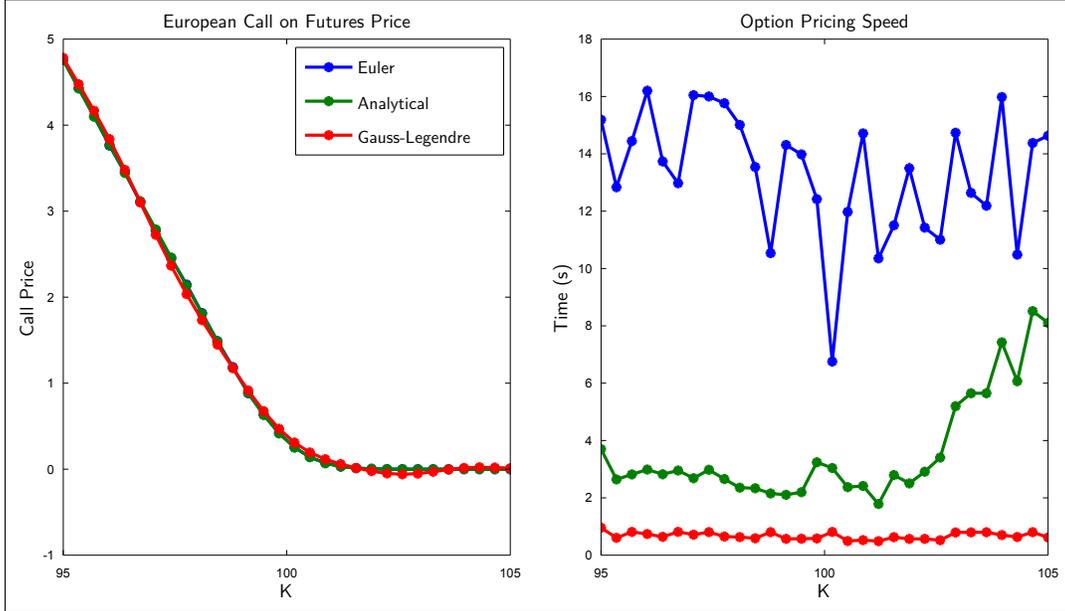
gained by attempting to reduce the size of the oscillations, which slow the convergence of the integral. Consider

$$E[(a + b^\top X)^+] = \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{\hat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2}\right] d\lambda. \tag{41}$$

A control variate technique would take the form

$$E[(a + b^\top X)^+] - E[Y^+] + E[Y^+] =$$
$$\frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{\hat{q}(\mu + i\lambda) - \hat{p}(\mu + i\lambda)}{(\mu + i\lambda)^2}\right] d\lambda + E[Y^+], \quad (42)$$

with Y an appropriately chosen scalar random variable. An option to be explored is choosing Y to be Gaussian with mean m and variance v, such that

$$m = E[a + b^\top X] \tag{43}$$
$$v = \mathrm{Var}(a + b^\top X) \tag{44}$$

and thus

$$\hat{p}(z) = e^{mz + \frac{1}{2}vz^2}. \tag{45}$$

17

Figure 7: The oscillations of the integrand
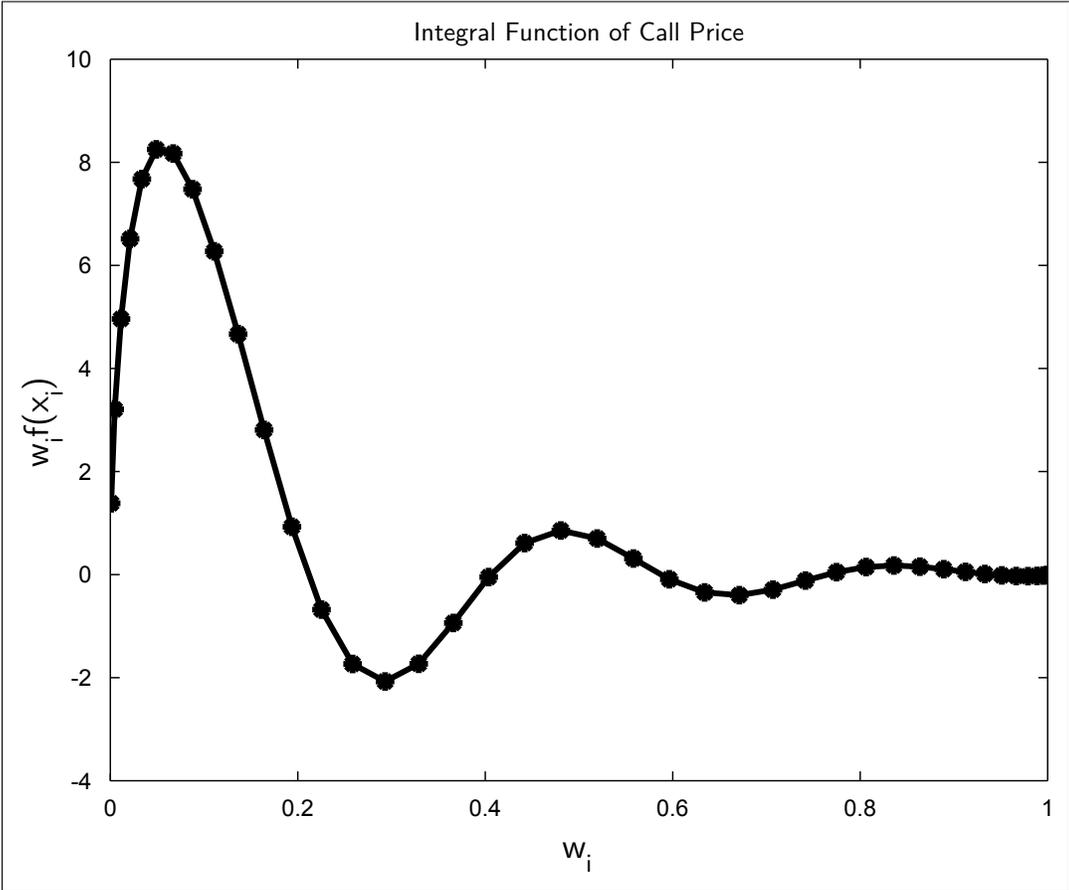
For this choice of Y, $E[Y^+]$ is the Bachelier option price, which is available in closed form.

4 Estimation

4.1 Quasi-maximum Likelihood Estimation

4.1.1 Overview

A quasi-maximum likelihood estimate (QMLE) is an estimate of a parameter vector Θ in a statistical model that is formed by maximizing a function that is related to the logarithm of the likelihood function, but it is not equal to it. In contrast, the

maximum likelihood estimate maximizes the actual log-likelihood function of the data and model.

The function that is maximized to form a QMLE is often a simplified form of the actual log-likelihood function. A common way to form such a simplified function is to use the log-likelihood function of a misspecified model that treats certain data values as being independent, even when in actuality they may not be. This removes any parameters from the model that are used to characterize these dependencies. This is only sensible if the dependency structure is a nuisance-parameter with respect to the goals of the analysis.

As long as the quasi-maximum likelihood function that is maximized is not overly simplified, the quasi-maximum likelihood estimate is consistent and asymptotically normal.

4.1.2   Estimation Procedure

The estimation approach is QML in conjunction with the extended Kalman filter (EKF). To apply the Kalman filter the model must be cast in state space form, which consists of a measurement equation and a transition equation.

The measurement equation describes the relationship between the state variables and the prices of futures and options, while the transition equation describes the discrete-time dynamics of the state variables.

Let X_t denote the vector of state variables and let Y_t denote the vector consisting of the price of futures and European call option prices. The general form of the transition equation is given by

$$X_{t+1} = \Phi_0 + \Phi_X X_t + w_{t+1} \tag{46}$$

with w_{t+1} and i.i.d random variable Φ_0, Φ_X can be computed in closed form. The measurement equation is given by

$$z_t = h(X_t; \Theta) + u_t, \tag{47}$$

with $u_t \sim N(0, \Omega)$ where z_t is the data vector, h is the pricing function and u_t is a vector of i.i.d Gaussian measurement errors with covariance matrix Ω.

Two assumptions are made to reduce the number of parameters in Ω:

1. All measurement errors are cross-sectionally uncorrelated, i.e Ω is a diagonal matrix.

2. A single variance applies to all measurement errors for the futures prices, and one more variance applies to all measurement errors for option prices.

This essentially states that futures and options prices of different terms are not driven by different sources of randomness.

4.1.3 Derivation of QML Function

Let $\hat{X}_t = E_t[X_t]$ and $\hat{X}_{t-1} = E_{t-1}[X_t]$ denote expectations of X_t and let P_t and $P_{t|t-1}$ denote the corresponding estimation error covariance matrices.

Linearizing the h-function in Equation 47 around $\hat{X}_{t-1}$, results in

$$z_t = \left(h\left((\hat{X}_{t-1}; \Theta) \right) - H_t' \hat{X}_{t-1} \right) + H_t' \hat{X}_t + u_t, \tag{48}$$

where

$$H_t' = \frac{\delta h(X_t)}{\delta X_t'} \big|_{X_t = \hat{X}_{t|t-1}}. \tag{49}$$

Assuming w_t in Equation 46 is i.i.d and Gaussian results in

$$X_t = \Phi_0 + \Phi_X X_{t-1} + w_t, \text{ with } w_t \sim N(0, Q_t). \tag{50}$$

The Kalman Filter applied to Equations 50 and 48 yields

$$\hat{X}_{t|t-1} = \Phi_0 + \Phi_X \hat{X}_{t|t-1}, \tag{51}$$

$$P_{t|t-1} = \Phi_X P_{t|t-1} \Phi_X' + Q_t, \tag{52}$$

$$\hat{X}_t = \hat{X}_{t|t-1} + P_{t|t-1} H_t' F_t^{-1} \epsilon_t, \tag{53}$$

and

$$P_t = P_{t|t-1} - P_{t|t-1} H_t' F_t^{-1} H_t P_{t|t-1}, \tag{54}$$

where the corresponding optimal predictor of z_t given information at $t-1$ is

$$z_{t|t-1} = h(\hat{X}_{t|t-1}) \tag{55}$$

and

$$\epsilon_t = z_t - \hat{z}_{t|t-1}$$
$$= z_t - h(\hat{X}_{t|t-1}) \tag{56}$$

is the prediction error. The dispersion matrix (i.e. the covariance matrix) of ϵ is

$$E[\epsilon_t \epsilon_t^\top] = F_t = H_t P_{t|t-1} H_t' + \Omega, \tag{57}$$

where the Kalman filter produces one-step-ahead forecasts for z_t, $\hat{z}_{t|t-1}$ and the corresponding error covariance matrices $F_{t|t-1}$.

If it assumed that X_t is Gaussian, imposing that $w_t \sim N(0, Q_t)$ and linearizing the h-function results in z_t following a Gaussian distribution with mean

$$E_{t-1}[z_t] = h(\hat{X}_{t|t-1}) - H_t' \hat{X}_{t|t-1} + H_t' E_{t-1}[X_t] \tag{58}$$

$$= h(\hat{X}_{t|t-1}) - H_t' \hat{X}_{t|t-1} + H_t' \hat{X}_{t|t-1} \tag{59}$$

$$= h(\hat{X}_{t|t-1}) \tag{60}$$

$$= \hat{z}_{t|t-1}, \tag{61}$$

and variance given by

$$\text{Var}(z_t) = \text{Var}\left(h(\hat{X}_{t-1}; \Theta) - H_t' \hat{X}_{t-1} + H_t' \hat{X}_t + u_t\right) \tag{62}$$

$$= \text{Var}\left(H_t' X_t + u_t\right) \tag{63}$$

$$= H_t \text{Var}(X_t) H_t' + \text{Var}(u_t) + 2\text{Cov}(H_t' X_t, u_t) \tag{64}$$

$$= H_t P_{t|t-1} H_t' + \Omega \tag{65}$$

$$= F_t. \tag{66}$$

Letting f_z and f_ϵ denote the probability density functions of z_t and ϵ_t respectively, it should be clear that

$$f_z(z_1, z_2..., , z_T) = \frac{1}{\sqrt{(2\pi)^T |F_t|}} \exp\left(-\frac{1}{2}(z_t - \hat{z}_t)^\top (F_t)^{-1}(z_t - \hat{z}_t)\right) \tag{67}$$

and

$$f_\epsilon(\epsilon_1, \epsilon_2..., , \epsilon_T) = \frac{1}{\sqrt{(2\pi)^T |F_t|}} \exp\left(-\frac{1}{2}(\epsilon_t)^\top (F_t)^{-1}\epsilon_t\right) \tag{68}$$

where T is the number of observation dates and $\epsilon_t = z_t - \hat{z}$ so that $\epsilon_t \sim N(0, F_t)$.

In order to compute the likelihood function of Θ note that the likelihood function associated with a dynamic time series model can be expressed in terms of a prediction error decomposition. The components from Equation 56 and Equation 57 form the prediction error decomposition of the log-likelihood function.

If Θ is defined to be the vector of the parameters of the state model, for a fixed value of Θ the Kalman filter produces the prediction errors ϵ and the prediction error variances $F_{t|t-1}$ from the prediction equations.

Considering Θ to be the vector variable the likelihood can be written as a prediction error decomposition of the form

$$L(\Theta; \epsilon_1, ..., \epsilon_T) = \prod_{t=1}^{T} f(\epsilon_i | \Theta). \tag{69}$$

By taking the logarithm it follows immediately that

$$\mathcal{L}(\Theta) = -\frac{1}{2}\log 2\pi \sum_{t=1}^{T} N_t - \frac{1}{2}\sum_{t=1}^{T} \log |F_t| - \frac{1}{2}\sum_{t=1}^{T} \epsilon_t' F_t^{-1} \epsilon_t, \tag{70}$$

where T is the number of observation dates and N_t is the dimension of ϵ_t. Approximating the true distribution of w_t in Equation 46 with a Gaussian distribution makes this a quasi-maximum likelihood procedure.

As a consequence the quasi-maximum likelihood estimator, $\hat{\Theta}$ is then

$$\hat{\Theta} = \arg\max_{\Theta} \mathcal{L}(\Theta)$$

which must be solved numericall.y

Note that although QML estimation has been shown to be consistent in many settings, it is in fact not consistent in a Kalman filter setting, but it has been shown that this effect is negligable in the term structure estimation framework.

4.2 Moment Derivation

Recall that the factor process dynamics under the pricing measure $\mathbb{Q}$ is specified by

$$dX_t = \kappa(\theta - X_t)dt + \text{Diag}\left(\sigma \circ \sqrt{X_t}\right) dW_t^{\mathbb{Q}} \tag{71}$$

where $\sigma_i \geq 0$ for $i = 1, \cdots, d$), $\kappa \in \mathbb{R}^{d \times d}$, $\theta \in \mathbb{R}^d$, and $W_t^{\mathbb{Q}}$ is d-dimensional Brownian Motion under $\mathbb{Q}$. It useful for both the parameter estimation and option pricing to have expressions for the first and second conditional moments of X_t. In this section the conditional expection of a random variable U_t under the $\mathbb{Q}$-measure will be denoted

$$E^{\mathbb{Q}}[U_t | \mathcal{F}_t] = E_t[U_t] \tag{72}$$

to ease notation.

The method of integrating factors can be used to determine the conditional expectation of X_t. Note

$$dX_t + \kappa X_t dt = \kappa\theta dt + \text{Diag}\left(\sigma \circ \sqrt{X_t}\right) dW_t^{\mathbb{Q}}. \tag{73}$$

This form suggests the integrating factor $e^{\kappa t}$. Multiplying both sides of the SDE by $e^{\kappa t}$ yields

$$d(e^{\kappa t} X_t) = e^{\kappa t}\kappa\theta dt + e^{\kappa t}\text{Diag}\left(\sigma \circ \sqrt{X_t}\right) dW_t^{\mathbb{Q}} \tag{74}$$

such that

$$e^{\kappa T} X_T = e^{\kappa t} X_t + \kappa^{-1}(e^{\kappa T} - e^{\kappa t})\kappa\theta + \int_t^T e^{\kappa s}\text{Diag}\left(\sigma \circ \sqrt{X_t}\right) dW_t^{\mathbb{Q}} \tag{75}$$

resulting in

$$X_T = e^{-\kappa(T-t)} X_t + (I - e^{-\kappa(T-t)})\theta + e^{-\kappa T} \int_t^T e^{\kappa s}\text{Diag}\left(\sigma \circ \sqrt{X_t}\right) dW_t^{\mathbb{Q}}. \tag{76}$$

Taking conditional expectations yields the first moment of X_t as

$$E_t[X_T] = e^{-\kappa(T-t)}X_t + (I - e^{-\kappa(T-t)})\theta,$$
$$= \theta + e^{-\kappa(T-t)}(X_t - \theta), \tag{77}$$

with $t \leq T$.

The expression for the second moment is unavailable in closed form, but can be solved numerically. Define

$$Y_t = X_t X_t^\top. \tag{78}$$

The SDE for Y_t can be found using Itō's lemma,

$$dY_t = d(X_t X_t^\top) \tag{79}$$
$$= dX_t X_t^\top + X_t dX_t^\top + d\langle X_t, X_t^\top \rangle, \tag{80}$$

where

$$dX_t^\top = (\theta^\top - X_t^T)\kappa^\top dt + (dW_t^{\mathbb{Q}})^\top \left[\text{Diag}\left(\sigma \circ \sqrt{X_t}\right)\right]^\top, \tag{81}$$

$$d\langle X_t, X_t^\top \rangle = \text{Diag}\left(\sigma \circ \sqrt{X_t}\right) dW_t^{\mathbb{Q}}(dW_t^{\mathbb{Q}})^\top \left[\text{Diag}\left(\sigma \circ \sqrt{X_t}\right)\right]^\top \tag{82}$$

and

$$dW_t^{\mathbb{Q}}(dW_t^{\mathbb{Q}})^\top = \mathbb{I}_d dt. \tag{83}$$

Thus

$$d(Y_t) = X_t\theta^\top\kappa^\top dt - Y_t\kappa^\top dt + \kappa\theta X_t^\top dt - \kappa Y_t^\top dt + \text{Diag}\left(\sigma^2 \circ X_t\right)dt$$
$$+ X_t(dW_t^{\mathbb{Q}})^\top \left[\text{Diag}\left(\sigma \circ \sqrt{X_t}\right)\right]^\top + \text{Diag}\left(\sigma \circ \sqrt{X_t}\right)dW_t^{\mathbb{Q}}X_t^\top. \tag{84}$$

Taking the conditional expectation yields

$$dE_t[Y_t] = -E_t[Y_t]\kappa^\top dt - \kappa E_t[Y_t]dt+$$
$$E_t[X_t]\theta^\top\kappa^\top dt + \kappa^\top\theta^\top E_t[X_t]dt + \text{Diag}(\sigma^2 \circ E_t[X_t])dt. \tag{85}$$

Defining

$$G(t) = E_t[Y_t Y_t^\top] \tag{86}$$

and

$$F(t) = E_t[X_t] \tag{87}$$

delivers the matrix ordinary differential equation

$$G'(t) + G(t)\kappa^\top + \kappa G(t) = F(t)\theta^\top\kappa^\top + \kappa^\top\theta^\top F(t) + \text{Diag}(\sigma^2 \circ F(t)). \tag{88}$$

Equation 88 can now be solved numerically to deliver the conditional variance of the X_t process.

Figure 8: A single simulated futures path with added low volatility noise
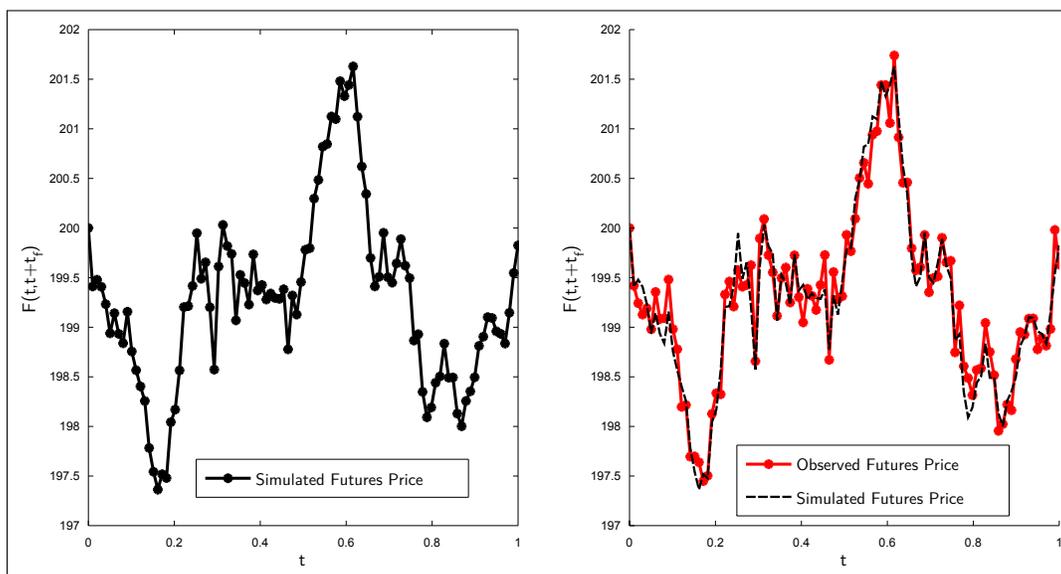
5 Implementation

The first step in implementing the QMLE is testing the Kalman Filter against simulated data. For data that consists only of the futures term structure realizations, the standard Kalman Filter is sufficient.

5.1 Filtering Futures

The first test case involves a single time-series of simulated futures data with the addition of low-volatility noise. Figure 8 illustrates the input observations and Figure 9 displays the corresponding filtered state. It is clear that in this simple case the standard Kalman Filter performs admirably.

The second test case involves three time-series of simulated futures data with differing terms and the addition of high-volatility noise. The input observations to the filter are displayed in Figure 10 whereas the corresponding filtered state is depicted in Figure 11.

Even though the true futures prices are almost dwarfed by noise, the Kalman Filter performs well. The added noise is somewhat offset by the increased number of paths as it is assumed (and simulated) that each path has the same noise variance. The final test requires option observations.

Figure 9: The output of the Kalman Filter application
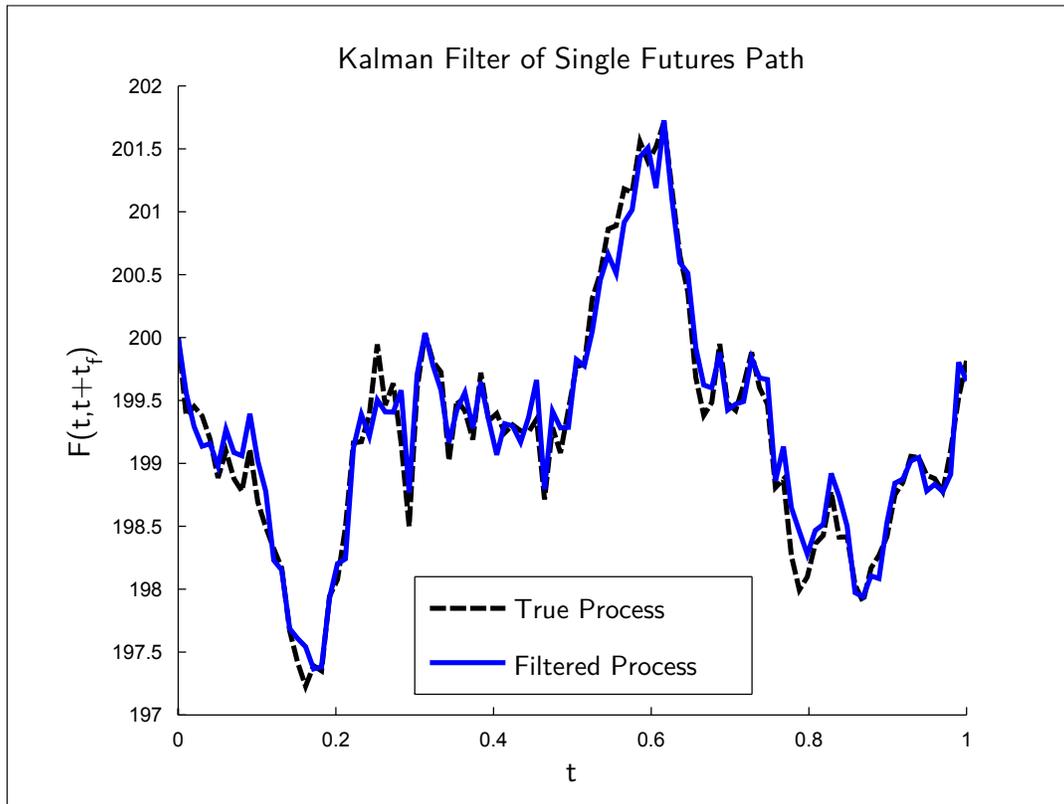
5.2 Filtering Options

When the input data consists of time-series of futures as well as options the Extended Kalman Filter performs poorly. Thus the Unscented Kalman Filter was implemented. A test case is illustrated in Figures 12 and 13, where it is again shown that the filter is performing well at recovering the underlying state.

5.3 The QMLE

Finding the QMLE requires performing an optimization over the Kalman Filter function calls. However, the Kalman Filter is numerically unstable: the required Cholesky decomposition of the transition covariance matrix will fail when small numerical errors accumulate and the matrix stops being positive-definite. A solution to this problem is to implement the square-root specification of the Kalman Filter which keeps the covariance matrices decomposed.

Figure 10: Multiple simulated futures paths with added high volatility noise
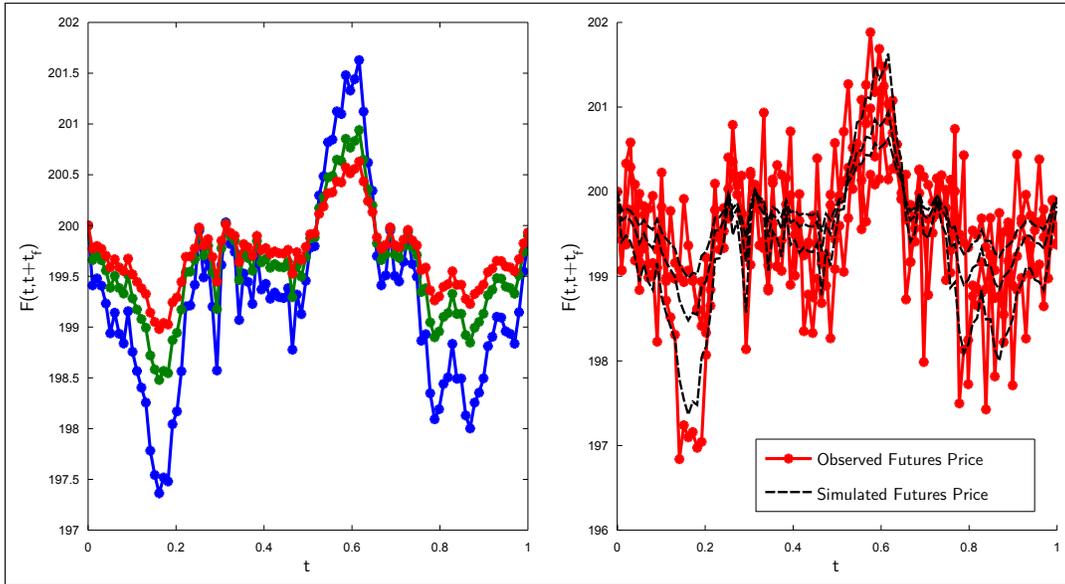
Figure 11: The output of the Kalman Filter application
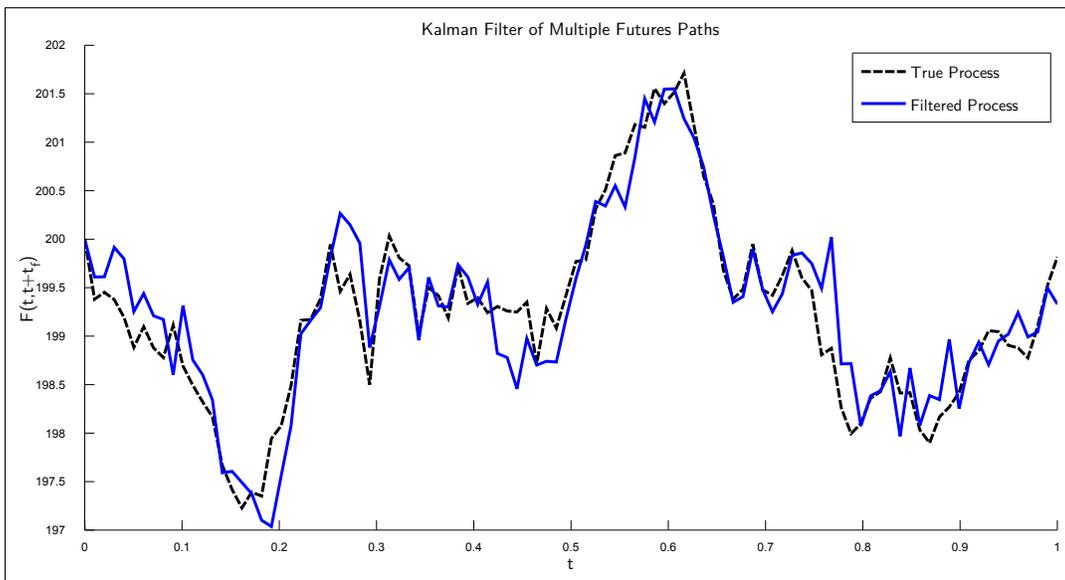
Figure 12: Simulated time-series data for futures and options with added noise
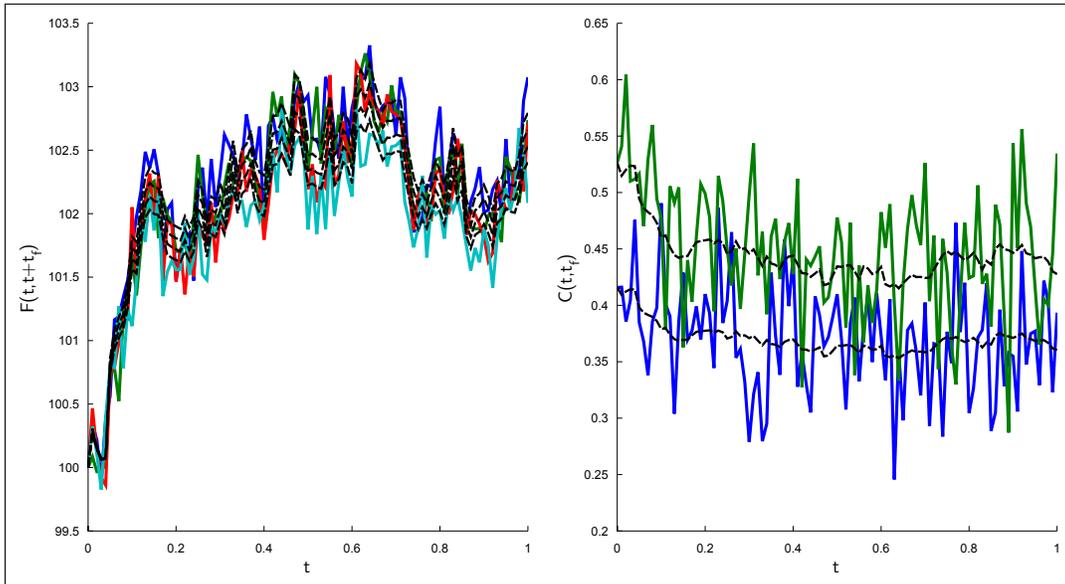
Figure 13: Applying the Unscented Kalman Filter to the time-series data of futures and options
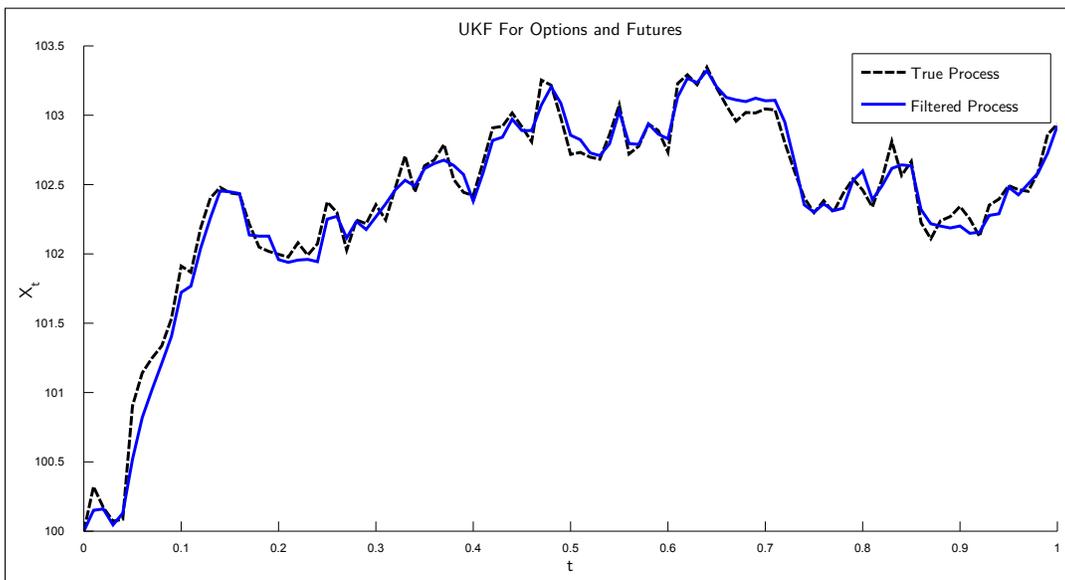
5.4 Conclusion

The investigation of the LSQ model proceeded well both analytically and numerically. The next step is to implement a numerically stable Kalman Filter such that real-world parameters can be estimated.

Bibliography

Back, J., Prokopczuk, M., 2013. Commodity price dynamics and derivative valuation: A review. International Journal of Theoretical and Applied Finance 16 (06), 1350032.

Barkoulas, J., Labys, W. C., Onochie, J., 1997. Fractional dynamics in international commodity prices. Journal of Futures Markets 17 (2), 161–189.

Duffie, D., Pan, J., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68 (6), 13431376.

Filipovic, D., 2009. Term-Structure Models: A Graduate Course. Springer Finance / Springer Finance Textbooks. Springer Berlin Heidelberg.
URL https://books.google.at/books?id=KqcSh6CavaAC

Filipovic, D., Larsson, M., Trolle, A. B., 2014. Linear-rational term structure models. Swiss Finance Institute Research Paper (14-15).

Fisher, M., Gilles, C., 1996. Term premia in exponential-affine models of the term structure. Manuscript, Board of Governors of the Federal Reserve System.

Geman, H., 2009. Commodities and commodity derivatives: modeling and pricing for agriculturals, metals and energy. John Wiley & Sons.

Nielsen, M. J., Schwartz, E. S., 2004. Theory of storage and the pricing of commodity claims. Review of Derivatives Research 7 (1), 5–24.

Pindyck, R. S., 2001. The dynamics of commodity spot and futures markets: a primer. The Energy Journal, 1–29.

Samuelson, P. A., 1965. Proof that properly anticipated prices fluctuate randomly. Industrial management review 6 (2), 41–49.

Trolle, A. B., Schwartz, E. S., 2009. Unspanned stochastic volatility and the pricing of commodity derivatives. Review of Financial Studies, hhp036.

Long-Dated Swaption Pricing in Single and Multi-Curve LIBOR Market Models

Team 4

LUKAS GONON, ETH Zurich
WEI GUAN, University College London
GARETH SCHUMANN, University of Cape Town
MICHAEL VAN GYSEN, University of Cape Town

Supervisor:
JOHAN DE KOCK, LibFin (Liberty Life)

Contents

1   Introduction

1.1   The market models and interest rate derivative pricing

The lognormal forward-LIBOR model (LFM or *LIBOR market model*) is one of the most popular interest rate models. The reason for this is simply how it prices caps, which along with swaptions are among the most liquidly traded interest rate derivatives. Firstly, there is an analytical expression for cap prices in this model, commonly referred to as the Black formula, see e.g. Brigo and Liinev (2005). This closed-form pricing rule has been market practice for pricing caps long before the LFM was introduced in Miltersen et al. (1997) and Brace et al. (1997). It also allows for efficient calibration of LFM to the market prices of caps. Secondly, although there is no simple formula for swaption prices in the LFM, these can be approximated very well by a closed-form expression also relying on the Black formula.

Similarly, under the lognormal forward-swap model (LSM), swaption prices are given exactly by an analogous Black formula and cap prices can be approximated suitably. Pricing swaptions with this analogous Black formula is again standard in the swaption-market. However, as shown in Brigo and Liinev (2005) and explained in more detail in Chapter 6.1 in Brigo and Mercurio (2006), there is no model under which swaptions and caps are *both* priced exactly with the Black formula.

In order to price interest rate derivatives that are not liquidly-traded, one can therefore calibrate the LFM (or LSM) to the market-prices of liquidly traded caps (swaptions) based on the corresponding Black formula and use Monte Carlo simulation to approximate the price of the derivative. If the non-liquidly traded derivative is itself a cap (swaption), of course, the exact formula can be used and for swaptions (caps) the approximate pricing formulas can be employed.

1.2   Research problem: Pricing of long-dated swaptions in single- and multi-curve LIBOR market models

Insurance companies pricing swaptions and calibrating to swaption data are faced with the following problem. In the Asset Liability Management (ALM) context the maturities and tenors (in the terminology of Chapter 1 in Brigo and Mercurio (2006)) of these swaptions are far larger than those of any swaptions traded in the market. Since the swaption market in South Africa is generally viewed not to be liquid enough for a sensible calibration of the LSM model, one could resort to the LFM. This brings us precisely to the setup outlined above: *we want to calculate swaption prices in the LFM*. Due to the large number of insurance policies that have to be valued in the ALM context, the number of sample paths that can be used for the Monte Carlo simulation is very limited. On the other hand, the Rebonato formula (see e. g. Section 6.15 in Brigo and Mercurio (2006)), a closed-form approximation for the implied volatility of a swaption price in the LFM model, has been found to be very accurate in empirical studies presented by several researchers, for example

in Brigo and Mercurio (2006). However, these empirical studies have been in the pricing context and not in the ALM context. In some parameter settings based on the calibration of the LFM to market data, however, the Rebonato volatility and the volatility implied from the price calculated using Monte Carlo seem to differ substantially. The problem examined by our team for the purposes of this challenge thus mainly comprehends the following questions:

> *Why are the two volatilities so different? Which of the two approximations should be trusted and how can we assess the accuracy of the calculated prices?*

As a further question we were asked to derive a similar approximation in a multi-curve setting. See Section 5 for a motivation of the problem.

1.3 Organisation of the report and proposed solutions

In Section 2 we give a review of the LIBOR market model, Monte Carlo pricing in this setup and derive the Rebonato formula. In Sections 3 and 4 we then discuss and assess different aspects that we consider possible solutions to the problem formulated above. These include

  (i) Consistency checks: Monte Carlo error bounds and bias control via exact caplet pricing.

 (ii) Sensitivity of the implied volatility to approximation error.

(iii) Improving the accuracy of Monte Carlo: reducing the step size, using predictor-corrector and switching to quasi-Monte Carlo.

In more detail, in Sections 3.1 and 3.2 we outline two simple checks that allow us to measure the quality of the Monte Carlo approximation for the price: Firstly, we obtain a confidence interval for the true price using the standard Monte Carlo error bounds and secondly, we show how the exact pricing formula for caps under the LFM model can be used to further assess the quality of a simulated Monte Carlo sample. In Section 3.3 we then look at the problem from a different point of view: given that swaption prices in the South African market are quoted in terms of implied volatility, this is the quantity of interest. However, the simulations give an approximation of the price, not the implied volatility. The inversion of the Black formula, mapping a price to its implied volatility, might in some cases be very sensitive to small price changes and thus amplify the Monte Carlo error. We illustrate how standard methods may be used to assess this sensitivity. In Section 3.4 we describe the predictor-corrector method which may allow for the improvement in the accuracy of Monte Carlo pricing for a fixed sample size. This will of course increase the computational cost required to generate the samples. In the ALM context, however, the number of policies valued with the same Monte Carlo sample

is large. Thus, the overall computational cost for valuing all these policies may be lower in comparison to standard Monte Carlo with larger sample size.

In Section 4 we then assess each of the above points in simulations. We consider different parameter scenarios, all of which have been obtained by calibrating the LFM to market data. We see that in cases of short maturities the approximation using the Rebonato formula indeed seems to be working quite well. For longer-dated swaptions, however, the two approximations can be quite far apart. In these cases, the consistency checks in (i) allow us to quantify the approximation quality of Monte-Carlo prices and we thus easily see that for long-dated maturities the Rebonato formula may not be a good approximation. Furthermore, we examine the different methods in (iii) and see, for example, that simulating forward rates with smaller time-steps do not seem to bring an improvement as opposed to only sampling them at the reset times of the swaption.

We conclude with Section 5, where we review the lognormal multi-curve LI-BOR market model introduced in Mercurio (2010) and the derivation of the analogous Rebonato formula. This provides an answer to the second question formulated above.

2 LIBOR Market Model

In this section we provide a review of the lognormal forward-LIBOR model, first introduced in Miltersen et al. (1997) and Brace et al. (1997). The exposition in this section is very close to Brigo and Mercurio (2006), but not fully self-contained for the sake of brevity. In Section 5, many of the notions are explained in more detail in a multi-curve setting.

2.1 Why Market Models are popular

The lognormal forward-LIBOR model (LFM or *LIBOR market model*) and the lognormal forward-swap model (LSM) are two of the most popular interest rate models. As outlined in the introduction, this is mainly due to the fact that they price caps and swaptions in agreement with the well-established market formulas. The Black cap formula is the standard formula used in the cap market and the prices computed from this formula coincide with the ones from the LFM. Moreover, the LSM prices swaptions with the Black swaption formula which again is standard in the swaption market. However, though these classic formulas for caps and swaptions can be derived rigorously separately in each of the two models, the two models are not compatible. As discussed in detail in Section 2.4, if forward rates are lognormal under each of their own forward measures, the swap rate is not exactly lognormal (although, it is not far away). There is empirical work on this matter, see Brigo and Liinev (2005) and Brigo and Mercurio (2006).

2.2 Definition

Let $T_{-1} = 0$ be the current time. Consider a set $\epsilon = \{T_0, ..., T_M\}$ of adjacent expiry-maturity pairs of dates. Let us denote by $\{\tau_0, ..., \tau_M\}$ the corresponding year fractions.

Consider the generic forward rate $F_k(t) := F(t; T_{k-1}, T_k)$, which is defined up to time T_{k-1}, where it coincides with the spot LIBOR rate $F_k(T_{k-1}) := L(T_{k-1}, T_k)$.

Consider now the probability measure $\mathbb{Q}^k$ associated with the numéraire $P(\cdot, T_k)$. As in Brigo and Mercurio (2006) we denote by $P(t, T)$ the price at time t of a zero-coupon-bond with maturity T, where $t \leq T$. $\mathbb{Q}^k$ is often called the forward adjusted measure for maturity T_k. The LFM assumes the following driftless geometric Brownian dynamics for F_k under $\mathbb{Q}^k$:

$$dF_k(t) = \sigma_k(t) F_k(t) dZ_k(t), \tag{1}$$

$t \leq T_{k-1}$, where dZ_k is a standard Brownian motion under $\mathbb{Q}^k$ and $\sigma_k(t)$ is a deterministic function representing the instantaneous volatility at time t for the forward LIBOR rate F_k. Throughout this report, we will always consider piecewise constant instantaneous volatility $\sigma_k(t) = \sigma_{k,\beta(t)}$, where $\beta(t) = m$ if $t \in (T_{m-2}, T_{m-1}]$. The noises in the dynamics of different forward rates are assumed to be instantaneously correlated according to

$$d\langle Z_i, dZ_j \rangle_t = \rho_{i,j} dt,$$

where the brackets denote the quadratic variation.

If we look at the dynamics of the forward LIBOR rate F_k under a measure $\mathbb{Q}^\alpha$ with $\alpha \neq k$, the forward rate $F_k(t)$ is not a martingale under $\mathbb{Q}^\alpha$ and a drift term appears:

$$dF_k(t) = \sigma_k(t) F_k(t) \sum_{j=\alpha+1}^{k} \frac{\rho_{k,j} \tau_j \sigma_j(t) F_j(t)}{1 + \tau_j F_j(t)} dt + \sigma_k(t) F_k(t) dZ_k(t), \tag{2}$$

for $k = \alpha + 1, ..., \beta$. For further details, see Section 5 of the report or Chapter 6 of Brigo and Mercurio (2006).

2.3 Pricing Caplets

Caps are collection of caplets. A caplet with maturity T_{i-1} has payoff $(F_i(T_{i-1} - K)^+$ paid at time T_i. Since in the LFM F_i follows a geometric Brownian motion under $\mathbb{Q}^i$, the price Cpl of a caplet is computed immediately as a classical Black

price:

$$
\begin{aligned}
\mathrm{Cpl} &= P(0, T_i)\mathrm{Bl}(K, F_i(0), v_i), \quad \text{with} \\
\mathrm{Bl}(K, F_i(0), v_i) &= F_i(0)\Phi(d_1(K, F_i(0), v_i)) - K\Phi(d_1(K, F_i(0), v_i)), \\
d_1(K, F_i(0), v_i) &= \frac{\ln(F_i(0)/K) + v_i^2/2}{v_i}, \\
d_2(K, F_i(0), v) &= \frac{\ln(F_i(0)/K) - v_i^2/2}{v_i},
\end{aligned}
\tag{3}
$$

Here Φ denotes the distribution function of a standard normal and v_i is given as

$$
\begin{aligned}
v_i^2 &= T_{i-1}v_{T_{i-1}}^2 \\
v_{T_{i-1}}^2 &= \frac{1}{T_{i-1}}\int_0^{T_{i-1}} \sigma_i(t)^2 dt
\end{aligned}
\tag{4}
$$

2.4 Limitations

The limitation of the LFM from a market perspective is that it does not price swaptions according to the Black swaption formula. The payoff of a swaption can be written as

$$
H = (S_{\alpha,\beta}(T_\alpha) - K)^+ \sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i),
\tag{5}
$$

where $S_{\alpha,\beta}$ is the forward swap rate. For example, in equation (6.33) in Brigo and Mercurio (2006)), we can express the forward swap rate $S_{\alpha,\beta}(T_\alpha)$ in terms of spanning forward rates $F_{\alpha+1}(T_\alpha), \ldots, F_\beta(T_\alpha)$,

$$
S_{\alpha,\beta}(t) = \frac{1 - \prod_{j=\alpha+1}^{\beta} \frac{1}{1+\tau_j F_j(t)}}{\sum_{i=\alpha+1}^{\beta} \tau_i \prod_{j=\alpha+1}^{i} \frac{1}{1+\tau_j F_j(t)}}.
\tag{6}
$$

The lognormal forward-swap model (LSM) assumes the following geometric Brownian driftless dynamics for a forward swap rate $S_{\alpha,\beta}(t)$ under the swap measure $\mathbb{Q}^{\alpha,\beta}$ (see Brigo and Mercurio (2006)):

$$
dS_{\alpha,\beta}(t) = \sigma^{(\alpha,\beta)}(t)S_{\alpha,\beta}(t)dW_t^{\alpha,\beta}
\tag{7}
$$

where $W^{\alpha,\beta}$ is a standard Brownian motion under $\mathbb{Q}^{\alpha,\beta}$.

Thus, basically we have two possibilities to calculate the price of a swaption: we can do this either under the LFM, which models the spanning forward rates, giving a swap rate as in equation (6) or under LSM which models forward swap rates directly. The two results are not the same.

This is because while the dynamics of the swap rate coming from the LSM (see Equation (7)) are lognormal, the dynamics of swap rate coming form the LFM (Equation (6)) are not lognormal.

Thus, under the LFM we can not get the Black formula for swaptions, which is the market standard. Instead, we need to price swaptions numerically under the LFM.

2.5   Monte Carlo Pricing of Swaptions with the LFM

The price p_0 at time 0 of a swaption can be obtained as the expectation of the payoff Equation (5) under the T_α-forward measure $\mathbb{Q}^\alpha$

$$p_0 = P(0, T_\alpha)\mathbb{E}^\alpha \left[(S_{\alpha,\beta}(T_\alpha) - K)^+ \sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i) \right] \tag{8}$$

where $\mathbb{E}^\alpha$ denotes the expectation under $\mathbb{Q}^\alpha$. Notice that for $k = \alpha + 1, \dots, \beta$,

$$P(T_\alpha, T_k) = \prod_{j=\alpha+1}^{k} \frac{1}{1 + \tau_j F_j(T_\alpha)} \tag{9}$$

and Equation (6) hold, so that the payoff inside the expectation in Equation (8) can be expressed only in terms of forward rates. For Monte Carlo simulation, we can thus proceed precisely as in Section 6.10 of Brigo and Mercurio (2006) and first simulate N realisations of $F_{\alpha+1}(T_\alpha), \dots, F_\beta(T_\alpha)$ under the measure $\mathbb{Q}^\alpha$. By taking logs and applying Itô's Lemma, we see that the dynamics (2) imply

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

To simulate the forward rates, we simply use the Euler scheme

$$\begin{aligned} \ln F_k^{\Delta t}(t + \Delta t) = \ln F_k^{\Delta t}(t) + \sigma_k(t) \sum_{j=\alpha+1}^{k} \frac{\rho_{k,j}\tau_j\sigma_j(t)F_j^{\Delta t}(t)}{1 + \tau_j F_j(t)^{\Delta t}} \Delta t \\ - \frac{\sigma_k(t)^2}{2}\Delta t + \sigma_k(t)(Z_k(t + \Delta t) - Z_k(t)) \end{aligned} \tag{10}$$

recalling that $Z_(t + \Delta t) - Z_(t) \sim \mathcal{N}(0, \rho)$.

We then evaluate the payoff inside the expectation in Equation (8) for each of these realisations (giving a scenario H^i), average and multiply by $P(0, T_\alpha)$, a quantity given by the initial data, to obtain the Monte Carlo approximation of the price

$$p_{MC} = \frac{P(0, T_\alpha)}{N} \sum_{i=1}^{N} H^i \tag{11}$$

9

The Monte Carlo implied volatility $v_{\alpha,\beta}^{MC}$ is the volatility, which has to be inserted into the Black swaption formula (see Brigo and Mercurio (2006)) in order to obtain p_{MC}. More precisely, it is the unique solution to

$$p_{MC} = C_{\alpha,\beta}(0)\mathrm{Bl}(K, S_{\alpha,\beta}(0), \sqrt{T_\alpha}v_{\alpha,\beta}^{MC}), \tag{12}$$

where Bl is as in Equation (3) and

$$C_{\alpha,\beta}(0) = \sum_{i=\alpha+1}^{\beta} \tau_i P(0, T_i).$$

2.6 The Rebonato formula for the swaption volatility

From equation (6), we can write

$$S_{\alpha,\beta}(t) = \sum_{k=\alpha+1}^{\beta} \omega_k(t) F_k(t), \tag{13}$$

where

$$\omega_k(t) = \frac{\tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}}{\sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}}. \tag{14}$$

Assuming that the variability of the weights ω is much smaller than the variability of the F_k, we can freeze the weights at their known value at time 0, leading to:

$$S_{\alpha,\beta}(t) \approx \sum_{k=\alpha+1}^{\beta} \omega_k(0) F_k(t).$$

We define this approximation of the swap rate as the Rebonato swap rate,

$$S_{\alpha\beta}^{Rebonato}(t) := \sum_{k=\alpha+1}^{\beta} \omega_k(0) F_k(t)$$

Taking differentials on both sides and equating the quadratic variations, we have

$$d\langle \ln S_{\alpha,\beta} \rangle_t \approx \sum_{i,j=\alpha+1}^{\beta} \frac{\omega_i(0)\omega_j(0)F_i(t)F_j(t)\rho_{i,j}\sigma_i(t)\sigma_j(t)}{S_{\alpha,\beta}(t)^2} dt.$$

We now introduce a further approximation by freezing all the forward rates F_k at time 0 in the above equation, which results in

$$d\langle \ln S_{\alpha,\beta} \rangle_t \approx \sum_{i,j=\alpha+1}^{\beta} \frac{\omega_i(0)\omega_j(0)F_i(0)F_j(0)\rho_{i,j}\sigma_i(t)\sigma_j(t)}{S_{\alpha,\beta}(0)^2} dt.$$

By the definition of Black swaption volatility under the LSM,

$$(v_{\alpha,\beta}^2) := \int_0^{T_\alpha} \sigma_{\alpha,\beta}^2(t)dt = \int_0^{T_\alpha} (d\ln S_{\alpha,\beta}(t))(d\ln S_{\alpha,\beta}(t))$$

Then we have

$$\int_0^{T_\alpha} d\langle \ln S_{\alpha,\beta} \rangle_t dt$$

$$\approx \sum_{i,j=\alpha+1}^{\beta} \frac{\omega_i(0)\omega_j(0)F_i(0)F_j(0)\rho_{i,j}}{S_{\alpha,\beta}(0)^2} \int_0^{T_\alpha} \sigma_i(t)\sigma_j(t)dt =: (v_{\alpha,\beta}^{LFM})^2$$

Then we have LFM Black-like swaption volatility

$$(v_{\alpha,\beta}^{Rebonato}) = \sqrt{\frac{1}{T_\alpha}(\frac{1}{S_{\alpha\beta}(0)})^2 \sum_{i,j=\alpha+1}^{\beta} \omega_i(0)\omega_j(0)F_i(0)F_j(0)\rho_{i,j} \int_0^{T_\alpha} \sigma_i(t)\sigma_j(t)dt}.$$

$$(15)$$

3 Implementation, Quality Assessment and Sensitivity Analysis

3.1 Standard Monte Carlo Error Bounds

Following Section 6.11 in Brigo and Mercurio (2006), the central limit theorem gives the following confidence interval for the Monte Carlo approximation of the swaption price[1]. The true value p_0 will lie inside the (random) interval

$$\left[p_{MC} - \Phi^{-1}\left(\frac{\alpha+1}{2}\right)\frac{\mathrm{Std}(p)}{\sqrt{N}}, p_{MC} + \Phi^{-1}\left(\frac{\alpha+1}{2}\right)\frac{\mathrm{Std}(p)}{\sqrt{N}} \right]$$

with probability α, where $\mathrm{Std}(p)$ is the (true) standard deviation of $P(0,T_\alpha)H$. Replacing $\mathrm{Std}(p)$ by the sample standard deviation

$$\widehat{\mathrm{Std}}^2 = \frac{1}{N}\sum_{i=1}^N (P(0,T_\alpha)H^i)^2 - \left(\frac{\sum_{i=1}^N P(0,T_\alpha)H^i}{N}\right)^2$$

we obtain an approximate confidence interval. Here p_{MC} and H^i are as in Equation (11).

---

[1]Note that here we have scaled the interval by $P(0,T_\alpha)$.

3.2 Validating the Monte Carlo Implementation Using the Black Formula for Caplets

To check if the Monte Carlo works correctly, let us look for a payoff H for which the price can be calculated exactly. Given a Monte Carlo sample used to approximate the price of a swaption, we can then use the same sample to calculate the Monte Carlo price of H. Comparing this to the known true price of H, this gives us an indication of how good the approximation of the swaption price with this sample is.

Let us consider a payoff at T_α given as

$$H_k = \tau_k D(T_\alpha, T_k)(F_k(T_\alpha) - K)^+,$$

where D denotes the stochastic discount factor associated to the risk-neutral measure $\mathbb{Q}$. The price c_0 of this contract at time 0 is thus given as $c_0 = \mathbb{E}[D(0, T_\alpha)H_k]$, where $\mathbb{E}$ denotes the expectation under the risk-neutral measure. By changing to the T_k forward-measure, we can rewrite this as

$$\mathbb{E}[D(0, T_\alpha)H_k] = \tau_k P(0, T_k)\mathbb{E}^k[(F_k(T_\alpha) - K)^+]$$
$$= \tau_k P(0, T_k)\mathrm{Bl}(K, F_k(0), v_k),$$

where Bl is as in Equation (3) and

$$v_k^2 = \int_0^{T_\alpha} \sigma_k(t)^2 dt = \sum_{j=0}^{\alpha} \sigma_{k,j+1}^2 \tau_j.$$

The second equality follows from the assumption that σ is piecewise constant.

However, to calculate this price using Monte Carlo, we cannot directly use the realisations generated above to approximate $\mathbb{E}^k[(F_k(T_\alpha) - K)^+]$; recall that these realisations were generated under $\mathbb{Q}^\alpha$ and not $\mathbb{Q}^k$. Nevertheless, we may obtain

$$\mathbb{E}[D(0, T_\alpha)H_\alpha] = \tau_k \mathbb{E}[D(0, T_\alpha)(F_k(T_\alpha) - K)^+ P(T_\alpha, T_k)]$$
$$= \tau_k P(0, T_\alpha)\mathbb{E}^\alpha[(F_k(T_\alpha) - K)^+ P(T_\alpha, T_k)]$$

by using Section 2.7 in Brigo and Mercurio (2006) for the first step and a change of numéraire for the second. For each $k = \alpha + 1, \ldots, \beta$ the expectation in the last step can now be calculated using the realisations of $F_k(T_\alpha)$ generated above (under $\mathbb{Q}^\alpha$) and Equation (9) for k.

3.3 Sensitivity Analysis

As we invert the Black formula to find the Monte Carlo implied volatility, see Equation (12), it might be informative to look at the sensitivity of the volatility to price changes.

The sensitivity of volatility to price change may be defined as

$$\zeta = \frac{\partial v_{\alpha,\beta}^{MC}}{\partial p_{MC}} = \frac{1}{\frac{\partial p_{MC}}{\partial v_{\alpha,\beta}^{MC}}}.$$

Note that ζ can be obtained by standard calculations as

$$\zeta = \frac{\sqrt{2\pi}}{S_{\alpha,\beta}(0)C_{\alpha,\beta}(0)\sqrt{T_\alpha}} \exp\left[\frac{1}{2}d_1^2(K, S_{\alpha,\beta}(0), \sqrt{T_\alpha}v_{\alpha,\beta}^{MC})\right] \qquad (16)$$

where the function d_1 is as in Equation (3).

By looking at the sensitivity, we can get an idea to what extent the error in the Monte Carlo approximation affects the volatility. More precisely, a large sensitivity will indicate that a even a small error $p_0 - p_{MC}$ may result in a large error $v_0 - v_{\alpha,\beta}^{MC}$ and vice versa. Here by v_0 we denote the true implied volatility, i.e. the solution of

$$p_0 = C_{\alpha,\beta}(0)\mathrm{Bl}(K, S_{\alpha,\beta}(0), \sqrt{T_\alpha}v_0)$$

analogously to (12).

3.4 Predictor-Corrector method

In the log discretisation used above, a discretisation error arises due to the fact that the drift is state dependent. Following an approach proposed by Hunter et al. (2001), it is possible to produce a more accurate estimate of the drift experienced over the update period.

The idea is to evolve the forward rates to the end of the period and then compute the terminal drift using the evolved rates. Using the same variates that were used to estimate the terminal drift, the initial forward rates are then evolved using a drift computed as the average of the initial and terminal drift.

Mathematically, we first compute

$$\ln \bar{F}_k^{\Delta t}(t + \Delta t) = \ln \hat{F}_k^{\Delta t}(t) + \sigma_k(t)\hat{\mu}_k(t)\Delta t$$
$$- \frac{\sigma_k(t)^2}{2}\Delta t + \sigma_k(t)(Z_k(t + \Delta t) - Z_k(t))$$

where

$$\hat{\mu}_k(t) = \sum_{j=\alpha+1}^{k} \frac{\rho_{k,j}\tau_j\sigma_j(t)F_j(t)}{1 + \tau_j F_j(t)}$$

for $k = \alpha + 1, ..., \beta$. Then we use these evolved rates to compute

$$\bar{\mu}_k(t) = \sum_{j=\alpha+1}^{k} \frac{\rho_{k,j}\tau_j\sigma_j(t + \Delta t)F_j(t + \Delta t)}{1 + \tau_j F_j(t + \Delta t)}.$$

13

Finally, using the same $(Z_k(t + \Delta t) - Z_k(t))$ as before, we compute the new rates,

$$
\begin{aligned}
\ln \hat{F}_k^{\Delta t}(t + \Delta t) = \ln \hat{F}_k^{\Delta t}(t) &+ \sigma_k(t) \frac{\hat{\mu}_k(t) + \bar{\mu}_k(t)}{2} \Delta t \\
&- \frac{\sigma_k(t)^2}{2} \Delta t + \sigma_k(t)(Z_k(t + \Delta t) - Z_k(t)).
\end{aligned}
\tag{17}
$$

4   Model Simulation

Simulating under the 2-Factor LMM

Under the calibrated 2-factor LMM, we analyse the discrepancy between $\upsilon_{\alpha,\beta}^{MC}$ and $\upsilon_{\alpha,\beta}^{Rebonato}$. We calculate $\upsilon_{\alpha,\beta}^{MC}$ as the Black implied volatility from p_{MC} and determine $\upsilon_{\alpha,\beta}^{Rebonato}$ from the approximation formula (15). Our simulations were conducted in MATLAB 2015a. Note that there are a number of varying parameters which produce the results below: most notably the total number of simulations (N), the number of times steps per year (nt) that are used to evolve the forward rates, whether quasi-Monte Carlo is implemented, and finally the use of the predictor-corrector method.

From our simulations we observe that in certain cases $\upsilon_{\alpha,\beta}^{Rebonato} \neq \upsilon_{\alpha,\beta}^{MC}$ and the approximation breaks down. In comparing the maturities and tenors of different swaptions, we note that the approximation holds for short-dated swaptions but deteriorates as we evaluate long-dated swaptions. We consider the sensitivity of swaption prices to a change in volatility, and illustrate the negative impact that a poor volatility approximation can have. We finally apply our simulations in the ALM context and suggest possible guidelines for appropriate implementing of the Rebonato approximation for swaption volatilities.

Quasi-Monte Carlo

For selected results below, we use the method of quasi-Monte Carlo (QMC) due to the well known fact that it has a faster rate of convergence when compared to Monte Carlo (MC) in two dimensions. For the majority of simulations the difference in accuracy between QMC and MC is neglible. Discrepancies only arise when the number of simulated time steps per year exceeds 10, i.e. for high dimensions. However, our results indicate that time steps greater than 4 per year do not yield better approximations. Therefore, for scenarios when the number of time steps is large, we resort back to MC simulations.

Our choice of low discrepancy sequence is the Sobol sequence. The method and list of initial direction numbers and primitive polynomials is provided by Joe and Kuo (2008).

Accuracy and efficiency in implementation

It is important to re-emphasise that our objective is not to improve the pricing of swaptions via the LFM model. Our simulations seek to determine the scenarios under which the approximation of volatility $\left(v_{\alpha,\beta}^{Rebonato} \right)$ becomes inaccurate. Our results are also based on data which has been produced using a particular calibration technique, and any conclusions made with respect to $v_{\alpha,\beta}^{Rebonato}$, will be specific to this data.

4.1 Short-dated versus Long-dated Swaptions

The short-dated swaptions we consider have maturities (T_α) of 1, 5 and 10 years and tenors $(T_\beta - T_\alpha)$ of 5, 10 and 25 years. These swaptions are not frequently traded in South Africa. A swaption labeled 1×5 illustrates a 1 year maturity (T_α) with a 5 year tenor $(T_\beta - T_\alpha)$. The approximations were considered with the use of 95% Monte Carlo error bounds. In Figure 1a, one can deduce that the Rebonato approximation performs relatively well for short-dated swaptions. Although slightly higher, the Rebonato volatility for the 1×5 swaption lies within the 95% error bounds of our Monte Carlo estimate for a sample of 50 000 paths. For the 10×15 swaption in Figure1b, the Rebonato approximation lies just within the error bounds. The tenor of both these swaptions is the same, yet for the swaption with the higher T_α, the Rebonato approximation overestimates volatility by 47 basis points.

One can also clearly identify the superior convergence of quasi-Monte Carlo compared to crude Monte Carlo. It is evident that when pricing according to QMC one should not need more than 20 000 samples as the change in the estimate is minimal.

Figure 1: Short-dated swaption pricing with increasing sample size

(a) Swaption volatility (1×5)

(b) Swaption volatility (10×15)
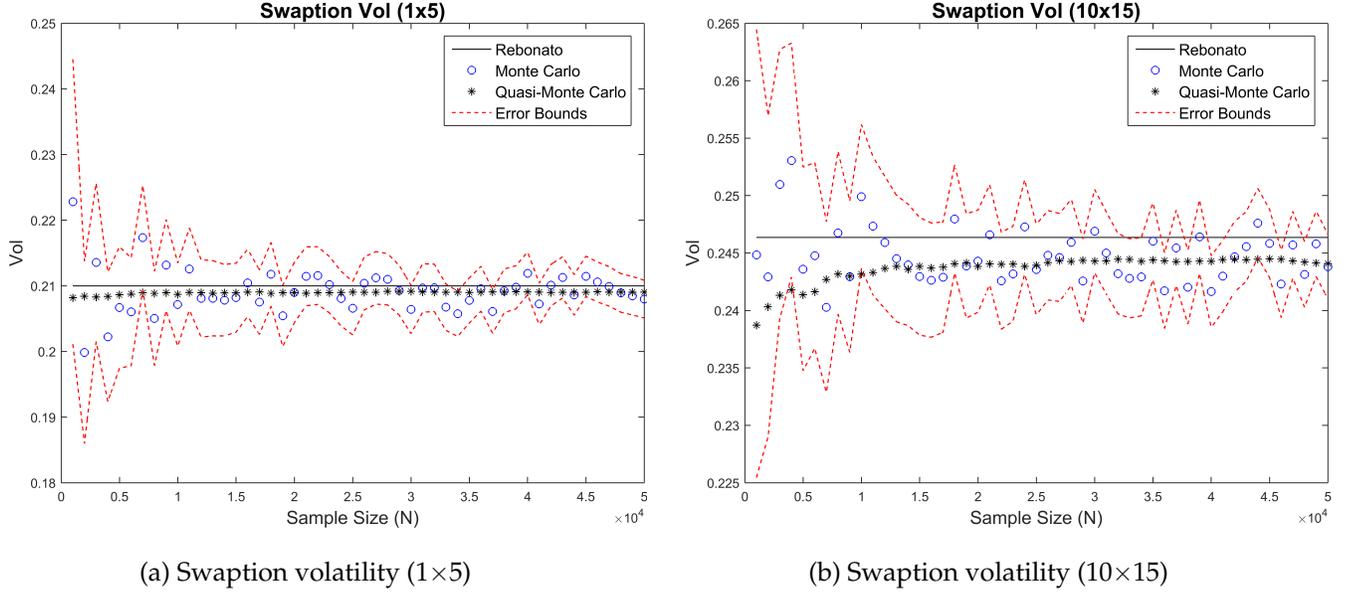
In Table 1 we display the relative volatility error (RVE), defined as

$$\text{RVE} = \frac{\upsilon_{\alpha,\beta}^{Rebonato} - \upsilon_{\alpha,\beta}^{MC}}{\upsilon_{\alpha,\beta}^{MC}},$$

for a flat yield curve (FYC) and observed yield curve (OYC).

Table 1: Short-dated relative volatility error for a FYC and the OYC (N = 10 000)

	1×5	1×10	1×15	5×10	5×15	10×10	10×15
FYC - Crude MC	0.46%	0.44%	0.78%	1.19%	1.81%	0.02%	2.04
FYC - Quasi-MC	0.05%	0.16%	0.22%	1.51%	1.66%	0.98%	1.93
OYC - Crude MC	0.97%	1.12%	3.96%	3.16%	4.28%	1.00%	1.04
OYC - Quasi-MC	0.46%	1.41%	2.93%	3.41%	4.16%	2.05%	0.93

The difference of 40 basis points in the volatilities of the longest maturity swaption does gives an indication of the effect of longer time horizons. The relative error on our caplet pricing validation was no higher than 2%, which supports the implementation of our Monte Carlo. In extending these simulations to long-dated swaption scenarios we found that the approximation began to break down the longer the time horizon. We consider a swaption with a fixed maturity of 25 years and

16

varying tenors ranging from 1 to 25 years. There appears to be a point around $T_\beta - T_\alpha = 10$ where the approximation and the simulation begin to diverge under this specific calibration.

Table 2: Long-dated relative volatility error (Crude Monte Carlo, $N = 10000$)

	Tenor ($T_\beta - T_\alpha$)						
	1	5	10	15	20	22	25
$T_\alpha = 25$	2.57%	1.05%	7.62%	20.31%	27.42%	35.35%	41.20%
	Maturity (T_α)						
	1	5	10	15	20	22	25
$T_\beta - T_\alpha = 25$	0.88%	0.08%	4.06%	14.41%	30.58%	34.49%	41.20%

Figure 2: Performance of $\upsilon_{\alpha,\beta}^{Rebonato} \neq \upsilon_{\alpha,\beta}^{MC}$ over long-dated swaptions
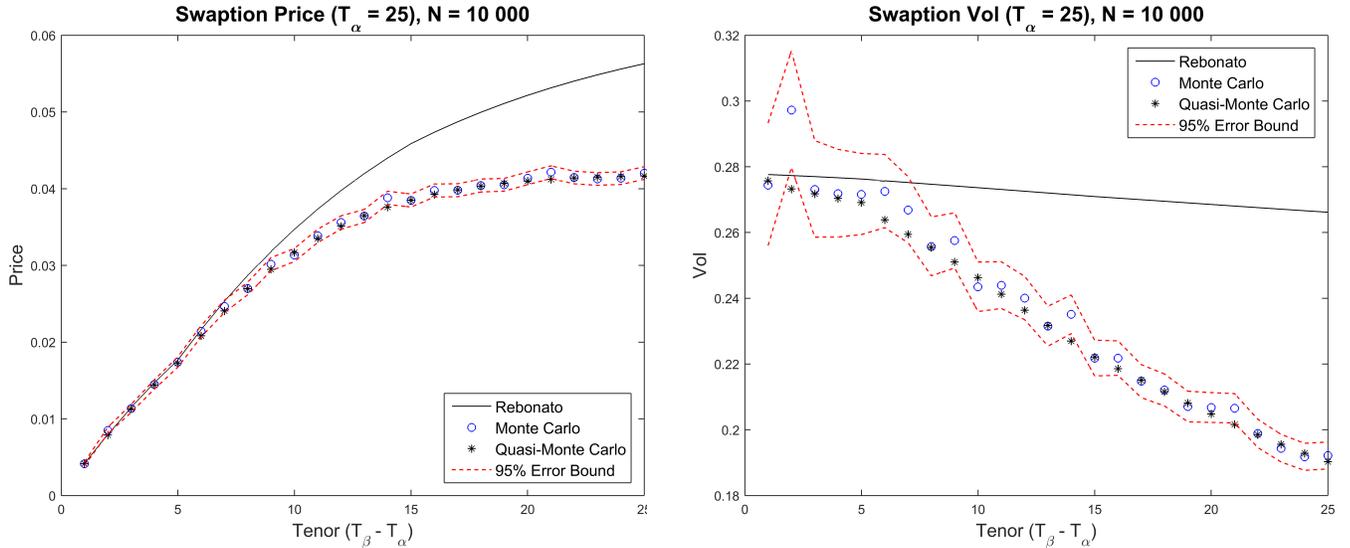
(a) Price discrepancy for fixed maturity (T_α)

(b) Volatility discrepancy for fixed maturity (T_α)

In Figure 1, for the longest time horizon $T_\beta - T_\alpha = 25$ there is a volatility discrepancy of 730 basis points, while the caplet relative error still supported Monte Carlo with levels no higher than 8%. These results highlight the limitation of Rebonato's approximation in pricing swaptions over long time horizons under this calibration. On a R10 000 000 25×25 swaption contract, the use of Rebonato's approximation would lead to a R730 000 price differential when compared to the quasi-Monte

Carlo estimate. However the differences between the prices and volatilities vary, which leads us to look at the sensitivities to further analyse the discrepancies.

4.2 Price and Volatility Sensitivity

In considering the same fixed maturity swaption as before ($T_\alpha = 25$), with N = 10 000, it may be of importance to consider price volatility sensitivities. The sensitivities can be calculated with the closed-form solution for the sensitivity in equation (3.3). This will give an indication of how sensitive the difference in price will be for a different volatility. Table 3 illustrates the sensitivity levels across short-dated swaptions. The sensitivity is calculated as $\frac{1}{\zeta}$ in order to consider the sensitivity of the volatility in terms of the price. Larger sensitivity and larger discrepancy of volatilities will, as predicted, cause larger discrepancies of prices.

Table 3: Basis point (BP) differentials (diff) and error sensitivities

	1×5	1×10	1×15	5×5	5×10	5×15	10×5	10×10	10×15
BP diff in Vol	9.675	27.770	56.940	42.355	78.470	94.909	32.184	48.323	22.618
BP diff in Price	1.178	5.911	15.740	8.991	28.615	43.394	6.697	16.168	9.042
$\frac{1}{\zeta}$	32.650	31.140	30.110	13.667	13.083	13.028	9.022	9.166	9.680

Figure 3: Consistency check with caplet price and swaption vega
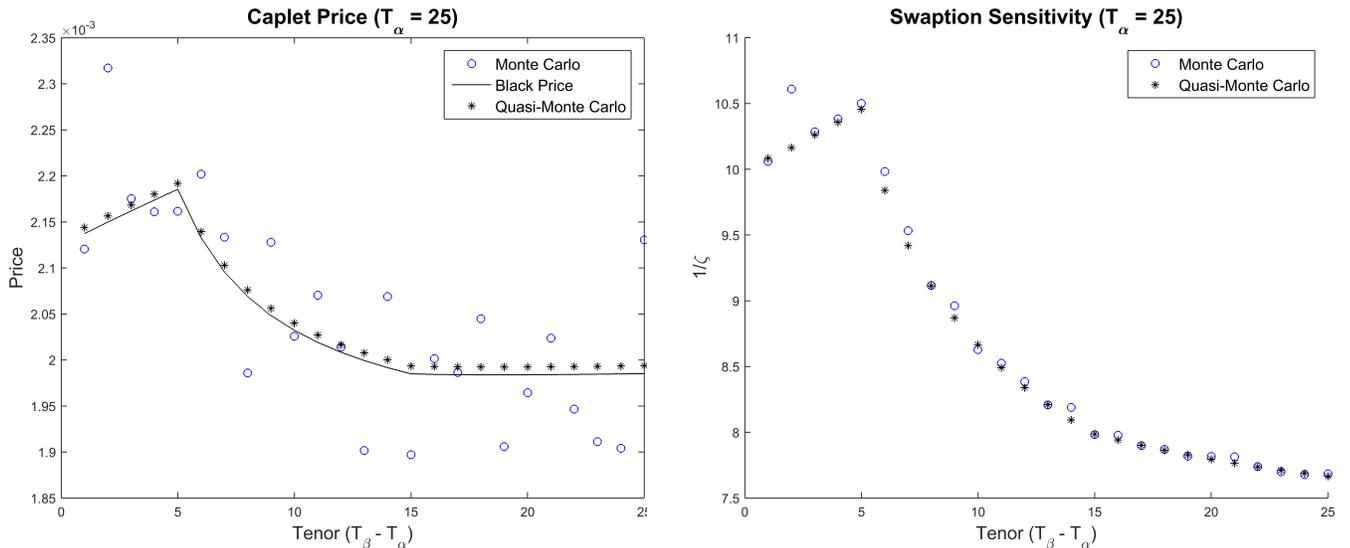
(a) Consistency check to the Black caplet price

(b) Swaption Vega as function of tenor

18

Plotting these sensitivities in Figure 3 appears to follow the movements of the caplet prices in this scenario. These sensitivities were calculated directly from the Black formula for ζ as a derivative of p^{MC}, which should explain the relationship.

What Figure 3 does give us is an indication of where a difference in volatilities will have a larger effect on the difference in price. It is not the absolute size of the sensitivity that matters, rather the relative size that indicates a scenario where the price is more/less sensitive to an erroneous volatility approximation.

4.3 Simulating in the ALM context

In the ALM context there are both restrictions on the number of simulated paths and greater interest in long-dated time horizons. We therefore consider our scenarios specific to these circumstances in order to understand the flexibility of the approximation and when it begins to stray.

We found that increasing the number of time steps per year in an attempt to improve the Euler approximation had little effect on the overall result, as shown in Figure 4. The minimum number or time steps per year was two, to match the reset dates of the calibrated yield instruments.

Figure 4: Swaption pricing with increased time steps per year
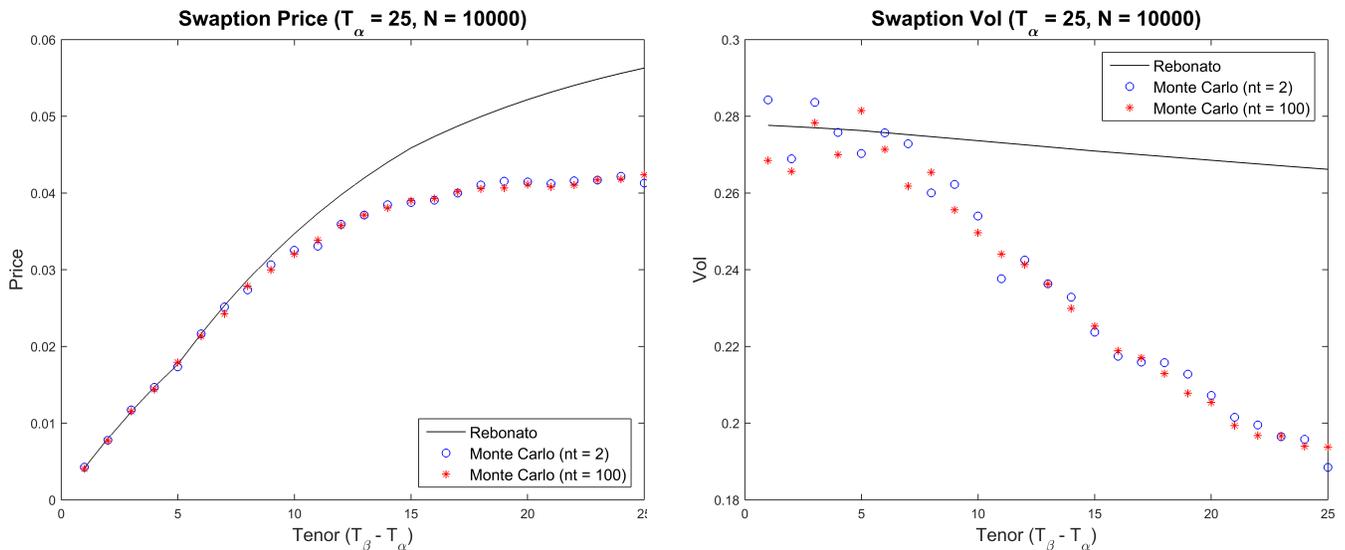
(a) Swaption price for $nt = 2$ and $nt = 100$ (b) Swaption volatility for $nt = 2$ and $nt = 100$

Implementing the predictor-corrector scheme in Figure 5 did not seem to improve results dramatically. Again, in the ALM context this scheme should optimise accuracy for a small number of samples. What we see here is how Rebonato's

approximation appears to lose accuracy under most circumstances as we move beyond a certain time horizon. This may be a result from the derivation itself, where the sample foward rates are "frozen" in their initial state (assumed to remain relatively constant through time).

Figure 5: Swaption price using the predictor-corrector (PC) method
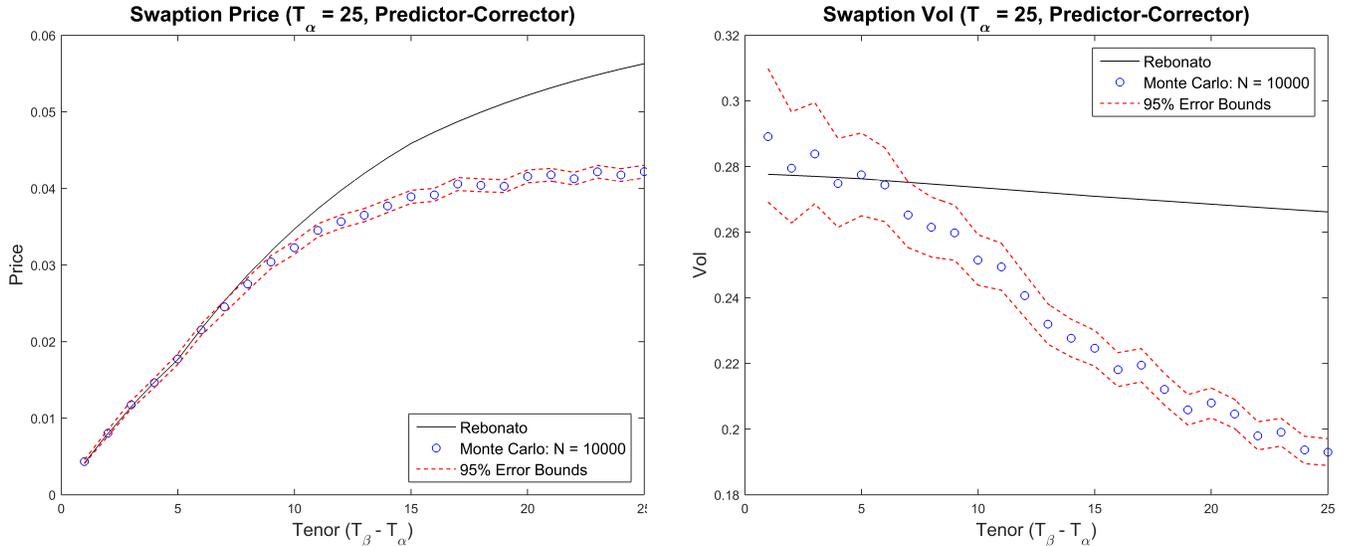
(a) Swaption price using the PC method

(b) Swaption volatility using the PC method

5 Multi-Curve Libor Market Model

In this Section, we will give a brief review of the extension of the LFM to a multi-curve setup as proposed in Mercurio (2010). This also leads to a Rebonato formula for approximating swaption prices in a multi-curve setup and therefore provides one possible answer to the second question formulated in the introduction. Due to the lack of data, we do not report on any simulations in the multi-curve setup.

5.1 From Single- to Multi-curve models

Already before the financial crisis in 2007, there was a slight discrepancy between standard interest rate models and observed market behaviour: There was a small spread between the overnight indexed swap (OIS) and the LIBOR rate. However, this spread was considered negligible. After 2007 the situation changed drastically and since then, the market has been showing a significant spread between the two rates. In the context of interest rate derivative pricing, this has thus created the need

for models in which the yield curve generating the cashflows and the curve with which these are discounted (the OIS rate) are distinct. For a detailed explanation, see Mercurio (2009).

Different approaches to incorporate this new feature have been taken in the literature, see for example Crépey et al. (2012) and further references therein. Notice that many of these models also tackle the problem of jointly modelling the OIS rate and a variety of LIBOR curves; in this review of Mercurio (2010) we restrict ourselves to a single LIBOR curve for notational simplicity.

5.2   Setup

Let us denote by $P_L(t, T)$ and $P_D(t, T)$ the zero-coupon-bond price at time t for maturity T associated to the LIBOR and OIS yield curve, respectively. For given times $T_0 < T_1, \cdots < T_M$, let us define as in the classical single-curve setup the (simply compounded) forward rate (associated to the respective curve $x \in \{L, D\}$) at time t with expiry T_{k-1} and maturity T_k as

$$F_k^x(t; T_{k-1}, T_k) = \frac{1}{\tau_k^x(T_{k-1}, T_k)} \left[\frac{P_x(t, T_{k-1})}{P_x(t, T_k)} - 1 \right],$$

for $t \le T_{k-1} < T_k$, where τ_k^x is the year-fraction associated with (T_{k-1}, T_k). Assuming the existence of a risk-neutral measure $\mathbb{Q}$, we define by $\mathbb{Q}_D^k$ the equivalent measure associated to the numéraire $P_D(\cdot, T_k)$, referred to as the T_k forward measure for the OIS curve.

Let us now consider the FRA (forward rate agreement) rate $L_k(t)$: this is the fixed rate K (at time t), that makes a contract paying $(F_k(T_{k-1}) - K)$ at time T_k fair at time t. Noting that the payoff of any traded asset discounted by $P_D(\cdot, T_k)$ is a martingale under $\mathbb{Q}_D^k$, by no-arbitrage we see that

$$L_k(t) = \mathbb{E}_D^k \left[F_k^L(T_{k-1}) | \mathcal{F}_t \right], \tag{18}$$

where $\mathcal{F}_t$ is the information available at time t. Thus, whereas in the single-curve setup the forward rate $F_k(t)^L$ coincides with the FRA rate L_k (and the forward rate F_k^D), here this does not hold anymore. Therefore, in contrast to the LFM, the stochastic model to be proposed will not be based on assumptions on the evolution of F_k, but rather on L_k. As can be seen from Equation (18), this is a martingale under $\mathbb{Q}_D^k$ and $L_k(T_{k-1}) = F_k(T_{k-1})$, which allows us to write the payoffs of e.g. swaptions and caplets on the LIBOR rate also in terms of the L_k. In this extended LIBOR market model, we now also want to impose a stochastic model for F_k^D. Denoting by $S_k(t)$ the spread

$$S_k(t) = L_k(t) - F_k^D(t),$$

we actually have three choices of what we want to be our fundamental modelling quantities:

- the rates L_k and F_l^D, or

- the rates L_k and spreads S_l, or

- the rates F_k^D and spreads S_l.

5.3 The lognormal Multi-Curve LFM

Still following Mercurio (2010), we choose the first option and impose lognormal dynamics on both rates. Thus, we assume that for each k the FRA rate evolves as a geometric Brownian motion under $\mathbb{Q}_D^k$,

$$dL_k(t) = \sigma_k(t) L_k(t) dZ_k(t)$$

for $t \le T_{k-1}$. Similarly, for the OIS forward rates we assume

$$dF_k^D = \sigma_k^D(t) F_k^D(t) dZ_k^D(t)$$

for $t \le T_{k-1}$. σ_k and σ_k^D are deterministic and Z_k^D, Z_k are Brownian motions under $\mathbb{Q}_D^k$ with correlation structure

$$d\langle Z_k, Z_j \rangle_t = \rho_{k,j} dt$$
$$d\langle Z_k^D, Z_j \rangle_t = \rho_{k,j}^{D,L} dt$$
$$d\langle Z_k^D, Z_j^D \rangle_t = \rho_{k,j}^{D,D} dt$$

chosen to ensure that the block matrix

$$R = \begin{bmatrix} \rho & \rho^{D,L} \\ (\rho^{D,L})' & \rho^{D,D} \end{bmatrix}$$

is positive semi-definite. Since this system of stochastic differential equations is formulated in terms of a different measure for each component, it is not obvious that a solution exists. However, we can fix a forward measure $\mathbb{Q}_D^\alpha$ and look at the dynamics that are implied by the above assumptions for *all* the FRA rates F_k and L_j under $\mathbb{Q}_D^\alpha$. This system can then be argued to admit a solution, see Section 5.2 of Mercurio (2010). For example, for $k = \alpha + 1, \ldots, \beta$ the dynamics under the T_α forward measure $\mathbb{Q}_D^\alpha$ are given as

$$
\begin{aligned}
dL_k(t) &= \sigma_k(t) L_k(t) \left[\sum_{h=\alpha+1}^{k} \frac{\rho_{k,h}^{L,D} \tau_h^D \sigma_h^D(t) F_h^D(t)}{1 + \tau_h^D F_h^D(t)} dt + dZ_k^\alpha \right] \\
dF_k^D(t) &= \sigma_k^D(t) F_k^D(t) \left[\sum_{h=\alpha+1}^{k} \frac{\rho_{k,h}^{L,D} \tau_h^D \sigma_h^D(t) F_h^D(t)}{1 + \tau_h^D F_h^D(t)} dt + dZ_k^{D,\alpha} \right].
\end{aligned}
\tag{19}
$$

A similar expression can be derived for $k = 1, \ldots, \alpha - 1$.

5.4 Pricing Caplets in the lognormal Multi-Curve LFM

Let us now examine how the formulas for caplet pricing reported in Section 2 change in this multi-curve setup. We are interested in the price at time 0 of a $T_{\alpha-1}$-caplet based on the LIBOR forward rate. As before, the payoff at time T_α is given as

$$H_{\alpha-1} = \tau_\alpha (F_\alpha^L(T_{\alpha-1}) - K)^+.$$

Changing the numéraire, we can see that the price at time 0 is therefore given as

$$\begin{aligned}
\mathrm{Cpl} &= \tau_\alpha P_D(0, T_\alpha) \mathbb{E}_D^\alpha[(F_\alpha^L(T_{\alpha-1}) - K)^+] \\
&= \tau_\alpha P_D(0, T_\alpha) \mathbb{E}_D^\alpha[(L_\alpha(T_{\alpha-1}) - K)^+] \\
&= \tau_\alpha P_D(0, T_\alpha) \mathrm{Bl}(K, L_\alpha(0), \sqrt{T_{\alpha-1}} v_\alpha),
\end{aligned}$$

where v_α^2 is given by equation (4) (with $i = \alpha$). This is completely analogous to the single-curve case: we only have to replace the discount factor by the one associated to the OIS-curve and replace the initial forward rate by the initial FRA-rate.

5.5 Swaption Pricing and Rebonato Formula for Lognormal FRA and OIS-Forward Rate Dynamics

As in the single-curve setup, we are interested in calculating swaption prices in this model.

5.5.1 Swap Rates

Firstly, let us consider an interest rate swap that exchanges at each $T_k \in \{T_{\alpha+1}, \ldots, T_\beta\}$ the LIBOR rate set at T_{k-1} for a fixed rate K. According to (22), the value at time $t \leq T_\alpha$ of receiving $\tau_k F_k(T_{k-1})$ at each time T_k is thus given as

$$\sum_{k=\alpha+1}^{\beta} P_D(t, T_k) \tau_k L_k(t) \tag{20}$$

whereas the value at time t of the fixed payments is given as

$$K \sum_{k=\alpha+1}^{\beta} \tau_k P_D(t, T_k). \tag{21}$$

The swap rate $S_{\alpha,\beta}(t)$ is defined as the fixed rate K which at time t renders the above contract fair. Combining equations (20) and (21), we obtain

$$S_{\alpha,\beta}(t) = \frac{\sum_{k=\alpha+1}^{\beta} \tau_k P_D(t, T_k) L_k(t)}{\sum_{k=\alpha+1}^{\beta} \tau_k P_D(t, T_k)},$$

which we can rewrite as

$$S_{\alpha,\beta}(t) = \sum_{k=\alpha+1}^{\beta} \omega_k(t)L_k(t) \quad , \quad \omega_k(t) = \frac{\tau_k P_D(t,T_k)}{\sum_{j=\alpha+1}^{\beta} \tau_j P_D(t,T_j)}. \tag{22}$$

5.5.2 Swaptions and the Rebonato Formula in the Multi-Curve Model

For a given rate K, a swaption is a financial contract that at time T_α gives its holder the right (but not the obligation) to enter a swap contract as specified above. The payoff of a swaption can thus be written as

$$H_{\alpha,\beta} = (S_{\alpha,\beta}(T_\alpha) - K)^+ \sum_{k=\alpha+1}^{\beta} \tau_k P_D(t,T_k).$$

Denoting by $\mathbb{Q}^{\alpha,\beta}$ the pricing measure associated with the numéraire $\sum_{k=\alpha+1}^{\beta} \tau_k P_D(\cdot,T_k)$, the value of a swaption at time 0 can be written as

$$\mathbb{E}^{\alpha,\beta}[(S_{\alpha,\beta}(T_\alpha) - K)^+] \sum_{k=\alpha+1}^{\beta} \tau_k P_D(0,T_k). \tag{23}$$

Assuming that we could write

$$dS_{\alpha,\beta}(t) = S_{\alpha,\beta}(t)v_{\alpha,\beta}(t)dZ_t^{\alpha,\beta} \tag{24}$$

for a $\mathbb{Q}^{\alpha,\beta}$-Brownian motion $Z^{\alpha,\beta}$ and a deterministic $v_{\alpha,\beta}$, we would obtain

$$d[S_{\alpha,\beta}, S_{\alpha,\beta}]_t = v_{\alpha,\beta}(t)^2 S_{\alpha,\beta}^2(t)dt. \tag{25}$$

On the other hand, if we froze the weights in Equation (22) at 0 and calculated the quadratic variation of the resulting expression, we should get

$$d[S_{\alpha,\beta}, S_{\alpha,\beta}]_t \approx d[\sum_{k=\alpha+1}^{\beta} \omega_k(0)L_k, \sum_{h=\alpha+1}^{\beta} \omega_h(0)L_h]_t$$

$$= \sum_{k,h=\alpha+1}^{\beta} \omega_k(0)\omega_h(0)L_k(t)L_h(t)\rho_{k,h}\sigma_k(t)\sigma_h(t)dt.$$

If we equate this with Equation (25) and replace $L_k(t), L_h(t)$ and $S_{\alpha,\beta}(t)$ by their values at time 0, we get that $v_{\alpha,\beta}(t)$ should be of the form

$$v_{\alpha,\beta}^2(t) = \sum_{k,h=\alpha+1}^{\beta} \frac{\omega_h(0)\omega_k(0)L_h(0)L_k(0)\rho_{k,h}}{(S_{\alpha,\beta}(0))^2}\sigma_h(t)\sigma_k(t).$$

24

If we now insert the assumed lognormal dynamics (Equation (24)) in the expectation in equation (23), as in the single-curve case we get a closed form approximation for the price of a swaption. Namely, we approximate the swaption price (Equation (23)) by

$$\sum_{j=\alpha+1}^{\beta} P_D(0, T_j) \text{Bl}(K, S_{\alpha,\beta}, \sqrt{T_\alpha} v_{\text{Reb}})$$

where

$$v_{\text{Reb}}^2 = \frac{1}{T_\alpha} \sum_{k,h=\alpha+1}^{\beta} \frac{\omega_h(0)\omega_k(0)L_h(0)L_k(0)\rho_{k,h}}{(S_{\alpha,\beta}(0))^2} \int_0^{T_\alpha} \sigma_h(t)\sigma_k(t) dt.$$

Note that again this is very similar to the derivation of the Rebonato formula (15) in Section 2.

6   Conclusions

Overall, we have proposed three different approaches to handle the problem formulated in the introduction. Firstly, we have proposed two consistency checks that allow us to evaluate the quality of the Monte Carlo approximations. For long-dated maturities the Rebonato formula seems to give a particularly good approximation, which makes this a very useful tool. Secondly, the standard sensitivity measure allows one to judge whether the Monte Carlo pricing error is amplified in the implied volatility or not. Thirdly, we have examined different variants of Monte Carlo pricing algorithms. We have seen that using smaller time-steps or Predictor Corrector does not seem to make a big difference, whereas using quasi-Monte Carlo instead of Monte Carlo does, as was to be expected.

Nevertheless, the Rebonato formula still has its virtues. This closed-form approximation to the implied volatility of a swaption is very accurate in some scenarios. However, one should always be aware of the limitations of this approximation.

Finally, we review the Multi-Curve LIBOR market model as introduced in Mercurio (2010) and derive the Rebonato formula in this setup. This provides one possible answer to the second question. It would be very interesting to extend the present report and examine the accuracy of the Rebonato formula in this multi-curve setup.

Open Questions

- In Proposition 6.15.2 in Brigo and Mercurio (2006) a second order approximation for the implied volatility is presented. According to their numerical results, the difference between the Rebonato formula and this formula is

practically negligible. Is this still true in the ALM context of long-dated swaptions? Could this second order approximation be used as a substitute for the Rebonato formula in the ALM context?

- Time constraints have only allowed us to numerically examine the Rebonato approximation in a single-curve setup. In Chapter 5, however, we have reviewed the analogous formula in a multi-curve LFM. Does the formula give good approximations when simulating in this setup? Does the approximation quality still deteriorate in the ALM context?

- The Rebonato formula approximates swaption prices by prices on a basket of forward rates. Can we exploit this connection to use existing results on basket options for a more through, mathematically rigorous analysis of the approximation quality?

Bibliography

Brace, A., Gatarek, D., Musiela, M., 1997. The market model of interest rate dynamics. Mathematical Finance 7(2), 127–154.

Brigo, D., Liinev, J., 2005. On the distributional distance between the lognormal libor and swap market models. Quantitative Finance 5, 433–442.

Brigo, D., Mercurio, F., 2006. Interest Rate Models: Theory and Practice - with Smile, Inflation and Credit. Springer Verlag.

Crépey, S., Grbac, Z., Nguyen, H. N., 2012. A multiple-curve hjm model of interbank risk. Mathematics and Financial Economics 6(3), 155–190.

Hunter, C. J., Jackel, P., Joshi, M. S., 2001. Drift approximations in a forward-rate-based libor market model. Market Model. Getting the Drift. Risk 14, 81–84.

Joe, S., Kuo, F., 2008. Constructing sobol sequences with better two-dimensional projections. SIAM Journal on Scientific Computing 30, 26352654.

Mercurio, F., 2009. Interest rates and the credit crunch: new formulas and market models. Technical report, Bloomberg Portfolio Research Paper No. 2010-01-Frontiers.

Mercurio, F., 2010. Modern libor market models: Using different curves for projection rates and for discounting. International Journal of Theoretical and Applied Finance 13, 1–25.

Miltersen, K., Sandmann, K., Sondermann, D., 1997. Closed form solutions for term structure derivates with log-normal interest rates. Journal of Finance 52(1), 409–430.

Valuation of callable floating rate notes with write down features

Team 5

MARIO GIURICICH, University of Cape Town
GABRIËL LE ROUX, University of Cape Town
KANGJIANAN XIE, University College London
DAVID STEFANOVITS, ETH Zurich

Supervisors:
OBEID MAHOMED, University of Cape Town
GIDEON VAN DER LINDE, Prescient Securities

Contents

1 Introduction and problem statement

The third variation of the Basel Accords (known as Basel III) arose because of the deficiencies in financial market regulation, noticed heavily during the 2008 financial crisis. One of the principal aims of Basel III is to strengthen and crystallise global capital and liquidity regulation, with the ultimate goal of encouraging a more resilient banking sector.

In January 2013, Basel III was implemented in South Africa. A significant revision of regulatory banking capital structures followed, leading to banks attempting to access capital via very innovative structures. These new innovative capital structures had to meet the new conditions and stipulations of Basel III. One of these new conditions was the classification of capital instruments into three tiers of subordination, ranging from the lowest priority to the highest - namely Tier 1, Tier 2 and Tier 3.

We now focus on Tier 2 instruments in the South African context. Within this context, most instruments slotting into Tier 2 take the form of callable floating-rate debt instruments. Various pricing methodologies may be applied to these floating-rate debt instruments, but it must be borne in mind that there is no clear market standard on their valuation. The opacity surrounding the pricing of these instruments, together with a lack of liquidity opens up potential arbitrage opportunities, as well as inconsistent mark-to-market processes. These shortfalls, as well as the penultimate fact that banks issue these instruments, more than justify the importance and the need for the construction of a robust and rigorous valuation framework.

Accordingly, our aim is to develop a reasonable yet robust methodology for valuing instruments of this nature within the South African market context.

In summary, the particular characteristics of the product that we consider is as follows. We consider a floating-rate note (FRN) issued by a defaultable entity with 10 years to maturity. We also introduce the option for the issuer to call back the bonds at year 5 or anytime after year 5 at the discretion of the issuer. Moreover, the issuer can activate during the whole time of the contract a write-down mechanism based on its solvency situation.

The first challenge of the problem is to understand the economics behind the contract and formulate a precise criterion based on which the issuer can judge whether or not to call back the bond or write down the principal.

The option to call back will be valuable to the issuer if he can refinance debt in the open market at a lower coupon (or credit spread). This means that for the issuer, it is optimal to call back the bond if its credit spread improves compared to the spread when the bond was originally issued. For this reason we decided to value the callability option based on market credit spreads and risk-free rates.

In the case of a bank issuing this instrument, the write-down feature is most likely tied to the issuer's capital ratios. We may expect the bank to write-down the principal if its Core Equity Tier 1 (CET1) ratio falls below a certain trigger. Therefore, the write-down feature should be most likely valued based on the issuer's book or market value of equity as well as the issuer's Risk Weighted Assets (RWA). Unfortunately, this is difficult to implement, especially in the South African context. See, for example, Brigo et al. (2015) for a firm value approach on this problem. In particular, a very large amount of data, which generally is not publicly or easily available, would be needed. For this reason we propose a simplified solution. We assume that the decision to write-down the principal is also tied to the credit spread of the issuer. In particular, we assume that if the credit spread of the issuer deteriorates

3

above a specific level during the life of the instrument, then the issuer will trigger a principal write-down. Under this simplifying assumption we are basically considering a high credit spread event, which is directly comparable to a low CET1 ratio event. To address the shortcoming of not considering the CET1 ratio directly, we develop a simple method to relate credit spread increases to equity devaluations.

The second challenge is to start with a model which is feasible for simulations of the underlying processes, in particular for the credit spread process. In Sections 5.1 and 5.2 below, we have considered tractable short-rate and reduced-form intensity models, which allow to simulate risk-free interest rates and credit spreads.

In order to address the above challenges, the report is structured as follows. Section 2 provides a brief background on the regulatory requirements for banks' capital, and brings to the fore further reasons as to why a valuation methodology for the instruments is necessary. A description of the valuation of vanilla FRNs in a South African context follows in Section 3, this being necessary because the basic underlying product of our instrument is essentially a vanilla FRN. Thereafter in Section 4, we propose the payoff structure of the product, and the discounted expected value thereof is presented. In order for the valuation to proceed, we require models for the underlying risk factors of this instrument and these are presented in Section 5. Section 6 heuristically relates the probability of default from the structural Merton model for equity to the reduced-form credit risk model we propose in Section 5. After that, Section 7 introduces the Kalman filter which we use in our work to estimate parameters. Sections 8 and 9 illustrate how we calibrated our models for the risk factors to market data and the valuation results respectively. Finally, section 10 concludes.

2 Hybrid capital securities and regulatory capital requirements

Hybrid capital securities such as CoCos or bonds with write-down features automatically absorb losses when the capital of the issuing bank falls below a certain level. These types of instruments have two defining characteristics: the mechanism of the loss absorption and the trigger that activates the mechanism. The former is either conversion into common equity or principal write-down. The trigger can be either mechanical or discretionary.

The most common book-value triggers are defined in terms of the book value of Common Equity Tier 1 (CET1) capital as a ratio of risk-weighted assets (RWA). The most common market-value triggers are set as a minimum ratio of the bank's stock market capitalisation to its assets. Figure 1 depicts the Basel III capital directives.
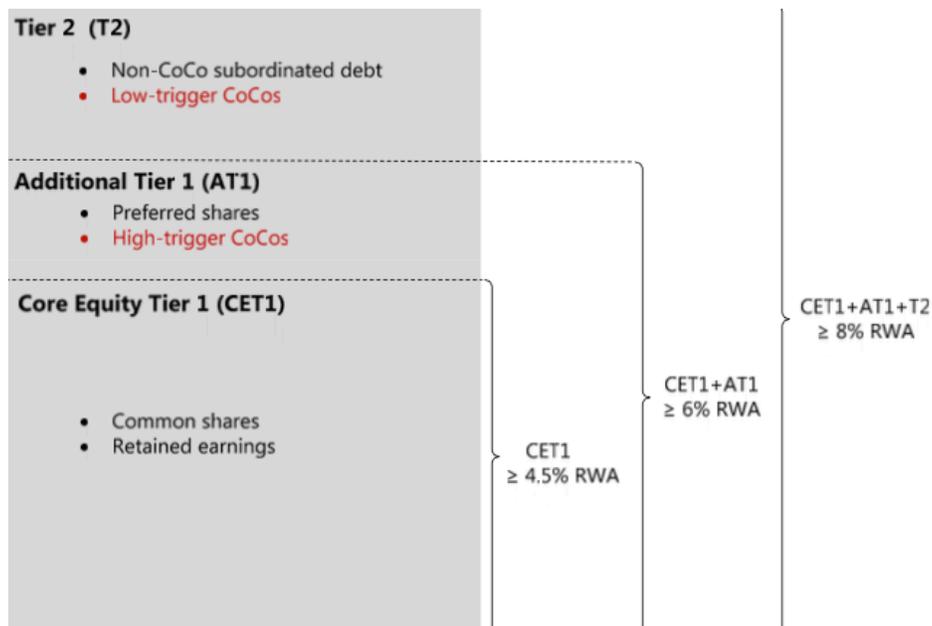
Figure 1: Basel III capital directives. *Source: BIS*

Under Basel III instruments such as CoCos or bonds with write-down features qualify as either Additional Tier 1 (AT1) or Tier 2 (T2) capital. As shown in the figures below such securities offer very attractive coupon rates to compensate investors for the risk of loss absorption. As a consequence of the pressure from markets and regulators to boost banks' capital, along with increasing coupon rates for investors, the volume of hybrid capital instruments issued for the purposes of AT1 or T2 provisions has increased considerably since 2012. For this reason it is critically important to have a valuation framework for these kinds of instruments. In particular, this report studies callable floating-rate bonds with write-down features as described in the problem statement. These types of bonds currently qualify as T2 instruments for South African banks.

Issuer	ISIN	Issue Date	Tier	Maturity	Ccy	Cpn	Reset	Amount (mn)	Coupon language	Basis of Trigger	Loss Absorption Mechanism	Trigger	Reg Call / Cap. event
Alternative Tier 1													
Banco Popular	XS0979444402	01-Oct-13	AT1	Perp-18	€	12%	5yr € m/s+1023.7	500	Fully opt.	B3 (Transitional) CET1 / RWAs	Full conversion to equity	B3 CET1 <5.125%, T1<6% AND Bank/Group losses 4MRQs AND Capital+Reserves down by 1/3rd	Par
Societe Generale	XS0867614595	29-Aug-13	AT1	Perp-18	$	8%	5y$ m/s+639.4	1250	Fully opt.	B3 (Transitional) CET1 / RWAs	Temporary principal write-down	B3 CET1 <5.125%	Current princ. amt.
BBVA	XS0926832907	26-Apr-13	AT1	Perp-18	$	9%	5y$ m/s+826.2	1500	Fully opt.	CET1 / Capital Principal / EBA CT1 / T1 Ratio	Full conversion to equity	1) CET1 < 5.125% 2) EBA CT1 < 7.0% 3) Cap. Principal < 7.0% 4) T1 < 6.0%	Par
Credit Suisse	XS0810846617	31-Jul-12	AT1 (BCN)	Perp-18	$	10%	6m $ m/s+665	1725	Fully opt.	B3 (Transitional) CET1 / RWAs	Full conversion to equity	1) B3 CET1 <7.0% or 2) non-viability declaration by FINMA	Par
Tier 2 / Senior													
Credit Agricole	US225313AC92	12-Sep-13	T2	33-18	$	8%	5y$ m/s+628.3	1000	Must pay	B3 (Transitional) CET1 / RWAs	Full permanent writedown	B3 CET1 < 7%	Par
Credit Suisse	XS0957135212	08-Aug-13	T2 (Low Trigger)	Aug-23	$	7%	-	2500	Must pay	B3 (Transitional) CET1 / RWAs	Full permanent writedown	1) B3 CET1 <5.0% or 2) non-viability declaration by FINMA	Par / 103 on capital event
UBS	CH0214139930	22-May-13	T2	23-18	$	5%	5y$ m/s+376.5	1500	Must pay	B3 (Transitional) CET1 / RWAs	Full permanent writedown	1) B3 CET1 <5.0% or 2) non-viability declaration by FINMA	Par / 101 on capital event
Barclays	US06739FHK03	04-Apr-13	T2	23-18	$	8%	5y$ m/s+683.3	1000	Must pay	B3 (Transitional) CET1 / RWAs	Full permanent writedown	B3 CET1 <7.0%	Par

Figure 2: Example of AT1 and T2 hybrid capital securities. *Source:* Citi Research Lorenzen (2015)

3 Pricing vanilla floating-rate notes in South Africa: expounded and applied

3.1 Introduction to vanilla floating-rate notes

Many variations of coupon-bearing bonds exist. One such variation is the floating-rate note (FRN), which is commonly issued by entities as a debt instrument which pays periodic coupon payments as well a nominal amount at maturity. In a vanilla FRN, the coupon payment is linked to a pre-specified floating reference rate. The first coupon of the FRN is typically known on the settlement date, but subsequent coupons are unknown because they are linked to future realisations of the reference rate. Each coupon has an associated coupon-paying date and coupon period, which are both generated prior to issue of the note. At the beginning of the coupon period, this date being known as the reset date, the reference rate determining the coupon is set, and at the end, the coupon is paid to the note holder. As an example of a vanilla FRN, consider the following: a note, with coupons determined by a (floating) reference rate such as the 3-month Johannesburg Interbank Agreed Rate (JIBAR), with coupons payable quarterly in arrears and coupon periods of 3 months.

Since FRNs are issued by entities quite separate from government, an element of credit risk is attached. The credit risk of the issuer is captured, in part, by attaching a pre-specified, fixed spread to the coupon payments. This spread also reflects the liquidity of the issued floating rate note i.e. the demand for this type of instrument in the market where it is traded, this market in South Africa being the Johannesburg Securities Exchange's (JSE) Debt Market.

Definition: The fixed spread, added to the coupon payments of a FRN, which is pre-specified at the outset of the FRN, is known as the *issuance spread*.

3.2 JSE valuation methodology for floating-rate notes

3.2.1 Assumptions

The JSE's model to value FRNs makes the following assumptions:

- Investors use FRN prices as deciding factors for entering into trades.

- Coupon rates vary, with the variation driven by the reference rate of the FRN.

- Coupon periods and coupon paying-dates are pre-specified.

- Dates are subject to the modified-following day-count convention: interest is payable on coupon-paying dates, unless that date falls on a weekend or public holiday, in which case the interest is paid on the next business day and the coupon includes these extra days.

3.2.2 Valuation methodology

The following methodology applies to valuing vanilla FRNs at dates during the term of the FRN. The setup of the vanilla floating-rate note is as follows - see Figure below for a diagrammatical depiction. For the FRN, let:

- N be the nominal amount of the FRN;

- t_s be the settlement date, which is assumed to fall between coupon-paying dates;

- t_n be the maturity date;

- t_r be the last reset date prior to the settlement date;

- $\{t_1, t_2, ..., t_n\}$ be the coupon-paying dates;

- $L(t_{i-1}, t_i)$ be the simple reference (floating) rate between coupon-paying dates t_{i-1} and t_i (for $i \in \{1, 2, ..., n\}$);

- $f(t_s, t_{i-1}, t_i)$ be the simple forward rate between coupon-paying dates t_{i-1} and t_i (for $i \in \{1, 2, ..., n\}$);

- $\tau_i = \left(\frac{t_i - t_{i-1}}{365}\right)$ be the period between coupon-paying dates;

- $B(t_s, t_i)$ be the default-free discount function at each coupon date and is defined to be the product of the forward period discount functions[1].

- $S(0, t_i - t_{i-1})$, for each i, be the fixed issuance spread agreed upon at issuance of the FRN, for the given tenor of the FRN, and

- $S(t, t_i - t_{i-1})$ be the market spread (defined below), agreed at any time $t > 0$ for the given coupon tenor.
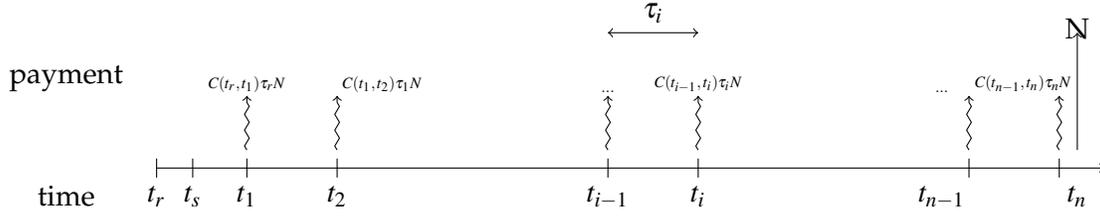
Figure 3: FRN cashflow depiction. For ease of notation, note that $C(t_{i-1},t_i) = L(t_{i-1},t_i) + S(0,t_i - t_{i-1})$. So, the coupon is dependent upon both the future reference rate and the issuance spread.

The following definition is also necessary: The spread, added to the forward rates in the forward period discount factors, is known as the *market spread*. The market spread varies with the valuation date $t > 0$ of the FRN, but is fixed for the remaining term of the FRN.

The following steps below follow the FRN pricing specifications given by the JSE - see Mavuso (2012) for further details. The JSE's methodology follows the steps outlined below. This methodology is commonly known as the *discount margin method*, and it involves discounting all cashflows to the settlement date. The issuance spread and market spread are taken as inputs to the valuation - these spreads are quoted in the market. The issuance spread is fixed from the issue date of the FRN, while the market spread changes with time.

1. Using the current bid and offer swap-zero curves, generate the mid swap-zero rate corresponding to the coupon dates. Use linear interpolation to interpolate the zero rates at each coupon dates.

2. Using no-arbitrage arguments, it is possible to justify the replacement of the future (floating) reference rates with their forward counterparts. Calculate the forward rates, for each coupon period, by using the discount factors for each of the future coupon-paying dates. Using the relation

$$\left(1 + y(t_s,t_{i-1})\frac{t_{i-1}-t_s}{365}\right)(1 + f(t_s;t_{i-1},t_i)\tau_i) = \left(1 + y(t_s,t_i)\frac{t_i-t_s}{365}\right), \tag{1}$$

an expression (see Equation 2 below) for the forward rate is obtained. $y(t_s,t_{i-1})$, for all i, represents the simple mid swap-zero rate, interpolated linearly off the market swap-zero curve observed at the settlement date t_s.

$$f(t_s;t_{i-1},t_i) = \frac{1}{\tau_i}\left(\frac{B(t_s,t_{i-1})}{B(t_s,t_i)} - 1\right) \tag{2}$$

Equation 2 above can be used to replace the future unknown reference rates $L(t_{i-1},t_i)$ for each i.

3. Predict the future coupon rates using the forward rates from step 2 as well as the fixed issuance spread. At a coupon-paying date t_i, the coupon of the FRN is given by

$$N(L(t_{i-1},t_i) + S(0,t_i - t_{i-1}))\tau_i, \tag{3}$$

---

[1]For example, the forward period discount function over the coupon period t_{i-1} to t_i is given by $(1 + f(t_s;t_{i-1},t_i)\tau_i)^{-1}$

and if the reference rate is replaced, Equation 3 becomes

$$N(f(t_s, t_{i-1}, t_i) + S(0, t_i - t_{i-1}))\tau_i, \tag{4}$$

4. Discount each future coupon payment, as well as the nominal received on maturity (time t_n), back to the settlement date t_s, using the simple mid swap-zero rates together with the market spread. That said, for a prototypical coupon-paying date t_i this means that the simple rate used to discount to time $t > 0$ is equal to $(Y(t, t_i) + S(t, t_i - t_{i-1}))$. Furthermore, in order to ensure that the FRN always prices at par on its issue date (and further assuming that the issue spread is equal to the market spread), the discount functions $\bar{B}(t_s, t_i)$, for each i, are constructed using the product of the forward period discount factors. Equation 5 below illustrates the generic discount function for the coupon-paying date t_i.

$$\bar{B}(t, t_i) = (1 + (y(t, t_1) + S(t, t_i - t_{i-1}))\tau_i)^{-1} \prod_{j=2}^{i} (1 + (f(t; t_{j-1}, t_j) + S(t, t_i - t_{i-1}))\tau_j)^{-1} \tag{5}$$

Therefore, the present value of the coupon paid at the coupon-paying date t_i is given by

$$N(f(t; t_{i-1}, t_i) + S(0, t_i - t_{i-1}))\tau_i \bar{B}(t, t_i). \tag{6}$$

5. Finally, sum all the present values of the coupons, as well as that of the nominal payable at maturity, in order to find the present value of the FRN, $V_{\text{FRN}}(t, N, \mathscr{T}, S(t, t_i - t_{i-1}))$, where $\mathscr{T} = \{t_0, t_1.t_2, ..., t_n\}$ is the vector representing the coupon-paying dates of the FRN. Equation 7 illustrates this sum.

$$V_{\text{FRN}}(t, N, \mathscr{T}, S(t, t_i - t_{i-1})) = \sum_{i=1}^{n} N(f(t; t_{i-1}, t_i) + S(0, t_i - t_{i-1}))\tau_i \bar{B}(t, t_i) + N\bar{B}(t, t_n) \tag{7}$$

However, the first coupon, due at time t_1, is typically known at time $t \in [t_0, t_1)$. Therefore, Equation 7 can be rewritten as follows:

$$V_{\text{FRN}}(t, N, \mathscr{T}, S(t, t_i - t_{i-1})) = N(y(t_0, t_1) + S(0, t_i - t_{i-1}))\tau_1 \bar{B}(t, t_1) \tag{8}$$

$$+ \sum_{i=2}^{n} N(f(t; t_{i-1}, t_i) + S(0, t_i - t_{i-1}))\tau_i \bar{B}(t, t_i) + N\bar{B}(t, t_n)$$

6. Steps 1 to 5 provide a methodology for the valuation of FRNs, but ignore the books close date for the coupons. The valuation of vanilla FRNs is now formalised more fully. Let t_{BCD} denote the books close date of the first coupon due - that coupon at t_1. Note that $t_{BCD} < t_1$. Define a cum/ex coupon function by

$$\phi = \begin{cases} 1 & \text{if } t_s < t_{BCD} \\ 0 & \text{if } t_s \geq t_{BCD}, \end{cases} \tag{9}$$

and let the number of days worth of accrued interest be given by

$$d = \begin{cases} t_s - t_r & \text{if } \phi = 1 \\ t_s - t_1 & \text{if } \phi = 0. \end{cases} \tag{10}$$

9

Therefore, the unrounded all-in price of the vanilla FRN at trade date t, with settlement date t_s, is given by

$$V_{\text{FRN}}^A(t_s,N,\mathscr{T},S(t,t_i-t_{i-1})) = \phi N(Y(t_r,t_1)+S(0,t_i-t_{i-1}))\tau_1\bar{B}(t;t_s,t_1) \tag{11}$$

$$+ \sum_{i=2}^{n} N(f(t;t_s,t_{i-1},t_i)+S(0,t_i-t_{i-1}))\tau_i\bar{B}(t;t_s,t_i) + N\bar{B}(t;t_s,t_n),$$

where

$$\bar{B}(t;t_s,t_i) = \frac{\bar{B}(t,t_i)}{\bar{B}(t,t_s)},$$

is the forward discount function over $[t_s,t_i]$ as seen at t. To calculate the clean price, it is necessary to first find the unrounded accrued interest, which is given by Equation 12 below.

$$\mathbb{I}_{\text{FRN}} = N(Y(t_r,t_1)+S(0,t_i-t_{i-1}))\tau_1\frac{d}{365}. \tag{12}$$

Ultimately, the unrounded clean price of the vanilla FRN is given by

$$V_{\text{FRN}}^C(t_s,N,\mathscr{T},S(t,t_i-t_{i-1})) = V_{\text{FRN}}^A(t_s,N,\mathscr{T},S(t,t_i-t_{i-1})) - \mathbb{I}_{\text{FRN}}. \tag{13}$$

As a concluding remark to the valuation of FRNs, it must be emphasised that the above steps apply to vanilla FRN valuation after the issue date. The value of the FRN will vary in the following three ways (note that these variations can be proved mathematically):

- if both $S(0,t_i-t_{i-1}) = S(t,t_i-t_{i-1})$ and the valuation date t falls on a reset date, then the FRN will trade at par. Furthermore, this observation speaks to the fact that the FRN will always be valued at par when it is originally issued if the issue date and settlement date coincide.

- if $S(0,t_i-t_{i-1}) < S(t,t_i-t_{i-1})$, the FRN will trade at a discount.

- if $S(0,t_i-t_{i-1}) > S(t,t_i-t_{i-1})$, the FRN will trade at a premium.

4 Instrument specification and valuation

This section will comprise four parts, so as to build up the valuation from its simplest case. We firstly consider valuing the callable floating rate note (CFRN) in the case when the issuer can only exercise at a single time point. At this time point, should there be a certain improvement in the credit spread, the issuer will recall the CFRN and reissue the underlying FRN at a lower spread - hence the European-style features of the CFRN. Secondly, we consider valuing this CFRN in the case where the issuer can exercise over a continuous range of times from a pre-specified time point onwards, i.e. we consider an American-style call feature. Therefore, the issuer can reissue the underlying FRN at a time point of his own choice (from the pre-specified time onwards, obviously). Finally, in both the European and American cases we allow for the possibility of write-down of the underlying FRN. The write-down can occur at any time during the life of the CFRN.

| Value of CFRN | = | Value of FRN | − | Value of issuer's callability | − | Value of write-down |

Figure 4: Heuristic valuation scheme of the CFRN: the instrument can be decomposed into a (deterministic) FRN, a "callability" option and a "write-down" option.

We value the CFRNs from the point of view of the investor. Therefore, from the perspective of the investor the ultimate value of the CFRN can be depicted in a heuristic fashion as illustrated in Figure 4 below.

Before proceeding, a distinction between default and write-down is needed. In our case, write-down of the nominal value of the FRN occurs when the credit quality of the issuer of the instrument worsens to a level below some pre-specified level. The nominal value can either be partially or fully written down in our model, however, South African practice is to fully write-down the nominal value to zero. It should be noted that the event of default is almost idiosyncratic to the write-down event, however both these events are systematically dependent. Default, in our case, is defined to be the first time the credit spread jumps to a very high level, which is high enough to infer that the company will no longer operate as a going-concern. This (high) level will not be specified in our contract, but it is given that this level is higher than the pre-specified level in the case of write-down.

4.1 Vanilla European callable floating-rate note valuation framework

4.1.1 The case without write-down

The first framework proposed is rather simple, and is useful for the purpose of estimating the value (or equivalently the excess spread) of the callability feature of the vanilla European callable FRN. It assumes that the underlying FRN cannot be written down, however default by the issuer is permitted. The callability feature of the FRN is, however, accounted for: at a contractually specified time T, the issuer of this FRN has the option to recall the FRN and re-issue it, should the issuer's creditworthiness deem this so. So for example, if the issuer's creditworthiness has improved at time T - this improvement being defined by a lower credit spread - then the issuer will exercise his option to recall the FRN and re-issue it to the investor at a lower credit spread (reflecting the improvement in the issuer's creditworthiness). Note that the option embedded in the FRN is a possible decrease in the coupon payments payable by the issuer after the exercise date. Therefore, because of the embedded optionality, one can deduce that this product should trade at a lower price than a standard FRN (or equivalently at a higher net fixed spread).

We now value the vanilla European callable FRN. Let $\mathcal{T} = \{t_0, t_1.t_2, ..., t_n\}$ be the vector representing the coupon payment dates of the FRN - the coupon payment dates are assumed to conform to a regular grid size, with intervals of length $t_i - t_{i-1} = \delta$. Consider now the value of the vanilla European callable FRN. The payoff of this FRN for the option holder (i.e. the issuer of this FRN) at time T is the gain in re-issuing the FRN at the time T market spread, $S(T, T + t_i - t_{i-1})$, only if this market spread is lower than a pre-specified lower threshold level. This threshold level (i.e. a floor of $\underline{S}$) is contractually set at a viable level lower than the original issuance spread $S(0, t_i - t_{i-1})$. Furthermore, the valuation of the option is performed under the risk-neutral probability measure, with the option payoff being discounted at the

11

risk-free rate.

Before we present the valuation formula, we define some further notation needed[2]. Define:

- $t_0 = 0$, so that the instrument is valued at time 0. Also, fix $t_n = \bar{T}$

- $S(t) = S(t, t + \delta)$.

- τ to be the time of default of the issuer.

- $V_{\text{FRN}}(t, N, \mathscr{T}, s)$ value of floating rate note at time t with notional N, time grid $\mathscr{T}$ and market spread s. We set $V_{\text{FRN}}(t, N, \mathscr{T}, \infty) = 0$ and $S(t) = \infty$ when $\tau \leq t$.

- $\bar{S}$ is the upper limit, set higher than the original issuance spread $S(0, t_i - t_{i-1})$. See Section 4.1.2 for further details.

- γ is the default time for the issuing entity.

- $\bar{\gamma} = \inf\{t \geq 0 | S(t) \geq \bar{S}\}$ and $\underline{\gamma}^T = \inf\{t \geq T | S(t) \leq \underline{S}\}$ where $\underline{S} < S(0) < \bar{S}$.

- $w(S(t)) = \begin{cases} w_0, & S(t) \geq \bar{S} \\ 0, & S(t) = \infty \end{cases}$ write-down for $w_0 \in [0,1]$.

With this all in mind, the valuation formula for the vanilla European callable FRN with no write-down features ($\overline{\text{ECFRN}}$), at time $t = 0$, can be written as

$$
\begin{aligned}
V_{\overline{\text{ECFRN}}}(0) = {} & V_{\text{FRN}}(0, N, \mathscr{T}, S(0)) \\
& - \mathbb{E}\left[B(T)^{-1} \left(V_{\text{FRN}}(T, N, \mathscr{T}, S(0)) - V_{\text{FRN}}(T, N, \mathscr{T}, S(T)) \right) 1_{S(T) \leq \underline{S}} 1_{\gamma > T} \right].
\end{aligned}
\tag{14}
$$

The motivation behind this rather simplistic approach lies within the realm of tractability and parsimony, and also to price the callability feature of the $\overline{\text{ECFRN}}$. Firstly, Equation 14 seems quite straightforward to evaluate and may require less computing power and time to evaluate. Secondly, Equation 14 could also be used as a simple and quick approximation to the vanilla European callable FRN value, when necessary.

Finally, the CFRNs are quoted on spread and not price. We therefore calculate $S_{\overline{\text{ECFRN}}}$ from Equation 17 below:

$$
V_{\overline{\text{ECFRN}}}(0) - V_{\text{FRN}}(0, N, \mathscr{T}, S_{\overline{\text{ECFRN}}}) = 0,
\tag{15}
$$

and imply the excess spread $\bar{\varepsilon} = S_{\overline{\text{ECFRN}}} - S(0, \delta)$, which is the value of the defaultable callability feature in spread terms.

---

[2]Note that some of the notation below is required in subsequent valuation formulae. We have defined all notation in a single place for ease of exposition.

4.1.2 The case with write-down

Extending the valuation framework in Section 4.1.1 above to include a write-down of the nominal value of the FRN yields a more complicated valuation formula. In the contract, the issuer of this FRN will write down its debt on the FRN if the market spread reaches a pre-specified upper threshold level. This threshold level, which is a ceiling of $\bar{S}$, is set at a level higher than the original issuance spread $S(0, \delta)$. The time of write-down, when the debt will be written down by the issuer, is given by $\bar{\gamma}$, which was defined in Section 4.1.1 above. Practice in the South African market is to assume that if $\bar{\gamma}$ is reached, a full write down occurs - this can be accounted for in our model by setting $w_0 = 0$.

On the other hand, if default occurs, then the nominal of the FRN is also written down to zero. Observe that in the definition of $\bar{\gamma}$, the case of default is implicitly allowed for in that at the time τ when the spread jumps to a very high value, the condition "$S(\tau) \geq \bar{S}$" is immediately satisfied, and therefore $\bar{\gamma} = \tau$.

Before presenting the valuation formula, it is useful, for illustrative purposes, to consider the cases when the callability option is exercised and when write-down occurs. These cases are illustrated in Figure 5 below.

Case 1: no default prior to maturity t_n, $\bar{\gamma} > \overline{T}$ and $S(T) \leq \underline{S}$

payment

$C(t_r,t_1)\tau_r N$ $C(t_1,t_2)\tau_1 N$ $C(t_2,t_3)\tau_2 N$ $C(t_{i-2},t_{i-1})\tau_{i-1}N$ $C^*(t_i,t_{i+1})\tau_{i+1}N$ $C^*(t_{n-1},t_n)\tau_n N$ N

$C^*(t_{i-1},t_i)\tau_i N$ $C^*(t_{n-2},t_{n-1})\tau_{n-1}N$

···   ···

time   0 t_1 t_2 $t_3\ t_3$ t_{i-1} T t_i t_{i+1} t_{n-1} $\overline{T}$

Option exercise date

Case 2: no default prior to maturity t_n, $\bar{\gamma} > \overline{T}$ and $S(T) > \underline{S}$

payment

$C(t_{i-1},t_i)\tau_i N$ $C(t_{n-2},t_{n-1})\tau_{n-1}N$

$C(t_r,t_1)\tau_r N$ $C(t_1,t_2)\tau_1 N$ $C(t_2,t_3)\tau_2 N$ $C(t_{i-2},t_{i-1})\tau_{i-1}N$ $C(t_i,t_{i+1})\tau_{i+1}N$ $C(t_{n-1},t_n)\tau_n N$ N

···   ···

time   0 t_1 t_2 $t_3\ t_3$ t_{i-1} T t_i t_{i+1} t_{n-1} $\overline{T}$

Option not exercised

Case 3: default or reaching $\bar{S}$ prior to time $\overline{T}$

payment $C(t_r,t_1)\tau_r N$ $C(t_1,t_2)\tau_1 N$ $C(t_2,t_3)\tau_2 N$ $C(t_{i-2},t_{i-1})\tau_{i-1}N$

···

time   0 t_1 t_2 $t_3\ t_3$ t_{i-1} $\bar{\gamma}$ T t_i t_{i+1} t_{n-1} $\overline{T}$

First time of write-down or default

Case 4: default between times T and $\overline{T}$, and $S(T) > \underline{S}$

payment

$C(t_{i-2},t_{i-1})\tau_{i-1}N$ $C(t_i,t_{i+1})\tau_{i+1}N$ $C(t_{j-1},t_j)\tau_j N$

$C(t_{i-1},t_i)\tau_i N$

$C(t_r,t_1)\tau_r N$ $C(t_1,t_2)\tau_1 N$ $C(t_2,t_3)\tau_2 N$ ···

···

time   0 t_1 t_2 $t_3\ t_3$ t_{i-1} T t_i t_{i+1} $t_j\ \bar{\gamma}$ $t_{j+1}\ t_{n-1}$ $\overline{T}$

Option not exercised   First time of write-down or default

Case 5: default between times T and $\overline{T}$, and $S(T) \leq \underline{S}$

payment

$C(t_{i-2},t_{i-1})\tau_{i-1}N$ $C^*(t_i,t_{i+1})\tau_{i+1}N$ $C^*(t_{j-1},t_j)\tau_j N$

$C^*(t_{i-1},t_i)\tau_i N$ ···

$C(t_r,t_1)\tau_r N$ $C(t_1,t_2)\tau_1 N$ $C(t_2,t_3)\tau_2 N$

···

time   0 t_1 t_2 $t_3\ t_3$ t_{i-1} T t_i t_{i+1} $t_j\ \bar{\gamma}$ $t_{j+1}\ t_{n-1}$ $\overline{T}$

Option exercise date   First time of write-down or default

Figure 5: Vanilla European callable floating-rate note cash flow depiction. For ease of notation, note that $C(t_{i-1},t_i) = L(t_{i-1},t_i) + S(0,\delta)$, and that $C^*(t_{i-1},t_i) = L(t_{i-1},t_i) + S(T,\delta)$, where δ is the (regular) time between the coupon dates, i.e. 3-month periods, and $S(T,\delta)$ is the time T issuance spread, equal to the time T market spread.

14

All these cases are accounted for by the use of indicator functions in the valuation formula for the vanilla European callable floating-rate note with write-down features (ECFRN). With the notation defined in Section 4.1.1 in mind the valuation formula at time 0 is therefore given by

$$
\begin{aligned}
V(0) = {} & V_{\text{FRN}}(0,N,\mathscr{T},S(0)) \\
& - \mathbb{E}\left[B(T)^{-1}\left(V_{\text{FRN}}(T,N,\mathscr{T},S(0)) - V_{\text{FRN}}(T,N,\mathscr{T},S(T))\right)1_{S(T)\leq\underline{S}}1_{\bar{\gamma}>T}\right] \\
& - \mathbb{E}\left[B(\bar{\gamma})^{-1}\left(V_{\text{FRN}}(\bar{\gamma},N,\mathscr{T},S(0)) - V_{\text{FRN}}(\bar{\gamma},Nw\left(S(\bar{\gamma})\right),\mathscr{T},S(0))\right)1_{\bar{\gamma}<\bar{T}}\right].
\end{aligned}
\tag{16}
$$

The motivation behind this more sophisticated formula is that both the callability and write-down features of the ECFRN are accounted for. Finally, the CFRNs are quoted on spread and not price. We therefore calculate S_{ECFRN} from Equation 17 below:

$$
V_{\text{ECFRN}}(0) - V_{\text{FRN}}(0,N,\mathscr{T},S_{\text{ECFRN}}) = 0,
\tag{17}
$$

and imply the excess spread, ε on the ECFRN by setting $\varepsilon = S_{\text{ECFRN}} - S(0,\delta)$, which again is the value of the defaultable callability features in spread terms. Obviously, ε is a more realistic spread for the type of vanilla European callable floating-rate notes that are market relevant, because it comprises three components - the spread for the default together with the spread for the write-down and the spread for the callability.

4.1.3 The case with write-down but without callability

Finally, it is also possible to value the instrument ignoring the callability. This is useful in order to calculate the spread, on the instrument, attributed to both default and write-down. With the notation defined in Section 4.1.1, the value of the instrument without callability is given by

$$
\begin{aligned}
V_{\underline{\text{ECFRN}}}(0) = {} & V_{\text{FRN}}(0,N,\mathscr{T},S(0)) \\
& - \mathbb{E}\left[B(\bar{\gamma})^{-1}\left(V_{\text{FRN}}(\bar{\gamma},N,\mathscr{T},S(0)) - V_{\text{FRN}}(\bar{\gamma},Nw\left(S(\bar{\gamma})\right),\mathscr{T},S(0))\right)1_{\bar{\gamma}<\bar{T}}\right].
\end{aligned}
\tag{18}
$$

Then calculate $S_{\underline{\text{ECFRN}}}$ from Equation 19 below

$$
V_{\underline{\text{ECFRN}}}(0) - V_{\text{FRN}}(0,N,\mathscr{T},S_{\underline{\text{ECFRN}}}) = 0,
\tag{19}
$$

and imply the excess spread $\underline{\varepsilon} = S_{\underline{\text{ECFRN}}} - S(0,\delta)$, which is the value of the write-down feature in spread terms.

4.1.4 Decomposition of the spread

As mentioned before, the spread for the vanilla ECFRN can be largely attributed to three dependent components - the callability of the FRN, the write-down feature and the defaultable value (although, it is difficult to disentangle defaultability from the value of the callability feature). Using our valuation model, it is possible, as demonstrated above, to imply the

component of S_{ECFRN} arising due to write-down risk (given by $\underline{\varepsilon}$) and the component of $S_{\overline{\text{ECFRN}}}$ arising due to default risk and the risk of the issuer calling back the FRN (given by $\overline{\varepsilon}$). From the values derived above, it is also possible to develop a decomposition for the spread by observing the following:

$$
\begin{aligned}
S_{\text{ECFRN}} &= S(0,\delta) + \varepsilon \\
&\approx S(0,\delta) + \overline{\varepsilon} + \underline{\varepsilon}.
\end{aligned}
$$

One would expect the final relation to be approximate, this relation holding because of the structural interdependence between the various components of the spread being omitted. However, one can deduce a considerable correlation between the callability feature and the write-down feature. Should the call option be exercised, the write down feature will essentially "fall away", thereby reducing the value of the call option relative to the write-down. Consequently, there appears to be a negative correlation between the value of the callability feature and the write-down feature. Therefore, we can reasonably deduce that 20 does not hold in approximation.

4.2 Vanilla American callable floating-rate note valuation framework

Because these CFRNs are not standardised, American-style features could also exist. Therefore, we allow for the exercise of the callability feature at any time between T and $\overline{T}$.

Allowing for the callability feature, default and write-down, the value of the instrument is given by (using the notation in Section 4.1.1),

$$
\begin{aligned}
V^A(0) = \; & V_{\text{FRN}}(0,N,\mathscr{T},S(0)) \\
& - \mathbb{E}\left[B(\underline{\gamma}^T)^{-1} \left(V_{\text{FRN}}(\underline{\gamma}^T,N,\mathscr{T},S(0)) - V_{\text{FRN}}(\underline{\gamma}^T,N,\mathscr{T},S(\underline{\gamma}^T)) \right) 1_{\overline{\gamma}>\underline{\gamma}^T} 1_{\underline{\gamma}^T<\overline{T}} \right] \\
& - \mathbb{E}\left[B(\overline{\gamma})^{-1} \left(V_{\text{FRN}}(\overline{\gamma},N,\mathscr{T},S(0)) - V_{\text{FRN}}(\overline{\gamma},Nw\left(S(\overline{\gamma})\right),\mathscr{T},S(0)) \right) 1_{\overline{\gamma}<\underline{\gamma}^T} 1_{\overline{\gamma}<\overline{T}} \right]. \quad (20)
\end{aligned}
$$

The spreads on this option can be calculated and attributed in precisely the same way as was done for the European version in Sections 4.1.1 to 4.1.4.

Now what becomes necessary is the modelling of the risk factors. This is covered in the following section.

5 Modelling the risk factors

5.1 Modelling default-free interest rates

$(\Omega,\mathscr{F},(\mathscr{F}(t))_{t\geq 0},\mathbb{Q})$ is a filtered probability space. The filtration satisfies the usual conditions. The measure $\mathbb{Q}$ plays the role of a risk neutral measure. All processes are defined on Ω, adapted to $(\mathscr{F}(t))_{t\geq 0}$, and cádlàg. $W_1 = (W_1(t))_{t\geq 0}$ and $W_2 = (W_2(t))_{t\geq 0}$ are two independent standard $(\mathscr{F}(t))_{t\geq 0}$-Brownian motions.

We assume that the dynamics of the short-rate process $(r(t))_{t\geq 0}$ under the risk-neutral measure $\mathbb{Q}$ is given by

$$
r(t) = X(t) + Y(t) + \varphi(t), \quad r(0) = r_0, \quad (21)
$$

for processes X and Y satisfying

$$
\begin{aligned}
dX(t) &= -aX(t)dt + \sigma dW_1(t), \quad X(0) = 0, \\
dY(t) &= -bY(t)dt + \eta\rho dW_1(t) + \eta\sqrt{1-\rho^2}dW_2(t), \quad Y(0) = 0,
\end{aligned}
\tag{22}
$$

where $r_0 \in \mathbb{R}$, $a, b, \sigma, \eta > 0$, $\rho \in [-1, 1]$ and φ is a deterministic function on $\mathbb{R}_+$. In particular $\varphi(0) = r_0$.

Lemma 5.1. *For $t \geq s$ the solution of the short-rate model (21)-(22) is given by*

$$
r(t) = X(s)e^{-a(t-s)} + Y(s)e^{-b(t-s)} + \sigma \int_s^t e^{-a(t-u)}dW_1(u)
$$

$$
+ \eta\rho \int_s^t e^{-b(t-u)}dW_1(u) + \eta\sqrt{1-\rho^2} \int_s^t e^{-b(t-u)}dW_2(u) + \varphi(t).
$$

The conditional distribution of $r(t)$ given $\mathscr{F}_s$ is Gaussian with mean

$$
\mathbb{E}[r(t) \mid \mathscr{F}(s)] = X(s)e^{-a(t-s)} + Y(s)e^{-b(t-s)} + \varphi(t),
$$

and variance

$$
\operatorname{Var}[r(t) \mid \mathscr{F}(s)] = \frac{\sigma^2}{2a}\left(1 - e^{-2a(t-s)}\right) + \frac{\eta^2}{2b}\left(1 - e^{-2b(t-s)}\right)
$$

$$
+ 2\rho\frac{\eta\sigma}{a+b}\left(1 - e^{-(a+b)(t-s)}\right).
$$

Proof. See Brigo and Mercurio (2007) Section 4.2. □

We denote by $P(t, T)$ the price at time t of a default-free zero-coupon bond maturing at time T with unit face value. By no-arbitrage pricing theory we have

$$
P(t, T) = \mathbb{E}\left[e^{-\int_t^T r(s)ds} \mid \mathscr{F}_t\right]
$$

Theorem 5.2. *In the short-rate model (21)-(22) zero-coupon bond prices for $t \leq T$ have the following (time-inhomogeneous) exponential affine form*

$$
P(t, T) = e^{-\int_t^T \varphi(u)du - \frac{1-e^{-a(T-t)}}{a}X(t) - \frac{1-e^{-b(T-t)}}{b}Y(t) + \frac{1}{2}\Phi(T-t)},
$$

where for $\tau \geq 0$

$$
\Phi(\tau) = \frac{\sigma^2}{a^2}\left(\tau + \frac{2}{a}e^{-a\tau} - \frac{1}{2a}e^{-2a\tau} - \frac{3}{2a}\right) + \frac{\eta^2}{b^2}\left(\tau + \frac{2}{b}e^{-b\tau} - \frac{1}{2b}e^{-2b\tau} - \frac{3}{2b}\right)
$$

$$
+ 2\rho\frac{\sigma\eta}{ab}\left(\tau + \frac{e^{-a\tau}-1}{a} + \frac{e^{-b\tau}-1}{b} - \frac{e^{-(a+b)\tau}-1}{a+b}\right).
$$

Proof. See Brigo and Mercurio (2007) Section 4.2. □

17

Corollary 5.3. *Let us denote by $f^M(0, \tau)$ the instantaneous forward rate at time 0 for a maturity τ implied by the term structure $\tau \mapsto P^M(0, \tau)$, i.e.,*

$$f^M(0, \tau) = -\frac{\partial \ln P^M(0, \tau)}{\partial \tau},$$

then the model (1.1) fits the currently-observed term structure of discounted factors if and only if, for each τ,

$$\varphi(\tau) = f^M(0, \tau) + \frac{\sigma^2}{2a^2}(1 - e^{-a\tau})^2$$

$$+ \frac{\eta^2}{2b^2}(1 - e^{-b\tau})^2 + \rho \frac{\sigma \eta}{ab}(1 - e^{-a\tau})(1 - e^{-b\tau}).$$

Proof. See Brigo and Mercurio (2007) Section 4.2. $\qquad\square$

In this model explicit pricing formulae for interest rate caps are available. These formulae are needed to calibrate the volatility parameters of the model to market data. First define the following quantity

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

$$+ \frac{\eta^2}{2b^3}\left[1 - e^{-b(S-T)}\right]^2 \left[1 - e^{-2b(T-t)}\right]$$

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

We denote by $\mathscr{T} = \{t_0, t_1, t_2, ..., t_n\}$ the set of the cap/floor payment dates, augmented with the first reset date t_0, and by $\tau_i = t_i - t_{i-1}$.

Theorem 5.4. *Caplets and floorlets values in the short-rate model (21)-(22) are given by*

$$\mathbf{Cpl}(t, T_1, T_2, N, X) = -N'P(t, T_2)\Phi\left(\frac{\ln \frac{NP(t,T_1)}{N'P(t,T_2)}}{\Sigma(t, T_1, T_2)} - \frac{1}{2}\Sigma(t, T_1, T_2)\right)$$

$$+ P(t, T_1)N\Phi\left(\frac{\ln \frac{NP(t,T_1)}{N'P(t,T_2)}}{\Sigma(t, T_1, T_2)} + \frac{1}{2}\Sigma(t, T_1, T_2)\right)$$

$$\mathbf{Fll}(t, T_1, T_2, N, X) = N'P(t, T_2)\Phi\left(\frac{\ln \frac{N'P(t,T_2)}{NP(t,T_1)}}{\Sigma(t, T_1, T_2)} + \frac{1}{2}\Sigma(t, T_1, T_2)\right)$$

$$- P(t, T_1)N\Phi\left(\frac{\ln \frac{N'P(t,T_2)}{NP(t,T_1)}}{\Sigma(t, T_1, T_2)} - \frac{1}{2}\Sigma(t, T_1, T_2)\right)$$

Proof. See Brigo and Mercurio (2007) Section 4.2. $\qquad\square$

The price of a cap (floor) is the sum of the prices of the underlying caplets (floorlets), the price at time t of a cap with cap rate (strike) X, nominal value N, set of times $\mathscr{T}$ and year fractions τ is then given by

$$\mathbf{Cap}(t, \mathscr{T}, \tau, N, X) = \sum_{i=1}^{n} \mathbf{Cpl}(t, t_{i-1}, t_i, N, X),$$

and the price of the corresponding floor is

$$\mathbf{Flr}(t, \mathscr{T}, \tau, N, X) = \sum_{i=1}^{n} \mathbf{Fll}(t, t_{i-1}, t_i, N, X).$$

5.2 Modelling credit spreads

Let τ be a $\mathbb{R}_+$-valued random variable on Ω which denotes the default time of an entity. We introduce the filtration $\mathscr{G}_t = \mathscr{F}_t \vee \sigma(\{\tau \leq t\}, u \leq t)$ which includes default monitoring. However, for pricing problems it is often more convenient to consider conditional expectations under the filtration $(\mathscr{F}(t))_{t \geq 0}$. The following result provides a relation between the conditional expectations with respect to the two filtrations, see Brigo and Mercurio (2007) Section 22.5.

Lemma 5.5 (Filtration switching formula). *Let H_T be a $\mathscr{G}_T$-measurable payoff then*

$$\mathbb{E}\left[1_{\tau > T} H_T \mid \mathscr{G}_t\right] = \frac{1_{\tau > t}}{\mathbb{Q}(\tau > t \mid \mathscr{F}_t)} \mathbb{E}\left[1_{\tau > T} H_T \mid \mathscr{F}_t\right]$$

We model τ as the first jump time of a Cox process with stochastic intensity process given by a $\mathbb{R}_+$-valued, $(\mathscr{F}_t)_{t \geq 0}$-adapted cádlág process $\lambda = (\lambda(t))_{t \geq 0}$. The assumption that λ is $(\mathscr{F}(t))_{t \geq 0}$-adapted means that the randomness we allow in the intensity is induced by the default-free market. We define the hazard process by $\Lambda(t) = \int_0^t \lambda(u) du$.

Lemma 5.6. *Under the modelling assumptions on τ the survival probability is given by*

$$\mathbb{Q}[\tau > t] = \mathbb{E}\left[e^{-\int_0^t \lambda(s) ds}\right] = \mathbb{E}\left[e^{-\Lambda(t)}\right].$$

Proof. See Brigo Brigo and Mercurio (2007) Section 22.2.3. We write the steps down since these are reused in the sequel. Note that $\Lambda(\tau)$ is $\exp(1)$-distributed and independent of $\mathscr{F}_t$ under $\mathbb{Q}$. Hence,

$$\mathbb{Q}[\tau > t] = \mathbb{Q}[\Lambda(\tau) > \Lambda(t)] = \mathbb{E}\left[\mathbb{E}\left[1_{\Lambda(\tau) > \Lambda(t)} \mid \mathscr{F}(t)\right]\right] = \mathbb{E}\left[\mathbb{E}\left[1_{\Lambda(\tau) > x}\right]\big|_{x = \Lambda(t)}\right]$$

$$= \mathbb{E}\left[\mathbb{Q}[\Lambda(\tau) > x]\big|_{x = \Lambda(t)}\right] = \mathbb{E}\left[e^{-\Lambda(t)}\right]$$

$\square$

We propose a model for the intensity process λ. Let $W^\lambda = \left(W^\lambda(t)\right)_{t \geq 0}$ be a standard $(\mathscr{F}(t))_{t \geq 0}$-Brownian motion with

$$d[W^\lambda, W_1](t) = v_1 dt, \quad \text{and} \quad d[W^\lambda, W_2](t) = v_2 dt,$$

where $v_{1,2} \in [-1, 1]$. Assume that λ evolves according to the following SDE

$$d\lambda(t) = (\lambda(t) - \lambda_0)\lambda(t) dt + \sigma^\lambda \lambda(t) dW^\lambda(t), \quad \lambda(0) = \lambda_0 \tag{23}$$

where $\lambda_0, \sigma^\lambda > 0$.

Theorem 5.7. *Let τ be the first jump time of a Cox process with stochastic intensity given by (23). Then, the following holds.*

(i) *The conditional density process* $\bar{\lambda}(t) = \partial_t \mathbb{Q}(\tau \le t | \mathscr{F}_t)$ *satifies the SDE*

$$d\bar{\lambda}(t) = -\lambda_0 \bar{\lambda}(t)dt + \sigma^\lambda \bar{\lambda}(t)dW_t^\lambda, \quad \bar{\lambda}(0) = \lambda_0,$$

i.e. $\bar{\lambda}$ *is geometric Brownian motion;*

(ii) $\lambda_t > 0$ $\mathbb{Q}$-*a.s. for all* $t \ge 0$; *and*

(iii) *The survival probability is given by*

$$\mathbb{Q}(\tau > t) = e^{-\lambda_0 t}.$$

Proof. We prove the four statements

(i) Since $\Lambda(\tau)$ is $\exp(1)$-distributed and independent of $\mathscr{F}(t)$ under $\mathbb{Q}$ we obtain

$$\bar{\lambda}(t) = \partial_t \mathbb{Q}(\tau \le t | \mathscr{F}(t)) = \partial_t \mathbb{E}[1_{\tau < t} | \mathscr{F}(t)] = \partial_t \mathbb{E}[1_{\Lambda(\tau) < \Lambda(t)} | \mathscr{F}(t)]$$
$$= \partial_t \left(1 - e^{-\Lambda(t)}\right) = \lambda(t)e^{-\Lambda(t)}.$$

Consider the C^2 function $f(x,y) = xe^{-y}$. Applying Itō's formula we obtain

$$d\bar{\lambda}(t) = d\left(f(\lambda(t), \Lambda(t))\right) = -\lambda(t)e^{-\Lambda(t)}d\Lambda(t) + e^{-\Lambda(t)}d\lambda(t)$$
$$= e^{-\Lambda(t)}\left(-\lambda_0\lambda(t)dt + \sigma^\lambda\lambda(t)dW_t\right) = -\lambda_0\bar{\lambda}(t)dt + \sigma^\lambda\bar{\lambda}(t)dW(t).$$

(ii) We have $\lambda(t) = \frac{\bar{\lambda}(t)}{e^{-\Lambda(t)}} > 0$ almost surely for each t since $\bar{\lambda}(t)$ is geometric Brownian motion with positive initial value.

(iii) The density of τ under $\mathbb{Q}$ is given by

$$\partial_t \mathbb{Q}(\tau \le t) = \partial_t \mathbb{E}\left[\mathbb{Q}(\tau \le t | \mathscr{F}_t)\right] = \mathbb{E}\left[\partial_t \mathbb{Q}(\tau \le t | \mathscr{F}_t)\right]$$
$$= E\left[\bar{\lambda}(t)\right] = \bar{\lambda}(0)e^{-\lambda_0 t} = \lambda_0 e^{-\lambda_0 t}.$$

Hence, for the survival probability we obtain

$$\mathbb{Q}(\tau > t) = \int_t^\infty \lambda_0 e^{-\lambda_0 s}ds = e^{-\lambda_0 t}.$$

$\square$

The price $\bar{P}(t,T)$ at time t of a defaultable zero coupon bond with maturity T and unit notional is given by

$$\bar{P}(t,T) = 1_{\tau > t}\mathbb{E}\left[\frac{B(t)}{B(T)}1_{\tau > T} \mid \mathscr{G}_t\right].$$

Define the spot rates of default-free and defaultable zero-coupon bonds by

$$Y(t,T) = -\frac{1}{T-t}\log P(t,T), \quad \text{and} \quad \bar{Y}(t,T) = \begin{cases} -\frac{1}{T-t}\log\bar{P}(t,T), & \tau > t \\ -\infty, & \tau \le t \end{cases}.$$

20

The credit spread at time t of a defaultable zero coupon bond with maturity T is defined by

$$S(t,T) = \begin{cases} \bar{Y}(t,T) - Y(t,T) = \frac{1}{T-t}\log\frac{P(t,T)}{\bar{P}(t,T)}, & \tau > t \\ \infty, & \tau \le t \end{cases}.$$

Theorem 5.8. *The price of a defaultable zero-coupon bond in the stochastic intensity model* (23) *is given by*

$$\bar{P}(t,T) = 1_{\tau>t}\mathbb{E}\left[e^{-\int_t^T (r(s)+\lambda(s))ds} \mid \mathscr{F}_t\right]$$

$$= 1_{\tau>t}\frac{\lambda(t)}{\bar{\lambda}(t)}\int_T^{+\infty} \mathbb{E}\left[e^{-\int_t^T r(s)ds}\bar{\lambda}(u) \mid \mathscr{F}_t\right]du. \tag{24}$$

In the case $v_2 = \sqrt{1-v_1^2}$ we have explicit formulas for the price of a defaultable zero-coupon bond

$$\bar{P}(t,T) = 1_{\tau>t}\frac{\lambda(t)}{\lambda_0}P(t,T)e^{\left(\tilde{\mu}(t,T) - \frac{\tilde{\sigma}(t,T)^2}{2}\right)}, \tag{25}$$

and for the credit spread

$$S(t,T) = \begin{cases} -\frac{1}{T-t}\log\left(\frac{\lambda(t)}{\lambda_0}e^{\left(\tilde{\mu}(t,T) - \frac{\tilde{\sigma}(t,T)^2}{2}\right)}\right), & \tau > t \\ \infty, & \tau \le t \end{cases}, \tag{26}$$

where

$$g_a(t,T) = \left(t + \frac{e^{-aT}}{a}\left(1 - e^{at}\right)\right),$$

$$g_b(t,T) = \left(t + \frac{e^{-bT}}{b}\left(1 - e^{bt}\right)\right),$$

$$\tilde{\mu}(t,T) = (-\lambda_0 - \frac{(\sigma^\lambda)^2}{2})(T - t) + \sigma^\lambda\left(v_1\frac{\sigma}{a}\left(g_a(t,T) - g_a(T,T)\right)\right.$$

$$\left. + v_1\rho\frac{\eta}{b}\left(g_b(t,T) - g_b(T,T)\right) + \sqrt{(1-v_1^2)(1-\rho^2)}\frac{\eta}{b}\left(g_b(t,T) - g_b(T,T)\right)\right),$$

$$\tilde{\sigma}(t,T)^2 = \left(\sigma^\lambda\right)^2(T - t)$$

In the case $v_1 = v_2 = 0$ explicit formulas are also available

$$\bar{P}(t,T) = 1_{\tau>t}P(t,T)\frac{\lambda(t)}{\lambda_0}e^{-\lambda_0(T-t)}, \tag{27}$$

$$S(t,T) = \begin{cases} -\frac{1}{T-t}\log\left(\frac{\lambda(t)}{\lambda_0}e^{-\lambda_0(T-t)}\right), & \tau > t \\ \infty, & \tau \le t \end{cases}. \tag{28}$$

Proof. The first equality in (24) is a standard result which can be found e.g. in Brigo and Mercurio (2007). Applying the filtration switching formula we obtain

$$
\bar{P}(t,T) = \frac{1_{\tau>t}}{\mathbb{Q}(\tau > t|\mathscr{F}_t)} \mathbb{E}\left[e^{-\int_t^T r(s)ds} \mid \mathscr{F}_t\right]
$$

$$
= \frac{1_{\tau>t}}{e^{-\Lambda(t)}} \mathbb{E}\left[e^{-\int_t^T r(s)ds}\mathbb{Q}(\tau > T|\mathscr{F}_T) \mid \mathscr{F}_t\right]
$$

$$
= \frac{1_{\tau>t}}{e^{-\Lambda(t)}} \mathbb{E}\left[e^{-\int_t^T r(s)ds}e^{-\Lambda(T)} \mid \mathscr{F}_t\right]
$$

$$
= 1_{\tau>t}\mathbb{E}\left[e^{-\int_t^T (r(s)+\lambda(s))ds} \mid \mathscr{F}_t\right].
$$

The first equality in (24) follows from

$$
\bar{P}(t,T) = \frac{1_{\tau>t}}{e^{-\Lambda(t)}} \mathbb{E}\left[e^{-\int_t^T r(s)ds}\mathbb{Q}(\tau > T|\mathscr{F}_T) \mid \mathscr{F}_t\right]
$$

$$
= \frac{1_{\tau>t}}{e^{-\Lambda(t)}} \mathbb{E}\left[\int_T^\infty e^{-\int_t^T r(s)ds}\partial_u\mathbb{Q}(\tau \leq u|\mathscr{F}_T) \mid \mathscr{F}_t\right]
$$

$$
= 1_{\tau>t}\frac{\lambda(t)}{\bar{\lambda}(t)}\int_T^\infty \mathbb{E}\left[e^{-\int_t^T r(s)ds}\bar{\lambda}(u) \mid \mathscr{F}_t\right]
$$

where we use Fubini's theorem.

Denote by $\mathbb{Q}^T$ the T-forward measure defined by the Radon-Nikodym derivative

$$
\xi_t = \left.\frac{d\mathbb{Q}^T}{d\mathbb{Q}}\right|_{\mathscr{F}_t} = \frac{P(t,T)}{P(0,T)B(t)}.
$$

By change of measure we obtain

$$
E\left[e^{-\int_t^T r(s)ds}\bar{\lambda}(u) \mid \mathscr{F}_t\right] = E\left[\frac{B(t)}{B(T)}\mathbb{E}[\bar{\lambda}(u)|\mathscr{F}_T] \mid \mathscr{F}_t\right]
$$

$$
= E^{\mathbb{Q}^T}\left[\frac{B(t)}{B(T)}E[\bar{\lambda}(u)|\mathscr{F}_T]\frac{\xi_t}{\xi_T} \mid \mathscr{F}_t\right]
$$

$$
= P(t,T)E^{\mathbb{Q}^T}\left[\bar{\lambda}(T)e^{-\lambda_0(u-T)} \mid \mathscr{F}_t\right].
$$

This means that the defaultable zero-coupon bond price is given by

$$
\bar{P}(t,T) = 1_{\tau>t}\frac{\lambda(t)}{\lambda_0\bar{\lambda}(t)}P(t,T)E^{\mathbb{Q}^T}\left[\bar{\lambda}(T) \mid \mathscr{F}_t\right]
$$

By Girsanov's theorem the following processes are independent Brownian motions under the forward measure $\mathbb{Q}^T$, see Brigo and Mercurio (2007) Lemma 4.2.2,

$$
dW_1^T(t) = dW^1(t) + \left(\frac{\sigma}{a}\left(1 - e^{-a(T-t)}\right) + \rho\frac{\eta}{b}\left(1 - e^{-b(T-t)}\right)\right)dt
$$

$$
dW_2^T(t) = dW^2(t) + \frac{\eta}{b}\sqrt{1 - \rho^2}\left(1 - e^{-b(T-t)}\right)dt.
$$

Hence, we obtain

$$W^\lambda(t) = \nu_1 W^1(t) + \sqrt{1 - \nu_1^2} W^2(t) = \nu_1 W_1^T(t) + \sqrt{1 - \nu_1^2} W_2^T(t)$$
$$- \nu_1 \frac{\sigma}{a} g_a(t,T) - \nu_1 \rho \frac{\eta}{b} g_b(t,T) - \sqrt{(1 - \nu_1^2)(1 - \rho^2)} \frac{\eta}{b} g_b(t,T).$$

The random variable $\log \frac{\bar{\lambda}(T)}{\bar{\lambda}(t)}$ is Gaussian given $\mathscr{F}_t$ under $\mathbb{Q}^T$ with mean $\tilde{\mu}(t,T)$ and variance $\tilde{\sigma}(t,T)^2$. This proves the formula. If we assume independence between W^λ and $W_{1,2}$ then we obtain

$$\bar{P}(t,T) = 1_{\tau>t} P(t,T) \frac{\lambda(t)}{\bar{\lambda}(t)} \int_T^\infty E\left[\bar{\lambda}(u) \mid \mathscr{F}_t\right] du$$
$$= 1_{\tau>t} P(t,T) \frac{\lambda(t)}{\bar{\lambda}(t)} \int_T^\infty \bar{\lambda}(t) e^{-(u-t)\lambda_0} du = 1_{\tau>t} \frac{\lambda(t)}{\lambda_0} P(t,T) e^{-\lambda_0(T-t)}.$$

The formula for the spread follows by definition. $\qquad\square$

6 Heuristic approach to estimate credit spreads from equity triggers

As we mentioned in the introduction, the write-down trigger should be more appropriately tied to a depreciation of the market or book value of the issuer rather then its credit spreads on the market. To address this issue we formulate a procedure to calculate the equity depreciation which corresponds to a certain credit spread. The market value of equity process of the issuer $(M(t))_{t\geq 0}$ follows the SDE

$$dM(t) = \mu^M M(t) dt + \sigma^M M(t) dW_t^M, \quad M(0) = M_0,$$

where $\mu^M, \sigma^M \geq 0$ are constants and $W^M = \left(W_t^M\right)_{t\geq 0}$ is a standard $(\mathscr{F}(t))_{t\geq 0}$-Brownian motion. Let τ^M denote the default time of this issuer, which is defined as

$$\tau^M := \inf\{0 \leq t \leq \overline{T} : M(t) \leq \underline{M}\},$$

where $\underline{M}$ is the flat barrier.
Then the default probability under this framework is given by

$$\mathbb{Q}\left[\tau^M \leq \overline{T}\right] = \left(\frac{\underline{M}}{M_0}\right)^{2\alpha} \Phi\left(\frac{\log\left(\frac{\underline{M}}{M_0}\right) + \overline{\mu}\overline{T}}{\sigma^M \sqrt{\overline{T}}}\right) + \Phi\left(\frac{\log\left(\frac{\underline{M}}{M_0}\right) - \overline{\mu}\overline{T}}{\sigma^M \sqrt{\overline{T}}}\right),$$

where $\overline{\mu} = \left(\mu^M - \frac{\sigma^{M2}}{2}\right)$, $\alpha = \frac{\overline{\mu}}{\sigma^{M2}}$ and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\phi^2/2} d\phi$ is the CDF of standard Normal random variable. See Bielecki and Rutkowski (2002) Section 3.2.

In order to obtain a closed-form expression we assume that $\nu_1 = \nu_2 = 0$. In the general case the calculation can be done numerically. According to Theorem 5.7 and formula (27), at time 0, we have

$$\mathbb{Q}\left[\tau \leq \overline{T}\right] = 1 - e^{-\lambda_0 \overline{T}} \quad and \quad S(0,\delta) = \lambda_0.$$

23

By choosing $S(0) = S(0, \delta) = \lambda_0$ such that $\mathbb{Q}\left[\tau \leq \overline{T}\right] = \mathbb{Q}\left[\tau^M \leq \overline{T}\right]$, we have

$$S(0) = \kappa\left(\frac{\underline{M}}{M_0}\right),$$

where

$$\kappa(x) = -\frac{1}{\overline{T}}\log\left(1 - x^{2\alpha}\Phi\left(\frac{\log x + \overline{\mu}\overline{T}}{\sigma^M\sqrt{\overline{T}}}\right) + \Phi\left(\frac{\log x - \overline{\mu}\overline{T}}{\sigma^M\sqrt{\overline{T}}}\right)\right)$$

for $x > 0$.

Credit Spread as a function of equity default trigger (%)
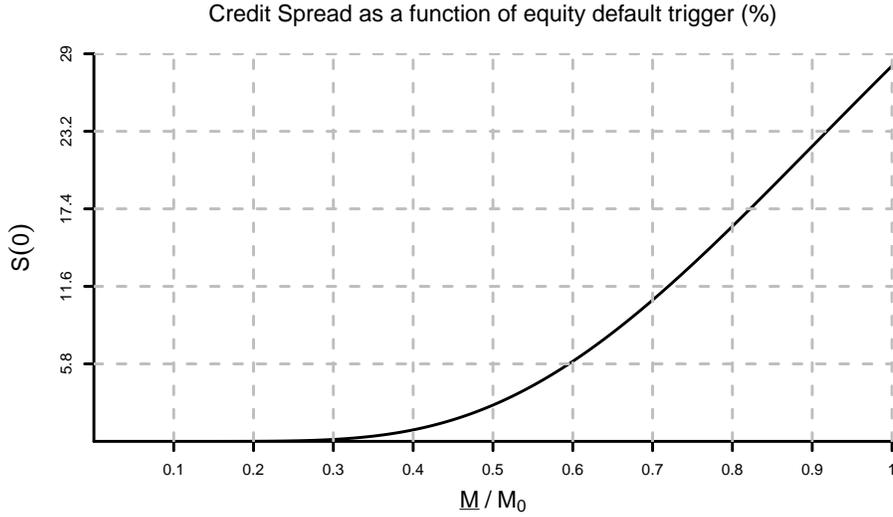
Figure 6: Credit Spread as a function of equity default trigger (%), with parameters $\mu^M = 0.03$, $\sigma^M = 0.15$ and $\overline{T} = 10$.

7 Maximum likelihood estimation and Kalman filter

We estimate our model for risk-free interest rates and credit spreads setting $\varphi \equiv 0$. Once we have estimated the model parameters we calibrate φ in order to perfectly match the initial risk-free yield curve.

For the Kalman filter we refer to Haykin (2004) Section 1.6. We write the procedure in full details assuming $v_{1,2} = 0$. For the case $v_2 = \sqrt{1 - v_1^2}$ proceed in the same way using (25).

7.1 Discretisation of the transition system

First, note that for estimation we need to consider the real-world dynamics of the underlying risk factors. We choose the market price of risk process such that the processes

$$dW^1 + \alpha X(t)dt, \quad dW^2 + \beta Y(t)dt, \quad dW^\lambda + \theta\lambda(t)dt, \quad \alpha, \beta, \theta \in \mathbb{R},$$

24

are independent Brownian motions under the real-world measure $\mathbb{P}$. We apply the first order discretisation to the SDEs for $X(t)$, $Y(t)$ and $\log \lambda(t)$, respectively. Denote the length of the time interval $[t_{i-1}, t_i]$ by δ. Note that the time grid in this section is for estimation and is not related to coupon payments and interest rate resets as in the previous section. Then, for $i \in \{1, 2, ..., n\}$ we have

$$X(t_i) = (1 - \delta a - \delta \sigma \alpha) X(t_{i-1}) + \sigma \sqrt{\delta} \varepsilon_i^1$$

$$Y(t_i) = (1 - \delta b - \delta \eta \sqrt{(1-\rho^2)} \beta) Y(t_{i-1}) - \delta \eta \rho \alpha X(t_{i-1}) + \eta \rho \sqrt{\delta} \varepsilon_i^1 + \eta \sqrt{(1-\rho^2)} \delta \varepsilon_i^2$$

$$\log \lambda(t_i) = \left(\lambda(t_{i-1}) - \lambda_0 - \sigma^\lambda \theta - \frac{1}{2}(\sigma^\lambda)^2 \right) \delta + \log \lambda(t_{i-1}) + \sigma^\lambda \sqrt{\delta} \varepsilon_i^3$$

where ε^1, ε^2 and ε^3 are independent standard Gaussian random variables under the real world measure $\mathbb{P}$.

Define $\mathbf{X}_i := \begin{pmatrix} X(t_i) \\ Y(t_i) \\ \log \lambda(t_i) \end{pmatrix} = \begin{pmatrix} X_{i,1} \\ X_{i,2} \\ X_{i,3} \end{pmatrix}$. Then, for $i \in \{1, 2, ..., n\}$

$$\mathbf{X}_i := \begin{bmatrix} 1 - \delta a - \delta \sigma \alpha & 0 & 0 \\ -\delta \eta \rho \alpha & 1 - \delta b - \delta \eta \sqrt{(1-\rho^2)} \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{X}_{i-1} + \delta \begin{pmatrix} 0 \\ 0 \\ e^{X_{i-1,3}} \end{pmatrix}$$

$$+ \begin{pmatrix} 0 \\ 0 \\ \left(-\lambda_0 - \sigma^\lambda \theta - \frac{1}{2}(\sigma^\lambda)^2 \right) \delta \end{pmatrix} + \begin{bmatrix} \sigma \sqrt{\delta} & 0 & 0 \\ \eta \rho \sqrt{\delta} & \eta \sqrt{(1-\rho^2)\delta} & 0 \\ 0 & 0 & \sigma^\lambda \sqrt{\delta} \end{bmatrix} \varepsilon_i,$$

where $\varepsilon_i := [\varepsilon_i^1 \; \varepsilon_i^2 \; \varepsilon_i^3]^\top$.

7.2 Discretisation of the measurement system

Define

$$\overline{\mathbf{Y}}_i = \left(\overline{Y}(t_i, t_i + \tau_1), \overline{Y}(t_i, t_i + \tau_m) \right)^\top,$$

$$\mathbf{1} = (1 \; ... \; 1)^\top,$$

$$\tau = (\tau_1 \; ... \; \tau_m)^\top, \quad \text{and}$$

$$\tau^{-1} = \left(\tau_1^{-1} \; ... \; \tau_m^{-1} \right)^\top.$$

For $i \in \{1, ..., n\}$ and $j \in \{1, ..., m\}$, we have

$$Y(t_i, t_i + \tau_j) = -\frac{1}{\tau_j} \log P(t_i, t_i + \tau_j)$$

$$= \frac{1}{\tau_j} \frac{1 - e^{-a\tau_j}}{a} X(t_i) + \frac{1}{\tau_j} \frac{1 - e^{-b\tau_j}}{b} Y(t_i) - \frac{1}{2\tau_j} \phi(\tau_j)$$

Hence, the measurement system (i.e. defaultable zero-coupon yields) satisfies the following equation

$$\overline{\mathbf{Y}}_i = \tau^{-1} \log \lambda_0 + \mathbf{Y}_i - \tau^{-1} X_{i,3} + \lambda_0 \tau + \mathbf{R}^{\frac{1}{2}} \upsilon_i$$

where

$$\mathbf{Y}_i = \psi_1 X_{i,1} + \psi_2 X_{i,2} - \Phi,$$

$$\psi_1 := \left(\frac{1}{\tau_j} \frac{1 - e^{-a\tau_j}}{a} \right)_{j=1,\dots,m},$$

$$\psi_2 := \left(\frac{1}{\tau_j} \frac{1 - e^{-b\tau_j}}{b} \right)_{j=1,\dots,m},$$

$$\Phi := \left(\frac{1}{2\tau_j} \phi(\tau_j) \right)_{j=1,\dots,m},$$

and the random term $\mathbf{R}^{\frac{1}{2}} v_i$ for non-singular $\mathbf{R}^{\frac{1}{2}} \in \mathbb{R}^{m \times m}$ denote an m-dimensional noise term which may indicate that defaultable zero-coupon yields are itself estimated from financial market data.

7.3 Extended Kalman filter

Anchoring:

$$\mathbf{X}_{0|0} = \begin{pmatrix} 0 \\ 0 \\ \log \lambda_0 \end{pmatrix} \quad \text{and} \quad \mathbf{P}_{0|0} = \delta \begin{bmatrix} \sigma^2 & \sigma\eta\rho & 0 \\ \sigma\eta\rho & \eta^2 & 0 \\ 0 & 0 & \sigma^{\lambda^2} \end{bmatrix}$$

The predicted state is given by

$$\mathbf{X}_{i|i-1} = \mathbf{f}(\mathbf{X}_{i-1|i-1}),$$

where

$$\mathbf{f}(\mathbf{X}) = \begin{bmatrix} 1 - \delta a - \delta\sigma\alpha & 0 & 0 \\ -\delta\eta\rho\alpha & 1 - \delta b - \delta\eta\sqrt{(1-\rho^2)}\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{X} + \begin{pmatrix} 0 \\ 0 \\ \delta e^{X_3} \end{pmatrix}$$

$$+ \begin{pmatrix} 0 \\ 0 \\ \left(-\lambda_0 - \sigma^\lambda\theta - \frac{1}{2}(\sigma^\lambda)^2 \right)\delta. \end{pmatrix}$$

The predicted covariance is given by

$$\mathbf{P}_{i|i-1} = \mathbf{F}_{i-1}\mathbf{P}_{i-1|i-1}\mathbf{F}_{i-1}^\top + \mathbf{Q},$$

where

$$\mathbf{F}_{i-1} = \left[\frac{\partial f_p}{\partial X_{i-1,q}} \right]_{p,q=1,2,3} = \begin{bmatrix} 1 - \delta a - \delta\sigma\alpha & 0 & 0 \\ -\delta\eta\rho\alpha & 1 - \delta b - \delta\eta\sqrt{(1-\rho^2)}\beta & 0 \\ 0 & 0 & 1 + \delta e^{X_{i-1,3}} \end{bmatrix},$$

$$\mathbf{Q} = \mathbf{P}_{0|0}$$

Measurement residuals are given by

$$\tilde{\mathbf{Y}}_i = \overline{\mathbf{Y}}_i - \mathbf{h}(\mathbf{X}_{i|i-1}),$$

where
$$\mathbf{h}(\mathbf{X}) = \tau^{-1} \log \lambda_0 + \psi_1 X_{i,1} + \psi_2 X_{i,2} - \Phi - \tau^{-1} X_{i,3} + \lambda_0 \tau.$$
Residuals covariance covariance is given by
$$\mathbf{S}_i = \mathbf{H} \mathbf{P}_{i|i-1} \mathbf{H}^\top + \mathbf{R},$$
where
$$\mathbf{H} := \left[\frac{\partial h_p}{\partial X_{i-1,q}} \right]_{p=1,\ldots,m; q=1,2,3} = \begin{bmatrix} \psi_{11} & \psi_{21} & -\tau_1^{-1} \\ \vdots & \vdots & \vdots \\ \psi_{1m} & \psi_{2m} & -\tau_m^{-1} \end{bmatrix} \in \mathbb{R}^{m \times 3}.$$

The near-optimal Kalman gain matrix and the updated covariance estimates and state estimates are given by
$$\mathbf{K}_i = \mathbf{P}_{i|i-1} \mathbf{H}^\top \mathbf{S}_i^{-1}$$
$$\mathbf{P}_{i|i} = (\mathbf{I} - \mathbf{K}_i \mathbf{H}) \mathbf{P}_{i|i-1},$$
$$\mathbf{X}_{i|i} = \mathbf{X}_{i|i-1} + \mathbf{K}_i \tilde{\mathbf{Y}}_i$$
where $\mathbf{I}$ is the 3×3 identity matrix.

7.4 Maximum likelihood estimation

For the underlying parameters $\Theta = (a, \sigma, \alpha, b, \eta, \rho, \beta, \lambda_0, \sigma^\lambda, \theta)$, we have the following log-likelihood function
$$l(\Theta) = \sum_{i=1}^n \left(-\frac{1}{2} \tilde{\mathbf{Y}}_i^\top \mathbf{S}_i^{-1} \tilde{\mathbf{Y}}_i - \frac{3}{2} \log(2\pi) - \frac{1}{2} \log(\det \mathbf{S}_i) \right)$$

The Kalman maximum likelihood estimator (KMLE) is found by maximising l. We select the noise term matrix $\mathbf{R}$ by numerical experiments. This noise term matrix can also be seen as the degree of information contained in the observations.

8 Estimation and calibration results

We select the model simultaneously from time series and prevailing market prices.

8.1 Calibration versus estimation

By calibration we mean solving an inverse problem for the parameters with respect to the prevailing market prices. Whereas, by estimation we determine the parameters applying statistical procedures to time series data.

Practical experience suggests that if all the parameters are determined by calibration, then the model poorly describes the dynamics of market risk factors. On the other hand, if all the parameters are obtained by estimation then typically the initial state of the model does not match market prices. For this reason we believe that the model should be selected from both calibration and estimation.

8.2 Our approach for determining the parameters under the risk neutral probability measure

In order to do valuation we need to determine the parameters under the risk neutral measure $a, \sigma, b, \eta, \rho, \lambda_0$ and σ^λ. These parameters are going to be estimated from times series of defaultable zero-coupon yields. For this purpose we apply the Kalman filter MLE procedure outlined above. Then, we calibrate the time dependent shift function φ in order to match the risk-free yield curve. The same idea of adding a deterministic time-dependent shift can also be implemented for the stochastic intensity λ which also allows to exactly calibrate the defaultable zero yields.

Alternatively, a, σ, b, η, ρ may as well be calibrated to market prices of interest rate caps/floors. See Theorem 5.4 and figures below. We do not follow this approach because interest rate caps and floors are based on LIBOR rates which are not a good proxy for risk-free rates. Since modelling credit spreads is very important for this project and better proxies for risk-free rates are available we use the Kalman filter approach which allow to input our proxies for risk-free and defaultable zero yields.
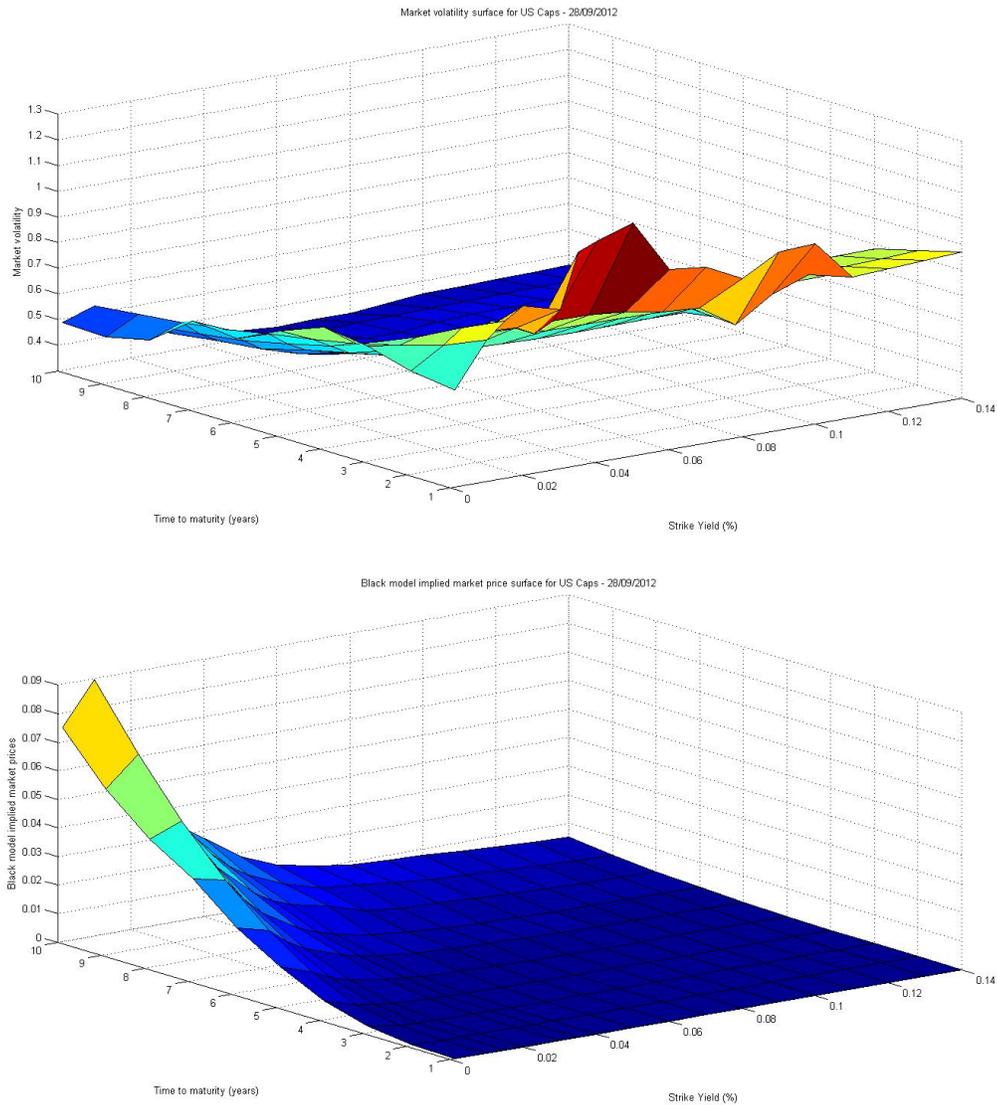
Figure 7: Implied volatility surfaces and market prices of interest rate caps for USD LIBOR.

8.3 Data used

We work with daily zero-coupon yields data. As proxy for the risk-free rates we use zero-coupon rates bootstrapped from OIS rates in USD (USD OIS ZC). As proxy for the credit spreads we use the credit spread on SA Government Bonds (SA Sov Sprd).

8.4 Estimation and calibration results

We now present some estimation and calibration results for some of the valuation dates.

Zero–coupon yields (%)
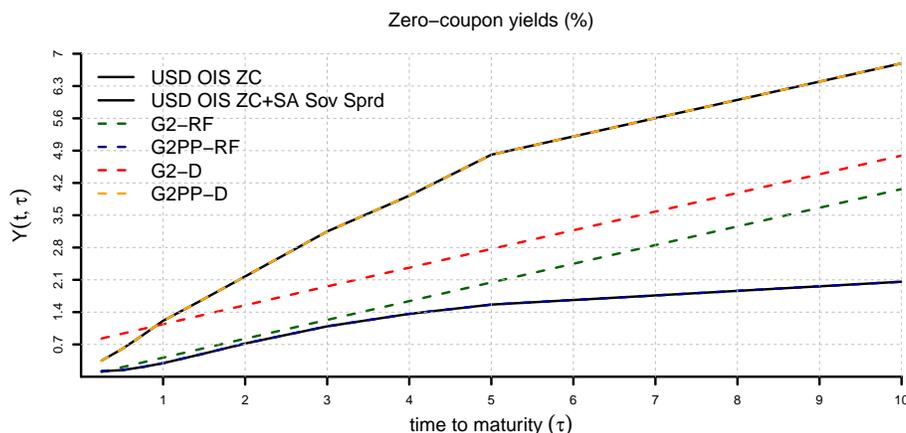
Figure 8: Calibration as of 31-12-2014. USD OIS ZC are risk-free zero coupon rates boot-strapped from OIS rates in USD. SA Sov Sprd is the credit spread on SA Government Bonds. G2-RF are the risk-free zero rates given by the model if we assume a constant shift. G2-D are the defaultable zero-rates given by the model if we assume a constant shift. G2PP-RF and G2PP-D are the corresponding models with time-dependent drift calibrated to the market rates. For the Kalman filter we set $R_0 = 10^{-4}$. We use 100 yield curve observations for the calibration.

| | G2++ | | | | | Credit | | $\mathbb{Q}(\tau \leq t) = 1 - \exp(-\lambda_0 t)$ | | |
date	a	σ	b	η	ρ	λ_0	σ^λ	$t = 1$	$t = 5$	$t = 10$
31-12-2014	0.70	7.5%	0.3	7.1%	-99.5%	0.60%	$2.9 \cdot 10^{-5}$	0.60%	2.95%	5.82%
31-12-2013	0.23	5.6%	1.16	5.7%	-98.6%	0.59%	$4.6 \cdot 10^{-5}$	0.60%	2.91%	5.73%
29-12-2012	0.05	0.35%	0.44	0.00%	20.1%	0.53%	$5.6 \cdot 10^{-5}$	0.53%	2.62%	5.16%
30-12-2008	0.34	14.16%	2.11	14.99%	-99%	0.68%	$9.5 \cdot 10^{-3}$	0.68%	3.34%	6.57%

Table 1: Model parameters for different valuation dates.

As we see in the figure and the table above the estimation and calibration results are reasonable. Volatility and speed of mean reversion parameters appear to be reasonable and consistent with other empirical studies. See e.g. Cuchiero (2006). One-, five- and ten-years default probabilities implied by the model also appear to be reasonable. We also observe that our simple credit risk model is able to capture only a constant credit spread. Of course, this is not realistic given the data. However, the result is not too bad in the context of the valuation problem, since for the callable floating rate note with write down features we only need a single credit spread.

9   Valuation results

In this section we implement the valuation models (Equations 14 to 20) for various parameters estimated and calibrated from the data, using the methodologies outlined in Section 8. No closed-form solutions to the valuation equations (see Equations 14, 16, 18 and 20) exist,

and these equations moreover contain first-hitting time random variables. Therefore, we implemented a Monte Carlo simulation.

Our Monte Carlo simulation exercise did not contain any nested simulation, therefore rendering our simulation exercise added efficiency from a computing point of view. We jointly simulated realisations of the short rate process $(r(t))_{t \geq 0}$ as well as the zero-coupon bond prices $P(t,T)$ (for t and $T \geq 0$), and the market spread $S(t,T)$ for maturity time T. We chose the particular form of interest rate model because the bond prices $P(t,T)$ are in closed form (see Theorem 5.2), so no nested Monte Carlo simulation is necessary.

From the point of view of the credit risk factors, the model for the underlying hazard process, $\Lambda(t) = \int_0^t \lambda(u)du$, was chosen on the grounds of tractability. The solution to the stochastic differential equation for the hazard rates follows a log-normal distribution. In addition, if independence is assumed between the short-rate risk factor and the hazard rate risk factor, explicit formulae for the spread can be computed - see Equation 26. This explicit formula was used in simulating the spread process, and furthermore made the computation simpler. We did realise that it is also possible to simulate assuming dependence between the short-rate and hazard rate risk factors, however, owing to time constraints, this was left as an opportunity for further refinement to the simulation.

Fifteen simulation exercises were executed in order to ascertain the change in market spread adjustment estimates to changes in the model parameters. We firstly analysed the behaviour of the estimate of the market spread adjustment to changes in the spread ceiling $\bar{S}$ - this may be an interesting relationship to consider, because issuers of these callable floating-rate notes are often at liberty to alter $\bar{S}$, and by using this chosen value of $\bar{S}$ they may be able to approximate a market price for the instrument. Secondly, we analysed the behaviour of the market spread adjustment with respect to (reasonable selected) changes in the parameters of the credit model. There were two parameters in this model - σ^λ and λ_0 - and each parameter was analysed in turn.

Because of time constraints, for each market spread adjustment simulation, we employed a sample size of 20 000. This is approximately one-fifth of that required by internal banking regulation, but because of time and computing constraints we were constrained to this value. We did attempt to calculate values with samples ranging in size from 10 000 paths to 20 000 paths, together with three standard deviation error bounds. Indeed, for each of the fifteen simulation exercises, we observed both the standard deviation error bounds as well as the Monte Carlo estimates to stabilise as the sample size increased from 10 000 paths to 20 000 paths.

Before proceeding to the results of the simulations for a FRN with coupon periods of 3 months, it is necessary to present the parameters used in the valuation model. We selected the parameters to be reasonable values reasonably representative of those calibrated and estimated from the data - and in some cases, the parameters employed were precisely those from the estimation and calibration. Table 2 below shows the values used for the fixed parameters.

Variable	Value
a	1
b	1
σ	0.01
η	0.01
ρ	-0.5
$\bar{T}$	10
T	5
δ	0.1
$\underline{S}/S_0$	0.75

Table 2: Fixed variables.

We now discuss the simulation results. In Figure 9 we observe that the value of the option to the issuer increases as credit spread volatility increases. Secondly, we observe that callability after time T is worth more than the one at time T. In Figure 10 we observe the value of the write-down to the issuer also increases if volatility increases. The value to the issuer of both the callability and the write-down is shown in Figure 10. In particular, we see that the value of the write-down is reduced considerably by callability after time T compared to callability at time T, because the issuer has a higher chance of closing the bond before the write-down may happen. In Figures 12 to 14 we study sensitivity with respect to the upper bound $\bar{S}$. Clearly the value of the write-down is decreasing in $\bar{S}$. We see that the option value is not considerably affected by this bound. Moreover, we observe that that the total value of the contract drops compared to the stand-alone value of the write-down, because when the option is exercised the write-down feature ceases to exist. This means that the value of the option and the write-down tend to be negatively correlated (see Section 4.1.4 for why we intuitively believe this to be the case). The dependence of the value on $\bar{S}$ can also be used conversely to imply a value of $\bar{S}$ from market prices, which may be useful for both issuers and investors in such instruments. In Figures 15-17 we observe that the value of the option (and write-down) tend to decrease (increase) as the probability of default increases. However, we see a mixed behaviour in the case of the option. This may be explained by the fact that higher value of the default probability means higher initial credit spread which makes the value of a FRN reissue higher to the issuer if the option is exercised.
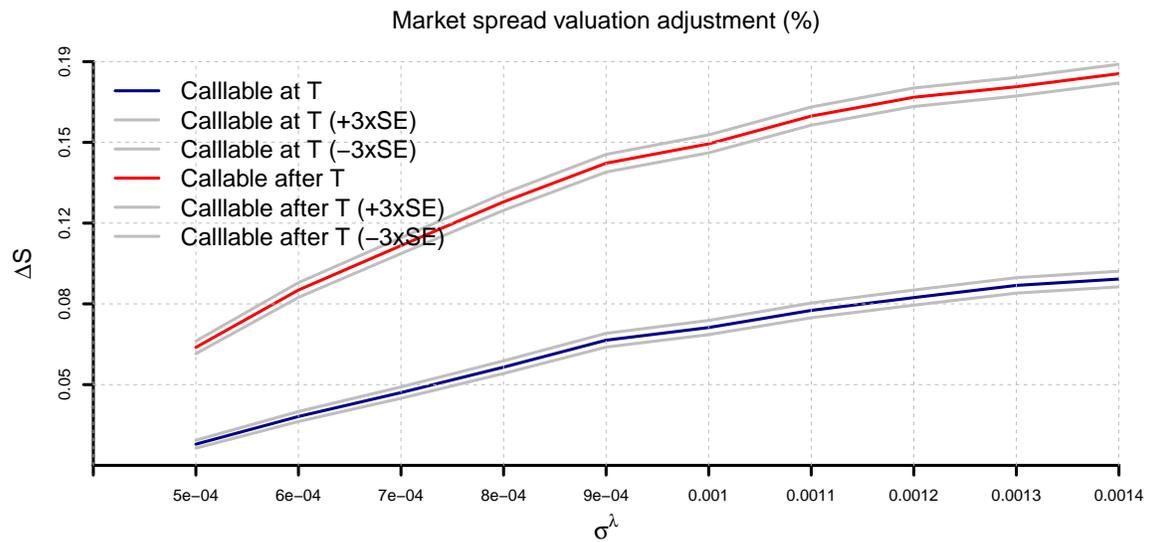
Figure 9: Monte Carlo market spread valuation adjustments (added to the time 0 market spread in order to compensate the investor for the risk of the issuer reissuing the FRN at a lower coupon), for a reasonable range of values of σ^λ. Each Monte Carlo estimate is based on a different sample of 20 000 paths, and plotted alongside each estimate is the 3 standard error bound. We set $\lambda_0 = 0.03$ and $\bar{S} = 2S_0$.
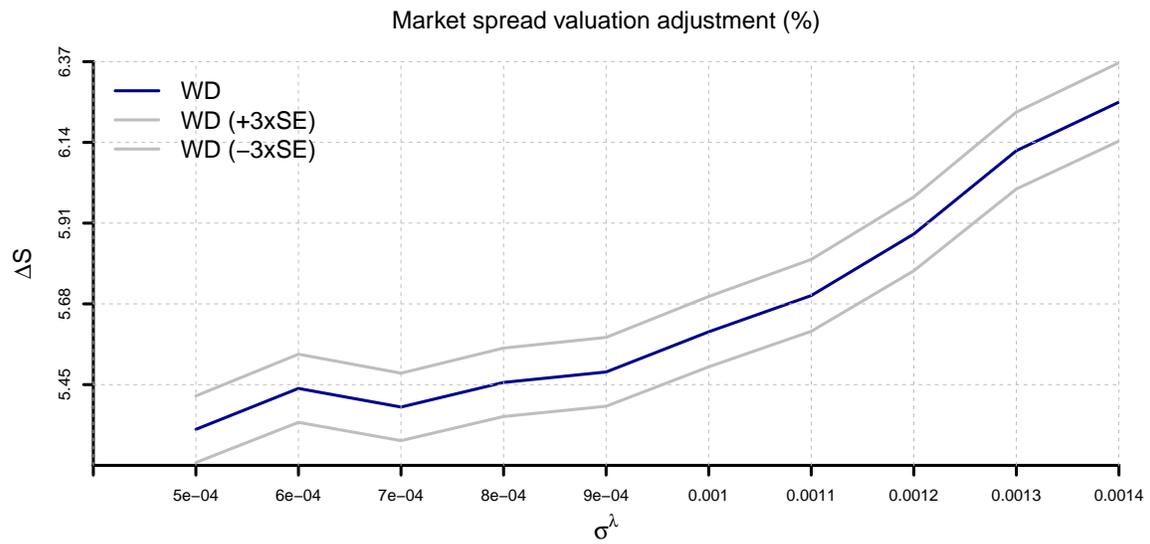
Figure 10: Monte Carlo market spread valuation adjustments (added to the time 0 market spread in order to compensate the investor for the risk of the issuer writing down the nominal of the FRN) for a reasonable range of values of σ^{λ}. Each Monte Carlo estimate is based on a different sample of 20 000 paths, and plotted alongside each estimate is the 3 standard error bound. We set $\lambda_0 = 0.03$ and $\bar{S} = 2S_0$.
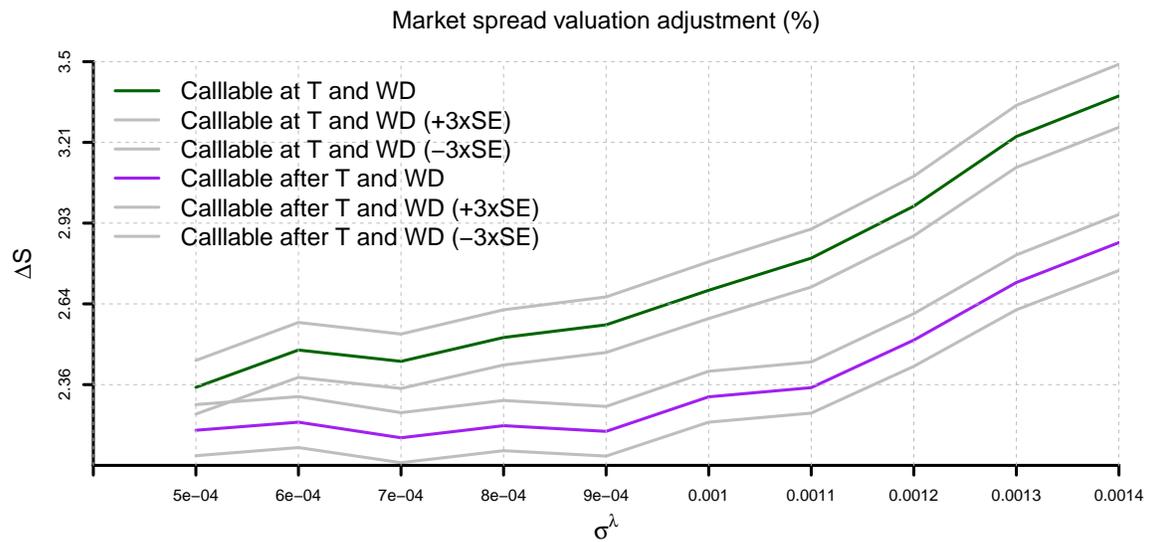
Figure 11: Monte Carlo market spread valuation adjustments (added to the time 0 market spread in order to compensate the investor for the risk of the issuer reissuing the FRN at a lower coupon and/or the issuer writing down the nominal of the FRN) for a reasonable range of values of σ^λ. Each Monte Carlo estimate is based on a different sample of 20 000 paths, and plotted alongside each estimate is the 3 standard error bound. We set $\lambda_0 = 0.03$ and $\bar{S} = 2S_0$.
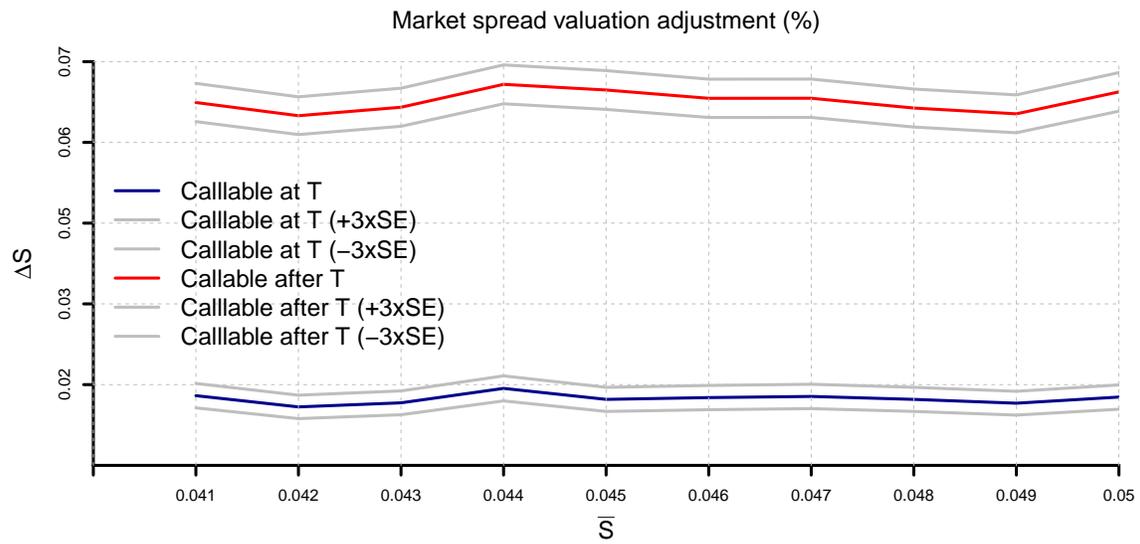
Figure 12: Monte Carlo market spread valuation adjustment estimates (added to the time 0 market spread in order to compensate the investor for the risk of the issuer reissuing the FRN at a lower coupon only), for a reasonable selected range of values of $\bar{S}$. Each Monte Carlo estimate is based on a different sample of 20 000 paths, and plotted alongside each estimate is the 3 standard deviation error bound. We set $\lambda_0 = 0.03$ and $\sigma^\lambda = 10^{-4}$.
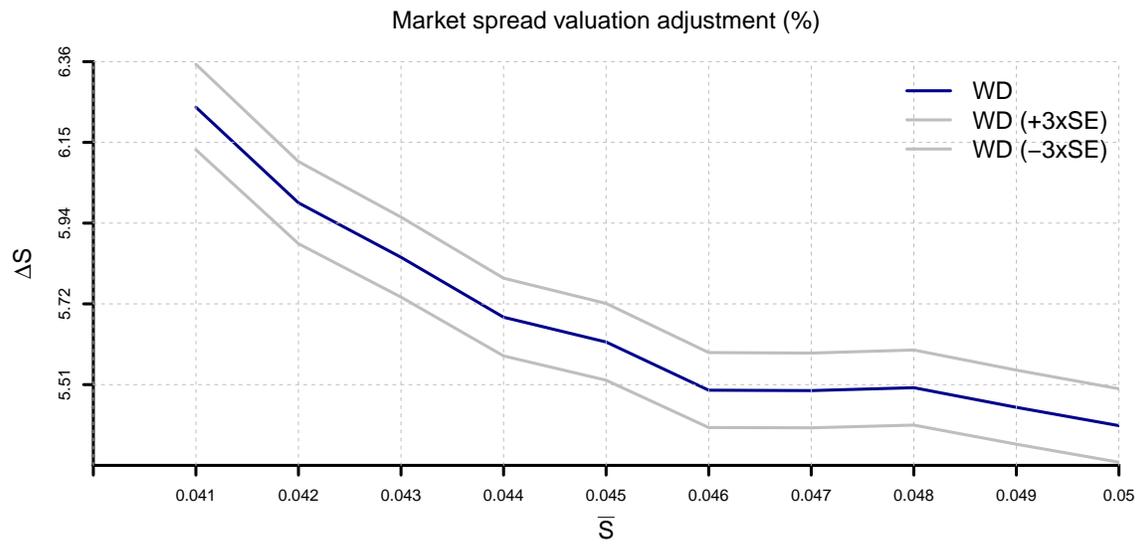
Figure 13: Monte Carlo market spread valuation adjustment estimates (added to the time 0 market spread in order to compensate the investor for the risk of the issuer writing down the nominal of the FRN only), for a reasonable selected range of values of $\bar{S}$. Each Monte Carlo estimate is based on a different sample of 20 000 paths, and plotted alongside each estimate is the 3 standard deviation error bound. We set $\lambda_0 = 0.03$ and $\sigma^\lambda = 10^{-4}$.
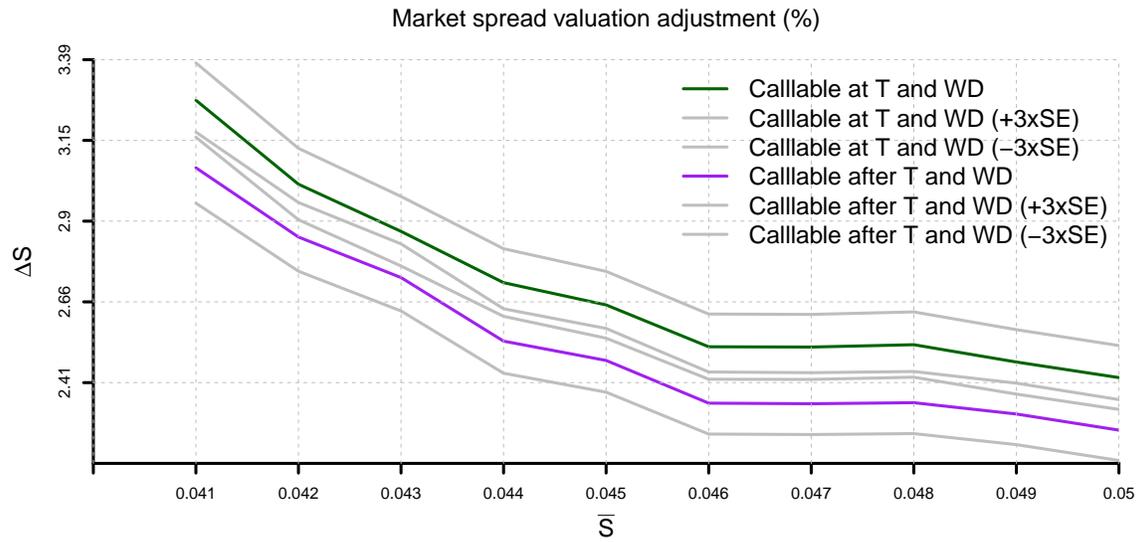
Figure 14: Monte Carlo market spread valuation adjustment estimates (added to the time 0 market spread in order to compensate the investor for the risk of the issuer reissuing the FRN at a lower coupon and/or the issuer writing down the nominal of the FRN), for a reasonable selected range of values of $\bar{S}$. Each Monte Carlo estimate is based on a sample of 20 000 paths, and plotted alongside each estimate is the 3 standard deviation error bound. We set $\lambda_0 = 0.03$ and $\sigma^\lambda = 10^{-4}$.
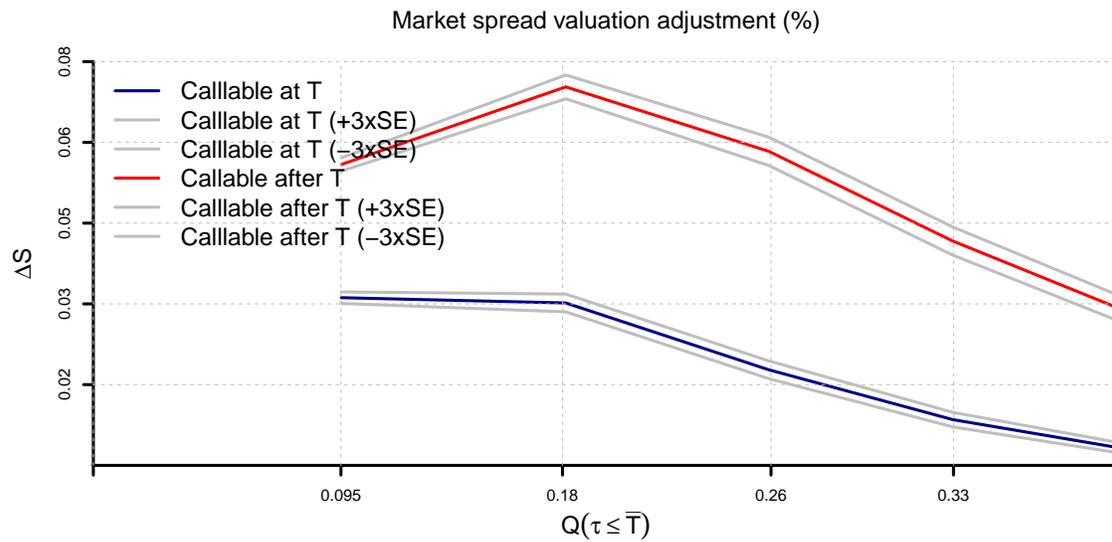
Figure 15: Monte Carlo market spread valuation adjustment estimates (added to the time 0 market spread in order to compensate the investor for the risk of the issuer reissuing the FRN at a lower coupon), for a selected range of values of the 10-year probability of default of the issuer. Each Monte Carlo estimate is based on a different sample of 20 000 paths, and plotted alongside each estimate is the 3 standard deviation error bound. We set $\sigma^\lambda = 10^{-4}$ and $\bar{S} = 2S_0$.
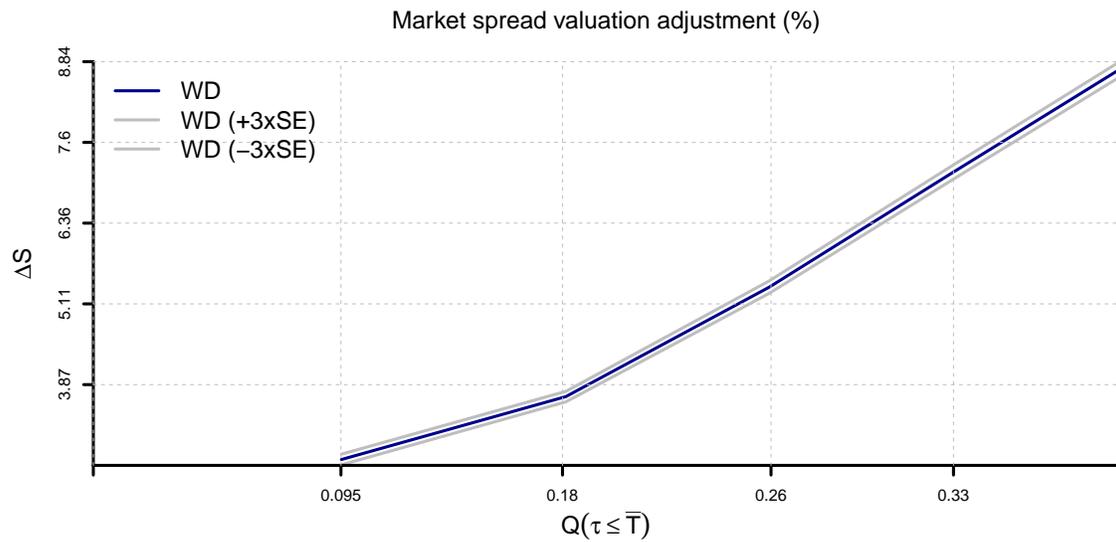
Figure 16: Monte Carlo market spread valuation adjustment estimates (added to the time 0 market spread in order to compensate the investor for the risk of the issuer writing down the nominal of the FRN), for a selected range of values of the 10-year probability of default of the issuer. Each Monte Carlo estimate is based on a different sample of 20 000 paths, and plotted alongside each estimate is the 3 standard deviation error bound. Note that by changing λ_0 we also modify the initial default probability and issuance spread. We set $\sigma^\lambda = 10^{-4}$ and $\bar{S} = 2S_0$.

Market spread valuation adjustment (%)
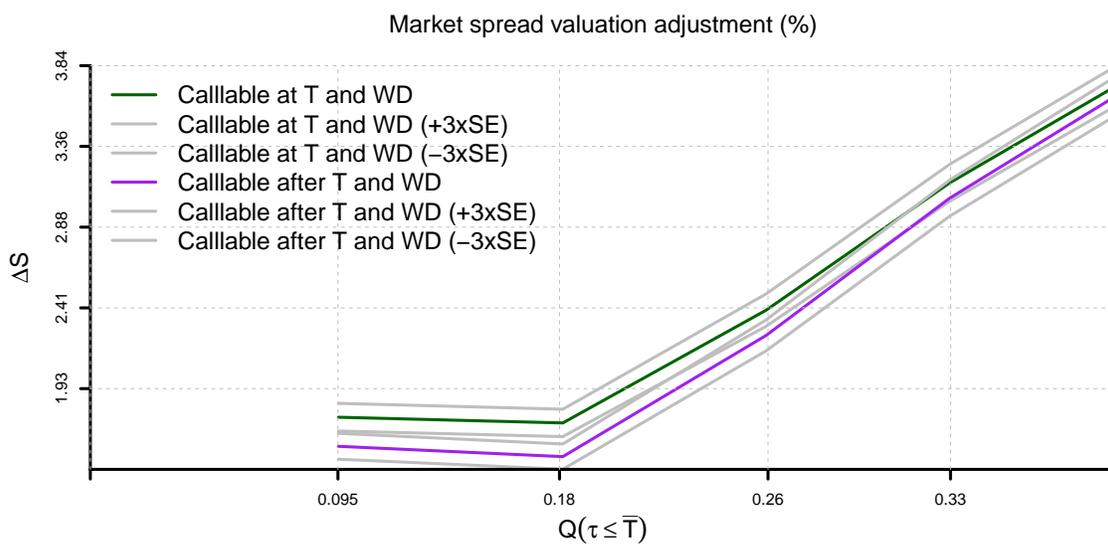
Figure 17: Monte Carlo market spread valuation adjustment estimates (added to the time 0 market spread in order to compensate the investor for the risk of the issuer reissuing the FRN at a lower coupon and/or the issuer writing down the nominal of the FRN), for a selected reasonable range of values of the 10-year probability of default of the issuer. for a selected range of values of the 10-year probability of default of the issuer. Each Monte Carlo estimate is based on a different sample of 20 000 paths, and plotted alongside each estimate is the 3 standard deviation error bound. We set $\sigma^{\lambda} = 10^{-4}$ and $\bar{S} = 2S_0$.

10   Conclusion

In this report, we developed a valuation methodology for callable floating-rate notes which allowed us to compute market spread adjustments for the issue of such instruments. This methodology, approached from the point of view of the issuer of the callable floating-rate note, essentially comprised a vanilla FRN valuation, the valuation of the callability feature and the valuation of the write-down feature. The valuation formulae did not admit analytical solutions, however, were straightforward to compute by Monte Carlo simulation. The tractability and parsimony of the processes assumed for the underlying risk factors added to the ease and efficiency of the Monte Carlo simulation.

From our valuation methodology, we were able to imply the components of the market spread adjustment which were attributed to the callability feature only and the write-down feature only. We found that the write-down component comprised the highest value to the issuer - this came as no surprise because of the fact that the issuer could write down his or her debt to zero.

In addition, we were able to develop a method to imply the upper barrier on the credit spread for the callable FRN, and managed to link this back to the equity of the issuer using the Merton Model. We believe that if one can imply the upper barrier on the callable FRN, there will be less opacity in the pricing of these instruments. Currently in the South African

market there is much opacity when it comes to valuing these such instruments. In sum, we hope that our research will shed some light on the current valuation practices on callable FRNs.

We end with a further research question. Given data on price-to-book and book values of assets from banks, can we find a link between the book value of equity over the book value of assets ratios and credit spreads using statistical models (e.g. regression)? In addition, is it possible to determine, based on the trigger for book value of equity over book value of assets ratios, a trigger value for the credit spread?

Bibliography

Bielecki, T. R., Rutkowski, M., 2002. Credit risk: modeling, valuation and hedging. Springer Science & Business Media.

Brigo, D., Garcia, J., Pede, N., 2015. Coco bonds pricing with credit and equity calibrated first-passage firm value models. International Journal of Theoretical and Applied Finance, 1550015.

Brigo, D., Mercurio, F., 2007. Interest rate models-theory and practice: with smile, inflation and credit. Springer Science & Business Media.

Cuchiero, C., 2007. Affine Interest Rate Models - Theory and Practice. Master thesis, Vienna University of Technology.

Haykin, S., 2004. Kalman filtering and neural networks. Vol. 47. John Wiley & Sons.

Lorenzen, H., 2015. Junior capital grows up a primer on new bank hybrid capital. Citi Research.

Mavuso, M., 2012. Money market subcommitee floating rate note specification. Johannesburg Stock Exchange.