Multi-Curve Cheyette-Style Models with Lower Bounds on Tenor Basis Spreads

Michael Konikov and Andy McClelland*
Numerix Quantitative Research and Development

December 18, 2019

Abstract

The modeling of tenor basis spreads is of central importance to CVA for tenor basis swaps. Such spreads are typically positive, suggesting a natural lower bound. We introduce a multi-curve Cheyette-style model with lower bounds enforced through level dependence in spread volatilities. The model is intuitive, easy to implement, and admits approximate swaption pricing formulae under affine specifications. We also discuss the importance of incorporating historical spread data into calibration criteria, and we formalize an according calibration strategy.

1 Overview and Motivation

Tenor basis swap spreads reflect the differences between forward reference rates of different tenor. For example, the 3M IBOR-*vs.*-6M IBOR spread reflects the difference between forward 3M and 6M IBORs. We use IBOR here to refer to an Inter-Bank Offered Rate, of which LIBOR (London IBOR) is a well-known example. Since the Financial Crisis of '07-'08, there has been significant volatility in observed spreads, and as discussed momentarily, how these spreads are modeled is of paramount importance to the CVA (Credit Valuation Adjustment) for a tenor basis swap.

LIBOR itself is being phased out in favor of risk-free reference rates, such as SOFR (Secured Overnight Financing Rate) in USD markets. The modeling framework developed here is of course applicable to similar reference rates which are not being phased out.[1] Examples include EURIBOR or TIBOR, which are being reformed to make heavier use of transaction data, or BBSW or CORA, which are already transaction based.[2] Similarly, the framework could be used for the BBI (Bank Bill Index) being developed by ICE (Intercontinental Exchange) as an alternative to USD LIBOR.[3] In short, the framework is applicable to any reference rate which resembles a money-market term rate, and which is thus sensitive to credit and liquidity concerns.

Denote the forward 3M and 6M IBORs by $R_3(t,T)$ and $R_6(t,T)$, where t is the calendar date and T is the fixing date. A 3M IBOR-*vs.*-6M IBOR tenor basis swap exchanges 3M IBOR-plus-spread at a 3M frequency for 6M IBOR-flat at a 6M frequency, and we will denote the spread by

*The authors would like to thank Greg Whitten for supporting research and development efforts at Numerix, as well as Ron Levin, Nader Rahman, Peter Jaeckel and Fabio Mercurio for providing valuable comments.

[1] A catalog of various reform efforts or terminations of references rates is provided in FSB ('18).

[2] EURIBOR and TIBOR are reference rates in Europe and Japan which use a LIBOR-like methodology. BBSW (Bank Bill Swap Rate) of Australia and CDOR (Canadian Dollar Offered Rate) rely on liquid bank-bill markets.

[3] See ICE ('19) for motivations for the BBI and a review of the proposed methodology.

$\phi_{3,6}$. Letting $P(t,T)$ denote discount factors computed off the ON (Overnight) curve, its value is

$$V_{3,6}(t) = \sum_j P(t,T_{6,j+1}) R_6(t,T_{6,j})\,\tau_{6,j} - \sum_i P(t,T_{3,i+1})(R_3(t,T_{3,i}) + \phi_{3,6})\,\tau_{3,i}\,, \tag{1}$$

assuming that the trade pays the 3M leg, and where we ignore convention- and calendar-related complications for notational convenience. In the above $T_{3,i}$ and $T_{6,j}$ refer to 3M and 6M fixing dates, respectively, and $\tau_{3,i}$ and $\tau_{6,j}$ are the relevant day-count fractions, *i.e.* $\tau_{3,i} \approx \frac{3}{12}$ and $\tau_6 \approx \frac{6}{12}$. The dominant risk factor for such a trade is the spread between the forward 6M IBOR curve and the forward 3M IBOR curve. Indeed, the value of the trade roughly satisfies $V_{3,6}(t) \propto R_6^{av}(t) - R_3^{av}(t) - \phi_{3,6}$, where $R_3^{av}(t)$ and $R_6^{av}(t)$ are average levels of the respective forward curves. The model selected for the 3M IBOR-*vs.*-6M IBOR tenor basis spread will thus having a greater bearing than any other factor on the CVA for such a basis swap. Indeed, computing the CVA for a tenor basis swap would resemble valuing a tenor basis swaption.

There is an existing literature on tenor basis modeling, and managing CVA for tenor basis swaps has indeed been an impetus for its growth. However, some important issues remain open. In particular, a well-documented feature of tenor basis swap spreads is that they rarely take negative values. Thus, a desirable property of tenor basis models is the ability to control the lower bounds of spreads, but there is an absence of specifications with this property within the popular "multiplicative" framework developed by Henrard ('13) and others.

To develop such a model, we build upon the existing work of Martinez ('09), Grbac and Runggaldier ('15) and Miglietta ('15), who recast multiplicative models in terms of instantaneous forward spreads. These authors used level-independent volatility functions for (Gaussian) spread processes, while we will use level-dependent volatility functions as a means for imposing lower bounds. As we shall see, the no-arbitrage drift restriction for the spread curve process is more complicated than it is for the discounting curve process. A solution for the drift function was not derived in the earlier works, and in turn, a Markov representation for the spread process was unavailable. Despite this, the earlier authors were able to make progress in characterizing IBOR dynamics, but as we will discuss this is *only* possible in the level-independent case. In this article we both solve the drift restriction and derive a Markov representation under the general case of level-dependent volatility. With this we arrive at an intuitive and tractable multi-curve specification in the spirit of Cheyette ('92) which supports lower bounds and is easy to implement.

We also discuss the calibration of multi-curve models, which is another issue underdeveloped in the literature. We argue for using a mix of historical and implied data, as opposed to relying on swaption volatilities alone. Specifically, we argue that loading the observed swaption skew onto the spread process can lead to spread dynamics which poorly match empirical behavior. A calibration strategy informed by this view is formalized and the results of a calibration exercise are presented.

2 The Model

2.1 IBOR Specification

We focus here on a two-curve setup consisting of the base curve, taken to be the ON curve, along with the 3M IBOR curve. We model the ON curve and the ON-*vs.*-3M IBOR spread curve, which links the ON curve and the 3M IBOR curve and governs the behavior of ON-*vs.*-3M IBOR tenor basis swap spreads. The analysis is readily extended to a larger set of curves including *e.g.* the 1M

and 6M IBOR curves, as is done in Appendix A. We begin with the ON curve, modeled via its instantaneous forward curve, $f_0(t,T)$. Discount factors are $P(t,T) = \exp(-\int_t^T f(t,u)\,du)$, and in a "single-curve world" the 3M IBOR off the ON curve would be

$$R_{0,3}(t,T) = \frac{1}{\tau_3}\Big(e^{\int_T^{T+\tau_3} f_0(t,u)\,du} - 1\Big)\,, \tag{2}$$

where we employ the subscript $0,3$ to emphasize that this is not the true 3M IBOR.

To compute the true 3M IBOR off the 3M IBOR curve, we follow Grbac and Runggaldier ('15), Miglietta ('15) and Martinez ('09), who employ an intuitive extension of (2),

$$R_3(t,T) = \frac{1}{\tau_3}\Big(e^{\int_T^{T+\tau_3} f_0(t,u)+s_3(t,u)\,du} - 1\Big)\,. \tag{3}$$

Here $s_3(t,T)$ is a fictitious process, playing the role of the instantaneous forward 3M IBOR spread curve over the ON curve. The specification can be motivated and justified theoretically by a simple credit-based argument. Let $P^*(t,T)$ denote the value of a hypothetical bond for a representative IBOR panel bank. For simplicity, we assume zero recovery and a deterministic hazard process $\lambda(t)$. It is straightforward to demonstrate under these conditions that the 3M IBOR would be equal to

$$\frac{1}{\tau_3}\mathbb{E}_t^{T+\tau_3}\left[\frac{1}{P^*(T,T+\tau_3)} - 1\right] = \frac{1}{\tau_3}\Big(e^{\int_T^{T+\tau_3} f_0(t,u)+\lambda(u)\,du} - 1\Big)\,, \tag{4}$$

which resembles (3) with $s_3(t,u)$ replaced by $\lambda(u)$. We note also that liquidity factors could easily be worked into $P^*(t,T)$ with similar effect.

The specification in (3) may be viewed as a special case of the "multiplicative" specification,

$$R_3(t,T) = \frac{1}{\tau_3}\Big(e^{\int_T^{T+\tau_3} f_0(t,u)\,du}\,S_3^{\times}(t,T) - 1\Big)\,, \tag{5}$$

as has been popularized by Henrard ('13) and others; here $S_3^{\times}(t,T)$ is a fictitious multiplicative discrete-tenor spread factor. The specification in (5) can be recast to that in (3) by setting

$$S_3^{\times}(t,T) = e^{\int_T^{T+\tau_3} s_3(t,u)\,du}\,. \tag{6}$$

The advantage of working with models of $s_3(t,T)$ is that the behavior of the observed tenor basis swap spread curve will roughly mimic that of $s_3(t,T)$. Indeed, a crude approximation of $dR_3(t,T)$ is simply $df_0(t,T) + ds_3(t,T)$, which follows from a simple Taylor expansion. Behaviors inherited from $s_3(t,T)$ include level dependence in volatility and lower bounds, which are of central interest to our objectives. We note that it is possible to control these aspects of spread behavior by manipulating the diffusion function of $S_3^{\times}(t,T)$, but with far less transparency than in the case of $s_3(T,t)$.[4] Moreover, extending the setup in (3) to a general multi-spread setting (*i.e.* for 6M IBORs) is much simpler than for that in (5), as discussed momentarily.

Another intuitive setup is the "additive" specification of Mercurio ('10), namely

$$R_3(t,T) = R_{0,3}(t,T) + S_3^{+}(t,T)\,, \tag{7}$$

where $R_{0,3}(t,T)$ is the single-curve 3M IBOR off the ON curve in (2), and where $S_3^{+}(t,T)$ is a fictitious additive discrete-tenor spread. We feel that modeling $s_3(t,T)$ retains the best of both worlds. It is as intuitive as the additive approach in that we simply add $s_3(t,T)$ to $f_0(t,T)$ and evaluate a standard IBOR formula as per (3), and we have the strong credit-based justification for the specification in (3) presented earlier, which is not available for the additive specification in (7).

[4] Eberlein *et al.* ('19) use a log-normal model of $S_3^{\times}(t,T)$ with a shift of one to enforce lower bounds at zero. Shifts allow for control over the lower bound but controlling level dependence requires more, *e.g.* a CEV specification.

2.2 Curve Dynamics and Drift Restrictions

We model $f_0(t,T)$ and $s_3(t,T)$ as each having separable HJM (Heath, Jarrow and Morton, '92) dynamics. This allows us to adapt the insights of Cheyette ('92) who explored such dynamics in the single-curve context. To ease exposition we use a simple specification wherein each curve has a loading onto a dedicated factor such that the curves are uncorrelated, and where each factor loading involves a single exponential. These assumptions are without any loss of generality. General factor structures allowing for nonzero correlations and composite factor loadings are treated fully in Appendix A.

Our system evolves under the spot measure as

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\,dt + \psi_0(t)\,e^{-\kappa_0(T-t)}\,dW_0(t) \\ ds_3(t,T) &= \mu_3(t,T)\,dt + \psi_3(t)\,e^{-\kappa_3(T-t)}\,dW_3(t)\,, \end{aligned} \tag{8}$$

with $dW_0(t)\,dW_3(t) = 0\,dt$. In order to enforce lower bounds on $s_3(t,T)$, we will adopt a diffusion function of the CEV (Constant Elasticity of Variance) type,

$$\psi_3(t) = \nu_3(t)\big(s_3(t,t) - l_3(t)\big)^{\beta}\,, \tag{9}$$

where $s_3(t,t)$ is the "short spread", and where $l_3(t)$ governs lower bounds. The use of the CEV specification is well understood in the single-curve context so we do not elaborate upon it here. We note only that when $s_3(t,t) = l_3(t)$, *i.e.* when the short spread attains its lower bound, the curve $s_3(t,T)$ is unable to diffuse to lower values. For the forthcoming analysis we do not impose (9), leaving $\psi_0(t)$ and $\psi_3(t)$ as general state-dependent quantities so as to obtain general results. We will also use the notation $\sigma_0(t,T) = \psi_0(t)\,e^{-\kappa_0(T-t)}$ and $\sigma_3(t,T) = \psi_3(t)\,e^{-\kappa_3(T-t)}$ where convenient.

Models such as (8) were proposed in Grbac and Runggaldier ('15), Miglietta ('15) and Martinez ('09). See also Geelhar *et al.* ('17) who extend such models for stochastic volatility. The typical no-arbitrage restriction applies for $\mu_0(t,T)$,

$$\int_t^T \mu_0(t,u)\,du = \frac{1}{2}\left(\int_t^T \sigma_0(t,u)\,du\right)^2 \implies \mu_0(t,T) = \left(\int_t^T \sigma_0(t,u)\,du\right)\sigma_0(t,T)\,, \tag{10}$$

where the latter follows from differentiation of the former against T. An analogous no-arbitrage restriction can be derived for $\mu_3(t,T)$. To this end, consider a contract paying $R_3(T,T)\,\tau_3$ at $T+\tau_3$, *i.e.* a standard 3M IBOR payment. Denoting its value by $V_3(t,T)$, we clearly have

$$V_3(t,T) = P(t,T+\tau_3)\,R_3(t,T)\,\tau_3 = e^{-\int_t^T f_0(t,u)\,du + \int_T^{T+\tau_3} s_3(t,u)\,du} - e^{-\int_t^{T+\tau_3} f_0(t,u)\,du}\,, \tag{11}$$

which must grow on average at the risk-free rate under the spot measure. Taking the differential of (11) and enforcing the restriction that $\mathbb{E}_t[dV_3(t,T)] = r(t)V_3(t,T)\,dt$ gives rise to the drift restriction

$$\int_T^{T+\tau_3} \mu_3(t,u)\,du = -\frac{1}{2}\left(\int_T^{T+\tau_3} \sigma_3(t,u)\,du\right)^2. \tag{12}$$

We note that under general conditions where $s_3(t,T)$ and $f_0(t,T)$ are correlated, covariance terms will appear in this restriction. Differentiating (12) against T yields the necessary condition

$$\mu_3(t,T+\tau_3) - \mu_3(t,T) = -\big(\sigma_3(t,T+\tau_3) - \sigma_3(t,T)\big)\left(\int_T^{T+\tau_3} \sigma_3(t,u)\,du\right), \tag{13}$$

which is a difference equation in $\mu_3(t,T)$. This stands in contrast to the restriction on $\mu_0(t,T)$ in (10) which resolves the nature of $\mu_0(t,T)$ conclusively. Grbac and Runggaldier ('15), Miglietta ('15) and Martinez ('09) encountered the restriction in (13), but did not obtain a solution to it for $\mu_3(t,T)$. In spite of this, Grbac and Runggaldier ('15) were able to make progress in characterizing IBOR dynamics. This was possible as they adopted a Gaussian specification, $\psi_3(t) = \nu_3(t)$, for a deterministic parameter curve $\nu_3(t)$. To appreciate the importance of this, note that all we require to describe 3M IBORs as per (3) is the quantity $\int_T^{T+\tau_3} s_3(t,u)\,du$, which evaluates to

$$\begin{aligned}\int_T^{T+\tau_3} s_3(t,u)\,du &= \int_T^{T+\tau_3} s_3(0,u)\,du + \int_0^t \int_T^{T+\tau_3} \mu_3(v,u)\,du\,dv + \int_0^t \int_T^{T+\tau_3} \sigma_3(v,u)\,du\,dW_3(v) \\ &= \int_T^{T+\tau_3} s_3(0,u)\,du - \frac{1}{2}\int_0^t \left(\int_T^{T+\tau_3} \sigma_3(v,u)\,du\right)^2 dv + \int_0^t \int_T^{T+\tau_3} \sigma_3(v,u)\,du\,dW_3(v)\end{aligned} \tag{14}$$

with use of (12). It is thus clear that the integrated spread has no direct dependence upon $\mu_3(t,T)$. However, for state-dependent $\sigma_3(t,T)$ such as our CEV specification, $\nu_3(t)(s_3(t,t) - l_3(t))^\beta$, the quantity $s_3(t,t)$ is present and this does depend on $\mu_3(t,T)$. It is only in the special case of a state-independent $\sigma_3(t,T)$ that $s_3(t,t)$ is not present and there is no need for $\mu_3(t,T)$.

To obtain a solution for $\mu_3(t,T)$ we return to (12). As was also noted by the earlier authors, closer inspection reveals some flexibility in choosing a solution. Indeed, we observe that $\mu_3(t,T)$ enters under an integral, and the region of integration is the IBOR coverage period, which slides with the fixing date T. Thus, we have a restriction on the behavior of $\mu_3(t,T)$ *over* an IBOR coverage period, but not *within* such a period. To understand this, note that constructing an IBOR could have been achieved with the discrete-tenor multiplier $S_3^\times(t,T)$. Introducing the infinitesimal-tenor $s_3(t,T)$ affords modeling advantages, but introducing an integral was ultimately artificial; we discuss alternative specifications momentarily. An immediate consequence of this is flexibility in specifying $s_3(t,T)$, *e.g.* we could initialize it with any $s_3(0,T)$ agreeing with $S_3^\times(0,T)$, though in practice $s_3(0,T)$ is of course determined by a financially-meaningful interpolation scheme. Flexibility in the drift $\mu_3(t,T)$ derives from the same source as does flexibility in $s_3(t,T)$.

For the separable volatility structure $\sigma_3(t,T) = \psi_3(t)\,e^{-\kappa_3(T-t)}$, there is in fact a solution for $\mu_3(t,T)$ which is "natural" in the sense that it satisfies (12) and requires no arbitrary parameterization. To obtain this solution, substitute the volatility structure into (13) for

$$\mu_3(t,T+\tau_3) - \mu_3(t,T) = \psi_3(t)^2\,\frac{\left(e^{-\kappa_3\tau_3}-1\right)^2}{\kappa_3}\,e^{-2\kappa_3(T-t)}\,. \tag{15}$$

The form of (15) motivates an ansatz of the type $\mu_3(t,T) = \psi_3(t)^2\,A_3\,e^{-2\kappa_3(T-t)}$. Substituting this into (15) and solving for A_3 yields the solution for $\mu_3(t,T)$,

$$\mu_3(t,T) = \psi_3(t)^2\,\frac{\left(e^{-\kappa_3\tau_3}-1\right)^2}{\kappa_3\left(e^{-2\kappa_3\tau_3}-1\right)}\,e^{-2\kappa_3(T-t)}\,, \tag{16}$$

which can be confirmed to satisfy the sufficient condition in (12). Importantly, this solution is separable in t and T, and thus a representation of $s_3(t,T)$ in terms of Markov states is available.

We can readily extend (3) to produce a 6M IBOR via the introduction of $s_6(t,T)$, which captures credit and liquidity factors driving the basis between the 3M IBOR and 6M IBOR curves,

$$R_6(t,T) = \frac{1}{\tau_6}\left(e^{\int_T^{T+\tau_6} f_0(t,u)+s_3(t,u)+s_6(t,u)\,du} - 1\right). \tag{17}$$

Note that we integrate $s_6(t,T)$ over an interval of length τ_6 along with both $f_0(t,T)$ *and* $s_3(t,T)$. This enforces that $\big(1+\tau_6\,R_6(t,T)\big) \geq \big(1+\tau_3\,R_3(t,T)\big)\big(1+\tau_3\,R_3(t,T+\tau_3)\big)$, or rather that a forward 6M IBOR is no less than the "average" of the two adjacent forward 3M IBORs spanning its coverage period. Further discussion of multiple spread curves, and the derivation of drift functions analogous to (16), *i.e.* for $\mu_6(t,T)$, are provided in Appendix A.

Before proceeding, we briefly consider an alternative to the formulation in (3). We have adopted this formulation, wherein the instantaneous spread curve $s_3(t,T)$ is modeled, as it has already been developed within the literature and adopted by others for practical applications. However, with minimal effort we could move to a formulation wherein one models the mean spread curve

$$\bar{s}_3(t,T) = \frac{1}{\tau_3}\int_T^{T+\tau_3} s_3(t,u)\,du\,, \tag{18}$$

which is akin to a discrete-tenor spread. Modeling $\bar{s}_3(t,T)$ with the same dynamics as those prescribed for $s_3(t,T)$, *i.e.* with $\bar{s}_3(t,t)$ appearing in the diffusion function, is an equivalent way of obtaining control over level dependence in spreads and enforcing lower bounds.

In most practical settings the two approaches will lead to near-identical implementations in terms of Markov representations, approximate swaption prices, *etc.* As such, one may be indifferent to which is adopted. Models of $s_3(t,T)$ are slightly more intuitive and accessible as we are simply analogizing instantaneous forward rates $f_0(t,T)$. However, the drift function of $\bar{s}_3(t,T)$ would be resolved explicitly, and while we are able to solve for the drift function of $s_3(t,T)$, some may prefer to avoid the issue entirely.

In more advanced settings involving *e.g.* jump discontinuities or time-varying reversion coefficients, recasting from $s_3(t,T)$ to $\bar{s}_3(t,T)$ offers more tangible advantages as complications arising from the restriction on the drift of $s_3(t,T)$ become more pronounced. These issues are discussed in greater detail in Appendix E, Appendix F and Appendix G where the formulation in (18) is developed further.

2.3 Markov State Processes, Swap Rates & Swaption Pricing

We present the Markov representation of $s_3(t,T)$ here for reference,[5]

$$s_3(t,T) = s_3(0,T) + e^{-\kappa_3(T-t)}X_3(t) + A_3\,e^{-\kappa_3(T-t)}\big(e^{-\kappa_3(T-t)}-1\big)Y_3(t)\,, \tag{19}$$

$$\begin{aligned} dX_3(t) &= \kappa_3\big(\tfrac{A_3}{\kappa_3}\psi_3(t)^2 - A_3Y_3(t) - X_3(t)\big)\,dt + \psi_3(t)\,dW_3(t) \\ dY_3(t) &= 2\kappa_3\big(\tfrac{1}{2\kappa_3}\psi_3(t)^2 - Y_3(t)\big)\,dt \end{aligned} \quad , \quad A_3 = \frac{\big(e^{-\kappa_3\tau_3}-1\big)^2}{\kappa_3\big(e^{-2\kappa_3\tau_3}-1\big)}\,. \tag{20}$$

In Appendix A there is a full derivation of this result, and there is a full derivation for the general framework with multiple spread curves, multiple factors and composite factor loadings. One interesting feature of this representation is that there is a convexity term in the drift of $X_3(t)$, namely $A_3\,\psi_3(t)^2$. For models of $f_0(t,T)$, such a term does not feature. Its presence here owes to the fact that $\mu_3(t,t) \neq 0$, which contrasts with $\mu_0(t,t) = 0$.

We briefly present approximate swap rate dynamics, and discuss how approximate basis swap spreads and swaption prices can be obtained by extension; derivations and results for the general

[5] Setting $Y_3'(t) = -A_3Y_3(t)$ would be more in line with the classical Cheyette ('92) approach. We separate $-A_3(t)$ so that $Y_3(t)$ can be reused in the representation for $s_6(t,T)$ *etc.* in the general case of multiple spread curves.

case are provided in Appendix B. Consider a forward-starting 3M IBOR-*vs.*-fixed swap, with fixing/payment dates of $T_i \in \{T_0, T_1, \dots\}$. The 3M IBOR fixes at T_i and exchanges against a fixed rate at T_{i+1}, with $T_{i+1} - T_i = \tau_{3,i}$. The ATM swap rate is

$$\omega_3(t\,;T_0) = \sum_i \frac{P(t,T_{i+1})R_3(t,T_i)\,\tau_{3,i}}{L(t\,;T_0)} \quad \text{where} \quad L(t\,;T_0) = \sum_j P(t,T_{j+1})\,\tau_{3,i}\,. \tag{21}$$

The well-known "frozen coefficients" technique for deriving approximate (annuity-measure) swap rate dynamics can be easily adapted for our modeling framework. For the model in (8) one obtains

$$d\omega_3(t\,;T_0) \approx \gamma_0(T_0)\,\psi_0(t)\,e^{-\kappa_0(T_0-t)}\,d\overline{W}_0^{T_0}(t) + \gamma_3(T_0)\,\psi_3(t)\,e^{-\kappa_3(T_0-t)}\,d\overline{W}_3^{T_0}(t)\,, \tag{22}$$

where the loadings $\gamma_0(T_0)$ and $\gamma_3(T_0)$ are functions of $P(0,T_i)$, $R_3(0,T_i)$ and $L(0\,;T_0)$. One can also derive approximate dynamics for tenor basis swap spreads, which have the same structure as those in (22), but the loading onto $d\overline{W}_0^{T_0}(t)$ is of course far smaller for spreads.

In general, pricing swaptions under (22) requires numerical techniques, complicating calibration. For affine specifications however, *e.g.* $\psi_0(t) = \nu_0(t)$ and $\psi_3(t) = \nu_3(t)(s_3(0,t) + X_3(t) - l_3(t))^{\frac{1}{2}}$, the vector process $[\omega_3(t;T_0), X_3(t), Y_3(t)]$ is affine and amenable to standard Fourier techniques.[6] Note that this is a general result, and holds for models of multiple spread curves and multiple factors, provided that the joint diffusion function is affine in the relevant short-rate factors and short-spread factors. For such models it is possible to approximate the volatilities of swap rates and basis swap spreads in closed form, which is useful to us as these quantities are used toward calibration.

3 Calibration

3.1 Historical *vs.* Implied Data

The use of historical data has already been considered by Henrard ('18), but this was not the central subject of his article and there remains much room for development. Of particular benefit would be further motivation and discussion of the benefits of such an approach *vis-á-vis* relying solely on swaption volatilities, and also a formalization of calibration strategies.

When discussing swaption (or capfloor) volatilities, we assume for simplicity that swaptions reference 3M IBOR-paying underlyings. In our setup, model-based swaption volatilities will be sensitive to both the ON volatility, $\nu_0(t)$, and the ON-*vs.*-3M IBOR spread volatility, $\nu_3(t)$. This is clear from (22) wherein swap rate innovations have a $d\overline{W}_3^{T_0}(t)$ contribution. As such, it is formally *impossible* to calibrate $\nu_0(t)$ from swaption volatilities without reference to $\nu_3(t)$. The latter must be accounted for when calibrating the former. If $\nu_3(t)$ is ignored (set to zero) during calibration of $\nu_0(t)$ and reintroduced afterward, the fit of the model to swaption volatilities will clearly be affected. More broadly, this would overstate the ON curve volatility, affecting things such as the calculation of ON futures convexity, CVA for OIS swaps, long-dated equity trades where discount-curve convexity is material, and so on.

Many authors calibrate multi-curve models to swaption volatilities in isolation, but we feel that for CVA purposes we are far better served by incorporating historical data. Indeed, under ideal settings the calibrated model will reproduce *historically-observed* spread behavior. To understand our position, consider a simple set of ATM (At-the-Money) co-terminal implied volatilities. It is

[6] A Feller condition can be derived for the case where $\beta = \frac{1}{2}$, as is presented in Appendix D.

known that with a deterministic $s_3(t,T)$, or rather just $s_3(T)$, we have sufficient information in this set to calibrate $\nu_0(t)$ at a given set of interpolation knots. Now consider volatilities for swaptions with the same tenor structure, but struck at ATM $\pm\, 0.5\%$. Technically, these additional implied volatilities will render the volatility of a stochastic $s_3(t,T)$, namely $\nu_3(t)$, identifiable also because $\nu_3(t)$ will affect the skew under CEV dynamics.[7]

However, as $f_0(t,T)$ is Gaussian and produces almost no skew, this would be akin to assuming that the factors driving observed skew in swaption volatilities, *i.e.* level dependence and stochastic volatility in the 3M IBOR curve, owe *entirely* to the behavior of the ON-*vs.*-3M IBOR spread curve. This is refuted empirically, and thus such a calibration is unlikely to be consistent with historically-observed ON-*vs.*-3M IBOR spreads. Of course, if we had supreme confidence in our model specification, we could possibly calibrate to swaption volatilities alone, but in practical CVA settings this is rarely true. Note also that the use of historical data is essentially unavoidable for the volatilities of other tenor basis spreads. Indeed, the 3M IBOR-*vs.*-6M IBOR spread curve has no bearing upon swaptions over 3M IBOR-*vs.*-fixed swaps.

3.2 Criterion

A method to incorporate historical data into our calibration criteria is now presented. A minimal setup is employed in order to convey the general concept. We assume the quantities to be calibrated are $\{\nu_0(t),\, \nu_3\}$. The rationale for a constant ν_3 is that it will be inferred from the historical basis swap spread volatility, *i.e.* a single number as opposed to a curve. This could be relaxed if we had a forecast of basis spread volatility, *e.g.* via EWMA (Exponentially-Weighted Moving Average). We only assume that $\{\kappa_0,\, \kappa_3\}$ are known so as to simplify the presentation; in practice these could be calibrated to (discrete-horizon) correlations between swap rates across tenors and between basis swap spreads across tenors, respectively, requiring only a mild extension of the calibration criteria. Similarly, we set the CEV coefficient to $\beta = \frac{1}{2}$, but it can of course be calibrated by incorporating other (higher-order) historical moments of basis swap spreads into the calibration criteria.

We require sufficient information to identify $\{\nu_0(t),\, \nu_3\}$. To this end we select an ON-*vs.*-3M basis swap spread of a given tenor τ^{bs}, *e.g.* 2Y, and compute its historical volatility $\hat{\varphi}_{0,3}(\tau^{bs})$. In general one could choose a set of such basis swaps with varying tenors and attempt a bestfit, but we keep matters simple here. We also select a set of observed swaption volatilities, $\hat{\upsilon}_3(\tau_i^{ex}, \tau_i^{sw}, k_i)$, where τ_i^{ex} is the exercise tenor, τ_i^{sw} is the underlying tenor, and k_i is the moneyness. Let $\varphi_{0,3}(\tau^{bs})$ and $\upsilon_3(\tau_i^{ex}, \tau_i^{sw}, k_i)$ denote the model based analogues of our calibration targets.

The primary role of ν_3 will of course be to match the historical basis swap spread volatility, and the primary role of ν_0 will be to bestfit the swaption volatilities. We reiterate however that this is a genuine *joint* calibration in that ν_3 will contribute to model-based swaption volatilities and $\nu_0(t)$ will contribute (mildly) to the model-based basis swap spread volatility. This being the case, the simple calibration strategy described here can be formalized as

$$\begin{aligned} \{\nu_0^*(t), \nu_3^*\} \quad = \quad & \underset{\{\nu_0(t),\nu_3\}}{\operatorname{argmin}} \sum_i w_i \big(\hat{\upsilon}_3(\tau_i^{ex}, \tau_i^{sw}, k_i) - \upsilon_3(\tau_i^{ex}, \tau_i^{sw}, k_i; \nu_0(t), \nu_3)\big)^2 \\ & \text{subject to} \\ & \hat{\varphi}_{0,3}(\tau^{bs}) = \varphi_{0,3}(\tau^{bs}; \nu_0(t), \nu_3)\,, \end{aligned} \tag{23}$$

[7] Though, in Appendix C we demonstrate that for values of $\nu_3(t)$ (or ν_3) which reproduce historical spread volatility, the skew produced is minimal relative to that typically observed in markets.

where w_i are weights. One may then fit a model for 3M IBOR-*vs*.-6M IBOR spreads with reference to their historical volatility, *i.e.* by solving $\hat{\varphi}_{3,6}(\tau^{bs}) = \varphi_{3,6}(\tau^{bs}; \nu_0^*(t), \nu_3^*, \nu_6)$ for ν_6, and so on.

3.3 A Worked Example

We calibrate the model in (8) with a Gaussian $f_0(t,T)$ and a square-root $s_3(t,T)$. Our historical window for basis swap spreads spans 11-Oct-'02 through 11-Oct-'19.[8] We present rolling volatilities for the spreads of 2Y tenor basis swaps in the top-left panel of Figure 1. As can be seen, volatility has only been significant since '07, and thus the historical sample for estimation begins there; the average volatility of the spread over this period was 22.60 bps. We use a market snapshot from 11-Oct-'19 to collect swaption volatilities and interest rate quotes for a complete calibration set.

Our calibrated model successfully reproduces our targeted quantities. Indeed, the simulated basis swap spread volatility was 22.36 bps, and market volatilities were easily fit to. The top-right panel of Figure 1 presents observed normal volatilities for 10Y co-terminal swaptions along with volatilities computed off the calibrated model. The analytical approximations discussed earlier were used for calibration, and thus they agree with market volatilities precisely. The approximation however suffers from inaccuracy owing to the freezing of swap rate coefficients, and thus we also present "true" model-based volatilities produced via simulation; the worst error is roughly 0.1 bps.

The bottom-left panel of Figure 1 presents simulated paths of the 2Y tenor basis swap spread. As desired, spreads generated by the model remained non-negative. Finally, we present exposure profiles (1%, mean, 99%) for a 5Y tenor basis swap with unit notional (paying the ON leg). We leave to future research a more thorough analysis of which specification, in terms of the number of factors, CEV coefficients and correlations structures, offers the best agreement with historical behavior, and which is most effective under a backtest of CVA hedging.

References

[1] Andersen, L. 2010. "Markov Models for Commodity Futures: Theory and Practice", *Quantitative Finance*, 10(8), 831-854.

[2] Andreasen, J. 2001. Turbo Charging the Cheyette Model, *SSRN*.

[3] Benhamou, E., Gobet, E & M. Miri. 2010. "Time-Dependent Heston Model", *SIAM Journal on Financial Mathematics*, 1(1), 289-325.

[4] Cheyette, O. 1992. "Markov Representation of the Heath-Jarrow-Morton Model", *SSRN*.

[5] Chiarella, C. & C. Sklibosios. 2003. "A Class of Jump-Diffusion Bond Pricing Models within the HJM Framework", *Asian-Pacific Financial Markets*, 10, 87-127.

[6] Cox, J., Ingersol, J. & S. Ross. 1985. "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53, 385-407.

[7] Duffie, D., Pan, J. & K. Singleton. 2000. "Transform Analysis & Asset Pricing for Affine Jump-Diffusions", *Econometrica*, 68(6), 1343-1376.

[8] The data used for this analysis was sourced from `Bloomberg`.

[8] Eberlein, E., Gerhart C. & Z. Grbac. 2019. "Multiple-Curve Levy Forward Price Model Allowing for Negative Interest Rates", *Mathematical Finance*.

[9] Grbac, Z. & W. Runggaldier. 2015. "Interest Rate Modeling: Post-Crisis Challenges and Approaches", *Springer*.

[10] Fang, F. & C. Oosterlee. 2008. "A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions", *SIAM Journal on Scientific Computing*, 31(2), 826-848.

[11] Feller, W. 1951. "Two Singular Diffusion Problems", *Annals of Mathematics*, 54, 173-182.

[12] Geelhaar, M., Hofmann, K., Buchel, P. & A. Papapantoleon. 2017. "An HJM-type Stochastic Spread Model: Theory and Implementation", 10^{th} *Bachelier Finance Society Congress*.

[13] FSB, 2018. "Reforming Major Interest Rate Benchmarks: Progress Report", Nov.

[14] Heath, D., Jarrow, R. & A. Morton X. 1992. "Bond Pricing and the Term Structure of Interest Rates", *Econometrica*, 60(1), 77-105.

[15] Henrard, M. 2013. "Multi-Curves Framework with Stochastic Spread: A Coherent Approach to STIR Futures and Their Options", *Open Gamma Quantitative Research*, 11.

[16] Henrard, M. 2018. "Hybrid Model: A Dynamic Multi-Curve Framework", *SSRN*.

[17] Heston, S. 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.

[18] ICE, 2019. "U.S. Dollar ICE Bank Yield Index Update", Jul.

[19] Martinez, T. 2009. "Drift Conditions on a HJM Model with Stochastic Basis Spreads", *Santander Internal Document*, Available at `RiskLab.es`.

[20] Mercurio, F. 2010. A LIBOR Market Model with a Stochastic Basis, *Risk*, Dec, 84-89.

[21] Mercurio, F. & M. Li. 2016. "The Basis Goes Stochastic: A Jump-Diffusion Model for Financial Risk Applications", *SSRN*.

[22] Miglietta, G. 2015. "Topics in Interest Rate Modeling", *PhD Thesis*, University of Padova.

[23] Schrager, D. & A. Pelsser. 2006. "Pricing Swaptions and Coupon Bond Options in Affine Term Structure Models", *Mathematical Finance*, 16(4), 673-694.

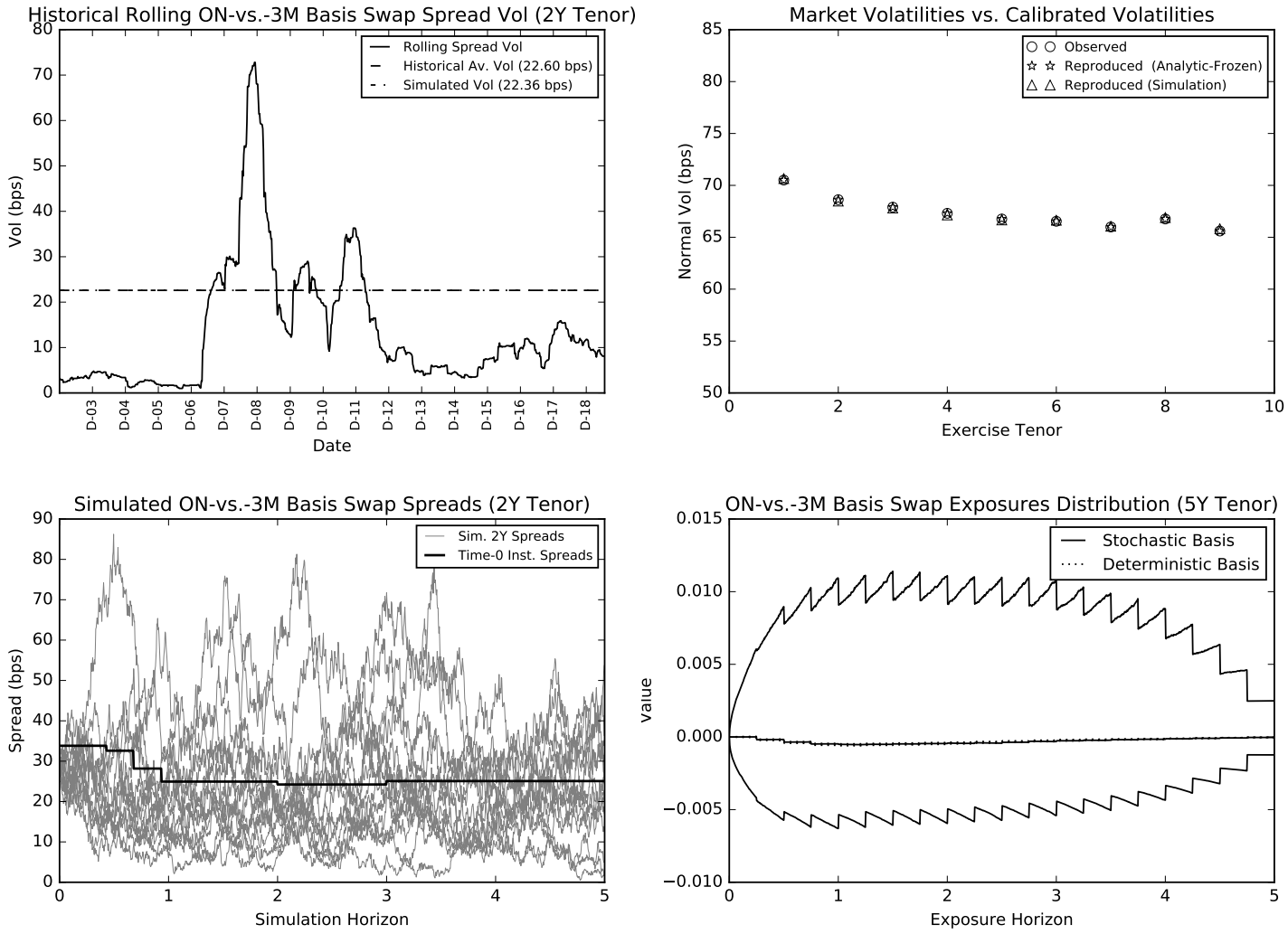

Figure 1:
Top-left panel: Historical rolling ON-*vs*.-3M IBOR tenor basis spread volatilities (2Y tenor), along with the average volatility and the fitted volatility.
Top-right panel: Calibrated model's fit to observed swaption volatilities, for both the analytic frozen-coefficient approximation and a full simulation under the true dynamics.
Bottom-left panel: Simulation of ON-*vs*.-3M IBOR basis swap spreads (2Y tenor) with the initial bootstrapped instantaneous spread curve ($s_3(0,T)$) for reference.
Bottom-right panel: Exposure percentiles for a unit-notional ON-*vs*.-3M IBOR tenor basis swap (5Y tenor) with zero basis volatility ($\nu_3 = 0$) profiles for reference.

A Derivations for the General Framework

In this appendix we consider the general case of N_{cv} spread curves, where the m^{th} spread curve has a diffusion function comprised of loadings onto N_{fc} factors, and where each loading is a composite of N_{sh} exponentials. The n^{th}-tenor IBOR is computed via

$$R_n(t,T) = \frac{1}{\tau_n}\left(\exp\left(\int_T^{T+\tau_n} f_0(t,u) + \sum_{m\in\mathcal{S}_{cv},\, m\leq n} s_m(t,u)\, du\right) - 1\right), \tag{24}$$

and the base curve $f_0(t,T)$ and the spread curves $s_m(t,T)$ evolve according to

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\, dt \; + \sum_{f\in\mathcal{S}_{fc}} \sum_{k\in\mathcal{S}_{sh}} \psi_{0,f,k}(t)\, e^{-\kappa_{0,f,k}(T-t)}\, dW_f(t) \\ ds_m(t,T) &= \mu_m(t,T)\, dt + \sum_{f\in\mathcal{S}_{fc}} \sum_{k\in\mathcal{S}_{sh}} \psi_{m,f,k}(t)\, e^{-\kappa_{m,f,k}(T-t)}\, dW_f(t)\,,\; m\in\mathcal{S}_{cv}\,. \end{aligned} \tag{25}$$

In the above, $\mathcal{S}_{cv}$ is a set of curve indices, *e.g.* $\mathcal{S}_{cv} = \{3, 6, 12, \dots\}$, $\mathcal{S}_{fc}$ is a set of factor indices, and $\mathcal{S}_{sh}$ is a set of shape indices. The dynamics in (25) of course admit a general correlation structure across all curves of interest. It is assumed throughout the coming derivations that $\kappa_{m,f,k} > 0$, $\forall\, m,\, f,\, k$.

Note that (24) is the generalization of (17) presented earlier in Section 2, which we repeat here for reference,

$$R_6(t,T) = \frac{1}{\tau_6}\left(e^{\int_T^{T+\tau_6} f_0(t,u)+s_3(t,u)+s_6(t,u)\, du} - 1\right), \tag{26}$$

It is clear from earlier discussion that $s_3(t,T)$ governs the ON-*vs.*-3M IBOR spread curve, and that $s_6(t,T)$ governs the 3M IBOR-*vs.*-6M IBOR spread curve. As such, one might more formally write these as $s_{0,3}(t,T)$ and $S_{3,6}(t,T)$, respectively. However, this clutters the presentation and we use our original, simpler notation. The interpretation of $s_3(t,T)$ and $s_6(t,T)$ in (26) of course extends to $s_m(t,T)$ in (24).

Recall also from earlier discussion that for the n^{th} IBOR, we choose to integrate all spreads $s_m(t,T)$, $m < n$, over the full coverage period τ_n, as opposed to their individual coverage periods τ_m. The rationale given for this was that we should expect the spot 12M IBOR to be greater than the "average" of the next four forward 3M IBORs, *i.e.* those which span its 12M coverage period. Indeed, if the forward 3M IBOR fixing in nine months' time were particularly high relative to the same forward when computed off the ON curve, it would be natural that the underlying credit and liquidity factors which drive this would manifest also in the spot 12M IBOR. This logic extends also to forward 12M IBORs, and (24) is a convenient way to formalize the relationship.

We will derive the drift restrictions and the Markov representation for the general framework presented above. We first consider the special case of a single spread curve, specifically the ON-*vs.*-3M spread curve $s_3(t,T)$, but we allow for a general diffusion function. We do this to introduce notation and build some intuition for the quantities involved in these calculations, before proceeding to the fully-general case. Note also that we begin with $s_3(t,T)$ because 3 is the smallest index in $\mathcal{S}_{cv}$, and is thus the shortest tenor under consideration. As such, it can be treated in isolation. Naturally, if the ON-*vs.*-1M IBOR spread curve were present, we would begin with $s_1(t,T)$, but as $s_3(t,T)$ was considered throughout the body of the article we begin with it to allow comparison with the simple case analyzed in Section 2.

We will introduce some notational conveniences that will simplify the expressions we encounter. This is necessary as the diffusion function is constructed from a 3-dimensional tensor,

$$\psi_{m,f,k}(t)\, e^{-\kappa_{m,f,k}(T-t)}\,, \tag{27}$$

and we will be computing products and sums involving this tensor. Representing such calculations can become cumbersome and complicate the exposition.

For a vector $X = [X_n]$, we will denote the sum $Y = \sum_n X_n$ as $Y = X_n$, with the summation being implicit in the presence of the n index on the RHS of the equality, but not on the LHS. Similarly, for the matrix $X = [X_{m,n}]$, we denote the vector $[Y_m]$, defined as $Y_m = \sum_n X_{m,n}$, as $Y_m = X_{m,n}$. By this we mean that we only sum over indices which appear on the RHS but do not appear on the LHS; we do not sum over indices on the LHS, rather we take them to denote specific elements of an array. With this convention in mind, we recast (24) and (25) as

$$R_n(t,T) = \frac{1}{\tau_n}\Big(e^{\int_T^{T+\tau_n} f_0(t,u)+1_{m\le n}\, s_m(t,u)\, du} - 1\Big)\,, \tag{28}$$

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\, dt \; + \psi_{0,f,k}(t)\, e^{-\kappa_{0,f,k}(T-t)}\, dW_f(t) \\ ds_m(t,T) &= \mu_m(t,T)\, dt + \psi_{m,f,k}(t)\, e^{-\kappa_{m,f,k}(T-t)}\, dW_f(t)\,,\; m \in \mathcal{S}_{cv}\,. \end{aligned} \tag{29}$$

It will prove useful to employ the compact notation

$$\begin{aligned} \sigma_{0,f,k}(t,T) &= \psi_{0,f,k}(t)\, e^{-\kappa_{0,f,k}(T-t)} \\ \sigma_{m,f,k}(t,T) &= \psi_{m,f,k}(t)\, e^{-\kappa_{m,f,k}(T-t)}\,, \end{aligned} \tag{30}$$

and

$$\begin{aligned} \sigma_{0,f}(t,T) &= \psi_{0,f,k}(t)\, e^{-\kappa_{0,f,k}(T-t)} \\ \sigma_{m,f}(t,T) &= \psi_{m,f,k}(t)\, e^{-\kappa_{m,f,k}(T-t)}\,, \end{aligned} \tag{31}$$

where we point out that the latter are the summed factor loadings, *i.e.* where we have summed over the shapes indexed by k on the RHS, as evidenced by the fact that k does not appear as an index on the LHS.

A.1 General Factor and Shape Specifications for a Single Spread

We begin with an analysis of the base curve $f_0(t,T)$ and the single ON-*vs.*-3M IBOR spread curve $s_3(t,T)$. In the following Appendix we will build upon this, adding *e.g.* the 3M IBOR-*vs.*-6M IBOR spread curve, and so on.

We repeat the formula for $R_3(t,T)$ from (3) and present the general form of the dynamics in (8),

$$R_3(t,T) = \frac{1}{\tau_3}\Big(e^{\int_T^{T+\tau_3} f_0(t,u)+s_3(t,u)\, du} - 1\Big)\,, \tag{32}$$

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\, dt + \psi_{0,f,k}(t)\, e^{-\kappa_{0,f,k}(T-t)}\, dW_{0,f}(t) \\ ds_3(t,T) &= \mu_3(t,T)\, dt + \psi_{3,f,k}(t)\, e^{-\kappa_{3,f,k}(T-t)}\, dW_{3,f}(t)\,. \end{aligned} \tag{33}$$

We consider again a contract paying $R_3(t,T)\,\tau_3$ at $T+\tau_3$. The value remains as per (11),

$$V_3(t,T) = e^{-\int_t^T f_0(t,u)\,du + \int_T^{T+\tau_3} s_3(t,u)\,du} - e^{-\int_t^{T+\tau_3} f_0(t,u)\,du}\,, \tag{34}$$

We require that the growth rate (or expected return) on this contract be equal to $r(t)$, which is satisfied when the growth rates of the two terms involved are both $r(t)$. It is clear that the second term in the above satisfies this as it is simply the value of a discount bond $P(t,T+\tau_3)$, and thus we focus on the first term. To this end, we first let

$$Z_3(t,T) = -\int_t^T f_0(t,u)\,du + \int_T^{T+\tau_3} s_3(t,u)\,du\,, \tag{35}$$

such that the first term of $V_3(t,T)$ is $e^{Z_3(t,T)}$,

$$V_3(t,T) = e^{Z_3(t,T)} - P(t,T+\tau_3)\,. \tag{36}$$

The differential of the first term is straightforwardly

$$de^{Z_3(t,T)} = e^{Z_3(t,T)}\left(dZ_3(t,T) + \frac{1}{2}\,dZ_3(t,T)\cdot dZ_3(t,T)\right). \tag{37}$$

The components $dZ_3(t,T)$ and $dZ_3(t,T)\cdot dZ_3(t,T)$ are easily computed to be

$$\begin{aligned} dZ_3(t,T) &= \left(r(t) - \int_t^T \mu_0(t,u)\,du + \int_T^{T+\tau_3} \mu_3(t,u)\,du\right)dt \\ &\quad + \left(-\int_t^T \sigma_{0,f}(t,u)\,du + \int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\right)dW_f(t)\,, \end{aligned} \tag{38}$$

and

$$\begin{aligned} dZ_3(t,T)\cdot dZ_3(t,T) &= \left(\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\right. \\ &\quad + \left(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\right)\left(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\right) \\ &\quad \left. -2\left(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\right)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\right)dt\,. \end{aligned} \tag{39}$$

Substituting (38) and (39) into (37), and applying the general form of the restriction on $\mu_0(t,T)$ from (10),

$$\int_t^T \mu_0(t,u)\,du = \frac{1}{2}\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right), \tag{40}$$

yields

$$\begin{aligned} \mathbb{E}_t\big[de^{Z_3(t,T)}\big] &= e^{Z_3(t,T)}\left(r(t) + \int_T^{T+\tau_3} \mu_3(t,u)\,du\right. \\ &\quad + \frac{1}{2}\left(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\right)\left(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\right) \\ &\quad \left. -\left(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\right)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\right)dt\,. \end{aligned} \tag{41}$$

To achieve a growth rate of $r(t)$ we clearly require that

$$\begin{aligned}\int_T^{T+\tau_3} \mu_3(t,u)\,du &= -\frac{1}{2}\Big(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\Big)\Big(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\Big) \\ &\quad +\Big(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\Big)\Big(\int_t^{T} \sigma_{0,f}(t,u)\,du\Big)\,.\end{aligned} \tag{42}$$

We next differentiate against T to obtain a difference equation analogous to (13),

$$\begin{aligned}\mu_3(t,T+\tau_3)-\mu_3(t,T) &= -\big(\sigma_{3,f}(t,T+\tau_3)-\sigma_{3,f}(t,T)\big)\Big(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\Big) \\ &\quad +\big(\sigma_{3,f}(t,T+\tau_3)-\sigma_{3,f}(t,T)\big)\Big(\int_t^{T} \sigma_{0,f}(t,u)\,du\Big) \\ &\quad +\Big(\int_T^{T+\tau_3} \sigma_{3,f}(t,u)\,du\Big)\,\sigma_{0,f}(t,T)\,.\end{aligned} \tag{43}$$

At this point it becomes necessary to use the explicit definitions of $\sigma_{0,f}(t,T)$ and $\sigma_{3,f}(t,T)$ in (31), which owes to the need to evaluate the integrals on the RHS of (43). It will prove useful to adopt some more simplifying notation,

$$\begin{aligned}\sigma_{0,f,k}(t,T) &= \psi_{0,f,k}(t)\,K_{0,f,k}(T-t)\,,\ K_{0,f,k}(T-t)=e^{-\kappa_{0,f,k}(T-t)} \\ \sigma_{3,f,k}(t,T) &= \psi_{3,f,k}(t)\,K_{3,f,k}(T-t)\,,\ K_{3,f,k}(T-t)=e^{-\kappa_{3,f,k}(T-t)}\,,\end{aligned} \tag{44}$$

and so on for $\sigma_{6,f}(t,T)$ *etc.*, as will become useful later. In preparation for this, we will use the general notation for the i^{th} curve of interest,

$$\sigma_{i,f,k}(t,T)=\psi_{i,f,k}(t)\,K_{i,f,k}(T-t)\,,\ K_{i,f,k}(T-t)=e^{-\kappa_{i,f,k}(T-t)}\,. \tag{45}$$

The benefits of this notation become clear almost immediately when recasting (43) as

$$\begin{aligned}&\mu_3(t,T+\tau_3)-\mu_3(t,T)= \\ &-\psi_{3,f,k}(t)\,\big(K_{3,f,k}(T+\tau_3-t)-K_{3,f,k}(T-t)\big)\,\psi_{3,f,l}(t)\int_T^{T+\tau_3} K_{3,f,l}(u-t)\,du \\ &+\psi_{3,f,k}(t)\,\big(K_{3,f,k}(T+\tau_3-t)-K_{3,f,k}(T-t)\big)\,\psi_{0,f,l}(t)\int_t^{T} K_{0,f,l}(u-t)\,du \\ &+\psi_{3,f,k}(t)\Big(\int_T^{T+\tau_3} K_{3,f,k}(u-t)\,du\Big)\,\psi_{0,f,l}(t)\,K_{0,f,l}(T-t)\,.\end{aligned} \tag{46}$$

We now note that the remaining integrals evaluate to

$$\begin{aligned}\int_T^{T+\tau_3} K_{3,f,k}(u-t)\,du &= \frac{K_{3,f,k}(T+\tau_3-t)-K_{3,f,k}(T-t)}{-\kappa_{3,f,k}}=\frac{K_{3,f,k}(T-t)\big(K_{3,f,k}(\tau_3)-1\big)}{-\kappa_{3,f,k}} \\ \int_t^{T} K_{0,f,l}(u-t)\,du &= \frac{K_{0,f,l}(T-t)-1}{-\kappa_{0,f,l}}\,,\end{aligned} \tag{47}$$

where we have used that fact that

$$K_{3,f,k}(T+\tau_3-t)-K_{3,f,k}(T-t) = e^{-\kappa_{3,f,k}(T+\tau_3-t)}-e^{-\kappa_{3,f,k}(T-t)} = K_{3,f,k}(T-t)\big(K_{3,f,k}(\tau_3)-1\big)\,. \quad (48)$$

With these observations (46) becomes

$$\begin{aligned}
&\mu_3(t,T+\tau_3)-\mu_3(t,T) = \\
&-\psi_{3,f,k}(t)K_{3,f,k}(T-t)\big(K_{3,f,k}(\tau_3)-1\big)\,\psi_{3,f,l}(t)\frac{K_{3,f,l}(T-t)\big(K_{3,f,l}(\tau_3)-1\big)}{-\kappa_{3,f,l}} \\
&+\psi_{3,f,k}(t)K_{3,f,k}(T-t)\big(K_{3,f,k}(\tau_3)-1\big)\,\psi_{0,f,l}(t)\frac{\big(K_{0,f,l}(T-t)-1\big)}{-\kappa_{0,f,l}} \\
&+\psi_{3,f,k}(t)\frac{K_{3,f,k}(T-t)\big(K_{3,f,k}(\tau_3)-1\big)}{-\kappa_{3,f,k}}\,\psi_{0,f,l}(t)\,K_{0,f,l}(T-t)\,.
\end{aligned} \quad (49)$$

At this point we are gaining some structure, with there being a mixture of quantities depending on t and quantities depending on exponentials in $T-t$. A similar structure will indeed be observed when we inspect the general case of N_{cv} spread curves. To capitalize on this, we adopt further notational conveniences for commonly-appearing products,

$$\begin{aligned}
\bar{\psi}_{i,j,f,k,l}(t) &= \psi_{i,f,k}(t)\,\psi_{j,f,l}(t) \\
\bar{K}_{i,j,f,k,l}(T-t) &= K_{i,f,k}(T-t)\,K_{j,f,l}(T-t)\,.
\end{aligned} \quad (50)$$

With this, (49) can be rewritten as

$$\begin{aligned}
&\mu_3(t,T+\tau_3)-\mu_3(t,T) = \\
&-\bar{\psi}_{3,3,f,k,l}(t)\frac{\big(K_{3,f,k}(\tau_3)-1\big)\big(K_{3,f,l}(\tau_3)-1\big)}{-\kappa_{3,f,l}}\,\bar{K}_{3,3,f,k,l}(T-t) \\
&+\bar{\psi}_{3,0,f,k,l}(t)\frac{\big(K_{3,f,k}(\tau_3)-1\big)}{-\kappa_{0,f,l}}\,\big(\bar{K}_{3,0,f,k,l}(T-t)-K_{3,f,k}(T-t)\big) \\
&+\bar{\psi}_{3,0,f,k,l}(t)\frac{\big(K_{3,f,k}(\tau_3)-1\big)}{-\kappa_{3,f,k}}\,\bar{K}_{3,0,f,k,l}(T-t)\,.
\end{aligned} \quad (51)$$

This can be written down compactly as

$$\begin{aligned}
&\mu_3(t,T+\tau_3)-\mu_3(t,T) = \\
&\bar{\psi}_{3,3,f,k,l}(t)\,C_{3,3,3,f,k,l}\,\bar{K}_{3,3,f,k,l}(T-t) \\
&+\bar{\psi}_{3,0,f,k,l}(t)\,C_{3,3,0,f,k,l}\,\bar{K}_{3,0,f,k,l}(T-t) \\
&+\bar{\psi}_{3,0,f,k,l}(t)\,D_{3,3,f,k,l}\,K_{3,f,k}(T-t)\,,
\end{aligned} \quad (52)$$

where the constants are defined as

$$\begin{aligned} C_{3,3,3,3,f,k,l} &= \frac{\big(K_{3,f,k}(\tau_3)-1\big)\big(K_{3,f,l}(\tau_3)-1\big)}{\kappa_{3,f,l}} \\ C_{3,3,3,0,f,k,l} &= -\big(K_{3,f,k}(\tau_3)-1\big)\left(\frac{1}{\kappa_{0,f,l}}+\frac{1}{\kappa_{3,f,k}}\right) \\ D_{3,3,3,f,k,l} &= \frac{\big(K_{3,f,k}(\tau_3)-1\big)}{\kappa_{0,f,l}}\,. \end{aligned} \tag{53}$$

One thing to note at this point, is that the constants have one more curve index than is strictly necessary. For example, $C_{3,3,3,3,f,k,l}$, which loads onto $\bar{\psi}_{3,3,f,k,l}(t)\,\bar{K}_{3,3,f,k,l}(T-t)$, may have been denoted $C_{3,3,f,k,l}$. The additional curve index owes to the fact that we are exploring the drift restriction for the $s_3(t,T)$ curve, and thus it is τ_3 which appears in the definitions of the constants. In the next section, we shall encounter similar constants multiplying with $\bar{\psi}_{3,3,f,k,l}(t)\,\bar{K}_{3,3,f,k,l}(T-t)$, but with *e.g.* τ_6 appearing in the definition, and we will thus require a way to distinguish between them. General notation (indices) will be presented in the following section as necessary. We note also that $D_{3,3,f,k,l}$ has no 0 subscript, as it will be the case that whenever $K_{3,f,k}(T-t)$ appears, it will do so multiplicatively with $\bar{\psi}_{3,0,f,k,l}(t)$, and thus the 0 subscript would be redundant.

Given (52), we follow the intuition in Section 2.2 and postulate an ansatz,

$$\begin{aligned} &\mu_3(t,T) = \\ &\quad \bar{\psi}_{3,3,f,k,l}(t)\,A_{3,3,3,f,k,l}\,\bar{K}_{3,3,f,k,l}(T-t) \\ &\quad + \bar{\psi}_{3,0,f,k,l}(t)\,A_{3,3,0,f,k,l}\,\bar{K}_{3,0,f,k,l}(T-t) \\ &\quad + \bar{\psi}_{3,0,f,k,l}(t)\,B_{3,3,f,k,l}\,K_{3,f,k}(T-t)\,. \end{aligned} \tag{54}$$

Substituting into (52) and solving for the coefficients by collecting like terms yields coefficients

$$\begin{aligned} A_{3,3,3,f,k,l} &= \frac{\big(K_{3,f,k}(\tau_3)-1\big)\big(K_{3,f,l}(\tau_3)-1\big)}{\kappa_{3,f,l}\,\big(\bar{K}_{3,3,f,k,l}(\tau_3)-1\big)} \\ A_{3,3,0,f,k,l} &= -\frac{\big(K_{3,f,k}(\tau_3)-1\big)}{\big(\bar{K}_{3,0,f,k,l}(\tau_3)-1\big)}\left(\frac{1}{\kappa_{0,f,l}}+\frac{1}{\kappa_{3,f,k}}\right) \\ B_{3,3,f,k,l} &= \frac{1}{\kappa_{0,f,l}}\,. \end{aligned} \tag{55}$$

It is readily verifiable that this solution satisfies the sufficient condition in (42), and thus we have the drift function for the dynamics in (33). We note that there is an obvious symmetry that may be taken advantage of,

$$\bar{\psi}_{3,3,f,k,l}(t)\,\bar{K}_{3,3,f,k,l}(T-t) = \bar{\psi}_{3,3,f,l,k}(t)\,\bar{K}_{3,3,f,l,k}(T-t)\,, \tag{56}$$

but leave discussion of this to the end of the next appendix, where a more general symmetry will appear for the case of multiple spread curves.

It will be useful in what follows to recast (54) as

$$\mu_3(t,T) = \bar{\psi}_{3,j,f,k,l}(t)\,A_{3,3,j,f,k,l}\,\bar{K}_{3,j,f,k,l}(T-t) + \bar{\psi}_{3,0,f,k,l}(t)\,B_{3,3,f,k,l}\,K_{3,f,k}(T-t)\,, \tag{57}$$

where the sum over j involves indices $j \in \{0, 3\}$. We will use essentially the same representation in the following section when handling the fully-general case of N_{cv} spread curves.

We turn now to the matter of a Markov representation. To this end, we write $s_3(t,T)$ as

$$\begin{aligned} s_3(t,T) &= s_3(0,T) + \int_0^t \mu_3(v,T)\,dv + \int_0^t \sigma_{3,f}(v,T)\,dW_f(v) \\ &= s_3(0,T) + A_{3,3,j,f,k,l}\,\bar{K}_{3,j,f,k,l}(T-t)\int_0^t \bar{\psi}_{3,j,f,k,l}(v)\,\bar{K}_{3,j,f,k,l}(t-v)\,dv \\ &\quad + B_{3,3,f,k,l}\,K_{3,f,k}(T-t)\int_0^t \bar{\psi}_{3,0,f,k,l}(v)\,K_{3,f,k}(t-v)\,dv \\ &\quad + K_{3,f,k}(T-t)\int_0^t \psi_{3,f,k}(v)\,K_{3,f,k}(t-v)\,dW_f(v)\,. \end{aligned} \tag{58}$$

From this we infer the "short" spread $s_3(t,t)$,

$$\begin{aligned} s_3(t,t) &= s_3(0,t) + A_{3,3,j,f,k,l}\int_0^t \bar{\psi}_{3,j,f,k,l}(v)\,\bar{K}_{3,j,f,k,l}(t-v)\,dv \\ &\quad + B_{3,3,f,k,l}\int_0^t \bar{\psi}_{3,0,f,k,l}(v)\,K_{3,f,k}(t-v)\,dv \\ &\quad + \int_0^t \psi_{3,f,k}(v)\,K_{3,f,k}(t-v)\,dW_f(v)\,. \end{aligned} \tag{59}$$

This allows us to assemble the "primary" Markov state processes $X_{3,f,k}(t)$,

$$\begin{aligned} X_{3,f,k}(t) &= A_{3,3,j,f,k,l}\int_0^t \bar{\psi}_{3,j,f,k,l}(v)\,\bar{K}_{3,j,f,k,l}(t-v)\,dv \\ &\quad + B_{3,3,f,k,l}\int_0^t \bar{\psi}_{3,0,f,k,l}(v)\,K_{3,f,k}(t-v)\,dv \\ &\quad + \int_0^t \psi_{3,f,k}(v)\,K_{3,f,k}(t-v)\,dW_f(v)\,, \end{aligned} \tag{60}$$

and we define the "auxiliary" Markov state processes $Y_{3,j,f,k,l}(t)$,

$$Y_{3,j,f,k,l}(t) = \int_0^t \bar{\psi}_{3,j,f,k,l}(v)\,\bar{K}_{3,j,f,k,l}(t-v)\,dv\,. \tag{61}$$

With these definitions, we are able to recast $s_3(t,t)$ from (59) as $s_3(t,t) = s_3(0,t) + X_{3,f,k}(t)$, and $s_3(t,T)$ from (58) as

$$\begin{aligned} s_3(t,T) &= s_3(0,T) + K_{3,f,k}(T-t)\,X_{3,f,k}(t) \\ &\quad + A_{3,3,j,f,k,l}\left(\bar{K}_{3,j,f,k,l}(T-t) - K_{3,f,k}(T-t)\right)Y_{3,j,f,k,l}(t)\,. \end{aligned} \tag{62}$$

The last remaining step is to write down the dynamics of the Markov processes,

$$\begin{aligned} dX_{3,f,k}(t) &= \kappa_{3,f,k}\bigg(\frac{A_{3,3,j,f,k,l}}{\kappa_{3,f,k}}\,\bar{\psi}_{3,j,f,k,l}(t) + \frac{B_{3,3,f,k,l}}{\kappa_{3,f,k}}\,\bar{\psi}_{3,0,f,k,l}(t) \\ &\quad -\frac{\kappa_{j,f,l}\,A_{3,3,j,f,k,l}}{\kappa_{3,f,k}}\,Y_{3,j,f,k,l}(t) - X_{3,f,k}(t)\bigg)\,dt + \psi_{3,f,k}(t)\,dW_f(t) \\ dY_{3,j,f,k,l}(t) &= (\kappa_{3,f,k}+\kappa_{j,f,l})\bigg(\frac{1}{(\kappa_{3,f,k}+\kappa_{j,f,l})}\,\bar{\psi}_{3,j,f,k,l}(t) - Y_{3,j,f,k,l}(t)\bigg)\,dt\,. \end{aligned} \tag{63}$$

As a final note, collapsing our general-factor general-shape model (25) of $f_0(t,T)$ and $s_3(t,T)$ to the workhorse model in (8) of Section 2.2 and evaluating (54) and (62) yields the Markov representation provided in (19) and (20) in Section 2.3, where the A_3 defined there plays the role of the $A_{3,3,3,1,1}$ defined here.

A.2 General Factor and Shape Specifications for Many Spreads

We now turn to the general setup where N_{cv} spread curves are necessary. To begin, we rewrite the formulation from (28) and (29) for ease of reference,

$$R_n(t,T) = \frac{1}{\tau_n}\Big(e^{\int_T^{T+\tau_n} f_0(t,u)+1_{m\le n}\,s_m(t,u)\,du} - 1\Big)\,, \tag{64}$$

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\,dt\; + \psi_{0,f,k}(t)e^{-\kappa_{0,f,k}(T-t)}\,dW_f(t) \\ ds_m(t,T) &= \mu_m(t,T)\,dt + \psi_{m,f,k}(t)e^{-\kappa_{m,f,k}(T-t)}\,dW_f(t)\,,\; m\in\mathcal{S}_{cv}\,. \end{aligned} \tag{65}$$

We will focus on deriving the dynamics of $s_n(t,T)$, assuming the dynamics of $s_m(t,T)$, $m<n$, are already available. As we shall demonstrate, there is a recurrence relation between drift functions across curves. This allows us to build the dynamics of the complete set of curves in a simple recursive fashion.

To proceed, we again seek a contract referencing the τ_n-tenor IBOR so as to derive a restriction on $\mu_n(t,T)$. We select a contract paying $R_n(t,T)\,\tau_n$ at $T+\tau_n$, and we denote its value by $V_n(t,T)$. Similar to the case for $V_3(t,T)$, we may represent the value as

$$V_n(t,T) = e^{Z_n(t,T)} - P(t,T+\tau_n)\,, \tag{66}$$

$$Z_n(t,T) = -\int_t^T f_0(t,u)\,du + \int_T^{T+\tau_n} 1_{m\le n}\,s_m(t,u)\,du\,. \tag{67}$$

To ease the exposition, we will drop $1_{m\le n}$ from further calculations, with the understanding that all sums over the curve index m will be over $m\in\mathcal{S}_{cv}$, $m\le n$. We will be careful to enforce this in later calculations, referencing it explicitly where necessary. With this we have

$$Z_n(t,T) = -\int_t^T f_0(t,u)\,du + \int_T^{T+\tau_n} s_m(t,u)\,du\,. \tag{68}$$

As in the previous appendix, we require that the first term of (66) grows on average at $r(t)$. Naturally,

$$de^{Z_n(t,T)} = e^{Z_n(t,T)}\bigg(dZ_n(t,T) + \frac{1}{2}\,dZ_n(t,T)\cdot dZ_n(t,T)\bigg)\,, \tag{69}$$

and

$$
\begin{aligned}
dZ_n(t,T) &= \left(r(t) - \int_t^T \mu_0(t,u)\,du + \int_T^{T+\tau_n} \mu_m(t,u)\,du\right) dt \\
&\quad + \left(-\int_t^T \sigma_{0,f}(t,u)\,du + \int_T^{T+\tau_n} \sigma_{m,f}(t,u)\,du\right) dW_f(t)\,, \\
dZ_n(t,T)\cdot dZ_n(t,T) &= \left(\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\right. \\
&\quad + \left(\int_T^{T+\tau_n} \sigma_{m,f}(t,u)\,du\right)\left(\int_T^{T+\tau_n} \sigma_{p,f}(t,u)\,du\right) \\
&\quad \left. -2\left(\int_T^{T+\tau_n} \sigma_{m,f}(t,u)\,du\right)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\right) dt\,.
\end{aligned}
\tag{70}
$$

Note that in computing $dZ_n(t,T)\cdot dZ_n(t,T)$ we are taking an inner product of $\int_T^{T+\tau_n} \sigma_{m,f}(t,u)\,du$ with itself, and this quantity is itself a sum over the spread curve index m. As such we are forced to introduce the spread curve index p when writing the product. Naturally, the sum over p is over $p \in \mathcal{S}_{cv}$, $p \le n$, as it is for m.

Proceeding as per the previous appendix and equating the drift of (69) with $r(t)$, we arrive at the drift restriction

$$
\begin{aligned}
&\int_T^{T+\tau_n} \mu_n(t,u)\,du = \\
&\quad -\int_T^{T+\tau_n} \mu_m(t,u)\,du && {\scriptstyle 0<m<n} \\
&\quad -\frac{1}{2}\left(\int_T^{T+\tau_n} \sigma_{m,f}(t,u)\,du\right)\left(\int_T^{T+\tau_n} \sigma_{p,f}(t,u)\,du\right) && {\scriptstyle 0<m\le n,\, 0<p\le n} \\
&\quad +\left(\int_T^{T+\tau_n} \sigma_{m,f}(t,u)\,du\right)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right) && {\scriptstyle 0<m\le n}\,.
\end{aligned}
\tag{71}
$$

We note that we have removed $\int_T^{T+\tau_n} \mu_n(t,u)\,du$ from $\int_T^{T+\tau_n} \mu_m(t,u)\,du$ and taken it to the LHS to isolate it as we seek $\mu_n(t,T)$. As such, $\int_T^{T+\tau_n} \mu_m(t,u)\,du$ on the RHS is a sum over $m \in \mathcal{S}_{cv}$, $m < n$, as opposed to $m \le n$. The diffusion components are involve sums which remain over $m, p \in \mathcal{S}_{cv}$, $m, p \le n$. For ease of reference, and to avoid confusion, in the above we have written out the ranges of the sums involved in each term. We will repeat this where useful in subsequent expressions.

Next we differentiate against T, as was done earlier in the case of $s_3(t,T)$. This yields the

restriction

$$
\begin{aligned}
&\mu_n(t,T+\tau_n)-\mu_n(t,T)= \\
&\quad -\big(\mu_m(t,T+\tau_n)-\mu_m(t,T)\big) && 0<m<n \\
&\quad -\big(\sigma_{m,f}(t,T+\tau_n)-\sigma_{m,f}(t,T)\big)\left(\int_T^{T+\tau_n}\sigma_{p,f}(t,u)\,du\right) && 0<m\le n,\,0<p\le n \\
&\quad +\big(\sigma_{m,f}(t,T+\tau_n)-\sigma_{m,f}(t,T)\big)\left(\int_t^{T}\sigma_{0,f}(t,u)\,du\right) && 0<m\le n \\
&\quad +\left(\int_T^{T+\tau_n}\sigma_{m,f}(t,u)\,du\right)\sigma_{0,f}(t,T) && 0<m\le n\,.
\end{aligned}
\tag{72}
$$

There is a clear analogy between (72) and (43) from the previous appendix, but in this case we must evaluate the sum $\mu_m(t,T+\tau_3)-\mu_m(t,T)$. As noted earlier, we assume that the $\mu_m(t,T)$ are available already and derive $\mu_n(t,T)$, demonstrating a recursive relationship between these drift functions, and more generally between the dynamics of the curves.

First, we postulate that the $\mu_m(t,T)$ have a form along the lines of $\mu_3(t,T)$ in (57),

$$
\mu_m(t,T)=\bar{\psi}_{i,j,f,k,l}(t)\,A_{m,i,j,f,k,l}\,\bar{K}_{i,j,f,k,l}(T-t)+\bar{\psi}_{i,0,f,k,l}(t)\,B_{m,i,f,k,l}\,K_{i,f,k}(T-t)\,,
\tag{73}
$$

where the curve indices on the RHS sum over $i,j\in\mathcal{S}_{cv}$, $0<i\le m$, $0\le j\le m$. We shall see that if this is true for all $\mu_m(t,T)$, $m<n$, then it is true also $\mu_n(t,T)$. Thus, given that we know it is true of $\mu_3(t,T)$, the form in (73) holds by induction. In demonstrating this, we will also arrive at recursion relations between $A_{m,i,j,f,k,l}$ and $A_{p,i,j,f,k,l}$, $p<m$, and $B_{m,i,f,k,l}$ and $B_{p,i,f,k,l}$, $p<m$. Substituting the form of (73) into (72), and invoking the definitions of the diffusion functions we arrive at

$$
\begin{aligned}
&\mu_n(t,T+\tau_n)-\mu_n(t,T)= \\
&\quad -\bar{\psi}_{i,j,f,k,l}(t)\,A_{m,i,j,f,k,l}\,\bar{K}_{i,j,f,k,l}(T-t)\left(\bar{K}_{i,j,f,k,l}(\tau_n)-1\right) && 0<m<n,\,0<i<n,\,0\le j<n \\
&\quad -\bar{\psi}_{i,0,f,k,l}(t)\,B_{m,i,f,k,l}\,K_{i,f,k}(T-t)\left(K_{i,f,k}(\tau_n)-1\right) && 0<m<n,\,0<i<n \\
&\quad -\psi_{m,f,k}(t)K_{m,f,k}(T-t)\big(K_{m,f,k}(\tau_n)-1\big)\,\psi_{p,f,l}(t)\frac{K_{p,f,l}(T-t)\big(K_{p,f,l}(\tau_n)-1\big)}{-\kappa_{p,f,l}} && 0<m\le n,\,0<p\le n \\
&\quad +\psi_{m,f,k}(t)K_{m,f,k}(T-t)\big(K_{m,f,k}(\tau_n)-1\big)\,\psi_{0,f,l}(t)\frac{\big(K_{0,f,l}(T-t)-1\big)}{-\kappa_{0,f,l}} && 0<m\le n \\
&\quad +\psi_{m,f,k}(t)\frac{K_{m,f,k}(T-t)\big(K_{m,f,k}(\tau_n)-1\big)}{-\kappa_{m,f,k}}\,\psi_{0,f,l}(t)\,K_{0,f,l}(T-t) && 0<m\le n\,.
\end{aligned}
\tag{74}
$$

Next we note that the drift-related terms contain exponential factors $K_{i,f,k}(T-t)$ and $K_{j,f,l}(T-t)$ (implicitly in $\bar{K}_{i,j,f,k,l}(T-t)$), while the diffusion-related terms contain exponential factors $K_{m,f,k}(T-t)$ and $K_{p,f,l}(T-t)$. These are of course the same factors, but indexed differently across the separate terms under consideration. As we will ultimately seek to collect like terms in these exponential factors, we simply convert the m, p indices in the diffusion-related terms to i, j

indices, taking care to maintain the ranges of summation for each term,

$$
\begin{aligned}
&\mu_n(t,T+\tau_n)-\mu_n(t,T)= \\
&-\bar{\psi}_{i,j,f,k,l}(t)\,A_{m,i,j,f,k,l}\left(\bar{K}_{i,j,f,k,l}(\tau_n)-1\right)\bar{K}_{i,j,f,k,l}(T-t) && 0<m<n,\,0<i<n,\,0\le j<n \\
&-\bar{\psi}_{i,0,f,k,l}(t)\,B_{m,i,f,k,l}\left(K_{i,f,k}(\tau_n)-1\right)K_{i,f,k}(T-t) && 0<m<n,\,0<i<n \\
&-\psi_{i,f,k}(t)K_{i,f,k}(T-t)\big(K_{i,f,k}(\tau_n)-1\big)\,\psi_{j,f,l}(t)\frac{K_{j,f,l}(T-t)\big(K_{j,f,l}(\tau_n)-1\big)}{-\kappa_{j,f,l}} && 0<i\le n,\,0<j\le n \\
&+\psi_{i,f,k}(t)K_{i,f,k}(T-t)\big(K_{i,f,k}(\tau_n)-1\big)\,\psi_{0,f,l}(t)\frac{\big(K_{0,f,l}(T-t)-1\big)}{-\kappa_{0,f,l}} && 0<i\le n \\
&+\psi_{i,f,k}(t)\frac{K_{i,f,k}(T-t)\big(K_{i,f,k}(\tau_n)-1\big)}{-\kappa_{i,f,k}}\,\psi_{0,f,l}(t)\,K_{0,f,l}(T-t) && 0<i\le n\,.
\end{aligned}
\tag{75}
$$

We now rearrange this to collect terms in the exponential factors $\bar{K}_{i,j,f,k,l}(T-t)$ and $K_{i,f,k}(T-t)$,

$$
\begin{aligned}
\mu_n(t,T+\tau_n)-\mu_n(t,T) \;=\; & \bar{\psi}_{i,j,f,k,l}(t)\,C_{n,i,j,f,k,l}\,\bar{K}_{i,j,f,k,l}(T-t) \\
& +\bar{\psi}_{i,0,f,k,l}(t)\,D_{n,i,f,k,l}\,K_{i,f,k}(T-t)\,,
\end{aligned}
\tag{76}
$$

where

$$
\begin{aligned}
C_{n,i,j,f,k,l} &= \frac{\big(K_{i,f,k}(\tau_n)-1\big)\big(K_{j,f,l}(\tau_n)-1\big)}{\kappa_{j,f,l}} && 0<i<n\,,\ 0<j<n, \\
&\quad -A_{m,i,j,f,k,l}\left(\bar{K}_{i,j,f,k,l}(\tau_n)-1\right), && \\
C_{n,i,j,f,k,l} &= \frac{\big(K_{i,f,k}(\tau_n)-1\big)\big(K_{j,f,l}(\tau_n)-1\big)}{\kappa_{j,f,l}}\,, && 0<i\le n\,,\ j=n\,, \\
& && i=n\,,\ 0<j\le n\,, \\
C_{n,i,0,f,k,l} &= -\big(K_{i,f,k}(\tau_n)-1\big)\left(\frac{1}{\kappa_{i,f,k}}+\frac{1}{\kappa_{0,f,l}}\right) && 0<i<n, \\
&\quad -A_{m,i,0,f,k,l}\left(\bar{K}_{i,0,f,k,l}(\tau_n)-1\right), && \\
C_{n,n,0,f,k,l} &= -\big(K_{n,f,k}(\tau_n)-1\big)\left(\frac{1}{\kappa_{n,f,k}}+\frac{1}{\kappa_{0,f,l}}\right), && \\
D_{n,i,f,k,l} &= \frac{\big(K_{i,f,k}(\tau_n)-1\big)}{\kappa_{0,f,l}}-B_{m,i,f,k,l}\left(K_{i,f,k}(\tau_n)-1\right), && 0<i<n\,, \\
D_{n,n,f,k,l} &= \frac{\big(K_{n,f,k}(\tau_n)-1\big)}{\kappa_{0,f,l}}\,, &&
\end{aligned}
\tag{77}
$$

where the sums over m in the above are over $m<n$, as in (73). We now substitute the form in (73) for $\mu_n(t,T)$, equate like terms in the exponential factors, and solve for $A_{n,i,j,f,k,l}$ and $B_{n,i,f,k,l}$. This yields

$$
\begin{aligned}
A_{n,i,j,f,k,l} &= \frac{C_{n,i,j,f,k,l}}{\big(\bar{K}_{i,j,f,k,l}(\tau_n)-1\big)} \\
B_{n,i,f,k,l} &= \frac{D_{n,i,f,k,l}}{\big(K_{i,f,k}(\tau_n)-1\big)}\,,
\end{aligned}
\tag{78}
$$

or more explicitly,

$$\begin{aligned}
A_{n,i,j,f,k,l} &= \frac{\big(K_{i,f,k}(\tau_n)-1\big)\big(K_{j,f,l}(\tau_n)-1\big)}{\kappa_{j,f,l}\left(\bar{K}_{i,j,f,k,l}(\tau_n)-1\right)} - A_{m,i,j,f,k,l}\,, && 0<i<n\,,\ 0<j<n\,,\\
A_{n,i,j,f,k,l} &= \frac{\big(K_{i,f,k}(\tau_n)-1\big)\big(K_{j,f,l}(\tau_n)-1\big)}{\kappa_{j,f,l}\left(\bar{K}_{i,j,f,k,l}(\tau_n)-1\right)}\,, && 0<i\le n\,,\ j=n\,,\\
& && i=n\,,\ 0<j\le n\,,\\
A_{n,i,0,f,k,l} &= -\frac{\big(K_{i,f,k}(\tau_n)-1\big)}{\big(\bar{K}_{i,0,f,k,l}(\tau_n)-1\big)}\left(\frac{1}{\kappa_{i,f,k}}+\frac{1}{\kappa_{0,f,l}}\right) - A_{m,i,0,f,k,l}\,, && 0<i<n\,,\\
A_{n,n,0,f,k,l} &= -\frac{\big(K_{n,f,k}(\tau_n)-1\big)}{\big(\bar{K}_{n,0,f,k,l}(\tau_n)-1\big)}\left(\frac{1}{\kappa_{n,f,k}}+\frac{1}{\kappa_{0,f,l}}\right),\\
B_{n,i,f,k,l} &= \frac{1}{\kappa_{0,f,l}} - B_{m,i,f,k,l}\,, && 0<i<n\,,\\
B_{n,n,f,k,l} &= \frac{1}{\kappa_{0,f,l}}\,.
\end{aligned} \tag{79}$$

Note that there is a clear recursive relation in the coefficients, *i.e.* $A_{n,i,j,f,k,l}$ depends on $A_{m,i,j,f,k,l}$, $m<n$, for the $\{0<i<n\,,\ 0<j<n\}$ case. We can actually simplify this further by noting that successive values of these loadings will cancel out, leaving only part of $A_{n-1,i,j,f,k,l}$ relevant to $A_{n,i,j,f,k,l}$.[9] To see this, note that for $A_{n,i,j,f,k,l}$ in the $\{0<i<n\,,\ 0<j<n\}$ case,

$$A_{n,i,j,f,k,l} = G_{n,i,j,f,k,l} - A_{m,i,j,f,k,l}\,,\ {}_{m<n}\,, \tag{80}$$

where

$$G_{n,i,j,f,k,l} = \frac{\big(K_{i,f,k}(\tau_n)-1\big)\big(K_{j,f,l}(\tau_n)-1\big)}{\kappa_{j,f,l}\left(\bar{K}_{i,j,f,k,l}(\tau_n)-1\right)}\,. \tag{81}$$

Using this recursion, we have that

$$\begin{aligned}
A_{n,i,j,f,k,l} &= G_{n,i,j,f,k,l} - \big(G_{n-1,i,j,f,k,l} - A_{m,i,j,f,k,l}\big) - A_{m,i,j,f,k,l}\,,\ {}_{m<n-1}\\
&= G_{n,i,j,f,k,l} - G_{n-1,i,j,f,k,l}\,.
\end{aligned} \tag{82}$$

Following similar logic for $A_{n,i,0,f,k,l}$ in the $\{0<i<n\}$ case,

$$A_{n,i,0,f,k,l} = G_{n,i,0,f,k,l} - G_{n-1,i,0,f,k,l}\,, \tag{83}$$

where

$$G_{n,i,0,f,k,l} = -\frac{\big(K_{i,f,k}(\tau_n)-1\big)}{\big(\bar{K}_{i,0,f,k,l}(\tau_n)-1\big)}\left(\frac{1}{\kappa_{i,f,k}}+\frac{1}{\kappa_{0,f,l}}\right), \tag{84}$$

and finally for $B_{n,i,f,k,l}$ in the $\{0<i<n\}$ case, similar logic yields

$$B_{n,i,f,k,l} = 0\,. \tag{85}$$

[9] We abuse notation here in that $n\in\{3,\,6,\,12,\,\dots\}$, and thus $n-1$ and n really refer to an element of this set and the next largest element of this set, *i.e.* 3 and 6, or 6 and 12. We use $n-1$ and n to avoid complex notation in the following arguments.

With this we can recast the drift loadings in (79) as

$$
\begin{aligned}
A_{n,i,j,f,k,l} &= \frac{\big(K_{i,f,k}(\tau_n)-1\big)\big(K_{j,f,l}(\tau_n)-1\big)}{\kappa_{j,f,l}\left(\bar{K}_{i,j,f,k,l}(\tau_n)-1\right)} && 0<i<n\,,\ 0<j<n,\\
&\quad -\frac{\big(K_{i,f,k}(\tau_{n-1})-1\big)\big(K_{j,f,l}(\tau_{n-1})-1\big)}{\kappa_{j,f,l}\left(\bar{K}_{i,j,f,k,l}(\tau_{n-1})-1\right)}\,,\\
A_{n,i,j,f,k,l} &= \frac{\big(K_{i,f,k}(\tau_n)-1\big)\big(K_{j,f,l}(\tau_n)-1\big)}{\kappa_{j,f,l}\left(\bar{K}_{i,j,f,k,l}(\tau_n)-1\right)}\,, && 0<i\le n\,,\ j=n\,,\\
& && i=n\,,\ 0<j\le n\,,\\
A_{n,i,0,f,k,l} &= -\frac{\big(K_{i,f,k}(\tau_n)-1\big)}{\big(\bar{K}_{i,0,f,k,l}(\tau_n)-1\big)}\left(\frac{1}{\kappa_{i,f,k}}+\frac{1}{\kappa_{0,f,l}}\right) && 0<i<n\,,\\
&\quad +\frac{\big(K_{i,f,k}(\tau_{n-1})-1\big)}{\big(\bar{K}_{i,0,f,k,l}(\tau_{n-1})-1\big)}\left(\frac{1}{\kappa_{i,f,k}}+\frac{1}{\kappa_{0,f,l}}\right),\\
A_{n,n,0,f,k,l} &= -\frac{\big(K_{n,f,k}(\tau_n)-1\big)}{\big(\bar{K}_{n,0,f,k,l}(\tau_n)-1\big)}\left(\frac{1}{\kappa_{n,f,k}}+\frac{1}{\kappa_{0,f,l}}\right),\\
B_{n,i,f,k,l} &= 0\,,\\
B_{n,n,f,k,l} &= \frac{1}{\kappa_{0,f,l}}\,.
\end{aligned}
\tag{86}
$$

With $\mu_n(t,T)$ in hand, we seek a Markov representation for $s_n(t,T)$. We follow the same procedure presented in the previous appendix, starting with

$$
\begin{aligned}
s_n(t,T) &= s_n(0,T)+\int_0^t \mu_n(v,T)\,dv+\int_0^t \sigma_{n,f}(v,T)\,dW_f(t)\\
&= s_n(0,T)+A_{n,i,j,f,k,l}\,\bar{K}_{i,j,f,k,l}(T-t)\int_0^t \bar{\psi}_{i,j,f,k,l}(v)\,\bar{K}_{i,j,f,k,l}(t-v)\,dv\\
&\quad +B_{n,n,f,k,l}\,K_{n,f,k}(T-t)\int_0^t \bar{\psi}_{n,0,f,k,l}(v)\,K_{n,f,k}(t-v)\,dv\\
&\quad +K_{n,f,k}(T-t)\int_0^t \psi_{n,f,k}(v)\,K_{n,f,k}(t-v)\,dW_f(v)\,.
\end{aligned}
\tag{87}
$$

The short spread for the n^{th} curve is

$$
\begin{aligned}
s_n(t,t) &= s_n(0,t)+A_{n,i,j,f,k,l}\int_0^t \bar{\psi}_{i,j,f,k,l}(v)\,\bar{K}_{i,j,f,k,l}(t-v)\,dv\\
&\quad +B_{n,n,f,k,l}\int_0^t \bar{\psi}_{n,0,f,k,l}(v)\,K_{n,f,k}(t-v)\,dv\\
&\quad +\int_0^t \psi_{n,f,k}(v)\,K_{n,f,k}(t-v)\,dW_f(v)\,.
\end{aligned}
\tag{88}
$$

In analogy with (60) and (61), we define the primary Markov processes for $s_n(t,T)$ as

$$\begin{aligned} X_{n,f,k}(t) &= A_{n,i,j,f,k,l} \int_0^t \bar{\psi}_{i,j,f,k,l}(v)\, \bar{K}_{i,j,f,k,l}(t-v)\, dv \\ &\quad + B_{n,n,f,k,l} \int_0^t \bar{\psi}_{n,0,f,k,l}(v)\, K_{n,f,k}(t-v)\, dv \\ &\quad + \int_0^t \psi_{n,f,k}(v)\, K_{n,f,k}(t-v)\, dW_f(v)\,, \end{aligned} \tag{89}$$

along with the auxiliary Markov processes

$$Y_{i,j,f,k,l}(t) = \int_0^t \bar{\psi}_{i,j,f,k,l}(v)\, \bar{K}_{i,j,f,k,l}(t-v)\, dv\,. \tag{90}$$

Note that the $Y_{i,j,f,k,l}(t)$ have no dependence on n. They are of precisely the same structure as $Y_{3,j,f,k,l}(t)$ in (61), and as such these can be reused across all $s_m(t,T)$, with curve-specific loadings $A_{n,i,j,f,k,l}$ as necessary. The dynamics of these Markov processes are

$$\begin{aligned} dX_{n,f,k}(t) &= \kappa_{n,f,k}\Bigg(\frac{A_{n,i,j,f,k,l}}{\kappa_{n,f,k}}\,\bar{\psi}_{i,j,f,k,l}(t) + \frac{B_{n,n,f,k,l}}{\kappa_{n,f,k}}\,\bar{\psi}_{n,0,f,k,l}(t) \\ &\quad - \frac{\big(\kappa_{i,f,k}+\kappa_{j,f,l}-\kappa_{n,f,k}\big)\,A_{n,i,j,f,k,l}}{\kappa_{n,f,k}}\,Y_{i,j,f,k,l}(t) - X_{n,f,k}(t)\Bigg)\,dt + \psi_{n,f,k}(t)\,dW_f(t) \\ dY_{i,j,f,k,l}(t) &= (\kappa_{i,f,k}+\kappa_{j,f,l})\Bigg(\frac{1}{\big(\kappa_{i,f,k}+\kappa_{j,f,l}\big)}\,\bar{\psi}_{i,j,f,k,l}(t) - Y_{i,j,f,k,l}(t)\Bigg)\,dt\,. \end{aligned} \tag{91}$$

With these definitions in place we have the Markov representation of $s_n(t,T)$,

$$\begin{aligned} s_n(t,T) &= s_n(0,T) + K_{n,f,k}(T-t)\,X_{n,f,k}(t) \\ &\quad + A_{n,i,j,f,k,l}\,\big(\bar{K}_{i,j,f,k,l}(T-t) - K_{n,f,k}(T-t)\big)\,Y_{i,j,f,k,l}(t)\,, \end{aligned} \tag{92}$$

For the *full* set of $s_m(t,T)$, $m \le n$, it is clear that we will need to simulate the following processes,

$$\begin{gathered} X_{i,f,k}(t)\,,\ 0 < i \le n\,, \\ Y_{i,j,f,k,l}(t)\,,\ 0 < i \le n\,,\ 0 < j \le n\,. \end{gathered} \tag{93}$$

Inspection of $Y_{i,j,f,k,l}(t)$ in (91) reveals a symmetry that can be used to reduce computational cost,

$$Y_{i,j,f,k,l}(t) = Y_{j,i,f,l,k}(t)\,, \tag{94}$$

Consider first the case that $i=j$, then we have that

$$Y_{i,i,f,k,l}(t) = Y_{i,i,f,l,k}(t)\,, \tag{95}$$

and thus for the "diagonal" processes $Y_{3,3,f,k,l}(t)$, $Y_{6,6,f,k,l}(t)$, *etc.* we do away with computing roughly half of them. More precisely, we need not compute $\frac{N_{cv}}{2}\,N_{sh}\,(N_{sh}-1)$ of a total of $N_{cv}\,N_{sh}^2$ of them. This is very similar to the standard result for single-curve Cheyette ('92) models. For the

case where $i \neq j$, the symmetry bears relevance for "off-diagonal" pairs $Y_{3,6,f,k,l}(t)$ and $Y_{6,3,f,k,l}(t)$, $Y_{3,9,f,k,l}(t)$ and $Y_{9,3,f,k,l}(t)$, *etc.* Indeed, once we have computed all of the $Y_{3,6,f,k,l}(t)$ processes, we can simply compute the $Y_{6,3,f,k,l}(t)$ processes via

$$Y_{6,3,f,k,l}(t) = Y_{3,6,f,l,k}(t)\,. \tag{96}$$

As such, we do away with computing half of the these also, or $\frac{N_{sh}^2}{2}\,N_{cv}\,(N_{cv}-1)$ of $N_{sh}^2\,N_{cv}\,(N_{cv}-1)$.

A.3 General Factor and Shape Specifications for the Base Curve

In the previous appendices we did not require the Markov representation of the base curve $f_0(t,T)$ as it was not relevant to the analysis. The Markov representation of $f_0(t,T)$ will however be of interest for swap rate approximations in the following appendix, and thus we derive it here quickly.

To begin, recall that the no-arbitrage restriction on $\mu_0(t,T)$ is

$$\int_t^T \mu_0(t,u)\,du = \frac{1}{2}\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right). \tag{97}$$

Differentiating this against T yields

$$\begin{aligned}\mu_0(t,T) &= \sigma_{0,f}(t,T)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right)\\ &= \bar{\psi}_{0,0,f,k,l}(t)\,A_{0,0,0,f,k,l}\,\bar{K}_{0,0,f,k,l}(T-t) + \bar{\psi}_{0,0,f,k,l}(t)\,B_{0,0,f,k,l}\,K_{0,f,k}(T-t)\,,\end{aligned} \tag{98}$$

where

$$A_{0,0,0,f,k,l} = -\frac{1}{\kappa_{0,f,l}} \quad \text{and} \quad B_{0,0,f,k,l} = \frac{1}{\kappa_{0,f,l}}\,. \tag{99}$$

With this, we can write the solution to $f_0(t,T)$ as

$$\begin{aligned}f_0(t,T) &= f_0(0,T) + \int_0^t \mu_0(v,T)\,dv + \int_0^t \sigma_{0,f}(v,T)\,dW_f(v)\\ &= f_0(0,T) + A_{0,0,0,f,k,l}\,\bar{K}_{0,0,f,k,l}(T-t)\int_0^t \bar{\psi}_{0,0,f,k,l}(v)\,\bar{K}_{0,0,f,k,l}(t-v)\,dv\\ &\quad + B_{0,0,f,k,l}\,K_{0,f,k}(T-t)\int_0^t \bar{\psi}_{0,0,f,k,l}(v)\,K_{0,f,k}(t-v)\,dv\\ &\quad + K_{0,f,k}(T-t)\int_0^t \psi_{0,f,k}(v)\,K_{0,f,k}(t-v)\,dW_f(v)\,.\end{aligned} \tag{100}$$

The short rate $f_0(t,t)$ is of the form

$$\begin{aligned}f_0(t,t) &= f_0(0,t) + A_{0,0,0,f,k,l}\int_0^t \bar{\psi}_{0,0,f,k,l}(v)\,\bar{K}_{0,0,f,k,l}(t-v)\,dv\\ &\quad + B_{0,0,f,k,l}\int_0^t \bar{\psi}_{0,0,f,k,l}(v)\,K_{0,f,k}(t-v)\,dv\\ &\quad + \int_0^t \psi_{0,f,k}(v)\,K_{0,f,k}(t-v)\,dW_f(v)\,,\end{aligned} \tag{101}$$

and we extract from this the primary and secondary Markov state processes driving $f_0(t,T)$,

$$\begin{aligned} X_{0,f,k}(t) &= A_{0,0,0,f,k,l} \int_0^t \bar{\psi}_{0,0,f,k,l}(v)\, \bar{K}_{0,0,f,k,l}(t-v)\, dv \\ &\quad + B_{0,0,f,k,l} \int_0^t \bar{\psi}_{0,0,f,k,l}(v)\, K_{0,f,k}(t-v)\, dv \\ &\quad + \int_0^t \psi_{0,f,k}(v)\, K_{0,f,k}(t-v)\, dW_f(v)\,, \end{aligned} \tag{102}$$

and

$$Y_{0,0,f,k,l}(t) = \int_0^t \bar{\psi}_{0,0,f,k,l}(v)\, \bar{K}_{0,0,f,k,l}(t-v)\, dv\,. \tag{103}$$

With these state processes in hand, $f_0(t,T)$ can be written as

$$\begin{aligned} f_0(t,T) &= f_0(0,T) + K_{0,f,k}(T-t)\, X_{0,f,k}(t) \\ &\quad + A_{0,0,0,f,k,l} \left(\bar{K}_{0,0,f,k,l}(T-t) - K_{0,f,k}(T-t)\right) Y_{0,0,f,k,l}(t)\,, \end{aligned} \tag{104}$$

where $X_{0,f,k}(t)$ and $Y_{0,0,f,k,l}(t)$ have the dynamics

$$\begin{aligned} dX_{0,f,k}(t) &= \kappa_{0,f,k}\left(-\frac{\kappa_{0,f,l}\, A_{0,0,0,f,k,l}}{\kappa_{0,f,k}}\, Y_{0,0,f,k,l}(t) - X_{0,f,k}(t)\right) dt + \psi_{0,f,k}(t)\, dW_f(t)\,, \\ dY_{0,0,f,k,l}(t) &= \left(\kappa_{0,f,k} + \kappa_{0,f,l}\right)\left(\frac{1}{\left(\kappa_{0,f,k} + \kappa_{0,f,l}\right)}\, \bar{\psi}_{0,0,f,k,l}(t) - Y_{0,0,f,k,l}(t)\right) dt\,. \end{aligned} \tag{105}$$

The preceding structure is of a similar form to that of the spread curves $s_m(t,T)$, which simplifies coming calculations. The typical representation for single-curve Cheyette ('92) models would replace the auxiliary Markov state processes with $Y'_{0,0,f,k,l}(t) = -A_{0,0,0,f,k,l}\, Y_{0,0,f,k,l}(t)$, but as noted earlier we have adopted the structure above in order to take advantage of the analogous processes $Y_{i,j,f,k,l}(t)$ appearing in the dynamics of all spread curves.

B Approximate Swap Rates and Basis Swap Spreads

In this appendix we derive the approximate swap rate dynamics in (22), or rather we derive their general form. We begin by writing down the discount bond prices $P(t,T)$ and the 3M IBORs $R_3(t,T)$ in terms of the Markov state processes introduced in the previous appendix. To this end, we have

$$\begin{aligned} P(t,T) &= \exp\left(-\int_t^T f_0(t,u)\, du\right) \\ &= \frac{P(0,T)}{P(0,t)} \exp\left(-\left(\int_t^T K_{0,f,k}(u-t)\, du\right) X_{0,f,k}(t) \right. \\ &\qquad \left. - A_{0,0,0,f,k,l} \left(\int_t^T \left(\bar{K}_{0,0,f,k,l}(u-t) - K_{0,f,k}(u-t)\right) du\right) Y_{0,0,f,k,l}(t)\right) \\ &= \frac{P(0,T)}{P(0,t)} \exp\left(F_{0,f,k}(t,T)\, X_{0,f,k}(t) + G_{0,0,0,f,k,l}(t,T)\, Y_{0,0,f,k,l}(t)\right)\,, \end{aligned} \tag{106}$$

where

$$F_{0,f,k}(t,T) = \frac{1}{\kappa_{0,f,k}}\big(K_{0,f,k}(T-t) - 1\big) \tag{107}$$

and

$$G_{0,0,0,f,k,l}(t,T) = \frac{A_{0,0,0,f,k,l}}{\big(\kappa_{0,f,k} + \kappa_{0,f,l}\big)}\left(\bar{K}_{0,0,f,k,l}(T-t) - \big(1 + \tfrac{\kappa_{0,f,l}}{\kappa_{0,f,k}}\big)\, K_{0,f,k}(T-t) + \tfrac{\kappa_{0,f,l}}{\kappa_{0,f,k}}\right). \tag{108}$$

For 3M IBORs $R_3(t,T)$ we will require the quantities $e^{\int_T^{T+\tau_3} s_3(t,u)\,du}$, and we thus invoke the Markov representation of $s_3(t,T)$ from (62),

$$\begin{aligned}
&\exp\left(\int_T^{T+\tau_3} s_3(t,u)\,du\right) \\
&= \exp\left(\int_T^{T+\tau_3} s_3(0,u)\,du + \left(\int_T^{T+\tau_3} K_{3,f,k}(u-t)\,du\right) X_{3,f,k}(t)\right. \\
&\quad \left. + A_{3,3,j,f,k,l}\left(\int_T^{T+\tau_3} \big(\bar{K}_{3,j,f,k,l}(u-t) - K_{3,f,k}(u-t)\big)\,du\right) Y_{3,j,f,k,l}(t)\right) \\
&= \exp\left(\int_T^{T+\tau_3} s_3(0,u)\,du\right)\, \exp\big(F_{3,f,k}(t,T)\, X_{3,f,k}(t) + G_{3,3,j,f,k,l}(t,T)\, Y_{3,j,f,k,l}(t)\big)\,,
\end{aligned} \tag{109}$$

where

$$F_{3,f,k}(t,T) = -\frac{1}{\kappa_{3,f,k}}\left(K_{3,f,k}(\tau_3) - 1\right) K_{3,f,k}(T-t) \tag{110}$$

and

$$\begin{aligned}
G_{3,f,k}(t,T) \;=\; & -\frac{A_{3,3,j,f,k,l}}{\big(\kappa_{3,f,k} + \kappa_{j,f,l}\big)}\left(\big(\bar{K}_{3,j,f,k,l}(\tau_3) - 1\big)\,\bar{K}_{3,j,f,k,l}(T-t)\right. \\
& \left. -\big(1 + \tfrac{\kappa_{j,f,l}}{\kappa_{3,f,k}}\big)\,\big(K_{3,f,k}(\tau_3) - 1\big)\,K_{3,f,k}(T-t)\right).
\end{aligned} \tag{111}$$

We note that the index j takes on the values $\{0,3\}$ in these sums. Combining the expressions in (106) and (109) allows us to write down the form of the 3M IBOR,

$$\begin{aligned}
R_3(t,T) \;=\; & \frac{1}{\tau_3}\left(\frac{P(0,T)}{P(0,T+\tau_3)}\,\exp\left(\int_T^{T+\tau_3} s_3(0,u)\,du\right)\right. \\
& \exp\big(\big(F_{0,f,k}(t,T) - F_{0,f,k}(t,T+\tau_3)\big)\,X_{0,f,k}(t) \\
& +\big(G_{0,0,0,f,k,l}(t,T) - G_{0,0,0,f,k,l}(t,T+\tau_3)\big)\,Y_{0,0,f,k,l}(t) \\
& \left. +F_{3,f,k}(t,T)\,X_{3,f,k}(t) + G_{3,3,j,f,k,l}(t,T)\,Y_{3,j,f,k,l}(t)\big) - 1\right).
\end{aligned} \tag{112}$$

It will become important to have the differentials of $P(t,T)$ and $R_3(t,T)$ available when taking the differentials of the swap rate. We will be interested only in the diffusion functions of these processes, as they will govern the diffusion of the swap rate. The drift functions will not be relevant as we will be working under the annuity measure associated with our swap rate of interest. As such, the swap rate will be a Martingale by construction.

Taking differentials of (106) and (112) yields

$$dP(t,T) = P(t,T)\,\frac{1}{\kappa_{0,f,k}}\left(K_{0,f,k}(T-t)-1\right)\psi_{0,f,k}(t)\,dW_f(t) \tag{113}$$

and

$$\begin{aligned} dR_3(t,T) &= -\left(\frac{1}{\tau_3}+R_3(t,T)\right)\left(\frac{1}{\kappa_{0,f,k}}\left(K_{0,f,k}(\tau_3)-1\right)\psi_{0,f,k}(t)\,K_{0,f,k}(T-t)+\right. \\ &\qquad \left.+\frac{1}{\kappa_{3,f,k}}\left(K_{3,f,k}(\tau_3)-1\right)\psi_{3,f,k}(t)\,K_{3,f,k}(T-t)\right)dW_f(t)\,. \end{aligned} \tag{114}$$

From here, we recall the forward-starting swap rate definition from (21),

$$\omega_3(t\,;T_0) = \frac{P(t,T_{i+1})R_3(t,T_i)\,\tau_3}{L(t\,;T_0)} \quad \text{where} \quad L(t\,;T_0) = P(t,T_{j+1})\,\tau_3\,, \tag{115}$$

where the dates $\{T_0,T_1,\dots\}$ are a series of fixing/payment dates spaced at 3M intervals, *i.e.* $T_{i+1}=T_i+\tau_3$. Note that in the above, we use the index notation developed in earlier appendices, meaning that we sum over i in the definition of $\omega_3(t\,;T_0)$, and over j in the definition of $L(t\,;T_0)$. Taking the differential of $\omega_3(t\,;T_0)$ will involve a composition of the differentials of the constituent processes,

$$\begin{aligned} d\omega_3(t\,;T_0) &= P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3\,d\left(\frac{1}{L(t\,;T_0)}\right) \\ &\qquad +\frac{dP(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)}+\frac{P(t,T_{i+1})\,dR_3(t,T_i)\,\tau_3}{L(t\,;T_0)}\,. \end{aligned} \tag{116}$$

We compute the terms of the above one at a time. First we compute $d\left(\frac{1}{L(t\,;T_0)}\right)$,

$$\begin{aligned} d\left(\frac{1}{L(t\,;T_0)}\right) &= \cdots dt-\frac{1}{L(t\,;T_0)^2}\,dL(t\,;T_0) \\ &= \cdots dt-\frac{1}{L(t\,;T_0)}\,\frac{P(t,T_{j+1})\,\tau_3}{L(t\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\left(K_{0,f,k}(T_{j+1}-t)-1\right)\psi_{0,f,k}(t)\,dW_f(t) \\ &= \cdots dt-\frac{1}{L(t\,;T_0)}\,\frac{P(t,T_{j+1})\,\tau_3}{L(t\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\,K_{0,f,k}(T_{j+1}-t)\,\psi_{0,f,k}(t)\,dW_f(t) \\ &\qquad +\frac{1}{L(t\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\,\psi_{0,f,k}(t)\,dW_f(t)\,, \end{aligned} \tag{117}$$

such that the first term can be written as a sum over

$$\begin{aligned} &P(t,T_{i+1})\,R_3(t,T)\,\tau_3\,d\left(\frac{1}{L(t\,;T_0)}\right) \\ &=\dots dt-\frac{P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)}\,\frac{P(t,T_{j+1})\,\tau_3}{L(t\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\,K_{0,f,k}(T_{j+1}-t)\,\psi_{0,f,k}(t)\,dW_f(t) \\ &\quad +\frac{P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\,\psi_{0,f,k}(t)\,dW_f(t)\,. \end{aligned} \tag{118}$$

The second can be written as a sum over

$$\begin{aligned}
&\frac{dP(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)} \\
&= \cdots\, dt + \frac{P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\,K_{0,f,k}(T_{i+1}-t)\,\psi_{0,f,k}(t)\,dW_f(t) \\
&-\frac{P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\,\psi_{0,f,k}(t)\,dW_f(t)\,.
\end{aligned} \tag{119}$$

Note that (118) and (119) share a similar structure with offsetting terms (the second term in each), and combining them reveals that the first two terms evaluate to a sum over

$$\begin{aligned}
\cdots\, dt + \frac{P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\,\Bigg(&K_{0,f,k}(T_{i+1}-t) \\
&-\frac{P(t,T_{j+1})\,\tau_3}{L(t\,;T_0)}\;K_{0,f,k}(T_{j+1}-t)\Bigg)\,\psi_{0,f,k}(t)\,dW_f(t)\,.
\end{aligned} \tag{120}$$

The third term evaluates to a sum over

$$\begin{aligned}
&\frac{P(t,T_{i+1})\,dR_3(t,T_i)\,\tau_3}{L(t\,;T_0)} \\
&= \cdots\, dt - \frac{P(t,T_{i+1})\,\big(1+R_3(t,T_i)\,\tau_3\big)}{L(t\,;T_0)}\Bigg(\frac{1}{\kappa_{0,f,k}}\big(K_{0,f,k}(\tau_3)-1\big)\,\psi_{0,f,k}(t)\,K_{0,f,k}(T_i-t) \\
&\qquad +\frac{1}{\kappa_{3,f,k}}\big(K_{3,f,k}(\tau_3)-1\big)\,\psi_{3,f,k}(t)\,K_{3,f,k}(T_i-t)\Bigg)\,dW_f(t)\,.
\end{aligned} \tag{121}$$

We can now combine (120) and (121) to write the dynamics of the swap rate as

$$\begin{aligned}
&d\omega_3(t\,;T_0) = \\
&\frac{P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\,\Bigg(K_{0,f,k}(T_{i+1}-t)-\frac{P(t,T_{j+1})\,\tau_3}{L(t\,;T_0)}\;K_{0,f,k}(T_{j+1}-t)\Bigg)\,\psi_{0,f,k}(t)\,d\overline{W}_f^{T_0}(t) \\
&-\frac{P(t,T_{i+1})\,\big(1+R_3(t,T_i)\,\tau_3\big)}{L(t\,;T_0)}\Bigg(\frac{1}{\kappa_{0,f,k}}\big(K_{0,f,k}(\tau_3)-1\big)\,\psi_{0,f,k}(t)\,K_{0,f,k}(T_i-t) \\
&\qquad +\frac{1}{\kappa_{3,f,k}}\big(K_{3,f,k}(\tau_3)-1\big)\,\psi_{3,f,k}(t)\,K_{3,f,k}(T_i-t)\Bigg)\,d\overline{W}_f^{T_0}(t)\,,
\end{aligned} \tag{122}$$

where the drift can only be ignored because we are operating under the annuity measure associated with $L(t\,;T_0)$, for which $\overline{W}^{T_0}(t)$ is a Brownian motion.

At this stage, we adopt the common convention of "freezing" components of the diffusion coefficients. Specifically, we freeze

$$\frac{P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)} \quad \text{at} \quad \frac{P(0,T_{i+1})\,R_3(0,T_i)\,\tau_3}{L(0\,;T_0)}\,, \tag{123}$$

$$\frac{P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)}\,\frac{P(t,T_{j+1})\,\tau_3}{L(t\,;T_0)} \quad \text{at} \quad \frac{P(0,T_{i+1})\,R_3(0,T_i)\,\tau_3}{L(0\,;T_0)}\,\frac{P(0,T_{j+1})\,\tau_3}{L(0\,;T_0)}\,, \tag{124}$$

and

$$\frac{P(t,T_{i+1})\left(1+R_3(t,T_i)\,\tau_3\right)}{L(t\,;T_0)} \quad \text{at} \quad \frac{P(0,T_{i+1})\left(1+R_3(0,T_i)\,\tau_3\right)}{L(0\,;T_0)}\,. \tag{125}$$

The standard rationale for freezing these quantities is that they have low variance relative to the diffusion functions of the driving processes, $\psi_{0,f,k}(t)$ and $\psi_{3,f,k}(t)$, where we note that the first and third quantities are Martingales but the second is not. Realistically however, these approximations work because the dominant term in (122) is the second term, owing to the presence of the coefficients

$$\frac{P(t,T_{i+1})\left(1+R_3(t,T_i)\,\tau_3\right)}{L(t\,;T_0)}\,. \tag{126}$$

Indeed, it is clear that

$$\frac{P(t,T_{i+1})\left(1+R_3(t,T_i)\,\tau_3\right)}{L(t\,;T_0)} \gg \frac{P(t,T_{i+1})\,R_3(t,T_i)\,\tau_3}{L(t\,;T_0)}\,. \tag{127}$$

Noting then that

$$\frac{P(t,T_{i+1})\left(1+R_3(t,T_i)\,\tau_3\right)}{L(t\,;T_0)} \approx \frac{P(t,T_{i+1})}{L(t\,;T_0)}\,, \tag{128}$$

which is indeed a low-variance Martingale relative to the diffusion functions $\psi_{0,f,k}(t)$ and $\psi_{3,f,k}(t)$, the rationale for the approximation becomes clear.

We might also add that the rationale offered here formed the basis for the crude approximation of $dR_3(t,T)$ given in Section 2, *i.e.*

$$\begin{aligned} dR_3(t,T) &= \cdots dt + \left(1+R_3(t,T)\,\tau_3\right)\left(\frac{1}{\tau_3}d\int_T^{T+\tau_3} f_0(t,u)\,du + \frac{1}{\tau_3}d\int_T^{T+\tau_3} s_3(t,u)\,du\right) \\ &\approx \frac{1}{\tau_3}d\int_T^{T+\tau_3} f_0(t,u)\,du + \frac{1}{\tau_3}d\int_T^{T+\tau_3} s_3(t,u)\,du \\ &\approx df_0(t,T) + ds_3(t,T)\,. \end{aligned} \tag{129}$$

The first approximations uses $1+R_3(t,T)\,\tau_3 \approx 1$ and ignores the drift contribution $\cdots dt$, and the second approximation assumes that variation in $df_0(t,u)$ and $ds_3(t,u)$ is largely homogeneous over intervals of τ_3.

The approach we have taken to deriving our approximations in this section bears a strong resemblance to the approach presented in Schrager and Pelsser ('06). These authors derive approximate swap rate dynamics for affine short-rate models in the single-curve context. When working with affine models, bond prices, and ratios thereof, are of course exponential-affine in the processes governing the short rate. We have demonstrated that the analogous quantities are also exponential-affine in a set of processes in the multi-curve Cheyette ('92) framework, and that this will be true irrespective of whether the dynamics of the curves involved are affine or not. As we are taking differentials of exponential-affine quantities analogous to those that Schrager and Pelsser ('06) encounter, it is natural that similar approximations of the resulting dynamics should ensue.

With the frozen coefficients in hand, we approximate the dynamics in (122) as

$$d\omega_3(t\,;T_0) \approx \left(\gamma_{0,f,k}(T_0)\,\psi_{0,f,k}(t)\,K_{0,f,k}(T_0-t) + \gamma_{3,f,k}(T_0)\,\psi_{3,f,k}(t)\,K_{3,f,k}(T_0-t)\right)d\overline{W}_f^{T_0}(t)\,, \tag{130}$$

where

$$\begin{aligned}
\gamma_{0,f,k}(T_0) &= \frac{P(0,T_{i+1})\,R_3(0,T_i)\,\tau_3}{L(0\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\left(K_{0,f,k}(T_{i+1}-T_0)-\frac{P(0,T_{j+1})\,\tau_3}{L(0\,;T_0)}\;K_{0,f,k}(T_{j+1}-T_0)\right)\\
&\quad -\frac{P(0,T_{i+1})\left(1+R_3(0,T_i)\,\tau_3\right)}{L(0\,;T_0)}\frac{1}{\kappa_{0,f,k}}\big(K_{0,f,k}(\tau_3)-1\big)\,K_{0,f,k}(T_i-T_0)\\
\gamma_{3,f,k}(T_0) &= -\frac{P(0,T_{i+1})\left(1+R_3(0,T_i)\,\tau_3\right)}{L(0\,;T_0)}\frac{1}{\kappa_{3,f,k}}\big(K_{3,f,k}(\tau_3)-1\big)\,K_{3,f,k}(T_i-T_0)\,.
\end{aligned} \tag{131}$$

Note that we are operating under the swap rate measure, and the transition to this measure gives rise to a drift adjustment to the dynamics of the primary Markov state variables, $X_{0,f,k}(t)$ and $X_{3,f,k}(t)$. The drift adjustment for each component Brownian motion is

$$\begin{aligned}
&dW_f(t)\,d\left(\frac{L(t\,;T_0)}{\exp\big(-\int_0^t r(v)\,dv\big)}\right)\Big/\left(\frac{L(t\,;T_0)}{\exp\big(-\int_0^t r(v)\,dv\big)}\right)\\
&\approx \frac{P(0,T_{j+1})\,\tau_3}{L(0\,;T_0)}\,\frac{1}{\kappa_{0,f,l}}\left(K_{0,f,l}(T_{j+1}-t)-1\right)\psi_{0,f,l}(t)\,dt
\end{aligned} \tag{132}$$

where we have frozen the quantity $\frac{P(0,T_{j+1})\,\tau_3}{L(0\,;T_0)}$ as per the swap rate coefficients. The drift adjustments for the Markov state processes, $X_{i,f,k}(t)$, are simply the products of (132) for each f and the volatility function of a given $X_{i,f,k}(t)$, which are simply sums of $\psi_{i,f,k}(t)$ over k. As such, the full drift adjustment for $X_{i,f,k}(t)$ is simply a sum in k and l over

$$\frac{P(0,T_{j+1})\,\tau_3}{L(0\,;T_0)}\,\frac{1}{\kappa_{0,f,l}}\left(K_{0,f,l}(T_{j+1}-t)-1\right)\bar{\psi}_{i,0,f,k,l}(t)\,. \tag{133}$$

At this point, it is useful to consider the pricing of swaptions. The value of swaption struck at ω_3^* will satisfy

$$\frac{\mathrm{swpn}(t\,;T_0,\omega^*)}{L(t\,;T_0)} = \overline{\mathbb{E}}_t^{T_0}\left[(\omega_3(T_0\,;T_0)-\omega^*)^+\right]. \tag{134}$$

Naturally, the expectation above is generally not available in closed form. It is however available when $\omega_3(t\,;T_0)$ is part of an affine system along with $X_{i,f,k}(t)$ and $Y_{i,j,f,k,l}(t)$. We note that this ultimately distills down to our choice of diffusion functions $\psi_{i,f,k}(t)$. These diffusion functions are said to be affine if

$$\bar{\psi}_{i,j,f,k,l}(t) = u_{i,j,f,k,l}(t) + v_{i,j,f,k,l,h,m,n}(t)X_{h,m,n}(t)\,, \tag{135}$$

where we note that $Y_{i,j,f,k,l}(t)$ will not enter diffusion functions for obvious reasons. Under these conditions, the covariance functions of the $X_{i,f,k}(t)$ processes will be affine in $X_{i,f,k}(t)$ by construction, and it is clear that the swap rate in (130) will have a variance function affine in $X_{i,f,k}(t)$. Moreover, the spot-measure drift functions of $X_{i,f,k}(t)$ and $Y_{i,j,f,k,l}(t)$ in (91) will clearly be affine in $X_{i,f,k}(t)$ and $Y_{i,j,f,k,l}(t)$ also. All that is necessary to check is that the drift adjustments owing to the transition to the annuity measure associated with $L(t\,;T_0)$ are affine. A quick inspection of (133) reveals this to be true, given that the only state dependence enters through through $\bar{\psi}_{i,0,f,k,l}(t)$. We discuss the use of Fourier techniques to price swaptions in affine cases in the following appendix.

We now consider a basis swap spread. Our goal is to derive approximate dynamics for this process such that we can compute its expected volatility. Such a volatility will be useful when

calibrating the parameters of the basis curve, as was discussed in Section 3.2. Indeed, this volatility can be aligned with its historical analogue through the choice of the parameters which govern it.

The spread under consideration is for an ON-*vs.*-3M IBOR tenor basis swap, exchanging the compounded ON rate over a 3M interval, plus a spread, against the 3M IBOR. For simplicity, the floating component of the cashflow on the ON leg is taken to be

$$\exp\left(\int_{T_i}^{T_{i+1}} r(t)\,dt\right) - 1\,, \tag{136}$$

i.e. where we assume that payment amount is continuously compounded at the ON rate $r(t)$. Given this, the par spread on this trade can straightforwardly be computed as

$$\phi_{0,3}(t\,;T_0) = \frac{P(t,T_{i+1})R_3(t,T_i)\,\tau_3}{L(t\,;T_0)} - \frac{P(t,T_i)-P(t,T_{i+1})}{L(t\,;T_0)}\,. \tag{137}$$

Note that the first term is simply the swap rate $\omega_3(t\,;T_0)$. We can in fact recast the second term such that it shares a similar structure. To do this, recall the formula for the classical single-curve 3M IBOR off the ON curve in (2), where we had denoted this quantity by $R_{0,3}(t,T)$,

$$R_{0,3}(t,T) = \frac{1}{\tau_3}\left(\exp\left(\int_T^{T+\tau_3} f_0(t,u)\,du\right) - 1\right) = \frac{1}{\tau_3}\left(\frac{P(t,T)}{P(t,T+\tau_3)} - 1\right). \tag{138}$$

Given these observations we can reformulate (137) as

$$\phi_{0,3}(t\,;T_0) = \omega_3(t\,;T_0) - \frac{P(t,T_{i+1})R_{0,3}(t,T_i)\,\tau_3}{L(t\,;T_0)}\,, \tag{139}$$

as one clearly has that $P(t,T_{i+1})R_{0,3}(t,T_i)\,\tau_3 = P(t,T_i) - P(t,T_{i+1})$, as $T_{i+1} = T_i + \tau_3$. With this in hand we may write the basis spread as

$$\phi_{0,3}(t\,;T_0) = \frac{P(t,T_{i+1})\big(R_3(t,T_i) - R_{0,3}(t,T_i)\big)\,\tau_3}{L(t\,;T_0)}\,. \tag{140}$$

Following precisely the same logic used to derive the approximate dynamics for $\phi_3(t\,;T_0)$ we arrive at

$$\begin{aligned} d\phi_{0,3}(t\,;T_0) \;\approx\; & \big(\gamma_{0,3,0,f,k}(T_0)\,\psi_{0,f,k}(t)\,K_{0,f,k}(T_0-t) \\ & \quad +\gamma_{0,3,3,f,k}(T_0)\,\psi_{3,f,k}(t)\,K_{3,f,k}(T_0-t)\big)\,d\overline{W}_f^{T_0}(t)\,, \end{aligned} \tag{141}$$

where

$$\begin{aligned} \gamma_{0,3,0,f,k}(T_0) \;=\; & \frac{P(0,T_{i+1})\big(R_3(0,T_i) - R_{0,3}(0,T_i)\big)\,\tau_3}{L(0\,;T_0)}\,\frac{1}{\kappa_{0,f,k}} \\ & \times\left(K_{0,f,k}(T_{i+1}-T_0) - \frac{P(0,T_{j+1})\,\tau_3}{L(0\,;T_0)}\;K_{0,f,k}(T_{j+1}-T_0)\right) \\ & -\frac{P(0,T_{i+1})\,\big(R_3(0,T_i) - R_{0,3}(0,T_i)\big)\,\tau_3}{L(0\,;T_0)}\,\frac{1}{\kappa_{0,f,k}}\big(K_{0,f,k}(\tau_3)-1\big)\,K_{0,f,k}(T_i-T_0)\,, \\ \gamma_{0,3,3,f,k}(T_0) \;=\; & -\frac{P(0,T_{i+1})\,\big(1+R_3(0,T_i)\,\tau_3\big)}{L(0\,;T_0)}\,\frac{1}{\kappa_{3,f,k}}\big(K_{3,f,k}(\tau_3)-1\big)\,K_{3,f,k}(T_i-T_0)\,. \end{aligned} \tag{142}$$

Noting that $\big(R_3(0,T_i) - R_{0,3}(0,T_i)\big)\,\tau_3$ will be very small relative to $1 + R_3(0,T_i)\,\tau_3$, it is clear that the coefficient $\gamma_{0,3,3,f,k}(T_0)$ will dominate $\gamma_{0,3,0,f,k}(T_0)$, and thus variation in the basis swap spread $\phi_{0,3}(t\,;T_0)$ will be driven primarily by variation in the ON-*vs.*-3M IBOR basis spread curve $s_3(t,T)$.

In the case of the swap rate, it is indeed proper to consider the rate for a swap referencing a fixed set of dates $\{T_0, T_1, \dots\}$, as we will seek to price swaptions written over such swaps. This is not precisely what we desire for basis swap spreads however. Our aim is to determine the volatility of a tenor basis swap with a given tenor, *e.g.* $\tau'_{bs} = 2Y$, so that we can compare it with what has been observed historically. These historical spreads are not for a basis swap with a given set of dates which are fixed over time, but rather each observed spread corresponds to a new basis swap, starting on the observation date, with a constant swap tenor of τ^{bs}.

We can easily modify the definition of the basis spread in (140) to obtain such a process,

$$\phi'_{0,3}(t\,;\tau'_{bs}) = \frac{P(t,t+\tau'_{i+1})\big(R_3(t,t+\tau'_i) - R_{0,3}(t,t+\tau'_i)\big)\,\tau_3}{L'(t\,;\tau'_{bs})}\,, \tag{143}$$

where $\{\tau'_0, \tau'_1, \dots, \tau'_{bs}\}$ is a series of offset tenors for fixing/payment dates as t advances, satisfying $\tau'_{i+1} = \tau'_i + \tau_3$. Note also that there is a necessary modification of the notation used for the annuity function,

$$L'(t\,;\tau'_{bs}) = P(t,t+\tau'_i)\,\tau_3\,, \tag{144}$$

where we sum over τ'_1, τ'_2, …, τ'_{bs}.

Following the logic used earlier, we are able to derive the spot-measure dynamics of the "constant-maturity" basis swap spread $\phi'_{0,3}(t\,;\tau'_{bs})$

$$d\phi'_{0,3}(t\,;\tau'_{bs}) \;=\; \cdots\,dt + \big(\gamma_{0,3,3,f,k}(t\,;\tau'_{bs})\,\psi_{0,f,k}(t) + \gamma_{0,3,0,f,k}(t\,;\tau'_{bs})\,\psi_{3,f,k}(t)\big)\,dW_f(t)\,, \tag{145}$$

where

$$\begin{aligned}
\gamma_{0,3,0,f,k}(t\,;\tau'_{bs}) \;=\;& \frac{P(t,t+\tau'_{i+1})\big(R_3(t,t+\tau'_i) - R_{0,3}(t,t+\tau'_i)\big)\,\tau_3}{L'(t\,;\tau'_{bs})}\,\frac{1}{\kappa_{0,f,k}} \\
&\times\left(K_{0,f,k}(\tau'_{i+1}) - \frac{P(t,t+\tau'_{j+1})\,\tau_3}{L'(t\,;\tau'_{bs})}\;K_{0,f,k}(\tau'_{j+1})\right) \\
&-\frac{P(t,t+\tau'_{i+1})\,\big(R_3(t,t+\tau'_i) - R_{0,3}(t,t+\tau'_i)\big)\,\tau_3}{L'(t\,;\tau'_{bs})}\,\frac{1}{\kappa_{0,f,k}}\big(K_{0,f,k}(\tau_3) - 1\big)\,K_{0,f,k}(\tau'_i) \\
\gamma_{0,3,3,f,k}(t\,;\tau'_{bs}) \;=\;& -\frac{P(t,t+\tau'_{i+1})\,\big(1 + R_3(t,t+\tau'_i)\,\tau_3\big)}{L'(t\,;\tau'_{bs})}\,\frac{1}{\kappa_{3,f,k}}\big(K_{3,f,k}(\tau_3) - 1\big)\,K_{3,f,k}(\tau'_i)\,.
\end{aligned} \tag{146}$$

We note at this point that the drift function cannot be ignored, as the quantity $L'(t\,;\tau'_{bs})$ is no longer a legitimate Numeraire. However, we will not be interested in the drift function in any case. Indeed, our goal in deriving the dynamics of $\phi'_{0,3}(t\,;\tau'_{bs})$ is to arrive at an approximate volatility for the processes, and by volatility, we mean the standard deviation of short-term (*i.e.* daily) changes, as it is the empirical analogue of this quantity that we will compute from historical data. This will of course begin with the calculation of running variation, *i.e.*

$$\begin{aligned}
\int_0^t d\phi'_{0,3}(v\,;\tau'_{bs})\,d\phi'_{0,3}(v\,;\tau'_{bs}) \;=\;& \int_0^t \big(\gamma_{0,3,0,f,k}(v\,;\tau'_{bs})\,\psi_{0,f,k}(v) + \gamma_{0,3,3,f,k}(v\,;\tau'_{bs})\,\psi_{3,f,k}(v)\big) \\
&\times\big(\gamma_{0,3,0,f,l}(v\,;\tau'_{bs})\,\psi_{0,f,l}(v) + \gamma_{0,3,3,f,l}(v\,;\tau'_{bs})\,\psi_{3,f,l}(v)\big)\,dv\,,
\end{aligned} \tag{147}$$

which is drift invariant. We will of course be required to take an expectation of this quantity however, as the diffusion functions $\psi_{0,f,k}(t)$ and $\psi_{3,f,k}(t)$ depend on the stochastic factors $X_{0,f,k}(t)$ and $X_{0,f,k}(t)$. Preferably, we would compute this under the physical measure, but for a practical calibration we will settle for the spot measure. Given these statements, what we formally seek is

$$\begin{aligned}\mathbb{E}\left[\int_0^t d\phi'_{0,3}(v\,;\tau'_{bs})\,d\phi'_{0,3}(v\,;\tau'_{bs})\right]^{\frac{1}{2}} &= \mathbb{E}\Big[\int_0^t \big(\gamma_{0,3,0,f,k}(v\,;\tau'_{bs})\,\psi_{0,f,k}(v)+\gamma_{0,3,3,f,k}(v\,;\tau'_{bs})\,\psi_{3,f,k}(v)\big)\\ &\quad\times\big(\gamma_{0,3,0,f,l}(v\,;\tau'_{bs})\,\psi_{0,f,l}(v)+\gamma_{0,3,3,f,l}(v\,;\tau'_{bs})\,\psi_{3,f,l}(v)\big)\,dv\Big]^{\frac{1}{2}},\end{aligned} \tag{148}$$

which we would then convert to an annualized figure. Given our earlier comments on affine models, it is clear that for such specifications, $\psi_{i,f,k}(t)\,\psi_{j,f,l}(t)$ will be affine in $X_{i,f,k}(t)$. Thus, once we have approximated $\gamma_{0,3,0,f,k}(t\,;\tau'_{bs})$ and $\gamma_{0,3,3,f,k}(t\,;\tau'_{bs})$ with deterministic functions, these expectations can be evaluated straightforwardly.

With regard to the matter of approximating $\gamma_{0,3,0,f,k}(t\,;\tau'_{bs})$ and $\gamma_{0,3,3,f,k}(t\,;\tau'_{bs})$, let us consider the dominant stochastic quantity involved in these,

$$\frac{P(t,t+\tau'_{i+1})\,\big(1+R_3(t,t+\tau'_i)\,\tau_3\big)}{L'(t\,;\tau'_{bs})}\,. \tag{149}$$

One could proceed as we did earlier and adapt the frozen coefficients technique accordingly. This involves a slight modification to the earlier application, which owes to the fact that the calendar date t drives *both* the valuation dates and the maturity dates (at offsets of τ'_i to t) of the quantities involved. To account for this, we take the quantity in (149) and approximate it as

$$\frac{\big(P(0,t+\tau'_{i+1})/P(0,t)\big)}{\big(L'(0\,;t,\tau'_{bs})/P(0,t)\big)}\left(1+\frac{1}{\tau_3}\left(\frac{\big(P(0,t+\tau'_i)/P(0,t)\big)}{\big(P(0,t+\tau'_{i+1})/P(0,t)\big)}\,e^{\int_{t+\tau_i}^{t+\tau_i+\tau_3} s_3(0,u)\,du}-1\right)\tau_3\right), \tag{150}$$

where we have defined

$$L'(t\,;T,\tau'_{bs}) = P(t,T+\tau'_{i+1})\,\tau_3\,. \tag{151}$$

To justify this approximation we note that, as in (128), we have that

$$\frac{P(t,t+\tau'_{i+1})\,\big(1+R_3(t,t+\tau'_i)\,\tau_3\big)}{L(t\,;\tau'_{bs})} \approx \frac{P(t,t+\tau'_{i+1})}{L(t\,;\tau'_{bs})}\,, \tag{152}$$

and it is reasonable to expect that that variation in $P(t,t+\tau'_{i+1})/L'(t\,;\tau'_{bs})$ will be minimal (relative to $\psi_{3,f,k}(t)$) given cancellation effects across the numerator and denominator. We note also that there will be similar cancellation effects for drift contributions, as both the numerator and denominator grow at *roughly* $r(t)$. However, we will not enjoy a perfect cancellation, as this quantity is not a Martingale under any standard pricing measure, let alone the spot measure. Indeed, the short rate $r(t)$ will appear in the growth rates of both the numerator and denominator and these effects will cancel, but there will remain the effect of the rolling maturities and the effect of convexity. Approximations with better accuracy can be obtained by replacing

$$\frac{P(t,t+\tau'_{i+1})\,\big(1+R_3(t,t+\tau'_i)\,\tau_3\big)}{L'(t\,;\tau'_{bs})} \quad\text{with}\quad \mathbb{E}_0\left[\frac{P(t,t+\tau'_{i+1})\,\big(1+R_3(t,t+\tau'_i)\,\tau_3\big)}{L(t\,;\tau'_{bs})}\right], \tag{153}$$

and approximating the expectation. A natural approach to this approximation is to write down the drift function of the quantity, replace any stochastic components with their spot-measure expectations, and integrate from 0 to t. For affine models, the expectations involved can be computed without difficulty.

Either of these approaches can of course be extended to the other quantities which contribute to the coefficients $\gamma_{0,3,0,f,k}(t\,;\tau'_{bs})$ and $\gamma_{0,3,3,f,k}(t\,;\tau'_{bs})$. We note however that the frozen-coefficients approach in (150) performs adequately for practical purposes, with approximate volatilities computed via this approach aligning well with simulated volatilities computed via the true dynamics; the relative error in the approximation is roughly 1%. In a similar vein, one can essentially ignore the base curve volatility when computing basis swap spread volatility as it contributes a trivial amount relative to basis spread curve volatility. This has practical implications for calibration, as it allows us to sequentialize the calibration; one may calibrate the parameters of the spread process to historical basis swap spread volatilities in isolation before then calibrating the parameters of the base curve to swaption volatilities.

The techniques described here can of course be adapted to computing correlations between basis swap spreads of different tenors, or swap rates of different tenors, so as to allow for the calibration of reversion coefficients $\kappa_{0,f,k}$ and $\kappa_{3,f,k}$. Note that what are of interest here may not be instantaneous correlations, but rather discrete-horizon correlations. Indeed, for one-factor models the instantaneous correlations between basis swap spreads at different tenors will of course be $+1$, but the discrete-horizon (*i.e.* 1Y) correlations will be $< +1$ (for non-zero reversions) and their values will respond to changing the reversion coefficients. Discrete-horizon correlations and related quantities can be computed by computing discrete-horizon co-moments of the form

$$\mathbb{E}\big[\big(\phi'_{0,3}(T;\,\tau'_{bs,1})\big)^m\,\big(\phi'_{0,3}(T;\,\tau'_{bs,2})\big)^n\big]\,, \tag{154}$$

which for affine models can be evaluated straightforwardly via repeated differentiation of the joint Fourier transform

$$\mathbb{E}\big[\exp\big(\theta_1\,\phi'_{0,3}(T;\,\tau'_{bs,1}) + \theta_2\,\phi'_{0,3}(T;\,\tau'_{bs,2})\big)\big]\,. \tag{155}$$

An alternative method for identifying reversion coefficients may be bestfit a set of basis swap spread volatilities across a range of tenors; reversion coefficients will of course govern the shape of this profile. Another alternative is to examine the auto-regression coefficients of tenor basis swap spreads.

In a similar spirit, one can use instantaneous correlations between swap rates and/or basis swap spreads at various tenors to calibrate correlations implicit in the diffusion function $\psi_{i,f,k}(t)\,\psi_{j,f,l}(t)$.

C Swaption Pricing

In this appendix we elaborate on the pricing of swaptions for affine model specifications, building upon the derivations and insights of the preceding appendix. To begin, we recall from (134) that the value of a swaption exercising at T_0 and struck at ω^* is given by

$$\frac{\mathrm{swpn}(t\,;T_0,\omega^*)}{L(t\,;T_0)} = \overline{\mathbb{E}}_t^{T_0}\big[(\omega_3(T_0) - \omega^*)^+\big] = \int_{\omega^*}^{\infty} (\omega - \omega^*)\,p_t(\omega)\,dw\,, \tag{156}$$

where $p_t(\cdot)$ is the time-t conditional density function of $\omega_3(T_0\,;T_0)$. Note that to avoid cluttered formulae, we denote $\omega_3(t\,;T_0)$ by the simpler $\omega_3(t)$ in this appendix, dropping the second argument corresponding to the start date of the underlying swap.

To evaluate the expectation in (156), we adopt the approach of Fang and Oosterlee ('08). They approximate $p_t(\cdot)$ as a cosine series with coefficients which depend upon the the Fourier Transform of $p_t(\cdot)$,

$$\hat{p}_t(\theta) = \overline{\mathbb{E}}_t^{T_0}\left[e^{i\theta\,\omega_3(T_0)}\right]. \tag{157}$$

To provide a concrete example of the technique, we consider the affine square-root formulation of the workhorse model in (8), as described in Section 2.2,

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\,dt + \nu_0(t)\,e^{-\kappa_0(T-t)}\,dW_0(t) \\ ds_3(t,T) &= \mu_3(t,T)\,dt + \nu_3(t)\left(s_3(t,t) - l_3(t)\right)^{\frac{1}{2}}\,e^{-\kappa_3(T-t)}\,dW_3(t)\,, \end{aligned} \tag{158}$$

where $\nu_0(t)$ and $\nu_3(t)$ are deterministic volatility profiles. Clearly, the ON curve has Gaussian dynamics while the ON-*vs.*-3M IBOR spread curve has square-root dynamics. Evaluating the formulae for the Markov representations in (104) and (62) from earlier appendices, one can readily derive Markov representations for $f_0(t,T)$ and $s_3(t,T)$, which will involve the state processes $X_0(t) = X_{0,0,1}(t)$ and $Y_0(t) = Y_{0,0,0,1,1}(t)$, and $X_3(t) = X_{3,3,1}(t)$ and $Y_3(t) = Y_{3,3,3,1,1}(t)$. We state these representations here for reference,

$$f_0(t,T) = f_0(0,T) + e^{-\kappa_0(T-t)}X_0(t) + A_0\,e^{-\kappa_0(T-t)}\big(e^{-\kappa_0(T-t)} - 1\big)Y_0(t)\,, \tag{159}$$

$$\begin{aligned} dX_0(t) &= \kappa_0\big(-A_0 Y_0(t) - X_0(t)\big)\,dt + \nu_0(t)\,dW_0(t) \\ dY_0(t) &= 2\kappa_0\big(\tfrac{1}{2\kappa_0}\nu_0(t)^2 - Y_0(t)\big)\,dt\,, \end{aligned} \qquad A_0 = -\frac{1}{\kappa_0}\,, \tag{160}$$

and

$$s_3(t,T) = s_3(0,T) + e^{-\kappa_3(T-t)}X_3(t) + A_3\,e^{-\kappa_3(T-t)}\big(e^{-\kappa_3(T-t)} - 1\big)Y_3(t)\,, \tag{161}$$

$$\begin{aligned} dX_3(t) &= \kappa_3\big(\tfrac{A_3}{\kappa_3}\,\nu_3(t)^2\left(s_3(0,t) - l_3(t)\right) - A_3 Y_3(t) - \big(1 - \tfrac{A_3}{\kappa_3}\,\nu_3(t)^2\big)X_3(t)\big)\,dt \\ &\qquad + \nu_3(t)\left(s_3(0,t) + X_3(t) - l_3(t)\right)^{\frac{1}{2}}\,dW_3(t) \\ dY_3(t) &= 2\kappa_3\big(\tfrac{1}{2\kappa_3}\nu_3(t)^2\left(s_3(0,t) - l_3(t)\right) + \tfrac{1}{2\kappa_3}\nu_3(t)^2 X_3(t) - Y_3(t)\big)\,dt\,, \end{aligned} \tag{162}$$

$$A_3 = \frac{\big(e^{-\kappa_3\tau_3} - 1\big)^2}{\kappa_3\big(e^{-2\kappa_3\tau_3} - 1\big)}\,.$$

Note that in the above we have invoked the fact that $s_3(t,t) = s_3(0,t) + X_3(t)$ in expressing the diffusion of $X_3(t)$. Note also that throughout the above we have used the simplified $\kappa_0 = \kappa_{0,0,1}$ and $\kappa_3 = \kappa_{3,3,1}$.

With the Markov dynamics in hand, we evaluate the approximate swap rate dynamics in (130),

$$\begin{aligned} d\omega_3(t\,;T_0) &\approx \gamma_0(T_0)\,\nu_0(t)\,e^{-\kappa_0(T_0-t)}\,d\overline{W}_0^{T_0}(t) \\ &\qquad + \gamma_3(T_0)\,\nu_3(t)\left(s_3(0,t) + X_3(t) - l_3(t)\right)^{\frac{1}{2}}\,e^{-\kappa_3(T_0-t)}\,d\overline{W}_3^{T_0}(t)\,, \end{aligned} \tag{163}$$

where

$$\begin{aligned}
\gamma_0(T_0) &= \frac{P(0,T_{i+1})\,R_3(0,T_i)\,\tau_3}{L(0\,;T_0)}\,\frac{1}{\kappa_0}\left(e^{-\kappa_0(T_{i+1}-T_0)} - \frac{P(0,T_{j+1})\,\tau_3}{L(0\,;T_0)}\,e^{-\kappa_0(T_{j+1}-T_0)}\right) \\
&\quad -\frac{P(0,T_{i+1})\,\big(1+R_3(0,T_i)\,\tau_3\big)}{L(0\,;T_0)}\,\frac{1}{\kappa_0}\big(e^{-\kappa_0\tau_3}-1\big)\,e^{-\kappa_0(T_i-T_0)}\,, \qquad (164)\\
\gamma_3(T_0) &= -\frac{P(0,T_{i+1})\,\big(1+R_3(0,T_i)\,\tau_3\big)}{L(0\,;T_0)}\,\frac{1}{\kappa_3}\big(e^{-\kappa_3\tau_3}-1\big)\,e^{-\kappa_3(T_i-T_0)}\,.
\end{aligned}$$

Noting that the process $[\omega_3(t), X_0(t), Y_0(t), X_3(t), Y_3(t)]^{\mathrm{T}}$ is clearly affine, we follow Duffie, Pan and Singleton ('00) and postulate an exponential-affine ansatz for the Fourier Transform in (157). Before proceeding, it is important to recognize that we require the system to be affine under the $\overline{\mathbb{P}}^{T_0}$-measure, but in the above we have presented the dynamics of $X_0(t)$, $Y_0(t)$, $X_3(t)$ and $Y_3(t)$ under the spot measure. Under the zero correlation case that we are considering, $dW_3(t) = d\overline{W}^{T_0}(t)$ and thus $X_3(t)$ does not undergo a drift adjustment. There is a (deterministic) drift adjustment for $X_0(t)$, but the process remains affine. We omit the details here because $X_0(t)$ and $Y_0(t)$ are in fact not relevant to the pricing of swaptions, owing to the fact that $X_0(t)$ does not appear explicitly in the dynamics of $\omega_3(t)$ in (163). Indeed, they can be left out of the analysis entirely, though the parameters of the ON curve, κ_0 and $\nu_0(t)$, of course still appear. In general, when $\psi_0(t)$ exhibits level dependence, $X_0(t)$ and $Y_0(t)$ would feature in the coming calculations.

Given the preceding discussion we postulate an ansatz for $\hat{p}_t(\theta)$ of the form

$$\hat{p}_t(\theta) = \exp\big(\beta_0(t\,;\theta) + \beta_1(t\,;\theta)\omega_3(t) + \beta_2(t\,;\theta)X_3(t) + \beta_3(t\,;\theta)Y_3(t)\big)\,. \qquad (165)$$

It can be demonstrated straightforwardly that the coefficient vector $[\beta_0(t\,;\theta), \beta_1(t\,;\theta), \beta_2(t\,;\theta), \beta_3(t\,;\theta)]^{\mathrm{T}}$ has a terminal condition of $[0, i\theta, 0, 0]$, and satisfies a system of ordinary differential equations,

$$\begin{aligned}
-d_t\beta_0(t) &= \frac{1}{2}\beta_1(t)^2\gamma_0(T_0)^2\nu_0(t)^2e^{-2\kappa_0(T_0-t)} + \Big(\beta_2(t)A_3 + \beta_3(t) \\
&\quad +\frac{1}{2}\beta_1(t)^2\gamma_3(T_0)^2\,e^{-2\kappa_3(T_0-t)} + \frac{1}{2}\beta_2(t)^2 \\
&\quad +\beta_1(t)\beta_2(t)\gamma_3(T_0)e^{-\kappa_3(T_0-t)}\Big)\big(s_3(0,t) - l_3(t)\big)\nu_3(t)^2 \\
-d_t\beta_1(t) &= 0 \qquad (166)\\
-d_t\beta_2(t) &= -\beta_2(t)\kappa_3 + \Big(\beta_2(t)A_3 + \beta_3(t) + \frac{1}{2}\beta_1(t)^2\gamma_3(T_0)^2e^{-2\kappa_2(T_0-t)} \\
&\quad +\frac{1}{2}\beta_2(t)^2 + \beta_1(t)\beta_2(t)\gamma_3(T_0)e^{-\kappa_3(T_0-t)}\Big)\nu_3(t)^2 \\
-d_t\beta_3(t) &= -\beta_2(t)A_3\kappa_3 - \beta_3(t)2\kappa_3\,.
\end{aligned}$$

Clearly $\beta_1(t) = i\theta$, and the remaining DEs can be solved by numerical methods (*eg.* Runge-Kutta). We note that it is trivial to parallelize the solution of these DEs across different frequencies θ, which can be useful in practice.

With $\hat{p}_t(\theta)$ in hand, we follow Fang and Oosterlee ('08) to use it toward an approximation of the valuation formula in (156).[10] First, we assume away the density outside of $\omega_3(T_0) \in [L, U]$ such that

$$i.\ \frac{\text{swpn}(t)}{L(t)} \approx \int_{\omega^*}^{U} (\omega - \omega^*) p_t(\omega)\, d\omega \quad \text{and} \quad ii.\ p_t(\omega) \approx \frac{\alpha_0}{2} + \sum_{j=1}^{N_J} \alpha_j \cos\left(\frac{j\pi(\omega - L)}{(U - L)}\right) \tag{167}$$

where

$$\alpha_j = \frac{2}{U-L}\text{Re}\left\{\exp\left(\frac{-ij\pi L}{U-L}\right)\hat{p}_t\left(\frac{j\pi}{U-L}\right)\right\} \approx \frac{2}{U-L}\int_{L}^{U} p_t(\omega)\cos\left(\frac{j\pi(\omega-L)}{(U-L)}\right)d\omega\,. \tag{168}$$

We combine these, substituting *ii.* into *i.*, and evaluate the remaining (simple) integrals for

$$\frac{\text{swpn}(t)}{L(t)} \approx \frac{\alpha_0}{2}\big((1/2)[U^2 - \omega^{*2}] - \omega^*[U - \omega^*]\big) + \sum_{j=1}^{N_J}\alpha_j I_j - \omega^* \sum_{j=1}^{N_J}\alpha_j J_j\,, \tag{169}$$

where

$$\begin{aligned} I_j &= \frac{U-L}{j\pi}\left[U\sin(j\pi) - \omega^*\sin\left(j\pi\frac{\omega^*-L}{U-L}\right)\right] + \frac{(U-L)^2}{(j\pi)^2}\left[\cos(j\pi) - \cos\left(j\pi\frac{\omega^*-L}{U-L}\right)\right], \\ J_j &= \frac{U-L}{j\pi}\left[\sin(j\pi) - \sin\left(j\pi\frac{\omega^*-L}{U-L}\right)\right]. \end{aligned} \tag{170}$$

In practice we choose a sensible starting value for the size of the restricted support, $U - L$, as some multiple of the swap rate volatility over the horizon $(0, T_0)$. We then fine tune $U - L$ and the number of frequencies N_J used via straightforward convergence analysis.

To assess the accuracy of the approximations presented here we present a brief, informal analysis. We consider a set of 5Y-into-5Y swaptions with absolute moneyness levels ranging from -200bps to +200bps, as such a range is what is typically quoted in practice. The model specification considered here is the working model presented in (158) and we use the market data and parameters calibrated for the example presented in Section 3.3. We present the calibrated parameters momentarily.

We compute swaption values produced via the approximations developed within this appendix, and we compare them with values produced off a simulation under the true dynamics. Regarding the latter, we do not simply simulate the swap rate off the approximate dynamics in (130). Rather, we perform a risk-neutral simulation of $[X_0(t), Y_0(t)]$ and $[X_3(t), Y_3(t)]$ according to their dynamics in (160) and (162). These simulated states are then used to produce swap values at expiry, which in turn are floored at zero and discounted off money-market account computed with use of the simulated $X_0(t)$ (*i.e.* $r(t) = f_0(0,t) + X_0(t)$).

Any inaccuracy in the approximate values will owe to *a*) the freezing of swap rate coefficients, *b*) the use of the COS method (*e.g.* truncation error and limiting the number of frequencies), and *c*) the use of a numerical scheme to solve the ODEs for the Fourier Transform coefficients. The only

[10] Direct Fourier inversion is possible, as in Duffie, Pan and Singleton ('00), and was used in a similar context by Schrager and Pelsser ('06). However, the COS method of Fang and Oosterlee ('08) enjoys favourable convergence properties, as discussed in their paper.

source of error we cannot control is that owing to the freezing of coefficients, as is standard for this method. To isolate this source of error, we use a high-quality COS approximation with $N_J = 251$ frequencies and a range $U - L$ of 15 times the approximate swap rate volatility (computed as per the previous appendix). Similarly, we us a high-quality simulation to obtain an accurate basis for comparison. The number of paths used was 1,000,000 and the timestep was set to $\Delta t = 0.02$, or roughly one business week (we also placed simulation steps at base-curve and spread-curve interpolation knots).

Figure 2 presents the two sets of swaption values. The values plotted with the dotted line are produced via simulation and those with the dashed line were produced via the frozen-coefficients COS approximation. The two valuation profiles are all but indistinguishable so we present also the relative error in implied volatilities in Figure 3. The greatest relative error is 0.5% (in magnitude) and this occurs at the most extreme strike considered. Such accuracy is sufficient for practical use, *i.e.* for calibrating the model and producing swaption exposures for CVA or PFE purposes. We note that for the ATM volatility the relative error is quite small at roughly -0.15%, which translates to an error of roughly -0.1 bps in the actual ATM implied volatility of 66.76 bps. As noted in the context of Figure 1 in Section 3.3, we found such accuracy to hold for all ATM volatilities across the 10Y co-terminal swaptions used for calibration.

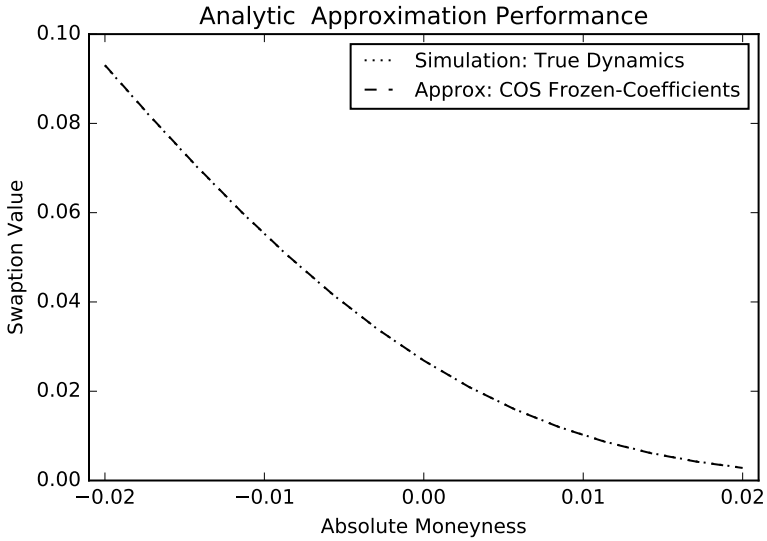

Figure 2: Accuracy of the analytical swaption pricing approximation: True and approximate values of 5Y-*vs*.-5Y swaptions with moneyness levels in [-200bs, +200bps].

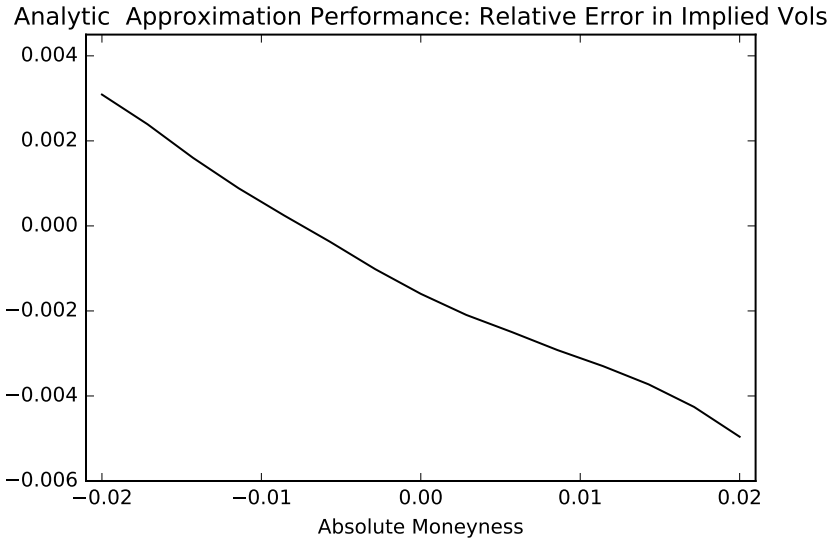

Figure 3: Accuracy of the analytical swaption pricing approximation: Relative implied volatility approximation errors of 5Y-*vs*.-5Y swaptions with moneyness levels in [-200bs, +200bps].

While we are examining the behavior of swaption prices under our model, we briefly examine the impact of the square-root specification on the implied volatility skew. To this end, we consider three cases.

1. The case where the historical volatility of the 2Y ON-*vs*.-3M IBOR tenor basis swap spread volatility was 22.60 bps.

2. A high-volatility case where this volatility was 1.5×22.60 bps.

3. A low-volatility case where this volatility was 0.5×22.60 bps.

We adopt the same calibration strategy for each case and use the resulting models to produce the corresponding volatility skews. These skews are of course generated via a full simulation to ensure that the resulting skews are not affected by the freezing of swap rate coefficients. Figure 4 presents the skews for the three cases along with the observed market skew for the 11-Oct-'19. As is seen, the skew for the base case is quite mild, and despite the fact that we have made dramatic changes to the target spread volatility, the resulting changes in the skew are small as compared with the different between the base-case skew and the observed market skew. This owes primarily to the fact that the volatility of the basis spread curve is dominated by that of the base curve in determining the swap rate volatility; the historical 2Y-tenor OIS rate volatility is 63.87 bps. As such, the effect upon swap rate tail behavior from adding a square-root $s_3(t, T)$ to a Gaussian $f_0(t, T)$ is muted. These observations are consistent with comments made in Section 3.3 regarding calibration regarding the use of skew data alone to calibrate the parameters of spread curves being problematic.

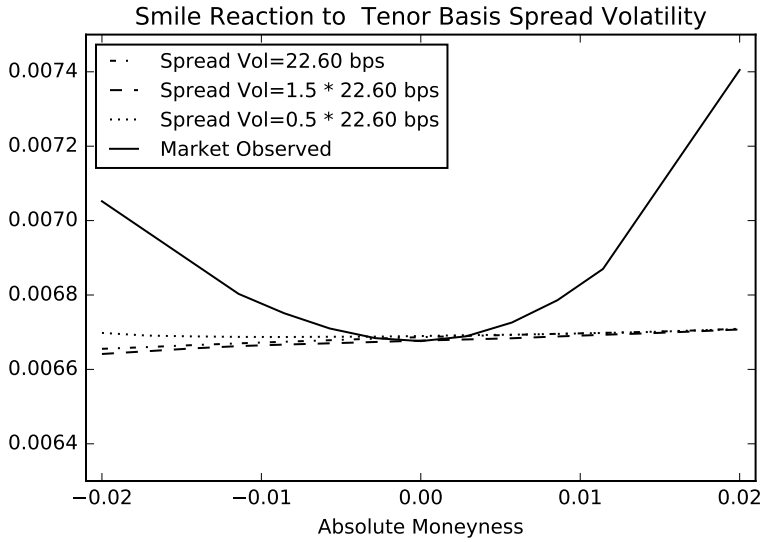

Figure 4: Skew profiles for the model calibrated under different values for historical basis swap spread volatility, with observed market skew for reference.

For reference, we present the calibrated parameters for the model of $s_3(t,T)$ in (158).

Parameter Estimate	Value
$\hat{\kappa}_3$	1.18
$\hat{\nu}_3$	0.0720
$\hat{l}_3^{av}$	-0.00394

Table 1: Calibrated parameters for the single-factor square-root model in (158).

We note that $\hat{l}_3^{av}$ is the target level of the lower bound $l_3(t)$, as is introduced and discussed thoroughly in the following appendix. Something else which is introduced in the following appendix is a Feller condition, and we confirm now that $\hat{\kappa}_3$ satisfies this condition.

D Lower Bounds & Feller Condition

In the appendix we focus on the lower bound of the ON-*vs.*-3M IBOR spread process under the case of CEV dynamics. We begin with a discussion of how the lower bound of the short spread $s_3(t,t)$ controls the lower bound of the associated ON-*vs.*-3M IBOR basis swap spreads, and discuss how this is relevant to calibration efforts. We then derive a condition for the lower bound which must be satisfied to ensure that the short spread is well behaved. Finally, we derive a Feller condition on the parameters of the spread process under the square-root specification.

For this analysis, we will work with the workhorse single-factor CEV model presented earlier,

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\,dt + \nu_0(t)\,e^{-\kappa_0(T-t)}\,dW_0(t) \\ ds_3(t,T) &= \mu_3(t,T)\,dt + \nu_3(t)\left(s_3(t,t) - l_3(t)\right)^{\beta} e^{-\kappa_3(T-t)}\,dW_3(t)\,. \end{aligned} \tag{171}$$

Recalling that the Markov representation for $s_3(t,T)$ under this model, as presented in (19), is

$$s_3(t,T) = s_3(0,T) + e^{-\kappa_3(T-t)}X_3(t) + A_3\, e^{-\kappa_3(T-t)}\big(e^{-\kappa_3(T-t)} - 1\big)Y_3(t)\,, \tag{172}$$

we may utilize the relation $s_3(t,t) = s_3(0,t) + X_3(t)$ to obtain a relation between $s_3(t,T)$ and $s_3(t,t)$,

$$s_3(t,T) = \big(s_3(0,T) - e^{-\kappa_3(T-t)}s_3(0,t)\big) + e^{-\kappa_3(T-t)}s_3(t,t) + A_3\, e^{-\kappa_3(T-t)}\big(e^{-\kappa_3(T-t)} - 1\big)Y_3(t)\,. \tag{173}$$

To derive a lower bound on $s_3(t,T)$, we first note that $s_3(t,t) \geq l_3(t)$ for a suitably-defined $l_3(t)$. We then note that $A_3Y_3(t) \leq 0$ (as demonstrated momentarily) and $\big(e^{-\kappa_3(T-t)} - 1\big) < 0$, such that the term involving $Y_3(t)$ is non-negative. As such, we have that

$$s_3(t,T) \geq \big(s_3(0,T) - e^{-\kappa_3(T-t)}s_3(0,t)\big) + e^{-\kappa_3(T-t)}l_3(t)\,. \tag{174}$$

We note in passing that $l_3(t)$ is a lower bound only for $s_3(t,t)$, and not for all of $s_3(t,T)$. Indeed, $s_3(t,T) \leq l_3(t)$ are clearly possible, *e.g.* when $s_3(0,T)$ is steeply downward sloping such that $s_3(0,T) \ll s_3(0,t)$.

Regarding the choice of $l_3(t)$, note that we are not interested in a bound on $s_3(t,T)$ *per se*, but rather a bound on the tenor basis swap spread, $\phi_{0,3}(t\,;\tau'_{bs})$, where we fix a particular swap tenor τ'_{bs}. Indeed, it is this quantity that we observe in practice and for which we can establish an empirical lower bound.

To analyze the bound for the basis swap spread, we use the approximation that

$$\phi_{0,3}(t\,;\tau'_{bs}) \approx \frac{1}{\tau'_{bs}}\int_t^{t+\tau'_{bs}} s_3(t,u)\,du\,, \tag{175}$$

and we are able to obtain an inequality for this latter quantity,

$$\frac{1}{\tau'_{bs}}\int_t^{t+\tau'_{bs}} s_3(t,u)\,du \geq \frac{1}{\tau'_{bs}}\left(\int_t^{t+\tau'_{bs}} s_3(0,u)\,du + \frac{\big(e^{-\kappa_3(T-t)} - 1\big)}{\kappa_3}s_3(0,t) - \frac{\big(e^{-\kappa_3(T-t)} - 1\big)}{\kappa_3}l_3(t)\right). \tag{176}$$

As we would like to enforce this over the entire simulation horizon (*i.e.* for arbitrary t) we use the approximation that $s_3(0,t) \approx s_3(0,T) \approx s_3^{av}$, where s_3^{av} is the average of the initial instantaneous forward spread curve ($s_3(0,t)$) over the entire simulation horizon. For similar reasons we use the same approximation for the lower bound, $l_3(t) \approx l_3^{av}$.

With this, we are have that

$$\frac{1}{\tau'_{bs}}\int_t^{t+\tau'_{bs}} s_3(t,u)\,du \geq \frac{1}{\tau'_{bs}}\left(\tau'_{bs} + \frac{\big(e^{-\kappa_3(T-t)} - 1\big)}{\kappa_3}\right)s_3^{av} - \frac{1}{\tau'_{bs}}\frac{\big(e^{-\kappa_3(T-t)} - 1\big)}{\kappa_3}l_3^{av}\,, \tag{177}$$

and by setting the RHS to zero we are able to solve for the ratio of the average lower bound and the average initial forward spread required for non-negative τ'_{bs}-tenor basis swap spreads,

$$\frac{l_3^{av}}{s_3^{av}} = \frac{1 - e^{-\kappa_3\tau'_{bs}} - \kappa_3\tau'_{bs}}{1 - e^{-\kappa_3\tau'_{bs}}}\,. \tag{178}$$

During calibration, this multiplier can be used to fix the average level lower bound. Such a calibration was used for the model which produced the simulated basis swap spreads of 2Y tenor in Figure 1. It is clear from this figure that a lower bound of (approximately) zero has been imposed.

We turn now to an additional restriction upon the behavior of the lower bound $l_3(t)$. First, we know from earlier discussion that $s_3(t,t) = s_3(0,t) + X_3(t)$ and thus $ds_3(t,t) = \dot{s}_3(0,t)\,dt + dX_3(t)$. As such, while we have that at $s_3(t,t) = l_3(t)$ there can be no diffusion to a lower value, $s_3(t,t)$ can be driven lower by $\dot{s}_3(0,t)$. At this point the diffusion function would not be well defined. Hence, there must be a relationship between $l_3(t)$ and $\dot{s}_3(0,t)$ to preclude this. Indeed, as we will establish here, the lower bound must satisfy

$$\dot{l}_3(t) \leq \dot{s}_3(0,t) + \kappa_3\big(s_3(0,t) - l_3(t)\big)\,. \tag{179}$$

Given this, it not possible to specify a constant lower bound of, say, l_3^{av}, using the notation introduced earlier; the lower bound must be flexible enough to track $s_3(0,t)$ when it is (sufficiently) downward sloping. However, one can obtain an $l_3(t)$ which respects (179), only deviating from l_3^{av} to track $s_3(0,t)$ downward when necessary, and returning to l_3^{av} as quickly as permitted by (179) when $s_3(0,t)$ is increasing.

To begin, we write down the dynamics of the short spread process, $s_3(t,t)$, where we again adopt the simple workhorse model. Recalling that $s_3(t,t) = s_3(0,t) + X_3(t)$ and invoking the dynamics for $X_3(t)$ and $Y_3(t)$ presented in (20) yields

$$\begin{aligned} ds_3(t,t) &= \kappa_3\big(\tfrac{1}{\kappa_3}\dot{s}_3(0,t) + s_3(0,t) + \tfrac{A_3}{\kappa_3}\nu_3(t)^2\big(s_3(t,t) - l_3(t)\big)^{2\beta} \\ &\qquad -A_3Y_3(t) - s_3(t,t)\big)\,dt + \nu_3(t)\big(s_3(t,t) - l_3(t)\big)^{\beta}\,dW_3(t) \\ dY_3(t) &= 2\kappa_3\big(\tfrac{1}{2\kappa_3}\nu_3(t)^2\big(s_3(t,t) - l_3(t)\big)^{2\beta} - Y_3(t)\big)\,dt\,, \end{aligned} \tag{180}$$

where A_3 is as per earlier,

$$A_3 = \frac{\big(e^{-\kappa_3\tau_3} - 1\big)^2}{\kappa_3\big(e^{-2\kappa_3\tau_3} - 1\big)}\,. \tag{181}$$

What we are concerned with here is whether the argument of the CEV diffusion function, $s_3(t,t) - l_3(t)$, remains non-negative. As such, we introduce a process $Z_3(t) = s_3(t,t) - l_3(t)$, which is simply the distance of the short spread from its intended lower bound, and investigate its behavior at $Z_3(t) = 0$. Clearly we have that $dZ_3(t) = ds_3(t,t) - \dot{l}_3(t)\,dt$, which evaluates to

$$\begin{aligned} dZ_3(t) &= \kappa_3\big(\tfrac{1}{\kappa_3}\big(\dot{s}_3(0,t) - \dot{l}_3(t)\big) + \big(s_3(0,t) - l_3(t)\big) + \tfrac{A_3}{\kappa_3}\nu_3(t)^2 Z_3(t)^{2\beta} \\ &\qquad -A_3Y_3(t) - Z_3(t)\big)\,dt + \nu_3(t)Z_3(t)^{\beta}\,dW_3(t) \\ dY_3(t) &= 2\kappa_3\big(\tfrac{1}{2\kappa_3}\nu_3(t)^2 Z_3(t)^{2\beta} - Y_3(t)\big)\,dt\,. \end{aligned} \tag{182}$$

If we inspect the drift function of $Z_3(t)$ at $Z_3(t) = 0$, we obtain

$$\kappa_3\big(\frac{1}{\kappa_3}\big(\dot{s}_3(0,t) - \dot{l}_3(t)\big) + \big(s_3(0,t) - l_3(t)\big) - A_3Y_3(t)\big)\,. \tag{183}$$

We can ignore the presence of $-A_3Y_3(t)$ by virtue of the fact that $Y_3(t) \geq 0$ as long as $l_3(s)$, $s \leq t$, has been chosen such that $Z_3(s) \geq 0$, $s \leq t$, and that $-A_3 > 0$ for $\kappa_3 > 0$. Given this, it is clear that the drift function of $Z_3(t)$ is non-negative at $Z_3(t) = 0$ provided that the inequality

$$\dot{l}_3(t) \leq \dot{s}_3(0,t) + \kappa_3\big(s_3(0,t) - l_3(t)\big) \tag{184}$$

is satisfied. We clearly expect to encounter situations where $\dot{s}_3(0,t) < 0$ for some values of t, and thus it is clear that $\dot{l}_3(t)$ must allow for occasional downward deviations to cater to this. Intuitively, if $s_3(t,t)$ has already reached $l_3(t)$, the presence of $\dot{s}_3(0,t)$ in its drift can drive it to lower values, but this is acceptable as long as $l_3(t)$ is decreasing at the same rate.

As alluded to earlier, while we cannot simply impose a constant lower bound of, say, l_3^{av}, we are able to select an $l_3(t)$ satisfying (184) with just enough flexibility to track $s_3(0,t)$ downward when necessary, but which reverts to a target level of l_3^{av} when $s_3(0,t)$ is not decreasing.

Before discussing this solution, consider the extreme case where we set $l_3(0) = l_3^{av}$ and let $l_3(t)$ evolve as

$$\dot{l}_3(t) = \dot{s}_3(0,t) + \kappa_3\big(s_3(0,t) - l_3(t)\big)\,, \tag{185}$$

i.e. where we enforce equality in (184). In this case, one would inevitably encounter a situation where $l_3(t)$ would track $s_3(0,t)$ when it is increasing, even if this means moving above l_3^{av}. Moreover, eventually via a reversion effect, one would observe $l_3(t) = s_3(0,t)$ at some value of t depending on κ_3. This is clearly not what we desire. Indeed, we must stop $l_3(t)$ tracking $s_3(0,t)$ upward or reverting toward $s_3(0,t)$ when $l_3(t) \geq l_3^{av}$.

To obtain such an $l_3(t)$, one can set $l_3(0) = l_3^{av}$ and let it evolve by computing

$$\dot{l}'_3(t) = \big(\dot{s}_3(0,t) + \kappa_3\big(s_3(0,t) - l_3(t)\big)\big)\,, \tag{186}$$

and then setting

$$\dot{l}_3(t) = \dot{l}'_3(t)\,\big(1 - 1\{l_3(t) = l_3^{av}\}\,1\{\dot{l}'_3(t) \geq 0\}\big)\,. \tag{187}$$

In essence, we allow $l_3(t)$ to evolve according to the RHS of (184), provided that when it is sitting at l_3^{av}, this evolution does not take it above l_3^{av}. This approach easily generalizes to cases where the initial instantaneous spread curve $s_3(0,t)$ is non-differentiable (or discontinuous), *i.e.* where piecewise linear (or constant) interpolation schemes are employed.

With a suitable $l_3(t)$ in hand, we now turn to the question of how the process $s_3(t,t)$ behaves in the vicinity of that lower bound. The question, for our purposes, is essentially one of whether the bound is attainable. It is well known that the analysis of Feller ('51) can be applied in square-root case, $\beta = \frac{1}{2}$, as is done *e.g.* for the well known Cox, Ingersol and Ross ('85) and Heston ('93) models. The same holds true for our model, though we do encounter one mild complication. This owes primarily to the presence of the state-dependent convexity term in the drift of the spread process, which for the case of $\beta = \frac{1}{2}$, is

$$\frac{A_3}{\kappa_3}\nu_3(t)^2\big(s_3(t,t) - l_3(t)\big)\,. \tag{188}$$

To begin, we transform the process $s_3(t,t)$ into the form of a classical square-root diffusion. The transformed process $Z_3(t) = s_3(t,t) - l_3(t)$ indeed satisfies this requirement, as becomes clear upon evaluating (182) at $\beta = \frac{1}{2}$,

$$\begin{aligned} dZ_3(t) &= \kappa_3\big(\tfrac{1}{\kappa_3}\big(\dot{s}_3(0,t) - \dot{l}_3(t)\big) + \big(s_3(0,t) - l_3(t)\big) + \tfrac{A_3}{\kappa_3}\nu_3(t)^2 Z_3(t) \\ &\qquad -A_3 Y_3(t) - Z_3(t)\big)\,dt + \nu_3(t) Z_3(t)^{\frac{1}{2}}\,dW_3(t) \\ dY_3(t) &= 2\kappa_3\big(\tfrac{1}{2\kappa_3}\nu_3(t)^2 Z_3(t) - Y_3(t)\big)\,dt\,. \end{aligned} \tag{189}$$

The dynamics of our $Z_3(t)$ specifically can now be written as

$$\begin{aligned} dZ_3(t) &= \kappa_3\big(\tfrac{1}{\kappa_3}\big(\dot{s}_3(0,t) - \dot{l}_3(t)\big) + \big(s_3(0,t) - l_3(t)\big) \\ &\qquad -A_3Y_3(t) - \big(1 - \tfrac{A_3}{\kappa_3}\nu_3(t)^2\big)Z_3(t)\big)\,dt + \nu_3(t)Z_3(t)^{\frac{1}{2}}\,dW_3(t)\,, \end{aligned} \tag{190}$$

or rather, as something closer to classical square-root formulation,

$$dZ_3(t) = \alpha_3(t)\big(\theta_3(t) - A_3Y_3(t) - Z_3(t)\big)\,dt + \nu_3(t)Z_3(t)^{\frac{1}{2}}\,dW_3(t)\,, \tag{191}$$

where

$$\begin{aligned} \alpha_3(t) &= \kappa_3\big(1 - \tfrac{A_3}{\kappa_3}\nu_3(t)^2\big) \\ \theta_3(t) &= \tfrac{1}{\kappa_3\left(1-\frac{A_3}{\kappa_3}\nu_3(t)^2\right)}\big(\big(\dot{s}_3(0,t) - \dot{l}_3(t)\big) + \kappa_3\big(s_3(0,t) - l_3(t)\big)\big)\,. \end{aligned} \tag{192}$$

As was done earlier, we can ignore the presence of $-A_3Y_3(t) \geq 0$, as this will only aid in forcing $Z_3(t)$ upwards, away from its lower bound of zero. With this, to ensure that the lower bound is unattainable, we are able to apply the Feller condition for square-root processes with time-varying coefficients, as presented in Benhamou, Gobet and Miri ('10),

$$\frac{2\,\alpha_3(t)\,\theta_3(t)}{\nu_3(t)^2} > 1\,,\ \ \forall t\,, \tag{193}$$

which evaluates to

$$\frac{2\,\big(\big(\dot{s}_3(0,t) - \dot{l}_3(t)\big) + \kappa_3\big(s_3(0,t) - l_3(t)\big)}{\nu_3(t)^2} > 1\,,\ \forall t\,, \tag{194}$$

in terms of our original parameters. In practice, we can simply select κ_3 according to

$$\kappa_3 > \max_t\ \left\{\frac{\nu_3(t)^2}{2\big(s_3(0,t) - l_3(t)\big)} - \frac{\big(\dot{s}_3(0,t) - \dot{l}_3(t)\big)}{\big(s_3(0,t) - l_3(t)\big)}\right\}. \tag{195}$$

E Developing the Model of Mean Spread Curves

As discussed in Section 2, an equivalent modeling approach to the one developed thus far is to work with mean spreads. To begin, we focus on the two-curve setup involving just the ON curve and the 3M IBOR curve. We will extend to the 6M IBOR curve later in the appendix. The mean spread curve for the ON-*vs.*-3M IBOR curve was defined earlier as

$$\bar{s}_3(t,T) = \frac{1}{\tau_3}\int_T^{T+\tau_3} s_3(t,u)\,du\,, \tag{196}$$

where we would use dynamics for $\bar{s}_3(t,T)$ along the lines of those used for $s_3(t,T)$ in (33),

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\,dt + \psi_{0,f,k}(t)\,e^{-\kappa_{0,f,k}(T-t)}\,dW_{0,f}(t) \\ d\bar{s}_3(t,T) &= \bar{\mu}_3(t,T)\,dt + \psi_{3,f,k}(t)\,e^{-\kappa_{3,f,k}(T-t)}\,dW_{3,f}(t)\,. \end{aligned} \tag{197}$$

Forward 3M IBORs would be constructed from this process via

$$R_3(t,T) = \frac{1}{\tau_3}\Big(e^{\int_T^{T+\tau_3} f_0(t,u)\,du + \tau_3\,\bar{s}_3(t,T)} - 1\Big)\,, \tag{198}$$

which is a direct analogy to (3).

Formally, $\bar{s}_3(t,T)$ is a fictitious discrete-tenor spread which enters additively with integrated ON rates, and its dynamics will govern observed ON-*vs*.-3M IBOR tenor basis swap spreads. Indeed, just as it was the case that $dR_3(t,T) \approx df_0(t,T) + ds_3(t,T)$, it is also true that $dR_3(t,T) \approx df_0(t,T) + d\bar{s}_3(t,T)$, as one might expect given that $d\bar{s}_3(t,T) \approx ds_3(t,T)$ for appropriately-modified parameters.

In this appendix we will sketch the derivation of the drift function of $\bar{s}_3(t,T)$, and we will briefly discuss the derivation of the corresponding Markov representation and approximate swap rate dynamics. There is a considerable overlap with the derivations already presented for the model of $s_3(t,T)$ and thus we do not linger long on the details. Indeed, implementations of the two models would be essentially the same. In the following appendices however we will explore extensions of the model, namely the addition of a jump component and the allowance for time-varying mean reversions. It will become apparent that moving from $s_3(t,T)$ to $\bar{s}_3(t,T)$ is advantageous under these advanced settings, as handling the drift function of $s_3(t,T)$ becomes more challenging.

Proceeding as per (34)-(42), it is straightforward to demonstrate the drift function $\mu_3(t,T)$ satisfies

$$\begin{aligned}\bar{\mu}_3(t,T) &= \bar{\psi}_{3,3,f,k,l}(t)\, A_{3,3,3,f,k,l}\, \bar{K}_{3,3,f,k,l}(T-t) \\ &\quad + \bar{\psi}_{3,0,f,k,l}(t)\, A_{3,3,0,f,k,l}\, \bar{K}_{3,0,f,k,l}(T-t) \\ &\quad + \bar{\psi}_{3,0,f,k,l}(t)\, B_{3,3,f,k,l}\, K_{3,f,k}(T-t)\,,\end{aligned} \tag{199}$$

for

$$A_{3,3,3,f,k,l} = -\frac{1}{2}\tau_3\,, \quad A_{3,3,0,f,k,l} = -\frac{1}{\kappa_{0,f,l}} \quad \text{and} \quad B_{3,3,f,k,l} = \frac{1}{\kappa_{0,f,l}}\,. \tag{200}$$

Note that in (42) we have that $\mu_3(t,T)$ enters under an integral, while we have here that $\bar{\mu}_3(t,T)$ is determined conclusively. This owes of course to the fact that $\bar{s}_3(t,T)$ enters (198) directly, as opposed to the case of $s_3(t,T)$ which enters (3) under an integral.

The form in (199) parallels that in (54) for $\mu_3(t,T)$. As such it is entirely trivial to to derive a Markov representation for $\bar{s}_3(t,T)$. Indeed, the manipulations presented in (58)-(62) can be applied straightforwardly to obtain a representation with the same structure. By extension, discount factors and 3M IBORs will clearly have the same structure as for the case of $s_3(t,T)$, being compositions of exponential-affine functions of the Markov states (discount factors will of course be identical). It follows immediately that the frozen-coefficients approach can be employed with minimal modifications. We omit the details are they are near-to identical to those presented in earlier appendices.

We proceed now to extend the model to multiple spread curves. To simplify matters, we consider only an extension to the 6M IBOR curve, avoiding the fully-general case. We do this to simplify the exposition and demonstrate the core result. It will be clear from our analysis that incorporating additional spread curves is straightforward, but would require the development of some additional notational conventions.

The equivalent extension of the IBOR formula from (3) to (17) becomes

$$R_6(t,T) = \frac{1}{\tau_3}\Big(e^{\int_T^{T+\tau_6} f_0(t,u)\,du + \tau_3\,(\bar{s}_3(t,T) + \bar{s}_3(t,T+\tau_3)) + \tau_6 \bar{s}_6(t,T)} - 1\Big)\,. \tag{201}$$

Note that we require the sum of $\bar{s}_3(t,T)$ and $\bar{s}_3(t,T+\tau_3)$ to ensure that the ON-*vs*.-3M IBOR spread for the entire coverage period of the 6M IBOR is present so as to ensure a natural alignment

of the 3M IBOR and 6M IBOR curves. Indeed,

$$\tau_3 \left(\bar{s}_3(t,T) + \bar{s}_3(t,T+\tau_3) \right) = \int_T^{T+\tau_6} s_3(t,u)\, du \,, \tag{202}$$

and it is this quantity which is present in the formulation for $s_3(t,T)$ and $s_6(t,T)$ in (17). To simplify the presentation, we introduce the notation $\bar{s}_{3,6}(t,T) = \bar{s}_3(t,T) + \bar{s}_3(t,T+\tau_3)$ to denote the relevant sum of spreads required for the construction of 6M IBORs, and we will use the notation $\bar{\mu}_{3,6}(t,T) = \bar{\mu}_3(t,T) + \bar{\mu}_3(t,T+\tau_3)$ and $\sigma_{3,6,f}(t,T) = \sigma_{3,f}(t,T) + \sigma_{3,f}(t,T+\tau_3)$ to denote the compositions of the drift and volatility functions; in a moment we will make use of the fact that

$$d\bar{s}_{3,6}(t,T) = \bar{\mu}_{3,6}(t,T)\, dt + \sigma_{3,6,f}(t,T)\, dW_{3,f}(t) \,. \tag{203}$$

For the dynamics of $\bar{s}_6(t,T)$ we will of course use dynamics analogous to those in (197) for $\bar{s}_3(t,T)$,

$$d\bar{s}_6(t,T) = \bar{\mu}_6(t,T)\, dt + \psi_{6,f,k}(t)\, e^{-\kappa_{6,f,k}(T-t)}\, dW_{6,f}(t) \,. \tag{204}$$

Proceeding as in (66)-(71) it is straightforward to derive a restriction on $\bar{\mu}_6(t,T)$. Indeed, it can easily be demonstrated that the drift of $\bar{\mu}_6(t,T)$ must satisfy

$$\begin{aligned} \bar{\mu}_6(t,T) \;=\; & -\frac{\tau_3}{\tau_6}\, \bar{\mu}_{3,6}(t,T) - \frac{1}{2}\frac{\tau_3^2}{\tau_6}\, \sigma_{3,6,f}(t,T)\, \sigma_{3,6,f}(t,T) - \frac{1}{2}\tau_6\, \sigma_{6,f}(t,T)\, \sigma_{6,f}(t,T) \\ & -\tau_3\, \sigma_{3,6,f}(t,T)\, \sigma_{6,f}(t,T) \\ & +\frac{\tau_3}{\tau_6}\, \sigma_{3,6,f}(t,T) \left(\int_t^T \sigma_{0,f}(t,u)\, du \right) + \sigma_{6,f}(t,T) \left(\int_t^T \sigma_{0,f}(t,u)\, du \right) . \end{aligned} \tag{205}$$

What is relevant here is that all of the variance-covariance terms reduce to quantities proportional to the combinations $\bar{\psi}_{i,j,f,k,l}(t)\, \bar{K}_{i,j,f,k,l}(T-t)$, including those of $\sigma_{3,6,f}(t,T)$ as they are simply sums over $\psi_{3,f,k}(t)\, e^{-\kappa_{3,f,k}(T-t)}$ and $\psi_{3,f,k}(t)\, e^{-\kappa_{3,f,k}(T+\tau_3-t)}$. Moreover, it has already been demonstrated the $\bar{\mu}_3(t,T)$ has this structure in (199), and thus so must $\bar{\mu}_{3,6}(t,T)$. As such, it becomes clear that $\bar{\mu}_6(t,T)$ has a solution of the form in (73),

$$\bar{\mu}_6(t,T) = \bar{\psi}_{i,j,f,k,l}(t)\, A_{6,i,j,f,k,l}\, \bar{K}_{i,j,f,k,l}(T-t) + \bar{\psi}_{i,0,f,k,l}(t)\, B_{6,i,f,k,l}\, K_{i,f,k}(T-t) \,, \tag{206}$$

for $i = 3,\, 6$ and $j = 0,\, 3,\, 6$, and where $A_{6,i,j,f,k,l}$ and $B_{6,i,f,k,l}$ depend upon $A_{3,i,j,f,k,l}$ and $B_{3,i,f,k,l}$ in a recursive manner. It is straightforward to demonstrate that this structure will persist as one proceeds to incorporate additional spread curves.

With this, we are ensured that Markov representations are available, as are other useful quantities (*e.g.* approximate basis swap spreads) that these afford.

F Time-Varying Reversion Coefficients

In this appendix we extend the dynamics of spread curves to allow for time-varying mean reversions. We begin with the mean spread curve $\bar{s}_3(t,T)$ and demonstrate that the extension is straightforward for this case. We then turn to the instantaneous spread curve $s_3(t,T)$ where, in contrast to the earlier case, a hurdle is encountered and we are forced to exert additional effort in obtaining a solution for the drift function.

With time-varying mean reversions, the dynamics in (197) become

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\,dt + \psi_{0,f,k}(t)\,e^{-\int_t^T \kappa_{0,f,k}(w)\,dw}\,dW_{0,f}(t) \\ d\bar{s}_3(t,T) &= \bar{\mu}_3(t,T)\,dt + \psi_{3,f,k}(t)\,e^{-\int_t^T \kappa_{3,f,k}(w)\,dw}\,dW_{3,f}(t)\,. \end{aligned} \tag{207}$$

Following the logic described earlier, we easily obtain a solution for the drift function. First, let us generalize the definitions of $K_{i,f,k}(T-t)$ and $\bar{K}_{i,j,f,k,l}(T-t)$ to

$$K_{i,f,k}(t,T) = e^{-\int_t^T \kappa_{i,f,k}(w)\,dw} \text{ and } \bar{K}_{i,j,f,k,l}(t,T) = e^{-\int_t^T (\kappa_{i,f,k}(w)+\kappa_{j,f,l}(w))\,dw}\,, \tag{208}$$

where we note that the earlier relationships $K_{i,f,k}(t,T) = K_{i,f,k}(t,S)\,K_{i,f,k}(S,T)$ of course still hold. In what follows it will be useful to have compact notation for the integral of $K_{0,f,k}(0,T)$ *w.r.t.* T, which enters the analysis via the volatility function for the integrated short rate. To this end we introduce

$$\Sigma_{i,f,k}(T) = \int_0^T K_{i,f,k}(0,u)\,du\,. \tag{209}$$

With these notational conveniences we are able to write the form of $\bar{\mu}_3(t,T)$ as

$$\bar{\mu}_3(t,T) = -\tau_3\,\bar{\psi}_{3,3,f,k,l}(t)\,\bar{K}_{3,3,f,k,l}(t,T) + \bar{\psi}_{3,0,f,k,l}(t)\,\bar{K}_{3,0,f,k,l}(t,T)\,\frac{\big(\Sigma_{0,f,l}(T) - \Sigma_{0,f,l}(t)\big)}{K_{0,f,l}(0,T)}\,. \tag{210}$$

It is straightforward to confirm that this collapses to the form in (199) which prevailed in the constant-reversion case. We state briefly the Markov representation for $\bar{s}_3(t,T)$ under time-varying reversion,

$$\begin{aligned} \bar{s}_3(t,T) &= \bar{s}_3(0,T) + K_{3,f,k}(t,T)\,X_{3,f,k}(t) - \frac{\tau_3}{2}\big(\bar{K}_{3,3,f,k,l}(t,T) - K_{3,f,k}(t,T)\big)\,Y_{3,3,f,k,l}(t) \\ &\quad + \frac{\big(\Sigma_{0,f,l}(T) - \Sigma_{0,f,l}(t)\big)}{K_{0,f,l}(0,t)}\,K_{3,f,k}(t,T)\,Y_{3,0,f,k,l}(t)\,, \end{aligned} \tag{211}$$

where

$$\begin{aligned} X_{3,f,k}(t) &= -\frac{\tau_3}{2}\int_0^t \bar{\psi}_{3,3,f,k,l}(v)\,\bar{K}_{3,3,f,k,l}(v,t)\,dv \\ &\quad + \int_0^t \bar{\psi}_{3,0,f,k,l}(v)\,\bar{K}_{3,0,f,k,l}(v,t)\,\frac{\big(\Sigma_{0,f,l}(t) - \Sigma_{0,f,l}(v)\big)}{K_{0,f,l}(0,t)}\,dv \\ &\quad + \int_0^t \psi_{3,f,k}(v)\,K_{3,f,k}(v,t)\,dW_{3,f,k}(t) \end{aligned} \tag{212}$$

and

$$Y_{3,j,f,k,l}(t) = \int_0^t \bar{\psi}_{3,j,f,k,l}(v)\,\bar{K}_{3,j,f,k,l}(v,t)\,dv\,. \tag{213}$$

Note that the factor $\big(\Sigma_{0,f,l}(T) - \Sigma_{0,f,l}(t)\big)$ appearing in the above appears also in Markov representations for single-curve Cheyette ('92) models under time-varying reversion, see *e.g.* Andreasen ('01).

It is straightforward but tedious to extend the analysis to the 6M IBOR curve and so on. Indeed, the analogy of (210) under the constant-reversion case will extend in a manner similar to (205) and

a Markov representation follows accordingly. We omit the details, but one can confirm that the quantities

$$\int_0^t \bar{\mu}_{3,6}(v,T)\,dv \quad \text{and} \quad \int_0^t \bar{\sigma}_{3,6,f,k}(v,T)\,\bar{\sigma}_{3,6,f,k}(v,T)\,dv\,, \tag{214}$$

which will feature in the representation of $\bar{s}_6(t,T)$, can be written in terms of the auxiliary state processes $Y_{3,j,f,k,l}(t)$ defined earlier. We note this as it ensures that additional state variables for the ON-*vs*.-3M IBOR curve dynamics are not required as we introduce other spreads, which is in line with what we encountered for the constant-reversion case.

Turning now to the instantaneous spread curve $s_3(t,T)$, we proceed under the simple case described in Section 2.2, wherein we use a single-factor specification with no correlation between the spread and the base curve. We do this to allow for a clean exposition of the issues at hand, but we note that the results we obtain extend to the general case.

To begin, we write down the dynamics of $s_3(t,T)$ under time-varying reversion,

$$ds_3(t,T) = \mu_3(t,T)\,dt + \psi_3(t)\,e^{-\int_t^T \kappa_3(w)\,dw}\,dW_3(t)\,, \tag{215}$$

and we note that under these simplified dynamics we require only that $S_3^{\times}(t,T)$ be a Martingale to ensure the absence of arbitrage. Recalling the sufficient condition from (12), we have that

$$\begin{aligned}\int_T^{T+\tau_3} \mu_3(t,u)\,du &= -\frac{1}{2}\left(\int_T^{T+\tau_3} \sigma_3(t,u)\,du\right)^2 \\ &= -\frac{1}{2}\psi_3(t)^2\,e^{2\int_0^t \kappa_3(w)\,dw}\left(\int_T^{T+\tau_3} e^{-\int_0^u \kappa_3(w)\,dw}\,du\right)^2.\end{aligned} \tag{216}$$

The corresponding necessary condition, *i.e.* the analogue of (13), is

$$\begin{aligned}&\mu_3(t,T+\tau_3) - \mu_3(t,T) \\ &\quad = -\frac{1}{2}\psi_3(t)^2\,e^{2\int_0^t \kappa_3(w)\,dw}\left(\int_T^{T+\tau_3} e^{-\int_0^u \kappa_3(w)\,dw}\,du\right)\left(e^{-\int_0^{T+\tau_3}\kappa_3(w)\,dw} - e^{-\int_0^T \kappa_3(w)\,dw}\right).\end{aligned} \tag{217}$$

There is no obvious solution for $\mu_3(t,T)$ to our no arbitrage restriction in terms of common functions, which stands in contrast to the case of constant mean reversion. We are able to proceed to obtain a solution which may be evaluated efficiently, but we must make use of the flexibility in $\mu_3(t,T)$ which was discussed in Section 2.2. Though, we are mindful not to abuse this flexibility such that arbitrary structures will emerge in spread dynamics and volatilities.

To begin, we first note that a separable (in t and T) solution for $\mu_3(t,T)$ is available to us, which is unsurprising given the separable nature of the volatility function. It we specify the drift function to have the form

$$\mu_3(t,T) = \psi_3(t)^2\,e^{2\int_0^t \kappa_3(w)\,dw}\,\mu_3'(T)\,, \tag{218}$$

and substitute this into (216), we reduce our problem to one of solving for the T-dependent function $\mu_3'(T)$,

$$\int_T^{T+\tau_3} \mu_3'(u)\,du = -\frac{1}{2}\left(\int_T^{T+\tau_3} e^{-\int_0^u \kappa_3(w)\,dw}\,du\right)^2. \tag{219}$$

With this, we seek out a solution for $\mu_3'(T)$ over the interval $[0, \tau_3)$, *i.e.* one which satisfies

$$\int_0^{\tau_3} \mu_3'(u)\, du \;=\; -\frac{1}{2}\left(\int_0^{\tau_3} e^{-\int_0^u \kappa_3(w)\, dw}\, du\right)^2, \tag{220}$$

and then build out the rest of $\mu_3(t,T)$, $T \geq \tau_3$, via the analogue of the necessary condition in (217),

$$\mu_3'(T+\tau_3) - \mu_3'(T) = -\left(\int_T^{T+\tau_3} e^{-\int_0^u \kappa_3(w)\, dw}\, du\right)\left(e^{-\int_0^{T+\tau_3} \kappa_3(w)\, dw} - e^{-\int_0^T \kappa_3(w)\, dw}\right). \tag{221}$$

To this end, we specify a functional form for $\mu_3'(T)$ on the interval $[0, \tau_3)$ of the type

$$\mu_3'(T) = A_3\, e^{-2a_3 T}. \tag{222}$$

Our rationale here in selecting this form is that *a*) it will collapse to the solution in (16) for constant reversions, and that *b*) it will also allow us to enforce that

$$\mu_3'(\tau_3) = \mu_3'(0) - \left(\int_0^{\tau_3} e^{-\int_0^u \kappa_3(w)\, dw}\, du\right)\left(e^{-\int_0^{\tau_3} \kappa_3(w)\, dw} - 1\right) \tag{223}$$

for appropriate choices of a_3 and A_3. This latter requirement is critical to ensuring that the solution for $\mu_3(t,T)$ is continuous in T. Indeed, we will be using (221) to build out $\mu_3(t,T)$, and as long as the solution $\mu_3'(T)$ is continuous on the *closed* interval $[0, \tau_3]$, then we are assured of a continuous solution as both terms on the RHS of (221) are continuous for all T. If we were not to enforce (223), there would (in general) be a discontinuity at $T = n\,\tau_3$, $n = 1, \ldots$.

Solving for the coefficients A_3 and a_3 in (222) is straightforward. Indeed, upon substitution the condition in (220) becomes

$$-A_3 \frac{\left(e^{-2a_3\tau_3} - 1\right)}{2a_3} = P_3\,, \quad P_3 = -\frac{1}{2}\left(\int_0^{\tau_3} e^{-\int_0^u \kappa_3(w)\, dw}\, du\right)^2, \tag{224}$$

and that in (223) becomes

$$A_3 e^{-2a_3\tau_3} = A_3 + Q_3\,, \quad Q_3 = -\left(\int_0^{\tau_3} e^{-\int_0^u \kappa_3(w)\, dw}\, du\right)\left(e^{-\int_0^{\tau_3} \kappa_3(w)\, dw} - 1\right). \tag{225}$$

We are able to solve for the coefficients directly as

$$a_3 = -\frac{1}{2}\,\frac{Q_3}{P_3} \quad \text{and} \quad A_3 = \frac{Q_3}{\left(e^{(Q_3/P_3)\,\tau_3} - 1\right)}. \tag{226}$$

We note the obvious result that, for constant κ_3, these reduce to

$$a_3 = \kappa_3 \quad \text{and} \quad A_3 = \frac{\left(e^{-\kappa_3\tau_3} - 1\right)^2}{\left(e^{-2\kappa_3\tau_3} - 1\right)}, \tag{227}$$

which aligns with our original result as presented in (16).

Building the solution for $\mu_3'(T)$ via (221) would be accomplished numerically at minimal cost, and these results can be extend to the general case. However, when an implementation requiring time-varying mean reversion is required, recasting the model in terms of $\bar{s}_3(t,T)$ would be more convenient as it avoids this issue entirely.

G Jump Processes in Tenor Basis Curves

In this appendix we extend the dynamics of spread curves to include jump discontinuities. See *e.g.* Mercurio and Li ('16) who provide motivation for such an extension. To simplify the presentation, we again focus only on the case of the ON-*vs*.-3M IBOR spread curve, noting that the extension to the 3M IBOR-*vs*.-6M IBOR spread curve, *etc.*, follows naturally. We begin with the case of the mean spread curve $\bar{s}_3(t,T)$ and then consider the case of the instantaneous spread curve $s_3(t,T)$. As was the case for time-varying mean reversions, adding jumps is relatively straightforward for $\bar{s}_3(t,T)$, while we encounter hurdles when treating $s_3(t,T)$. These hurdles again owe to complications associated with the drift restriction on $\mu_3(t,T)$. Note that for simplicity, we only allow for jumps in the spread curve process, omitting them from the base curve process. This is done only for presentational purposes and is with no loss of generality. Moreover, we only allow for a single jump component in the spread process; it is straightforward to generalize to multiple jump processes, each with its own response profile.

Extending the dynamics in (197) to include a jump component may be accomplished via

$$\begin{aligned} df_0(t,T) &= \mu_0(t,T)\,dt + \psi_{0,f,k}(t)\,e^{-\kappa_{0,f,k}(T-t)}\,dW_{0,f}(t) \\ d\bar{s}_3(t,T) &= \bar{\mu}_3(t,T)\,dt + \psi_{3,f,k}(t)\,e^{-\kappa_{3,f,k}(T-t)}\,dW_{3,f}(t) + k_3(t)\,e^{-\alpha_3(T-t)}\,dN_3(t)\,. \end{aligned} \tag{228}$$

The only aspect of these dynamics which is new is the jump term

$$k_3(t)\,e^{-\alpha_3(T-t)}\,dN_3(t)\,, \tag{229}$$

which is simply the differential of a marked point process. The jump size $k_3(t)$ has a distribution function $F_{k_3}(\cdot)$ which we leave unspecified for now, noting only that we will restrict it to be time-homogeneous. The counting process $N_3(t)$ has an intensity $\lambda_3(t)$ which satisfies $\mathbb{P}(dN_3(t)=1) = \lambda_3(t)\,dt$, and we take $\lambda_3(t)$ to be deterministic. Finally, we use a decay function with reversion coefficient α_3 to control the effect of jumps at the shorter end *vis-á-vis* the effect at the longer end.

Following the logic espoused earlier it is entirely straightforward to derive a restriction on $\bar{\mu_3}(t,T)$ analogous to that in (199), with the only difference being that the jump-modified version of Ito's Lemma must be applied,

$$\begin{aligned} \bar{\mu}_3(t,T) &= -\frac{1}{2}\tau_3\,\sigma_{3,f}(t,T)\,\sigma_{3,f}(t,T) + \sigma_{3,f}(t,T)\left(\int_t^T \sigma_{0,f}(t,u)\,du\right) \\ &\quad -\lambda_3(t)\,\mathbb{E}_t\big[e^{k_3(t)\,\tau_3\,e^{-\alpha_3(T-t)}} - 1\big]\,. \end{aligned} \tag{230}$$

To simplify this expression, we introduce the characteristic function of the jump distribution,

$$\Psi_{k_3}(\theta) = \mathbb{E}_t\big[e^{\theta\,k_3}\big]\,, \tag{231}$$

such that the jump-related term in (230) can be written as $-\lambda_3(t)\big(\Psi_{k_3}\big(\tau_3\,e^{-\alpha_3(T-t)}\big) - 1\big)$. What we face in this term is directly analogous to what one would face in the context of a single-curve Cheyette model extended for jumps; see *e.g.* Chiarella and Sklibosios ('03) or Andersen ('10) who explores a closely-related model of commodity curves.

Let us henceforth ignore the diffusion coefficients in (230) as they can be handled straightforwardly. With this, the restriction in (230) becomes

$$\bar{\mu}_3(t,T) = -\lambda_3(t)\big(\Psi_{k_3}\big(\tau_3\,e^{-\alpha_3(T-t)}\big) - 1\big)\,. \tag{232}$$

In general, this function is not separable in t and T, which can be problematic from the perspective of one seeking a Markov representation for $\bar{s}_3(t,T)$. However, given that $\lambda_3(t)$ is deterministic it follows that $\bar{\mu}_3(t,T)$ is deterministic also, and as such this will not cost us Markovianity. Indeed, with minimal effort we are able to derive a representation of $\bar{s}_3(t,T)$ of the form

$$\bar{s}_3(t,T) = \bar{s}_3(0,T) + \int_0^t \bar{\mu}_3(v,T)\,dv + e^{-\alpha_3(T-t)} \int_0^t k_3(v)\,e^{-\alpha_3(t-v)}\,dN_3(v)\,, \tag{233}$$

where we are still ignoring diffusion for simplicity. Upon setting

$$M_3(t,T) = -\int_0^v \lambda_3(v)\big(\Psi_{k_3}\big(\tau_3\,e^{-\alpha_3(T-v)}\big)-1\big)\,dv \quad \text{and} \quad X_3(t) = \int_0^t k_3(v)\,e^{-\alpha_3(t-v)}\,dN_3(v)\,, \tag{234}$$

this becomes

$$\bar{s}_3(t,T) = \bar{s}_3(0,T) + M_3(t,T) + e^{-\alpha_3(T-t)}\,X_3(t)\,. \tag{235}$$

The state process $X_3(t)$ accumulating the dump innovations is clearly Markov and the integrated drift $M_3(t,T)$, which is deterministic, can be evaluated numerically.

As noted in Andersen ('10), under special conditions of a constant intensity, λ_3, and an exponential distribution for the jump size, we are able to evaluate $\Psi_{k_3}(\cdot)$ in closed form. These conditions are likely adequate for practical implementations, which is especially true of using an exponential distribution as it would allow us to maintain a lower bound on spreads. Moreover, anything more complex, *e.g.* a time-varying intensity, would require a very involved calibration.

To see what such a specification affords us, let $k_3 \sim \text{Exp}(\bar{k}_3)$, such that

$$\Psi_{k_3}\big(\tau_3\,e^{-\alpha_3(T-t)}\big) = \frac{1}{1-\tau_3\,\bar{k}_3\,e^{-\alpha_3(T-t)}}\,, \tag{236}$$

and that

$$-\lambda_3\,\big(\Psi_{k_3}\big(\tau_3\,e^{-\alpha_3(T-t)}\big)-1\big) = -\lambda_3\,\frac{\tau_3\,\bar{k}_3\,e^{-\alpha_3(T-t)}}{1-\tau_3\,\bar{k}_3\,e^{-\alpha_3(T-t)}}\,. \tag{237}$$

With this, the integrated drift $M_3(t,T)$ may be evaluated in closed form to produce

$$M_3(t,T) = -\lambda_3 \int_0^t \frac{\tau_3\,\bar{k}_3\,e^{-\alpha_3(T-v)}}{1-\tau_3\,\bar{k}_3\,e^{-\alpha_3(T-v)}}\,dv = \frac{\lambda_3}{\alpha_3}\log\!\left(\frac{1-\tau_3\,\bar{k}_3\,e^{-\alpha_3(T-t)}}{1-\tau_3\,\bar{k}_3\,e^{-\alpha_3 T}}\right). \tag{238}$$

As such, to evaluate $\bar{s}_3(t,T)$ we simulate the Markov $X_3(t)$ as per usual and multiply by $e^{-\alpha_3(T-t)}$, and then add to it $M_3(t,T)$ which is readily computed via the above.

We turn now to the case of the instantaneous spread $s_3(t,T)$. Following precisely the same logic as for $\bar{s}_3(t,T)$, we obtain a drift restriction of the form

$$\begin{aligned}
\int_T^{T+\tau_3} \mu_3(t,T)\,du &= -\lambda_3\,\mathbb{E}\left[e^{k_3(t)\,\left(\frac{-1}{\alpha_3}\right)\left(e^{-\alpha_3\tau_3}-1\right)\,e^{-\alpha_3(T-t)}} - 1\right] \\
&= -\lambda_3\,\Psi_{k_3}\!\left(-\frac{\left(e^{-\alpha_3\tau_3}-1\right)}{\alpha_3}\,e^{-\alpha_3(T-t)}\right) \\
&= -\lambda_3\left(\frac{-\left(e^{-\alpha_3\tau_3}-1\right)\bar{k}_3\,e^{-\alpha_3(T-t)}}{\alpha_3+\left(e^{-\alpha_3\tau_3}-1\right)\bar{k}_3\,e^{-\alpha_3(T-t)}}\right),
\end{aligned} \tag{239}$$

where we have ignored diffusion-related terms as they have no bearing on the analysis, and where we have enforced the same simplifying distributional assumptions for the jump component as considered for $\bar{s}_3(t,T)$.

The corresponding necessary condition for $\mu_3(t,T)$ is

$$\mu_3(t,T+\tau_3) - \mu_3(t,T) = -\lambda_3 \frac{\alpha_3^2\big(e^{-\kappa_3\tau_3}-1\big)\,e^{-\alpha_3(T-t)}\,\bar{k}_3}{\big(\alpha_3+\big(e^{-\kappa_3\tau_3}-1\big)\,e^{-\alpha_3(T-t)}\,\bar{k}_3\big)^2}\,. \tag{240}$$

In this restriction we face a similar hurdle to that encountered in the context of (217), where we were extending the model to accommodate time-varying mean reversions. Namely, it is not obvious that $\mu_3(t,T)$ has a solution in terms of common functions. In contrast to the earlier case however, we cannot even assert that the solution is separable, meaning that if we were to proceed numerically we would need to solve for each (t,T) combination. There may exist efficient ways of doing this, but we do not pursue the possibility here. As was the case for time-varying reversions, it is preferable to recast in terms of $\bar{s}_3(t,T)$ when seeking to allow for jump innovations in tenor basis spreads.