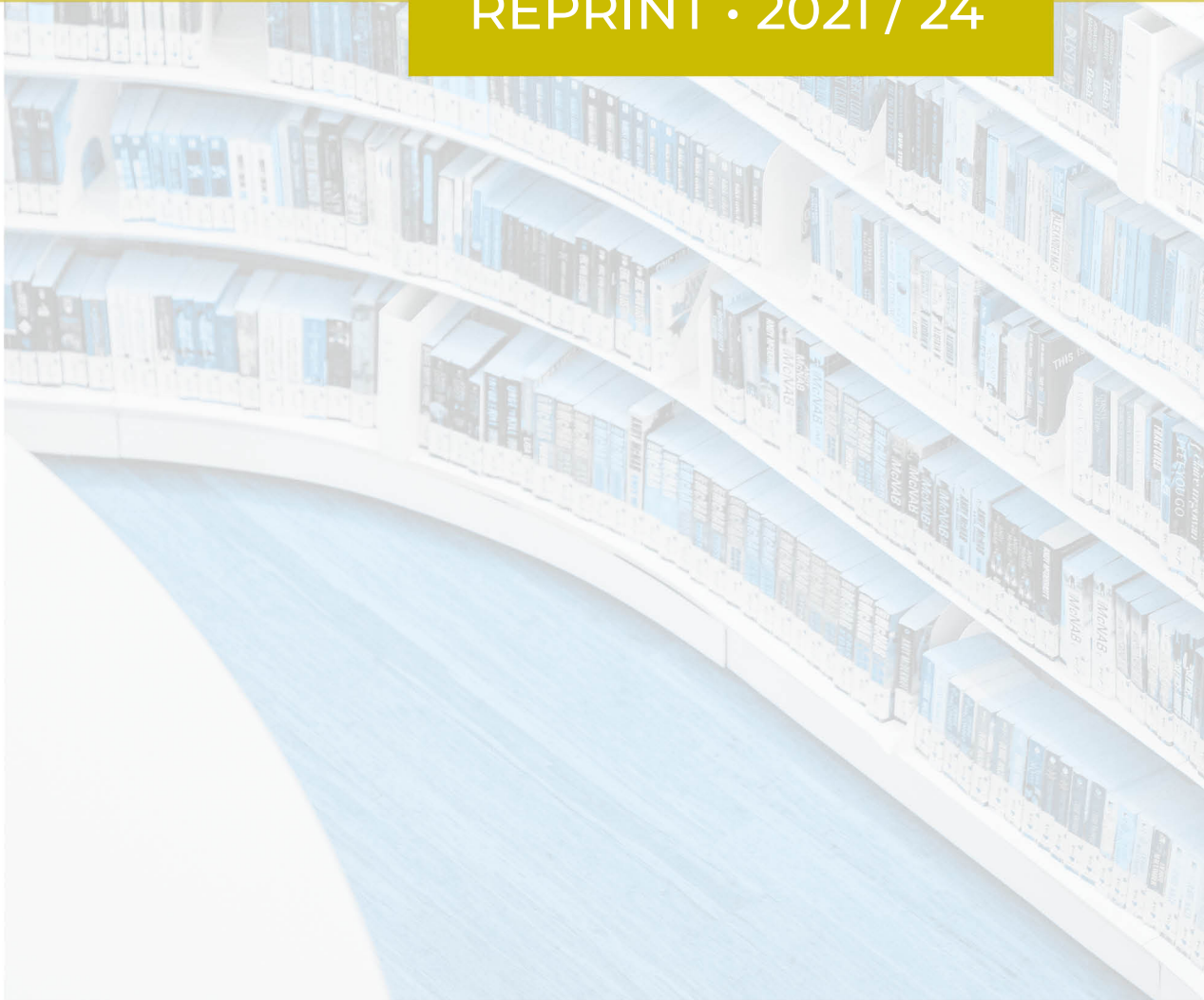


AFFINE TERM STRUCTURE MODELS: A TIME-CHANGE APPROACH WITH PERFECT FIT TO MARKET CURVES

Cheikh Mbaye, Frédéric Vrins

REPRINT · 2021 / 24



LFIN

Voie du Roman Pays 34, L1.03.01 B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: lidam-library@uclouvain.be

<https://uclouvain.be/en/research-institutes/lidam/lfin/publications.html>

Affine term structure models: a time-change approach with perfect fit to market curves

Cheikh Mbaye^a & Frédéric Vrins^{a,b*}

^a UCLouvain, LFIN/LIDAM, Belgium

^b HEC Montréal, Departement of Decision Sciences, Canada

Abstract

We address the so-called *calibration problem*, which consists of fitting in a tractable way a given model to a specified term structure such as yield, prepayment or default probability curves. Time-homogeneous affine jump diffusions are tractable processes but have limited flexibility; they fail to perfectly replicate actual market curves. Applying a deterministic shift to the latter is a simple but efficient solution that is widely used by both academics and practitioners. However, the shift approach may not be appropriate when positivity is required, a common constraint when dealing with credit spreads or default intensities. In this paper, we address this problem by adopting a time-change technique. Specific attention is paid to the Cox-Ingersoll-Ross model with compound Poisson jumps (JCIR), which remains standard for modelling intensities. Our time-changed JCIR (TC-JCIR) is compared to the shifted JCIR (JCIR++) in various credit applications such as credit default swap, credit default swaption and credit valuation adjustment under wrong-way risk. The TC-JCIR model is able to generate much larger implied volatilities and covariance effects than JCIR++ under positivity constraints and represents an appealing alternative to the latter.

Keywords: model calibration, credit risk, stochastic intensity, jump diffusions, term structure models, time-change techniques

1 Introduction

Model calibration is a standard problem in many areas of finance. It consists of tuning a model such that it “best fits” market quotes at a given time. For example, financial markets provide a set of prices associated with liquid instruments that openly trade on the market. Alongside risk management (hedging), the main purpose of a model here is to act as an “interpolation/extrapolation” tool, i.e., to obtain the value of products at a given time t for which the market does not disclose prices in a transparent way. This could happen because either the product to be priced is “exotic” (quoted on a bilateral basis) or because its cashflow schedule is not in line with the products that currently trade openly at t (a situation that commonly happens since products that were “standard” at inception may have time-to-expiry or moneyness levels that are no longer “standard” afterwards).

*Voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium. E-mail: frederic.vrins@uclouvain.be.

Mathematically, model calibration is nothing but an optimization problem. Starting from a set of prices quoted on the market (called *market prices*) for a set of specific financial products (*calibration instruments*), it consists of computing the model parameters such that the prices generated by the model (*model prices*) best fit the market prices, according to some cost function. Model calibration is crucial in finance; it is strongly related to arbitrage opportunities. In practice, only models that are able to reproduce the market prices of “simple instruments” (either in a perfect way or at least up to the bid-ask spread) are trustworthy enough in regard to pricing other instruments. We refer to (Brigo and Mercurio, 2006; Joshi, 2003; Veronesi, 2010) and references therein. As a simple example, one can price exotic derivatives (such as barrier options) using a stochastic volatility model such as Heston in a semi-analytical way (Carr and Madan, 1999; Heston, 1993). The parameters of the Heston model will be obtained by “calibration” to a volatility surface, i.e., to a set of liquid (“plain vanilla”) options, such as European calls and puts of various strikes and maturities. The justification for this is that in a no-arbitrage, complete market setup, the price of an option can be obtained by computing the cost of setting up a self-financing hedging strategy. This cost depends on the prevailing prices of the hedging instruments. If the model fails to correctly price the latter, there is no chance it can correctly price the option.

In this work, we focus on financial calibration problems arising in other asset classes: interest rates and credit (Brigo and Mercurio, 2006; Duffie and Singleton, 2003). When specifying an interest rate model to price a derivative on, say, the Libor 3M index, one needs to make certain that the model generates a discount curve that is in line with that extracted from market quotes of simpler Libor 3M-indexed products. In this case, the set of calibration instruments could be forward rate agreements (FRAs), interest rate swaps (IRSs), and vanilla cap/floors or swaptions. The calibration step guarantees that the model price associated with every calibration instrument will agree with its market price. Similarly, adjusting the value of derivatives for counterparty risk (a problem known as *credit valuation adjustment*, or CVA) generally involves a stochastic model to represent the default of the counterparty with whom the trade is executed. The default probability of the counterparty can be extracted from a set of calibration instruments, the prices of which are driven by the default likelihood of the counterparty, e.g., corporate bonds or credit default swaps (CDSs). In this context, the default model must be *calibrated* in such a way that the default probability curve generated by the stochastic model agrees with that implied by the prices of the corresponding instruments (see, e.g., (Gregory, 2010) and (Stein and Pong, 2011) for a general overview of CVA and (Brigo et al., 2014) for a discussion of bilateral CVA in the presence of collateralization agreements).

The calibration constraint raises practical issues. Indeed, the models that are actually used in the industry must have a tractability level that is compatible with real-time pricing but, as explained above, must be flexible enough to match the information conveyed by the calibration instruments. Moreover, depending on the application at hand, they have to satisfy additional constraints; for instance, a stochastic intensity must be non-negative. Affine term-structure models (ATSM) have been extensively used in fixed income because of their analytical tractability; see, e.g., (Duffie et al., 2003) and (Duffie and Kan, 1996) for an excellent review and mathematical treatment of those processes. In practice, homogeneous affine jump diffusions (HAJDs) are extremely popular (Duffie et al., 1999). For instance, the Vasicek (Ornstein-Uhlenbeck) model is a short-rate model widely used in both industry and academia (Vasicek,

1977). Cox-Ingersoll-Ross (also known as CIR, with mean-reverting square-root diffusion dynamics, SRD) (Cox et al., 1985) can be preferred when non-negativity is required, possibly combined with independent compounded Poisson jumps (JCIR or SRJD). The latter is also widely used to model volatilities or default intensities.

However, it is in general impossible with either model (even in a multi-factor setup) to achieve a perfect fit: the flexibility of HAJD models is limited, and they are in general unable to generate a specified discount curve. The same problem arises when dealing with credit derivatives: it is generally impossible to make sure that the default intensity process, modelled with HAJD dynamics, generates a default probability curve that is in line with the corresponding curve, exogenously given by the market via the calibration instruments.

Several routes can be followed to address this issue. The first consists of disregarding this lack of flexibility. Nevertheless, working with a model that fails to yield a perfect fit to the market is often unacceptable in practice. Indeed, as explained above, the models are used to value derivative positions, and a mismatch with the market can introduce tremendous bias in the valuation of the book of companies or financial institutions. Another possibility is to significantly increase the complexity of the models (e.g., by relaxing time homogeneity). This is often to be avoided in practice for computational, identification or overfitting issues. For instance, it is easy to generate a flexible non-negative process by considering a flexible shadow process floored at zero. This is the approach adopted in (Black, 1995) to build non-negative short-rate processes. However, as Black pointed out, this “may kill any chance of having analytic solutions for the prices of derivatives”. Although, according to Black, “it hardly complicates the numerical solutions at all” and does not trigger issues for pricing, it becomes a severe limitation for calibration purposes. Indeed, solving for the parameters that best fit the market prices of a set of instruments requires multiple iterations, which, for this kind of models, may lead to, e.g., Monte Carlo squared. A trade-off consists of extending “simple models” in a “simple way”, such that these extensions can fit the market while remaining analytically tractable, at least as far as plain vanilla instruments are concerned.

A first approach (called the *shift approach*) is to shift a tractable shadow process with a deterministic function, φ . For instance, (Dybvig, 1997) shows that the term structure of interest rates can be reproduced by shifting a Ho-Lee (Ho and Lee, 1986) or a Vasicek process. Later, (Brigo and Mercurio, 2001) extended this idea to a broader class of models, thereby providing a simple but clever solution to the calibration problem. For instance, shifting the Vasicek, CIR and JCIR models yields the so-called Hull-White, CIR++ or JCIR++ (also known as SSRJD) models, respectively (Brigo and Mercurio, 2006). This trick is actually very powerful: it solves the calibration problem at no cost, since a deterministic shift does not impact the model’s tractability. Moreover, the shift function is known analytically as a function of the parameters of the shadow HAJD and the market curve to be fitted by the model. A second possibility (called the *time-change approach*) could consist of a shadow process time-changed with a clock Θ . Time-change techniques were first studied in the sixties (Dambis, 1965; Dubins and Schwartz, 1965). The first application to finance dates back to the early 2000s. Geman et al. used Lévy processes and interpreted the new time scale as business time, in contrast with calendar time (Geman et al., 2001). This was then applied to stochastic volatility models (Carr et al., 2003). Thanks to subordinated Lévy models, the authors introduced the leverage effect, as well as a long-term skew. Many other financial applications of time-change techniques can

be found in the review (Swishchuk, 2016). More recently, stochastic time-change processes have been considered to introduce two-sided jumps in non-negative processes (Mendoza-Arriaga and Linetsky, 2014). The analytical tractability of the resulting model is preserved to some extent. This model has recently been applied to counterparty credit risk (Mbaye and Vrins, 2018). In this work, we exploit the time-change idea in yet another way to solve a completely different problem. The intuition is that by slowing down or speeding up the time of a shadow model, at the appropriate rate, one would obtain a model that could fit most discount curves. If, moreover, the chosen shadow process is quite tractable (such as an HAJD) and if the model for the clock is not too involved (e.g., deterministic, as in the shift approach), it is possible to obtain a model with appealing properties. The same trick was proposed in (Rogers, 1995) to fit the term structure of interest rates with a non-negativity constraint, taking a CIR as the shadow process.¹

The shift approach is straightforward, extremely simple and therefore widely used in the industry. However, it suffers from an important limitation. Suppose, for instance, that one needs to model non-negative interest rates. One possibility would be to use a CIR++ process. While CIR++ is the shifted version of a non-negative (plain CIR) shadow process, there is no guarantee that the resulting CIR++ itself would still be positive: this depends on whether the shift is a non-negative function. However, the shift fills the gap between the information conveyed by the calibration instruments on the one hand (e.g., the discount curve implied by the market) and the parameters of the underlying model (the discount curve implied by the plain CIR dynamics) that, without specific measures, might become negative. Of course, one could think of adding a non-negativity constraint to the shift in the calibration step. However, clearly, this requires playing with the parameters of the underlying model. Moreover, as we will show, this drastically restricts the parameters of the underlying CIR, hence the randomness embedded in the model. By contrast, as explained in (Rogers, 1995), the time-change approach does not suffer from this problem: time-changing a non-negative process leads to another non-negative process. One could think that, compared to the shift approach, the expression of the time-change function is much more difficult to infer and the dynamics of the time-changed model harder to grasp. This is actually not the case.

Whereas the non-negativity constraint is obviously obsolete in the interest rate context prevailing nowadays, it remains a primary concern in other applications. In credit risk, for instance, one is led to simulate a default intensity that is consistent with the term structure of credit spreads (or, in other words, with the default probability curve implied by the market). Despite its theoretical inconsistencies (allowing for negative “intensities” with non-zero probability), it turns out that the shift approach remains standard market practice even in applications that, theoretically speaking, require non-negativity.

In this paper, we exploit the same idea as in (Rogers, 1995) but make several additional contributions. First, we clarify when perfect fit can be achieved using the time-change approach. Second, we focus on the non-negativity constraint and show that the time-change function providing the perfect fit is easily found via a simple ordinary differential equation or, equivalently, using a function inversion. Third, we show that when the market curve and the chosen non-negative model are such that there exists a non-negative shift function allowing for a perfect fit,

¹We are grateful to an anonymous referee for suggesting this reference. We were not aware of the latter and proposed the time-change approach independently.

the time-change approach features a larger implied volatility than the shift approach when the same set of parameters is used for the shadow process; we prove this for both CIR and JCIR shadow processes. Fourth, we illustrate the power of our approach in two applications taken from credit: the pricing of CDS options and computation of derivative pricing while accounting for counterparty risk under exposure-credit dependence (wrong-way risk, WWR). In either case, all the considered default models perfectly fit the risk-neutral default probability curve extracted from the term structure of CDS spreads of the reference entity. The results illustrate the desirable feature of large implied volatilities inherent to the time-change approach: they are able to generate more realistic option prices and WWR impacts than the shift approach calibrated on the same probability curve under a non-negativity constraint.

Notwithstanding the fact that we restrict our analysis to reduced-form models when pricing credit-sensitive instruments, our approach is of potential use for other models in many areas of finance and insurance. Ongoing work suggests that it can be applied to other default models, including firm-value (structural) models (Merton, 1974), as well as to linear-rational (polynomial) models (Filipovic et al., 2017). Other models could also be considered, such as the Φ -martingale approach of (Jeanblanc and Vrins, 2018) or (Crépey et al., 2012). In terms of applications, the proposed method can be used in life insurance to calibrate the mortality rate to mortality tables. Our time-changed process could also be used to model prepayment rates in mortgage-backed securities (MBS). These products naturally exhibit negative convexity due to the negative relationship between interest and prepayment rates: householders tend to refinance their loans when interest rates drop. This calls for a stochastic prepayment (i.e., non-negative) rate that will be correlated with interest rates and the parameters of which could be calibrated to agree with the averaged values given in the PSA measure, the indicator attached to MBS that characterizes the prepayment speed in the pool (Veronesi, 2010; Gorovoy and Linetsky, 2007). The proposed method could also be useful in engineering, such as the modelling of device or material performance degradation over time, for which the average outstanding performance evolution is given according to some quality standards.

The paper is organized as follows. The calibration problem is introduced in Section 2, and two specific cases (cashflow discounting and probability curves) are discussed. Some properties of ATSM and AJD models that will serve as shadow processes are recalled. The deterministic shift approach is recalled in Section 3, with a focus on one-dimensional HAJD shadow processes. We discuss the limits of this technique when the resulting process needs to satisfy a positivity constraint. As a solution to this specific issue, we introduce the time-change technique in Section 4. Although our approach can be applied to any tractable shadow process, we derive specific results associated with the JCIR shadow process in Section 5, due to its importance for solving the calibration problem under positivity constraint, and because of its central role in credit risk modelling. Finally, we compare in Section 6 our model's performance to that of the shift approach on three different pricing problems taken from credit risk: CDS curve calibration, the pricing of CDS options and the pricing of credit valuation adjustment under wrong-way risk. For the sake of brevity, most proofs and peripheral results are relegated to the appendix.

2 The calibration problem

The calibration problem consists of finding, for a given time- s market curve $P_s(t)$, $t \geq s$, and a parametric family of models $x = x(\Xi)$, the specific set of parameters Ξ^* (hence, the specific model $x^* := x(\Xi^*)$ in this family) such that the corresponding model curve $P_s^{x^*}$ best fits the market curve, according to some criterion. Mathematically speaking,

$$\Xi^* := \arg \min_{\Xi} \left\langle P_s^{x(\Xi)} \mid P_s \right\rangle, \quad (1)$$

where $P_s^{x(\Xi)}$ is the time- s curve implied by $x(\Xi)$ and $\langle f|g \rangle$ represents a divergence measure between two functions f, g . In practice, one often computes the mean-square error (MSE) between f and g on a set of maturities $\mathcal{T} := \{T_1, \dots, T_n\}$:

$$\langle f|g \rangle := \frac{1}{n} \sum_{i=1}^n (f(T_i) - g(T_i))^2. \quad (2)$$

A model x is said to *perfectly fit* the market curve at time s up to horizon T , $T > s$, whenever $P_s^x(t) = P_s(t)$ for all $t \in [s, T]$.²

2.1 Setup

We consider a frictionless market free of arbitrage opportunities in which trading takes place continuously over time. Uncertainty is modelled through a probability space $(\Omega, \mathcal{G}, \mathbb{Q})$ and an information flow $\mathbb{G} := \{\mathcal{G}_t, t \geq 0\}$, corresponding to the filtration generated by the stochastic market variables (risk factors, prices, interest rates, default intensities, default event, etc.), and $\mathbb{Q}$ denotes the risk-neutral probability measure, referred to as the *pricing measure*. In the sequel, we focus on the case in which P are *discount curves*.

Definition 1 (Discount curve). *A time- s discount curve is any differentiable function of the form*

$$P_s : [s, \infty) \rightarrow (0, \infty), \quad t \mapsto P_s(t)$$

satisfying $P_s(s) = 1$.

In the specific $s = 0$ case, a *time-0 discount curve* $P_0(t)$ is simply called a *discount curve* and is denoted $P(t)$, assuming implicitly that $t \geq 0$. Any time- s discount curve admits an exponential integral from:

Lemma 1. *Every time- s discount curve P_s admits a representation in terms of the time- s instantaneous forward rate curve f_s :*

$$P_s(t) = e^{-\int_s^t f_s(u) du}, \quad t \geq s.$$

If moreover $P_s(t)$ is strictly decreasing on (s, ∞) , then $f_s(t)$ is strictly positive for all $t > s$.

Proof. Since $P_s(t) > 0$ for all $t \geq s$ and $P_s(t)$ is differentiable with respect to t on (s, ∞) , one can define the time- s instantaneous forward rate curve as $f_s(t) := \frac{-1}{P_s(t)} \frac{d}{dt} P_s(t) = -\frac{d}{dt} \ln P_s(t)$ for all $t > s$; the value $f_s(s)$ is not identified but can be defined by, e.g., the limit as $t \downarrow s$. Moreover, if P_s is strictly decreasing on (s, ∞) , then $f_s(t) > 0$ for all $t > s$. $\square$

²To ease the notations, we write P^x instead of $P^{x(\Xi)}$ when no confusion is possible.

Discount curves are of paramount importance in finance. As suggested by the name, $P_s(t)$ allows one to compute the time- s value of a cashflow paid at time $t \geq s$, both in a credit risk-free and credit-risky setup. As a special case of the second framework, they encompass survival probability curves, defined as one minus cumulative distribution functions. This is briefly recalled in the next two subsections.

2.1.1 Discounting in a default-free market

In this application, we are interested in the time- s price of a risk-free zero-coupon bond (ZCB) with maturity t and face value 1, denominated in a given currency. In particular, $P_s(t)$ and $P_s^r(t)$ provide the market and the model prices of that instrument, respectively. Adopting a short-rate model, $P_s^r(t)$ corresponds to the $\mathbb{Q}$ -expectation of the stochastic discount factor $D_s(t) := e^{-\int_s^t r_u du}$ conditional upon the available information (Brigo and Mercurio, 2006):

$$\mathbb{E}[1D_s(t)|\mathcal{G}_s] = \mathbb{E}\left[e^{-\int_s^t r_u du} \middle| \mathcal{G}_s\right] =: P_s^r(t) .$$

In this context, our aim is to find a realistic parametric model $x = x(\Xi)$ that would depict the dynamics of the risk-free rate r . The model needs to be tractable and flexible enough, in the sense that for any given yield curve P , one can easily find a specific set of parameters Ξ^* such that $P^{x^*} = P$.

2.1.2 Discounting in a defaultable market

Adopting the same framework as above, the time- s price of a zero-coupon bond paying one unit of currency at time $t \geq s$ *contingent on the survival of the issuer up to the payment date* is given by a similar expression as above. Mathematically, this corresponds to

$$\mathbb{E}\left[\mathbb{1}_{\{\tau > t\}} D_s(t) \middle| \mathcal{G}_s\right] ,$$

where τ represents the random default time of the issuer and $\mathbb{1}_{\{A\}}$ is the indicator function.

In the sequel, we consider a *reduced-form* (or *intensity*) default model of Cox type. We refer the reader to (Duffie and Singleton, 1999) and (Lando, 2004) for an extensive exposition of this class of models. In this setup, we work with a sub-filtration $\mathbb{F} := \{\mathcal{F}_t, t \geq 0\}$ of $\mathbb{G}$, such that all risk factors associated with default-free assets are $\mathbb{F}$ -adapted. The default time $\tau = \tau(\lambda)$ is then defined as the first-passage time of the process $\Lambda := \{\Lambda_t, t \geq 0\}$, $\Lambda_t := \int_0^t \lambda_s ds$ above a unit-mean exponential random variable $\mathcal{E}$. The process λ is non-negative and $\mathbb{F}$ -progressively measurable such that Λ is $\mathbb{F}$ -adapted. By contrast, $\mathcal{E}$ is taken independent from $\mathcal{F}_\infty$. Hence, the default indicator process $H := \{H_t, t \geq 0\}$, $H_t := \mathbb{1}_{\{\tau \leq t\}}$ is *not* adapted to $\mathbb{F}$. The total filtration is recovered by enlarging $\mathbb{F}$ with $\mathbb{H} := \{\mathcal{H}_t, t \geq 0\}$, the natural filtration of H ; $\mathcal{G}_t := \mathcal{H}_t \vee \mathcal{F}_t$. Observe that $\mathbb{F}$ collects partial information regarding the default: τ is a $\mathbb{G}$ -stopping time but not an $\mathbb{F}$ -stopping time.

In this framework, λ is the $\mathbb{F}$ -intensity of τ . Indeed, Λ is the $\mathbb{F}$ -compensator of H in the sense that the process $M := \{M_t, t \geq 0\}$ defined as

$$M_t := H_t - \Lambda_{\tau \wedge t} , \quad \Lambda_t := \int_0^t \lambda_u du$$

is a $\mathbb{G}$ -martingale. The non-negativity constraint on λ results from the fact that H is non-decreasing. In this model, the default event $\{\tau \leq t\}$ is modelled as $\{\Lambda_t \geq \mathcal{E}\}$, and the survival probability is given by

$$\mathbb{Q}(\tau > t) = \mathbb{Q}(\Lambda_t < \mathcal{E}) = \mathbb{Q}(U < e^{-\Lambda_t}) = \mathbb{E}[e^{-\Lambda_t}]$$

where $U := e^{-\mathcal{E}}$ is a random variable uniformly distributed on $[0, 1]$.

A central result in stochastic calculus is the so-called *Key lemma*. This fundamental theorem allows one to write the $\mathcal{G}_s$ -conditional expectation of $X\mathbb{1}_{\{\tau > t\}}$ as the $\mathcal{F}_s$ -conditional expectation of $Xe^{-\int_s^t \lambda_u du}$, rescaled by $\mathbb{1}_{\{\tau > s\}}$, for every integrable and $\mathcal{F}_t$ -measurable random variable X (see, e.g., Corollary 5.1.1, formula (5.13) in (Bielecki and Rutkowski, 2002)). It is due to Dellacherie and Meyer (Dellacherie and Meyer, 1980), although its use in financial applications was proposed later (see, e.g., (Collin-Dufresne et al., 2004), (Bielecki et al., 2011) for numerous examples in credit risk and (Brigo and Vrins, 2018) for a specific application to counterparty credit risk). Applying the Key lemma to $X \leftarrow D_s(t)$ yields

$$\mathbb{E}[\mathbb{1}_{\{\tau > t\}} D_s(t) | \mathcal{G}_s] = \mathbb{1}_{\{\tau > s\}} \mathbb{E}\left[e^{-\int_s^t (\lambda_u + r_u) du} \middle| \mathcal{F}_s\right] =: \mathbb{1}_{\{\tau > s\}} P_s^{\lambda+r}(t).$$

The survival probability function associated with τ conditional on $\mathcal{G}_s$ is obtained as a special case where $r = 0$:

$$\mathbb{Q}(\tau > t | \mathcal{G}_s) = \mathbb{E}[\mathbb{1}_{\{\tau > t\}} | \mathcal{G}_s] = \mathbb{1}_{\{\tau > s\}} P_s^\lambda(t).$$

This curve plays a central role in credit-risk applications, such as those considered in Section 6.

In this context, our aim is to find a realistic parametric model $x = x(\Xi)$ that would depict the dynamics of the $\mathbb{F}$ -intensity process λ . The model needs to be tractable and flexible enough, in the sense that for any given survival probability curve G extracted from the prices of defaultable instruments such as corporate bonds or CDSs, one can easily find a specific set of parameters Ξ^* such that $P^{x^*} = G$.

Remark 1. Notice that in contrast to interest rates, which can – and some currently do – take negative values, non-negativity is a formal requirement when x represents a stochastic intensity. Of course, one could relax the positivity constraint on λ and still model the event as the first passage time of Λ above $\mathcal{E}$. However, λ would no longer correspond to the $\mathbb{F}$ -intensity of τ , and the analytical expression of the model-implied survival probability would become $\mathbb{Q}(\tau > t) = \mathbb{Q}(\sup_{s \leq t} \Lambda_s < \mathcal{E}) = \mathbb{E}[e^{-\sup_{s \leq t} \Lambda_s}]$. This corresponds to a different class of models, for which analytical tractability is likely to be lost.

2.2 The perfect fit problems

Equation (1) suggests that the calibration problem consists of finding the parameters of a given model to minimize the discrepancies between the market and model curves up to a time horizon. However, it is clear that the choice of model class will also have a substantial impact. Indeed, depending on the chosen family of models, the minimum of the error function could be large, small or even zero (in which case the *perfect fit* is obtained).

Inspired by the financial problems mentioned in Section 2.1, we consider the following problems directly related to the perfect fit constraint. The first does not impose any constraint on the process x under consideration.

Problem 1. Find a tractable family of processes $x = x(\Xi)$ such that, for any given discount curve P , there exists a specific set of parameters Ξ^* satisfying

$$P_s^{x^*}(t) = P_s(t) , \quad x^* := x(\Xi^*)$$

for all $t \in [s, T]$.

Depending on the application at hand, one may need to impose additional constraints on x . As suggested by the risky discounting example, non-negativity is crucial. This leads us to consider a second (constrained) problem.

Problem 2. Find a tractable family of non-negative processes $x = x(\Xi)$ such that, for any given strictly decreasing discount curve P , there exists a specific set of parameters Ξ^* satisfying

$$P_s^{x^*}(t) = P_s(t) , \quad x^* := x(\Xi^*)$$

for all $t \in [s, T]$.

In either problem, *tractability* refers to the fact that model calibration (1) – which features an optimization over the parameter space – is not too computationally cumbersome. Solving this optimization problem typically requires many iterations and hence numerous evaluations of the objective function. This suggests that a highly desirable feature of the model is that it admits a closed-form expression for $P^{x(\Xi)}$ or, at least, that the latter can be computed without having to rely on time-consuming numerical methods such as Monte Carlo simulations, for example. In the next section, we briefly recall what is probably the most tractable family of models, namely, affine term structure models (ATSM) and affine jump diffusions (AJD), in particular.

2.3 Affine jump diffusions

ATSMs are extensively used in finance because they offer an appealing modelling framework: they are scarce, and empirical evidence suggests that they depict market dynamics relatively well. Let us consider a process x written as an affine function of an m -dimensional Markov shadow process y

$$x_t := \rho(t) \cdot y_t + \varphi(t) , \tag{3}$$

for some bounded time-dependent functions $\rho : \mathbb{R}^+ \rightarrow \mathbb{R}^m$ and $\varphi : \mathbb{R}^+ \rightarrow \mathbb{R}$. Such models correspond to the ATSM class whenever the log-price of the zero-coupon bond is an affine function of y , i.e.,

$$P_s^x(t) = P_s^{x(\Xi)}(t) = \mathbb{E} \left[e^{-\int_s^t x_u du} \middle| \mathcal{F}_s \right] = \exp \{ A_s^x(t; \Xi) - B_s^x(t; \Xi) \cdot y_s \} \tag{4}$$

where Ξ is the set of parameters governing y and A_s^x, B_s^x are differentiable functions satisfying $A_s^x(s; \Xi) = B_s^x(s; \Xi) = 0$. ATSMs are tractable in the sense that the functions in (4) are found by solving Riccati's equations. Those models provide an interesting route to solve calibration problems. For instance, P_s^r and P_s^λ in the above two examples are time- s discount curves whenever r and λ are ATSMs, respectively.³

³As shown in Section 8.2, it is clear that other dynamics than AJDs can lead to discount curves, but we restrict ourselves to AJDs in this paper.

An interesting special case is to consider a single-factor ($m = 1$) homogeneous affine jump-diffusion (HAJD) for the shadow process y , with stochastic differential equation (SDE) :

$$dy_t = (a + by_t)dt + \sqrt{c + dy_t}dW_t + dJ_t \quad (5)$$

where W is a Brownian motion, J is a compounded Poisson process with jump size mean α and intensity ω , and the coefficients are assumed to be compatible with the existence of a unique solution. Setting $\rho = 1$ and $\varphi = 0$, the process x in (3) coincides with y in (5), and embeds many variants of Vasicek and JCIR, for which the A^y, B^y functions are known in closed form.⁴ The analytical form of the model discount curve substantially facilitates calibration procedures such as (1), as illustrated in risk-free and risky discounting applications.

Remark 2. *As explained above, the risky discount curve P_s^y is a time- s discount curve when y is an AJD. Intuitively, one expects $P_s^{\lambda+r}$ to be a discount curve when λ and r are univariate AJDs under mild additional conditions. For instance, $P_s^{\lambda+r}$ is a discount curve when r and λ are obtained from the linear transform of a same (multivariate) AJD, as shown in Duffie et al. (1999). Other sufficient conditions are proposed in the Appendix (Section 8.2).*

3 A deterministic shift extension

HAJD models such as Vasicek, CIR and JCIR seem appropriate to solve problems 1 and 2. Unfortunately, they do not allow for a perfect fit to a given discount curve P , except in very special cases. One way to address this problem is to consider more complex dynamics for y , such as time-inhomogeneous or multi-factor diffusions. However, considering time-inhomogeneous versions might give rise to identification issues and, in general, the analytical expressions of A^x and B^x cannot be found. Multi-factor processes (usually, $m = 2$ or $m = 3$) are popular in the interest rate literature, but one-factor ($m = 1$) models remain widely used in credit risk applications, essentially due to the scarcity of option data.⁵ Alternatively, one could consider a transform of y in (5). Of course, the transform must be chosen to deliver the required flexibility without entailing analytical tractability.

A first route would be to take (3) with $\varphi = 0$ and vary ρ to obtain the perfect fit. This approach, however, may not be the preferred option in the sense that while it belongs to the class of ATSMs, there is no guarantee that the A^x, B^x coefficients are still known in closed form even if A^y, B^y are. On the other hand, another approach consists of taking $\rho = 1$ and varying φ instead. This is precisely the *shift approach* and leads to a tractable model equipped with the perfect fit feature. In this section, we briefly recall the main features of this technique and stress that it may not be appropriate for solving Problem 2.

3.1 Model setting

The shift approach was introduced in the seminal paper by (Brigo and Mercurio, 2001) precisely to address calibration issues such as Problem 1. This method actually corresponds to taking $x := x^\varphi$, defined as

$$x_t^\varphi := y_t + \varphi(t) . \quad (6)$$

⁴We recall (and derive) some properties of these processes in the Appendix (Section 8.1) for further reference.

⁵As explained in Remark 3, our approach can be easily extended to $m > 1$ and to time-inhomogeneous AJD

where the shadow model y is an HAJD of the form (5).

Clearly, x^φ is an ATSM provided that φ is bounded (which will be the case in practice). This framework has three key advantages for calibration purposes. First, in the specific case in which the A^y, B^y coefficients of the shadow model are known in closed form, the curve $P^{x^\varphi(\Xi)}$ is analytically tractable in terms of calibration since

$$P^{x^\varphi(\Xi)}(t) = \exp \{A^{x^\varphi}(t; \Xi) - B^{x^\varphi}(t; \Xi)x_0\}$$

with $A^{x^\varphi}(t; \Xi) = A^y(t; \Xi) - \int_0^t \varphi(u)du + B^y(t; \Xi)\varphi(0)$ and $B^{x^\varphi}(t; \Xi) = B^y(t; \Xi)$. Second, as already noted in (Brigo and Mercurio, 2001), whatever the shadow process y , the parameter Ξ and the market discount curve P , there always exists a shift function $\varphi(t) = \varphi^*(t; \Xi)$ that provides a perfect fit between the x^φ model and the market. Moreover, the corresponding shift function is available in closed form from the instantaneous forward rate curves. This is summarized in the next lemma.

Lemma 2. *The x model defined according to (6) where y is an HAJD solves Problem 1 provided that*

$$\varphi(t) \leftarrow \varphi^*(t; \Xi) := \frac{d}{dt} \ln \frac{P^{y(\Xi)}(t)}{P(t)} = f(t) - f^{y(\Xi)}(t) . \quad (7)$$

where f and f^y are the instantaneous forward rates associated with P and P^y , respectively.

The celebrated Hull-White and (J)CIR++ models, extensively used in industry, actually correspond to the shift approach when considering Vasicek and (J)CIR as shadow processes, respectively.

The fact that, for a given shadow process, perfect fit can be attained for every set of parameters Ξ is appealing. However, it also suggests that the calibration problem (1) is ill posed because of identification issues. Indeed, the choice of Ξ is completely arbitrary since the error between P^{x^φ} and P can be set to zero for any Ξ , provided that one chooses $\varphi(t) \leftarrow \varphi^*(t; \Xi)$. In particular, one could take the null process for y and $\varphi = f$. This trivial choice makes x^φ deterministic, which is most likely not the desired result. A common practice to circumvent this indeterminacy is thus to either (i) extend the set of calibration instruments, incorporating products that are sensitive to volatility (such as interest-rate or credit options), or (ii) require the shadow process to fit the market *as best as possible* (to obtain a specific set Ξ^* of parameters) and then take $\varphi(t; \Xi^*)$ as a shift function:

$$x_t^{\varphi^*} = y_t^* + \varphi^*(t) \quad \text{where} \quad y^* := y(\Xi^*) , \quad \varphi^*(t) := \varphi(t; \Xi^*) , \quad \Xi^* := \arg \min_{\Xi} \left\langle P^{y(\Xi)} | P \right\rangle . \quad (8)$$

The role of the shift is thus merely to compensate for the remaining discrepancies between P (the market curve) and P^{y^*} (the curve generated by the *best* shadow process, y^*).

3.2 Dealing with the positivity constraint

As discussed above, the deterministic shift extension nicely solves Problem 1. To solve Problem 2, one could simply proceed similarly but consider a non-negative shadow process y . However, it is clear from (7) that there is no reason for the shifted process x^φ to remain non-negative even when considering a non-negative shadow process (this issue will be illustrated using a real

example in Section 6.1).⁶ Therefore, this strategy does not provide a valid solution to Problem 2.

To satisfy the non-negativity constraint mentioned in Problem 2, one needs to impose the non-negativity constraint on the shift *at the optimal parameters*. The shift function under the positivity constraint is referred to with the notation $\varphi^{*,+}$ (to stress the difference with the unconstrained counterpart, φ^*).

Lemma 3. *Let y be an AJD that is non-negative on $[0, T]$ with parameters $\Xi^{*,+}$ given by*

$$\Xi^{*,+} := \arg \min_{\Xi} \left\langle P^{y(\Xi)} | P \right\rangle \quad \text{subject to} \quad f^{y(\Xi)}(t) \leq f(t), \quad \forall 0 \leq t \leq T. \quad (9)$$

Then, the x^φ model (6) with $\varphi(t) \leftarrow \varphi^{*,+}(t) := \varphi^*(t; \Xi^{*,+})$ solves Problem 2.

Note that there always exists a set of parameters Ξ such that the constraint is satisfied. Indeed, all parameters Ξ associated with the deterministic case $y = 0$ yield $f^{y(\Xi)} = 0$. Clearly, the constraint is satisfied since f is strictly positive given that P is strictly decreasing, by assumption. The shift is simply given by the market forward rate $\varphi \leftarrow \varphi^{*,+} = f$. However, the trivial process parameter is likely not to be satisfactory.

4 A deterministic time-change extension

To circumvent the drawbacks of the deterministic shift extension with regard to Problem 2, we propose a different approach but still rely on framework (3). The only difference is that we consider an affine transform of a time-changed AJD.

The geometry of the perfect fit problem suggests that a given decreasing discount curve P can be reproduced by “stretching” the x -axis of another decreasing P^y discount curve in an appropriate way. This, in turn, suggests that the perfect fit problem could be addressed by adequately time-changing a shadow process. We show in this section that if the discount curve P^y of the shadow model is known analytically, the time-change function achieving the perfect fit is easy to obtain. Moreover, if the time-change function does not differ excessively from calendar time, the dynamics of the time-changed process will resemble those of the original function. Because time-changing a positive process results in another positive process, the time-change approach offers an appealing solution to the perfect fit problem under positivity constraints.

The x model is obtained by time-changing an AJD using a specific (but deterministic) *clock* Θ that may differ from the calendar clock.

Definition 2 (Clock). *A clock is an application*

$$\Theta : \mathbb{R}^+ \rightarrow \mathbb{R}^+, \quad t \mapsto \Theta(t)$$

that is grounded, increasing and differentiable. In other words, a clock is any function Θ of the form

$$\Theta(t) := \int_0^t \theta(u) du \quad \text{where} \quad \theta(u) > 0, \quad \forall u \geq 0.$$

⁶Even worse, (8) seems contradictory with non-negative shift functions: from (8), P^{y^*} passes through P . Consequently, φ^* will lead to a perfect fit but will correct for both negative and positive errors. In other words, φ^* will change in sign.

We define our model as $x = x^\theta$, obtained from the following transform of the shadow process y :

$$x_t^\theta := \theta(t)y_{\Theta(t)} . \quad (10)$$

Defining the process $y^\theta := \{y_{\Theta(t)}, t \geq 0\}$, the above model fits into (3) with $m = 1$. The next lemma provides the dynamics of x^θ in the general case where the coefficients in the shadow process (5) are time-dependent. It shows, in particular, that the time-change model belongs to the class (3) of ATSMs under the new filtration.⁷

Lemma 4. *Let Θ be a clock and y be a process solving (5) where, here, the $a, b, c, d, \alpha, \omega$ coefficients are allowed to be time-dependent. Then, x^θ in (10) is an ATSM under the filtration $\mathbb{F}^\theta := \{\mathcal{F}_{\Theta(t)}, t \geq 0\}$. Furthermore, y^θ solves a stochastic differential equation*

$$dy_t^\theta = \left(a^\theta(t) + b^\theta(t)y_t^\theta \right) dt + \sqrt{c^\theta(t) + d^\theta(t)y_t^\theta} dB_t + dJ_t^\theta$$

where $g^\theta(u) := g(\Theta(u))\theta(u)$ for $g \in \{a, b, c, d\}$, B is an $\mathbb{F}^\theta$ -Brownian motion and J^θ is a jump process under $\mathbb{F}^\theta$ with known time- t probability distribution $\nu_t^\theta := \nu_{\Theta(t)}$ and intensity $\omega^\theta(t, x) = \omega(\Theta(t), x)\theta(t)$.

Proof. It is obvious from (10) that x^θ takes the form (3) with $\rho \leftarrow \theta$, $\varphi \leftarrow 0$ and $y \leftarrow y^\theta$. It remains to be shown that the zero-coupon bond price under x^θ has the exponential affine form in (4). This is straightforward, using a change of variable:

$$\begin{aligned} P_s^{x^\theta(\Xi)}(t) &= \mathbb{E} \left[e^{-\int_s^t x_u^\theta du} \middle| \mathcal{F}_{\Theta(s)} \right] \\ &= \mathbb{E} \left[e^{-\int_s^t \theta(u)y_{\Theta(u)}^\theta du} \middle| \mathcal{F}_{\Theta(s)} \right] \\ &= \mathbb{E} \left[e^{-\int_{\Theta(s)}^{\Theta(t)} y_u^\theta du} \middle| \mathcal{F}_{\Theta(s)} \right] \\ &= \exp \left\{ A_{\Theta(s)}^y(\Theta(t); \Xi) - B_{\Theta(s)}^y(\Theta(t); \Xi)y_s^\theta \right\} . \end{aligned} \quad (11)$$

Hence, the zero-coupon bond price under model x^θ admits an exponential affine form with $A_s^{x^\theta}(t; \Xi) = A_{\Theta(s)}^y(\Theta(t); \Xi)$ and $B_s^{x^\theta}(t; \Xi) = B_{\Theta(s)}^y(\Theta(t); \Xi)$. It follows that x^θ is an ATSM.

To compute the dynamics of y^θ , let us first observe that

$$y_t^\theta := y_{\Theta(t)} = y_0 + \int_0^{\Theta(t)} (a(u) + b(u)y_u) du + \int_0^{\Theta(t)} \sqrt{c(u) + d(u)y_u} dW_u + \int_0^{\Theta(t)} dJ_u .$$

The integral of the drift term yields

$$\int_0^{\Theta(t)} (a(u) + b(u)y_u) du = \int_0^t (a(\Theta(u)) + b(\Theta(u))y_{\Theta(u)}) \theta(u) du = \int_0^t \left(a^\theta(u) + b^\theta(u)y_u^\theta \right) du ,$$

where $a^\theta(t) := a(\Theta(t))\theta(t)$ and similarly for b^θ . With regard to the diffusion coefficient,

$$\int_0^{\Theta(t)} \sqrt{c(u) + d(u)y_u} dW_u = \int_0^t \sqrt{c(\Theta(u)) + d(\Theta(u))y_{\Theta(u)}} dW_{\Theta(u)} .$$

⁷Note that this property is specific to deterministic time changes. For instance, Example 6.5 in [Filipovic and Larsson \(2020\)](#) shows that a Vasicek process $dx_t = -\kappa x_t dt + \sigma dW_t$ time-changed with an independent Poisson subordinator is not affine.

Now, observe that Θ is a clock; it is an increasing function of time, and the process $W^\theta := \{W_{\Theta(t)}, t \geq 0\}$ is a local martingale with quadratic variation $\langle W^\theta, W^\theta \rangle_t = \Theta(t)$. From (Jeanblanc et al., 2009), the process $B := \{B_t, t \geq 0\}$ defined as

$$B_t := \int_0^t \frac{1}{\sqrt{\theta(u)}} dW_{\Theta(u)} \quad (12)$$

is a Brownian motion. Hence,

$$\int_0^{\Theta(t)} \sqrt{c(u) + d(u)y_u} dW_u = \int_0^t \sqrt{c(\Theta(u)) + d(\Theta(u))y_u^\theta} \sqrt{\theta(u)} dB_u = \int_0^t \sqrt{c^\theta(u) + d^\theta(u)y_u^\theta} dB_u,$$

where c^θ, d^θ are defined in a similar way as before and B is an $\mathbb{F}^\theta$ -Brownian motion. With regard to the jump term, note that

$$\int_0^{\Theta(t)} dJ_u = J_{\Theta(t)} =: J_t^\theta$$

where J^θ is a jump process with probability distribution $\nu_t^\theta := \nu_{\Theta(t)}$ and intensity $\omega^\theta(t, \cdot) := \omega(\Theta(t), \cdot)\theta(t)$. Because $\omega(t, x)$ is affine in x , $\omega^\theta(t, x) = \omega_0^\theta(t) + \omega_1^\theta(t)x$ where $\omega_0^\theta(t) := \omega_0(\Theta(t))\theta(t)$ and similarly for ω_1^θ . The proof is concluded by differentiating y^θ . $\square$

Remark 3. The above lemma shows that whenever y is a one-factor ATSM (under $\mathbb{F}$) with known A^y, B^y functions, x^θ is an ATSM (under $\mathbb{F}^\theta$) with known $A^{x^\theta}, B^{x^\theta}$ functions. However, more generally, even if y is not exponential affine, the ZCB price of x^θ is a simple transform of the ZCB of y : $P_s^{x^\theta(\Xi)}(t) = P_{\Theta(s)}^{y(\Xi)}(\Theta(t))$, showing that the x^θ model is tractable as soon as y is. Hence, both the shift and the time-change approach could be used to calibrate many other models, such as linear-rational term-structure models, for example Filipovic et al. (2017). Similarly, the results shown in Lemma 4 can be extended to the multi-factor case provided that we time-change the ATMS (3) instead of the corresponding AJD (for instance, we could time-change $e \cdot y$ with a single clock, where y is an n -dimensional AJD and e is a conformable vector of ones). The resulting model will also be an ATMS under $\mathbb{F}^\theta$.

This model is appealing for several reasons. First, just as the shift extension, it is an ATSM; hence, it is a tractable model. Second, because x_t^θ is a positive rescaling of the process y sampled at time $\Theta(t)$, the range of x^θ is linked to that of y . In particular, if the range of y is $\mathbb{R}$, as for the Vasicek model, then so is the range of x^θ . However, if y is non-negative as in the JCIR case, then so is x^θ . This solves the drawback of the shift approach related to Problem 2. Finally, the time-dependent feature of the clock rate θ is expected to provide additional flexibility in the calibration properties of x^θ with respect to that of the homogeneous model y .

Two questions remain open in this respect. First, we need to clarify the circumstances under which the model provides a perfect fit. Second, in the case where the perfect fit can be achieved, we need to provide an efficient procedure to compute the resulting *optimal clock*, Θ^* . The price to pay is that the time-change model is not as flexible as the deterministic shift approach. Indeed, starting with a given model y , the x^θ model can only generate specific shapes for discount curves. We are thus more dependent on the initial choice of the base model y . Fortunately, it turns out that a perfect fit is achievable for a wide set of market curves, including all decreasing discount curves, considered in Problem 2. This is clearly the

most important case since (i) it corresponds to the case in which the shift approach fails to provide a convincing solution, and (ii) it is probably the most common case in practice, since it encompasses the class of discount curves with non-negative rates (or, more generally, with non-negative instantaneous forward rates), as well as the set of all continuous survival probability curves. Moreover, even if the mathematical expression for the clock Θ^* is not available in closed form, its numerical computation turns out to be easy.⁸

Theorem 1. *Let P be a discount curve and y be a shadow model such that P^y is a discount curve. Define x^θ as in (10). Then, $P^{x^\theta} = P$ provided that $\Theta \leftarrow \Theta^*$, where Θ^* satisfies the first-order ODE*

$$\theta^*(t) := \frac{d}{dt}\Theta^*(t) = \frac{f(t)}{f^y(\Theta^*(t))}, \quad (13)$$

with f, f^y the associated instantaneous forward curves. Moreover, if P and P^y are strictly decreasing, the solution to (13) exists, is a clock, and is given by

$$\Theta^*(t) := Q^y(P(t)), \quad (14)$$

where Q^y is the inverse of the shadow-process discount curve, P^y .

Proof. See Section 8.3. □

Observe that the optimal clock Θ^* actually depends on the parameters Ξ of the shadow model. Similar to the shift, we actually have $\Theta^*(t) = \Theta^*(t; \Xi)$. Although other frameworks are possible, we set $\Xi = \Xi^*$ as in (8). Similar to the function φ in the shift approach, the purpose of the clock Θ is then to absorb the remaining errors between P^{y^*} and P .

5 Focus on the JCIR shadow process

In this section, we provide specific results associated to the case where we consider a JCIR as shadow process. This process is positive and hence can help solving Problem 2. Moreover, it is very tractable and, therefore, extra analytical results can be found.⁹

The JCIR dynamics are

$$dy_t = \kappa(\beta - y_t)dt + \delta\sqrt{y_t}dW_t + dJ_t, \quad y_0 > 0, \quad (15)$$

where κ, β, δ are strictly positive constants and J is a compounded Poisson process with non-negative parameters ω, α (the CIR is recovered as a special case by choosing (α, ω) such that $\alpha\omega = 0$). This SDE corresponds to (5) provided that $(a, b, c, d) = (\kappa\beta, -\kappa, 0, \delta^2)$ for which many quantities (including the functions A^y, B^y , moments and moment generating functions) are known in closed-form; see Appendix (Section 8.1.3). As a result, the process x^θ in (10) is also tractable. For example, the k -th moment of x_t^θ is $m^{x^\theta}(k, t) = \theta(t)^k m^y(k, \Theta(t))$, and the moment generating function is $\psi^{x^\theta}(u, t) = e^{u\theta(t)} \psi^y(u, \Theta(t))$. More interestingly, the k -th moment and the moment generating function of $X_t^\theta := \int_0^t x_u^\theta du$ coincide with those of $Y_{\Theta(t)} := \int_0^{\Theta(t)} y_u du$.

⁸When no confusion is possible, the explicit reference to the model parameters Ξ is avoided to simplify notation.

⁹The Vasicek process is also very tractable but is not appropriate for Problem 2. Hence, the results are relegated in the Appendix.

The instantaneous forward curve is given by $f_0^{JCIR}(t)$ in (31):

$$f^y(t) = \frac{2\kappa\beta(e^{t\gamma} - 1)}{2\gamma + (\kappa + \gamma)(e^{t\gamma} - 1)} + y_0 \frac{4\gamma^2 e^{t\gamma}}{[2\gamma + (\kappa + \gamma)(e^{t\gamma} - 1)]^2} + \frac{2\omega\alpha(e^{t\gamma} - 1)}{2\gamma + (\kappa + \gamma + 2\alpha)(e^{t\gamma} - 1)}, \quad (16)$$

where $\gamma := \sqrt{\kappa^2 + 2\delta^2}$.

5.1 The shifted JCIR

The implied shift $\varphi^*(t, \Xi)$ is given by eq.(7). The dynamics of x^{φ^*} thus take the form

$$dx^{\varphi^*} = d\varphi^*(t, \Xi) + dy_t. \quad (17)$$

When φ^* admits a derivative, x^{φ^*} takes a similar form as (15), except that it becomes time-inhomogeneous (the long-term mean β becomes $\beta(t) := \beta + (\varphi^*(t, \Xi))'/\kappa$). Note, however, that in many applications, φ is usually not continuous w.r.t. t , such that $\varphi^{*'} does not exist everywhere.$

Two specifications of the shifted JCIR model can be considered, corresponding to two calibration procedures for the JCIR shadow model. The first specification, called S-JCIR, consists of using as the shadow model the JCIR with parameters Ξ^* (the parameters found without imposing a positivity constraint, calibrating the JCIR using (8)). The resulting shift function is thus $\varphi^*(t) = \varphi^*(t; \Xi^*)$. The second specification, called PS-JCIR, consists of using JCIR parameters $\Xi^{*,+}$ (the parameters found when imposing a positivity constraint, calibrating the JCIR using (9)). The resulting shift function is thus $\varphi^{*,+}(t) = \varphi^*(t; \Xi^{*,+})$. Note that in contrast to S-JCIR, the PS-JCIR model ensures non-negativity.

Although the PS-JCIR allows both for a perfect fit and the non-negativity constraint, one may argue that it is not as tractable as the JCIR. Indeed, the optimization problem (9) is more difficult than (8) due to the constraint on the instantaneous forwards, even if sufficient conditions on the parameters can be found. Second, and probably more important, this constraint is binding, in the sense that it often profoundly impacts the optimal parameter set.

In practice, the PS-CIR model often yields dynamics associated with shadow processes featuring “little randomness”. This will be illustrated in Section 5.2, first by comparing the variance of the integrated S-JCIR and PS-JCIR processes, then by considering financial applications. This was already noted in (Brigo and Mercurio, 2006, sec. 3.9.3, p.107-109). To circumvent this issue, the authors suggest dropping the positivity constraint, considering S-CIR instead of PS-CIR. While this is perfectly fine in an interest rate framework (where positivity might be desirable in some cases but is not theoretically required), this is much more problematic when dealing with default intensities: such applications typically require large volatility but also strict positivity. Increasing the variance of the shifted CIR process without breaking Feller’s constraint can be achieved by incorporating compound Poisson jumps (considering JCIR instead of CIR as the shadow model). Unfortunately, increasing the jump activity while maintaining the calibration to a given market curve f is difficult under the positivity constraint. Indeed, the minimum of the implied shift function is driven down when increasing the jump activity because the difference in the instantaneous forward rates, $f^{JCIR}(t) - f^{CIR}(t)$, is non-negative and increases with ω, α for $\alpha, \omega > 0$ (see the Appendix, Section 8.1.3). This observation combined with (7) leads to a lower shift function φ for the shifted JCIR than for the corresponding shifted

CIR. For this reason, there is a need for an alternative to S-JCIR and PS-JCIR that combines (i) tractability, (ii) the perfect fit feature, (iii) the large implied volatility and (iv) positivity.

5.2 The time-changed JCIR

It is clear from the previous results that time-changing a JCIR provides a solution to Problem 2.

Corollary 1. *Let y be a JCIR with parameters Ξ . Then, model x^θ defined in (10) with $\Theta \leftarrow \Theta^*(t; \Xi)$ solves Problem 2.*

Proof. Because y is a JCIR, P^y is a discount curve and is analytically tractable. Moreover, the latter is strictly decreasing since y is almost certainly positive. We conclude the proof by relying on Theorem 1. $\square$

The optimal clock Θ^* leading to the perfect fit to a given strictly decreasing curve P is given by (13), where the forward curve f^y is given in (16). From Lemma 4, x^θ is clearly an ATSM with dynamics:

$$dx_t^\theta = x_t^\theta \frac{d\theta(t)}{\theta(t)} + \kappa(t) (\beta(t) - x_t^\theta) dt + \delta(t) \sqrt{x_t^\theta} dB_t + dJ_t^\theta \quad (18)$$

where $\beta(t) := \beta\theta(t)$, $\kappa(t) := \kappa\theta(t)$, $\delta(t) = \delta\theta(t)$ and J^θ is a pure jump process whose time- t jump-size distribution ν_t is exponential with mean $\alpha(t) = \alpha\theta(t)$ and jump intensity $\omega(t) := \omega\theta(t)$. The dynamics of TC-JCIR consist of two parts: a first part, which corresponds to that of a JCIR with time-dependent coefficients and an extra drift correction.¹⁰

In contrast to S-JCIR x^{φ^*} , the positivity constraint on x^{θ^*} is automatically satisfied for every (strictly decreasing) market curve and every Ξ (such that y is not trivially equal to 0). The PS-JCIR model $x^{\varphi^*,+}$ avoids this problem but at the expense of having a process that is, to a large extent, deterministic (i.e., $x_t^{\varphi^*,+}$ varies in a small neighbourhood around $f(t)$; see Section 6). Consequently, the TC-JCIR model is expected to feature higher volatility than the corresponding PS-JCIR, at least up to some time horizon. This is summarized in the next theorem.

Theorem 2. *Let P be a strictly decreasing discount curve with instantaneous forward rate curve f and y be a JCIR process with parameter $\Xi^{*,+}$ (such that the perfect fit shifted-JCIR model $x_t^{\varphi^*}$ actually corresponds to PS-JCIR and is positive). Then, the ODE (13) with f^y given by (16) admits a solution that satisfies $\Theta^*(t) = \Theta^*(t; \Xi^{*,+}) \geq t$. Moreover, the variance of the corresponding perfect fit TC-JCIR model $x_t^{\theta^*}$ satisfies:¹¹*

$$1) \mathbb{V} [X_t^{\theta^*}] \geq \mathbb{V} [X_t^{\varphi^*}], \forall t \geq 0,$$

$$2) \mathbb{V} [x_t^{\theta^*}] \geq \mathbb{V} [x_t^{\varphi^*}] \text{ if one of the following holds:}$$

$$i) y_0 = \beta + \frac{\omega\alpha}{\kappa},$$

¹⁰Note that when θ is differentiable, $x_t^\theta \frac{d\theta(t)}{\theta(t)} = x_t^\theta \frac{\theta'(t)}{\theta(t)} dt$ can be absorbed in the affine drift.

¹¹The sufficient conditions (i)-(iv) are related to cases where $\mathbb{V}[y_t]$ is increasing and $f^{JCIR}(t)$ decreasing with respect to t .

- ii) f constant and $y_0 \leq \beta + \frac{\omega\alpha}{\kappa}$,
- iii) $y_0 > \beta + \frac{\omega\alpha}{\kappa}$ and $t < \Theta^{\star-1}(t_1)$, and
- iv) $(\kappa\beta + \omega\alpha)/\gamma < y_0 < \beta + \frac{\omega\alpha}{\kappa}$ and $t > \Theta^{\star-1}(t_2)$

where

$$t_1 := \frac{1}{\kappa} \ln \left(1 + \frac{y_0 + 2\omega\alpha^2/\gamma^2}{y_0 - \beta - \omega\alpha/\kappa} \right) \quad \text{and} \quad t_2 := \frac{1}{\gamma} \ln \frac{(\gamma - \kappa)(\kappa\beta + y_0\gamma + \omega\alpha) - 2\omega\alpha^2}{(\kappa + \gamma)(y_0\gamma - \kappa\beta - \omega\alpha) - 2\omega\alpha^2}.$$

Proof. See Section 8.4. □

In summary, the TC-JCIR model provides an elegant solution to Problem 2: it is non-negative (in contrast with S-JCIR), is almost as tractable as the simple JCIR (in contrast with PS-JCIR), provides a perfect fit to every strictly decreasing discount curve (as both S-JCIR and PS-JCIR) and, to some extent, exhibits a larger variance (than PS-JCIR). Therefore, when a positivity constraint is required, the TC-JCIR avoids the drawbacks of the S-JCIR models. The only price to pay is that the clock is not available in closed form but requires a (simple) numerical inversion. These features are illustrated in the next section on various applications taken from credit risk modelling.

We conclude this section with several observations about the variance of the integrated TC-JCIR. These relationships will indeed prevail for the calibrated intensity processes involved in the application section (the corresponding figures can be found in the Appendix, Section 8.5).

- First, the integrated JCIR process with parameter Ξ^* is expected to feature a larger variance than that with parameter $\Xi^{\star,+}$:

$$\mathbb{V}[Y_t(\Xi^*)] \geq \mathbb{V}[Y_t(\Xi^{\star,+})], \text{ in general.} \quad (19)$$

Indeed, the aim of the shift is to “compensate” the trend of the shadow process to satisfy the calibration constraint. However, because of the zero lower bound, increasing the volatility of a JCIR will push the trend upward, and this cannot be “compensated” with a non-negative shift. The only way for $y(\Xi^{\star,+})$ to satisfy both constraints is to embed limited randomness. By contrast, there is no limit on the volatility of $y(\Xi^*)$ since an arbitrarily large upward trend can be compensated thanks to a negative shift. Consequently, the volatility of $y(\Xi^{\star,+})$ is expected to be smaller than that of $y(\Xi^*)$, in general. The same applies to the integrated processes.

- Second, the variance of the integrated TC-JCIR is larger than that of the equivalent “valid” integrated PS-JCIR:

$$\mathbb{V}[Y_{\Theta^{\star}(t)}(\Xi^{\star,+})] \geq \mathbb{V}[Y_t(\Xi^{\star,+})], \text{ always.} \quad (20)$$

Indeed, it is known from Theorem 2 that when working under the positivity constraint (i.e., when y is driven by $\Xi^{\star,+}$), $\Theta^{\star}(t) := \Theta(t; \Xi^{\star}) \geq t$. The variance inequality then results from the fact that, for any parameter, the variance of Y is an increasing function of time (Lemma 6 in the Appendix, Section 8.1.2).

- Third, the variance of the TC-JCIR using Ξ^* is expected to be comparable to the variance of the S-JCIR with the same parameter:

$$\mathbb{V}[Y_{\Theta^*(t)}(\Xi^*)] \approx \mathbb{V}[Y_t(\Xi^*)]. \quad (21)$$

This can be understood intuitively as follows. As explained above, the parameter $\Xi \leftarrow \Xi^*$ computed using (8) leads HAJD y to best fit the market curve, and the clock is used to absorb the remaining discrepancies. Therefore, if the fit between P and P^y is not too poor, one expects the clock not to deviate much from the actual time, i.e., $\Theta^*(t) \approx t$. In this case, $x_t^{\theta^*} = \theta^*(t)y_t^{\theta^*} \approx y_{\Theta^*(t)} \approx y_t$.

6 Application to credit risk modelling

We consider a reduced-form default model as in Section 2.1.2. The stochastic intensity will be modelled by using a CIR as a shadow process. We consider three models for the $\mathbb{F}$ -intensity process: S-CIR, PS-CIR, and TC-CIR. Each model will be perfectly calibrated to the same survival probability curve. For the reader's convenience, the acronyms and descriptions of these models are summarized in Table 1.¹²

Acronym	Positivity	Definition	Ξ
S-CIR	No	$\lambda^\varphi = y + \varphi^*$	Ξ^*
PS-CIR	Yes	$\lambda^{\varphi,+} = y + \varphi^{*,+}$	$\Xi^{*,+}$
TC-CIR	Yes	$\lambda^\theta = \theta^* y^{\theta^*}$	Ξ^*

Table 1: Intensity models considered in this section. In all cases, the shadow model (y) is a CIR, but the set of parameters Ξ may differ. All models rely on the market-implied shift or time-change functions and, therefore, lead to a perfect fit to a given survival probability curve.

Bear in mind that in this context, the S-CIR model is actually flawed because there is a non-zero probability of observing negative intensities. However, we present the results of the model as a benchmark since, as explained in the introduction, it remains a common approach despite its drawbacks.

We compare the shifted CIR (S-CIR and PS-CIR) to the time-changed CIR (TC-CIR) on several aspects related to a real-world example where the reference entity is *Ford Inc.* We also discuss the TC-JCIR case when relevant. We first analyse the perfect fit feature of both types of models, as well as the non-negativity property of λ . A comparison of the variance of the various integrated processes Λ is provided in Appendix 8.5. We then analyse their behaviours in two different applications, namely, the pricing of various credit default swaptions (otherwise known as CDS options, or CDSO) with Ford as the reference entity, or on the credit valuation adjustment (CVA) of prototypical FRA and IRS exposures where Ford is the trade's counterparty.

It is widely acknowledged that “pure credit instruments” such as CDS or CDSO are quite insensitive to the stochasticity of interest rates in realistic conditions. This has been discussed

¹²Note that to simplify notation, we drop the $\star$ superscript in λ^φ , $\lambda^{\varphi,+}$ and λ^θ . However, in the sequel, we always assume that the shift function φ in λ^φ and the clock function θ in λ^θ correspond to those given by the perfect fit constraint (i.e., to φ^* , $\varphi^{*,+}$ and θ^* for the S-CIR, PS-CIR and TC-CIR, respectively).

explicitly for the CIR base model in (Brigo and Alfonsi, 2005) and (Brigo and Cousot, 2006). Hence, we consider a deterministic short-rate process, which is stressed by the notation $r_u = r(u)$. In this case, one simply obtains $P_s^r(t) = D_s(t) = e^{-\int_s^t r(u)du}$.¹³

6.1 Perfect fit of CDS term structure

A CDS is a financial instrument used by two parties – called the protection buyer and the protection seller – allowing to transfer to the protection seller the financial loss that the protection buyer would suffer if a particular default event happened to a third party called the *reference entity*. We let τ be the default time of the latter. In a default swap contracted at time t , started at time T_a with maturity T_b , the protection buyer pays a coupon (of spread) k at a set of payment dates $T_a, \dots, T_b$ as long as the reference entity does not default. The protection seller agrees to make a single payment LGD to the protection buyer if the default occurs between T_a and T_b . When applicable, the protection buyer makes a final payment corresponding to the spread accrued since the last payment date before default. For further details on the mechanics of this product, we refer to (Brigo and Alfonsi, 2005) and (Brigo and El-Bachir, 2010).¹⁴

The CDS term structure consists of a set of par spreads associated with CDSs of various maturities. The time- t par spread $s_t(T_i)$ of a CDS contract of maturity T_i is defined as the contract spread k that sets the value of the CDS contract to 0 at time t . Table 2 gives the CDS term-structure of Ford taken from Bloomberg on November 12, 2018.

Maturity (years)	1	3	5	7	10
Spread (bps)	18.3	136.6	191.9	267.6	280.6

Table 2: CDS spread term structure of Ford on November 12, 2018. Source: Bloomberg.

In this context, the market curve to be fitted is the risk-neutral survival probability curve, defined as $G(t) := \mathbb{Q}(\tau > t)$ associated with the default time τ of a given reference entity (here, Ford Inc.). It can be extracted from the CDS term structure by inverting the no-arbitrage pricing formulae of the corresponding contracts. In practice, one only has a few calibration equations, say n , given by the number of market quotes (here, $n = 5$). It is therefore not possible to estimate the full (i.e., infinite-dimensional) market curve G without further assumptions. It is common market practice to consider the CDS model from the International Swap and Derivative Association (ISDA) – otherwise known as the JP Morgan model – (Markit, 2004), which provides a slightly simplified version of the actual no-arbitrage pricing formula applied to CDSs. In this approach, the survival probability G is parametrized via a positive hazard rate function h playing a similar role as the instantaneous forward rate f :

$$G(t) := e^{-\int_0^t h(s)ds}.$$

In this expression, h is itself parametrized by n constants $h_1, h_2, \dots, h_n$ bootstrapped from the spreads $s_1, s_2, \dots, s_n$ associated with the maturities $T_1, T_2, \dots, T_n$. We follow market practice

¹³Given that the interest rates have little impact on the figure and that our main objective is to discuss the impact of the default model, we considered the zero risk-free rate in the numerical applications below.

¹⁴For further details on actual market conventions, we refer the interested reader to (Markit, 2004) and (Markit, March 13, 2009).

and assume that h is piecewise constant between the maturities with flat extrapolation, i.e.,

$$h(t) = \sum_{i=1}^n \mathbb{1}_{\{T_{i-1} \leq t < T_i\}} h_i + \mathbb{1}_{\{t \geq T_n\}} h_n ,$$

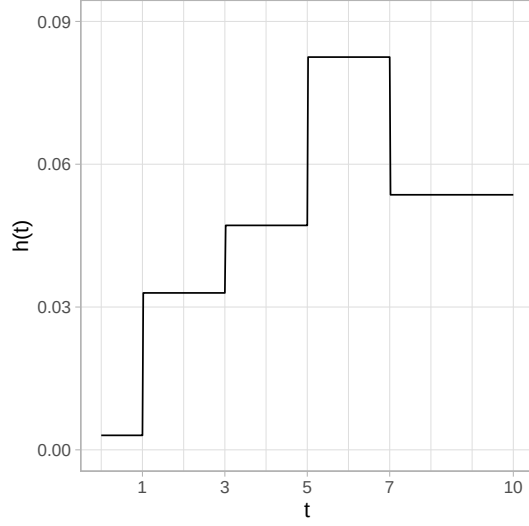
where $T_0 := 0$, $LGD := 1 - R$ with $R = 40\%$ the assumed recovery rate of the firm and h_i s are positive constants. In the sequel, we set our calibration horizon to the last observable tenor: $T = T_n$. The piecewise constant hazard rate function h associated with our Ford example is shown in Fig. 1(a). The corresponding survival probability curve G is shown in Fig. 1(d).

Calibrating CIR shadow model y yields Ξ^* shown in Table 3. In line with market practice, we used (1) with $P^x \leftarrow P^y$ considering (2) as the error function and $\mathcal{T}$ as the set of available liquid CDS maturities available. Fig. 1(b) and Fig. 1(c) show the implied shift (φ^* given by (8)) and implied clock (Θ^* given by (14)) leading to calibrated S-CIR and TC-CIR, respectively. By construction, the survival probability curves implied by either model are perfectly in line with the market curve G shown in Fig. 1(d).

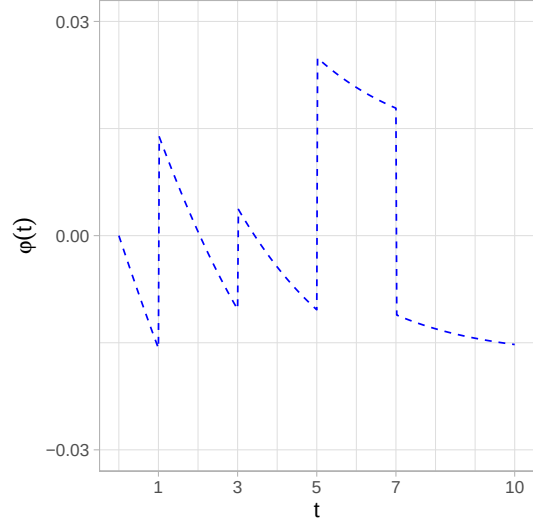
Ξ	κ	β	δ
Ξ^*	0.0555	0.3018	0.2939
$\Xi^{*,+}$	0.2118	0.0030	0.0006

Table 3: Calibration parameters in the Ford example. Parameters Ξ^* and $\Xi^{*,+}$ correspond to the parameters of CIR model y with and without a positivity constraint, where y_0 is set exogenously to the first level of the piecewise hazard rate function, $h_0 = 0.0030$. The parameters Ξ_0^* and $\Xi_0^{*,+}$ (corresponding to the similar cases but where y_0 is a parameter that enters the optimization procedure) are given in Table 5 in the Appendix (Section 8.7.1).

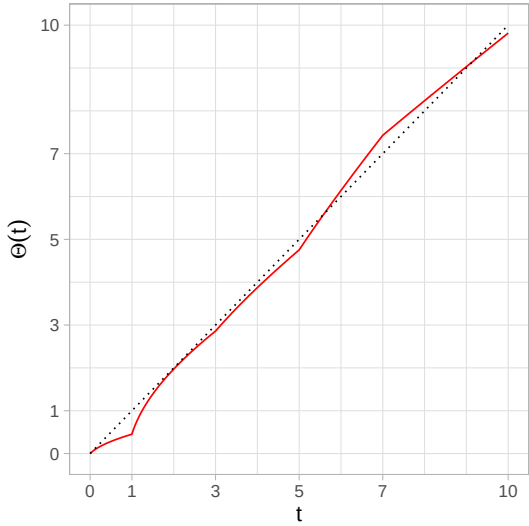
It is clear from Fig. 1(b) that the shift function φ^* can take negative values, so that the S-CIR yields negative “default intensities” λ^φ and is thus flawed. This contrasts with the TC-CIR approach since λ^θ is a positive process if y is. To fix this issue in the deterministic shift framework, one can rely on the PS-CIR model, $\lambda^{\varphi,+}$, which consists of taking $\Xi^{*,+}$ as parameters of the shadow CIR (see Table 3). As illustrated in Figure 2 with our Ford example, this procedure is very restrictive: it leads to a curve $P^{y(\Xi^{*,+})}$ that is decreasing at a very low rate (blue dotted curve in Fig. 2(a)). In particular, the shape of $P^{\lambda^{\varphi,+}}$ essentially results from the shift, not from the shadow process. This is problematic: it basically amounts to stating that $h \approx \varphi^{*,+}$ (compare Fig. 1(a) to Fig. 2(b)), i.e., that the PS-CIR process $\lambda^{\varphi,+}$ is essentially deterministic. This will impose strong limitations on the resulting default model and will be further discussed in the remaining subsections.



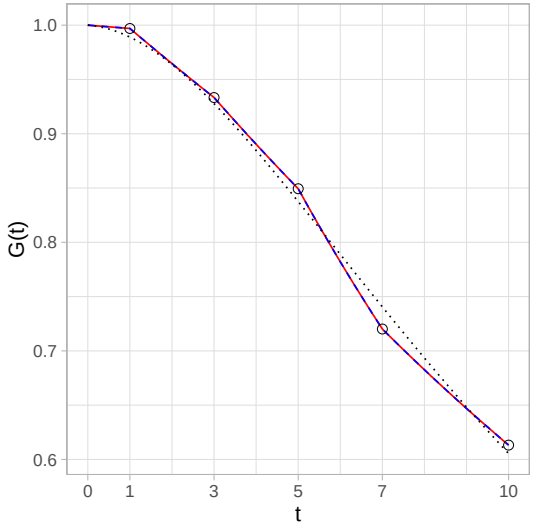
(a) Hazard rate function (h)



(b) Optimal shift ($\varphi^* = \varphi(\Xi^*)$)



(c) Optimal clock ($\Theta^* = \Theta(\Xi^*)$)



(d) Survival probability ($P^{y(\Xi^*)}, G = P^{\lambda^\varphi} = P^{\lambda^\theta}$)

Figure 1: Fitting Ford CDS term structure with adjusted CIR models. The survival probability curve is parametrized with a piecewise constant hazard rate function h extracted from market prices taken from Bloomberg on November 12, 2018, panel (a). Shadow model y is a CIR with parameters $\Xi = \Xi^*$ given in Table 3, obtained from (8). The shift function φ^* of the S-CIR is shown in panel (b). Panel (c) gives the clock Θ^* of the TC-CIR, while panel (d) yields the survival probability curves given by the market or associated with $\mathbb{Q}(\tau(\lambda) > t)$ for various intensity models: the best shadow model (dotted black), the S-CIR and the TC-CIR (dashed blue-red). By construction, the curves of the last two models coincide with G (circles drawn at CDS maturities).

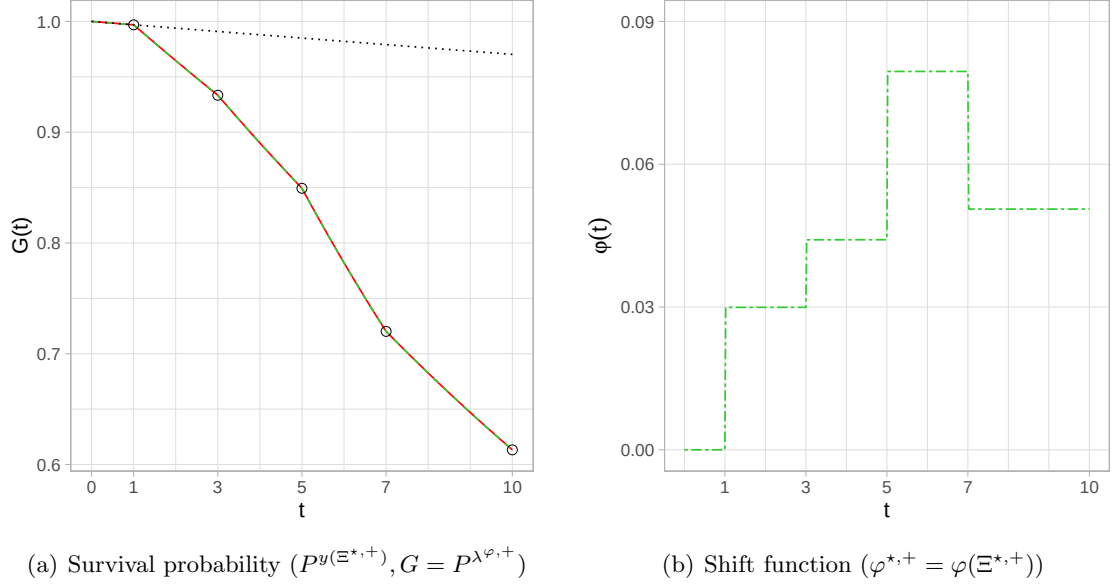
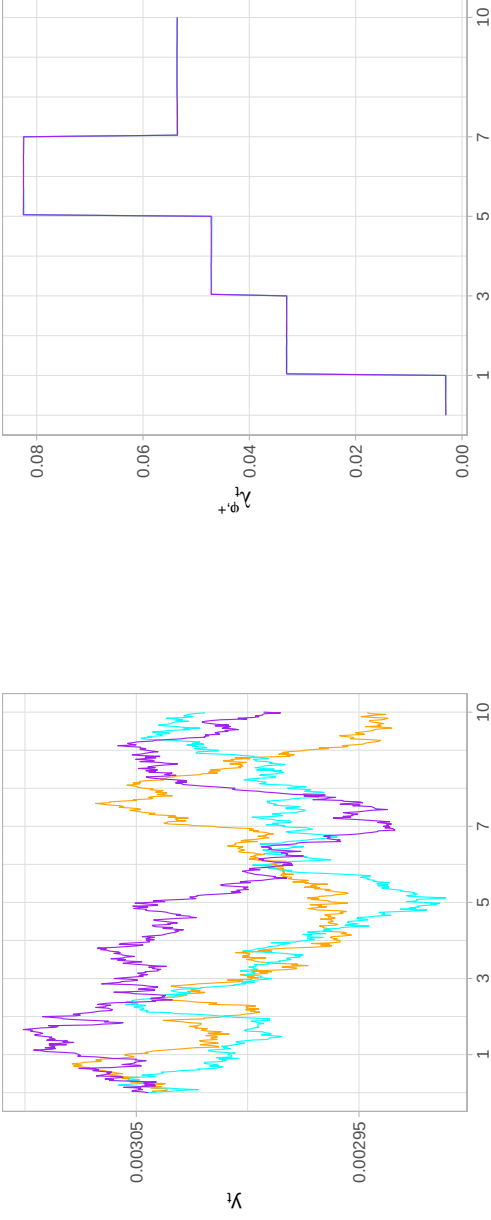
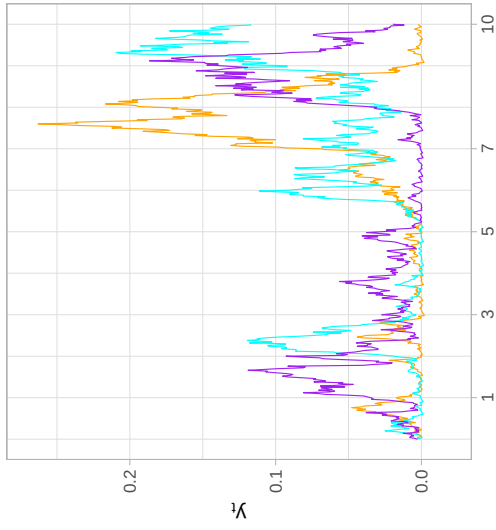


Figure 2: Fitting Ford CDS term structure using PS-CIR. The CIR shadow model has parameters $\Xi^{*,+}$. Panel (a) shows the survival probability curve G (circles drawn at CDS maturities) which, by construction, corresponds to $P^{\lambda^{\varphi^{*,+}}}$ (dashed green-red) and the curve $P^{y(\Xi^{*,+})}$ associated with the shadow model (dotted black). The implied shift function (panel (b)) is non-negative and roughly corresponds to the hazard rate function associated with G (see Fig. 1(a)).

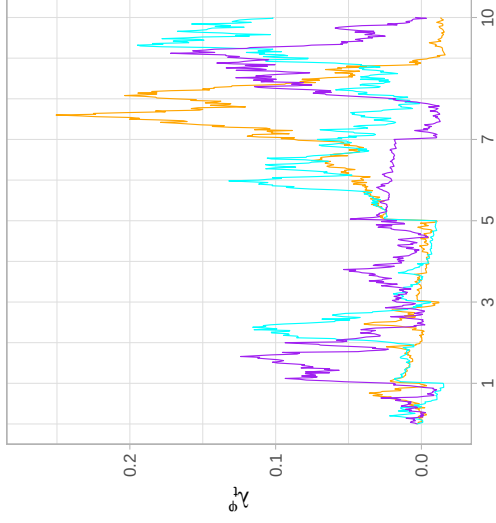
Figure 3 provides a comparative illustration of the dynamics of the S-CIR, PS-CIR and TC-CIR models. Three sample paths of the CIR($\Xi^{*,+}$) model are shown in Fig. 3(a). Applying the shift function $\varphi^{*,+}$ to the latter yields the trajectories of the PS-CIR, Fig. 3(b). In this example, the sample paths of the PS-CIR model essentially collapse to the shift function. This illustrates the fact that calibrating a CIR model using (9) may lead to a model that is almost deterministic. Alternatively, one could relax the positivity constraint and consider as the shadow model a CIR but now calibrated according to (1), i.e., using the set of parameters Ξ^* . Three sample paths of this model are shown in Fig. 3(c). Applying the shift function φ^* to those leads to the S-CIR model; see Fig. 3(d). Comparing Fig. 3(b) and Fig. 3(d), one can see that relaxing the positivity constraint in the calibration of the shadow model tends to increase the featured “randomness”. However, this gain in volatility comes at the expense of obtaining a shift function φ^* that will become negative; hence, the S-CIR yields “intensity” trajectories that can be negative, in contrast with those of the PS-CIR. Note that in both cases, the shift function introduces “jumps” at the CDS tenors resulting from the piecewise constant hazard rate parametrization adopted by the market. Finally, Fig. 3(e) shows the sample paths of the TC-CIR obtained by scaling and time-changing the CIR(Ξ^*) trajectories in Fig. 3(c) using clock Θ^* . Only TC-CIR rules out negative trajectories while maintaining a decent volatility level. In addition, the trajectories of the TC-CIR are closer to those of the shadow model (Fig. 3(c)) than those of the shifted CIR (not all the trajectories jump by the same magnitude at the CDS tenors).



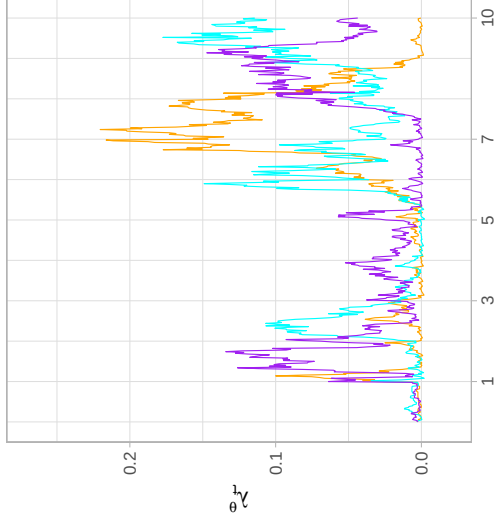
(a) Shadow model $\text{CIR}(\Xi^{*,+})$



(c) Shadow model $\text{CIR}(\Xi^*)$



(d) Shifted model $\text{S-CIR}(\text{CIR}(\Xi^*) + \varphi^*)$



(e) $\text{TC-CIR}(\text{CIR}(\Xi^*) \text{ adjusted with } \Theta^*)$

(b) Shifted model $\text{PS-CIR}(\text{CIR}(\Xi^{*,+}) + \varphi^{*,+})$

Figure 3: Three sample paths of various models associated with Ford. Imposing non-negativity leads to models featuring little randomness (a and b). This problem is mitigated by using CIR parameters found without the positivity constraint (c). Applying the shift leads to the calibrated S-CIR model (d). The trajectories can be negative and exhibit large jumps of deterministic size at the CDS tenors. TC-CIR (e) behaves better: it is calibrated, and the paths are non-negative, do not exhibit large deterministic jumps, and feature decent volatility levels.

6.2 Pricing CDS options

Because a CDSO is an option on a CDS, we start by recalling the no-arbitrage pricing equation of a CDS. We denote by t the valuation time and assume $\tau > t$ because it is pointless to price a CDS (or a CDSO) post-default. From the perspective of the protection buyer, the time- t value of a 1 dollar notional CDS $CDS_t(a, b, k)$ starting at time T_a with maturity T_b , $t \leq T_a < T_b$, a spread k and (known) loss given default LGD is given by the difference between the conditional risk-neutral expectations of the protection and premium cashflows:

$$\begin{aligned} CDS_t(a, b, k) &= \mathbb{E} \left[(1 - R) \mathbb{1}_{\{T_a \leq \tau \leq T_b\}} P_t^r(\tau) | \mathcal{G}_t \right] \\ &\quad - k \mathbb{E} \left[\sum_{i=a+1}^b \left(\mathbb{1}_{\{\tau \geq T_i\}} \alpha_i P_t^r(T_i) + \mathbb{1}_{\{T_{i-1} \leq \tau < T_i\}} \alpha_i \frac{\tau - T_{i-1}}{T_i - T_{i-1}} P_t^r(\tau) \right) \middle| \mathcal{G}_t \right] \end{aligned}$$

with α_i being the day count fraction between dates T_{i-1} and T_i , which, in a standard CDS, is approximately 0.25 (quarterly payment dates). In a reduced-form setup, when the default is triggered by the first jump of a Cox process with intensity λ , this expression can be developed explicitly thanks to the Key lemma:

$$CDS_t(a, b, k) = \mathbb{1}_{\{\tau > t\}} \left(-(1 - R) \int_{T_a}^{T_b} P_t^r(u) \partial_u P_t^\lambda(u) du - k C_t(a, b) \right), \quad (22)$$

where $C_t(a, b)$ is the risky duration, i.e., the time- t value of the CDS premia paid during the life of the contract when the spread is 1:

$$C_t(a, b) := \sum_{i=a+1}^b \alpha_i P_t^r(T_i) P_t^\lambda(T_i) - \int_{T_{i-1}}^{T_i} \frac{u - T_{i-1}}{T_i - T_{i-1}} \alpha_i P_t^r(u) \partial_u P_t^\lambda(u) du.$$

The spread that, at time t , sets the forward start CDS at 0, called the par spread, is given by:

$$\mathbb{1}_{\{\tau > t\}} s_t(a, b) := \mathbb{1}_{\{\tau > t\}} \frac{-(1 - R) \int_{T_a}^{T_b} P_t^r(u) \partial_u P_t^\lambda(u) du}{C_t(a, b)}. \quad (23)$$

The no-arbitrage price of a call option on such a contract at time $t = 0$ becomes

$$\begin{aligned} PSO(a, b, k) &= \mathbb{E} \left[(CDS_{T_a}(a, b, k))^+ P^r(T_a) \right] \\ &= P^r(T_a) \mathbb{E} \left[e^{-\Lambda_{T_a}} \left((1 - R) - \sum_{i=a+1}^b \int_{T_{i-1}}^{T_i} g_i(u) P_{T_a}^{r+\lambda}(u) du \right)^+ \right], \end{aligned}$$

where $g_i(u) := (1 - R)(r(u) + \delta_{T_b}(u)) + k \frac{\alpha_i r(u)}{T_i - T_{i-1}} (1 - (u - T_{i-1}))$, with $\delta_s(u)$ being the Dirac delta function centred on s .

Considering as the intensity process the S-CIR, PS-CIR and TC-CIR leads to model prices denoted $PSO^\varphi(a, b, k)$, $PSO^{\varphi,+}(a, b, k)$ and $PSO^\theta(a, b, k)$, respectively. Interestingly, these models are equally tractable, as they feature similar expressions that can be written in terms of the shadow process y or its time integral, Y . For instance,

$$\Lambda_{T_a}^\varphi = Y_{T_a} + \int_0^{T_a} \varphi(u) du, \quad \Lambda_{T_a}^\theta = Y_{\Theta(T_a)},$$

and

$$\begin{aligned} P_{T_a}^{\lambda^\varphi}(u) &= P_{T_a}^y(u) e^{\int_{T_a}^u \varphi(s) ds} = \exp \left\{ A_{T_a}(u) - B_{T_a}(u) y_{T_a} + \int_{T_a}^u \varphi(s) ds \right\} , \\ P_{T_a}^{\lambda^\theta}(u) &= P_{\Theta(T_a)}^y(u) = \exp \left\{ A_{\Theta(T_a)}(u) - B_{\Theta(T_a)}(u) y_{\Theta(T_a)} \right\} . \end{aligned}$$

Recall that these expressions have a closed form when y is, for example, a (J)CIR process.

Such options have little liquidity. Models are then often compared in terms of their capabilities to generate large *implied volatilities*. Indeed, empirical evidence shows that this is a typical feature of CDS option quotes when disclosed. Therefore, we compare the models in terms of their *Black volatilities* (the Black model for PSO is recalled in the Appendix, Section 8.7.2). The Black volatility associated with a model price $PSO^{model}(a, b, k)$ is thus the volatility $\bar{\sigma}$ satisfying $PSO^{model}(a, b, k) = PSO^{Black}(a, b, k, \bar{\sigma})$. Recall that in all cases, the intensity process λ is calibrated to the market, i.e., $P^\lambda(t) = G(t)$. In other words, choosing, e.g., a CIR process for the base intensity process y combined with the correct shift with (φ^+) or without (φ) a positivity constraint or, finally, using the correct clock rate θ , all three models yield the same survival probability curve ($P^{\lambda^\varphi}(t) = P^{\lambda^{\varphi,+}}(t) = P^{\lambda^\theta}(t) = G(t)$). Hence, all these models agree on the par spread:

$$s_0(a, b) = \frac{(1 - R) \int_{T_a}^{T_b} P^r(u) h(u) G(u) du}{C_0(a, b)} .$$

We compare the S-CIR, the PS-CIR and the TC-CIR. The base HAJD process y in TC-CIR is taken to be the same as that of the S-CIR. One can see from Table 4 that the S-CIR features large implied volatilities. Recall, however, that it allows for negative intensities and hence is not appropriate. The PS-CIR model is not capable of generating large volatility levels, in line with the previous discussion. The TC-CIR fits in between: it rules out negative intensities while maintaining substantial volatility levels.

One might be concerned by the fact that the implied volatilities of the TC-CIR remain relatively small. This can be addressed in two ways. First, one can vary the parameter Ξ . However, the Feller constraint is often required to hold, which imposes limits on the volatility of the process. Another approach consists of considering a JCIR model as the HAJD. Indeed, JCIR is often considered when large volatilities are required. However, as explained in Section 3.2, increasing the volatility by boosting the jump activity while maintaining the calibration to a given market curve G reinforces the positivity issue. Fortunately, TC-JCIR does not suffer from this limitation; one can drastically increase the jump activity without impacting the positivity of TC-JCIR. As a consequence, the TC-JCIR seems very appropriate when one needs a positive but yet high-volatility process. This is illustrated in the last two columns of Table 4 using the same jump parameters as those given in (Brigo and Mercurio, 2006). We maintain the same parameter Ξ as before for the diffusion part and vary the jump rate (ω) and jump size (α) in the compound Poisson process J . In every case, the clock Θ is chosen such that the model perfectly fits Ford's survival probability curve. Interestingly, the (positive) TC-JCIR model can feature much larger implied volatility levels than the PS-CIR. The results for the S-JCIR are also larger than those for the PS-CIR, but they are not shown because the negative intensity problem is magnified.

T_a	T_b	S-CIR	PS-CIR	TC-CIR	TC-JCIR(0.1,0.1)	TC-JCIR(0.15,0.15)
1	3	67.12%	1.03%	43.68%	79.04%	100.17%
1	5	45.10%	1.25%	26.92%	48.07%	69.69%
1	7	27.72%	1.64%	16.86%	30.43%	44.00%
1	10	21.52%	1.65%	12.86%	23.33%	33.75%
3	5	61.30%	0.85%	57.16%	65.50%	82.60%
3	7	34.88%	1.15%	36.33%	42.85%	53.17%
3	10	27.65%	1.08%	27.17%	33.17%	41.36%
5	7	34.81%	1.16%	42.60%	49.13%	60.11%
5	10	30.96%	0.93%	31.19%	37.34%	45.78%
7	10	45.37%	0.63%	38.93%	40.59%	46.02%

Table 4: Black volatilities for at-the-money ($k = s_0(a, b)$) CDS options with start and end dates (T_a, T_b) (first two columns) implied by S-CIR, PS-CIR and TC-CIR with $y_0 = h_0$ (next three columns) using Monte Carlo simulation (500K paths with time step 0.01). The case in which y_0 enters the optimization procedure is given in Table 6. The S-CIR is flawed in all cases since the minimum of the shift function is negative in all cases. Of the two valid intensity models, PS-CIR and TC-CIR, the latter exhibits a much higher implied volatility. The last two columns refer to TC-JCIR. The parameters correspond to (ω, α) , the jump arrival rate ω and the jump size, respectively. For the models featuring jumps, we use 10^6 paths.

6.3 Wrong-way risk impact in credit valuation adjustments

A major concern of the post-crisis regulation is the modelling of the capital requirement for firms while accounting for some credit adjustment to the valuation under credit risk. Counterparty credit risk is defined as the risk that the counterparty of an over-the-counter (OTC) deal will default before the maturity of the contract. The latter can be seen as an option given to the counterparty and can be priced in a risk-neutral setup by adjusting the OTC derivative, leading to CVA. The latter is nothing but the expected losses due to the missed payments associated with the OTC portfolio. In a risk-neutral specification and assuming that $\tau > 0$, the current ($t = 0$) value of the CVA is expressed as:

$$\text{CVA} = \mathbb{E} \left[(1 - R) V_\tau^+ \mathbb{1}_{\{\tau \leq T\}} \right]$$

where V stands for the discounted exposure (i.e., the exposure process rescaled by the stochastic discount factor D). A straightforward application of the Key lemma (under some technical conditions that are valid here) yields

$$\text{CVA} = \mathbb{E} \left[(1 - R) \int_0^T V_u^+ \lambda_u e^{-\Lambda_u} du \right]. \quad (24)$$

The CVA of the shifted and the time-changed models, CVA^φ and CVA^θ , correspond to the above expression, replacing (λ, Λ) with $(\lambda^\varphi, \Lambda^\varphi)$ and $(\lambda^\theta, \Lambda^\theta)$, respectively. The purpose of this section is to illustrate the order of magnitude of CVA figures that can be obtained with either model. In particular, our aim is not to represent a specific exposure. Instead, we simplify the

analysis by considering two prototypical dynamics:

$$\begin{aligned} dV_t &= \nu dW_t^V, \\ dV_t &= \left(\gamma(T-t) - \frac{V_t}{T-t} \right) dt + \nu dW_t^V. \end{aligned}$$

where W^V is an $\mathbb{F}$ -Brownian motion. The first SDE is that of a martingale and can depict the evolution of the discounted price of a forward contract prior to its cashflow date. The second SDE corresponds to a Brownian bridge with drift, and mimics the dynamics of the discounted price of an asset paying continuous dividends. These two models were previously used in (Vrins, 2017) and (Brigo and Vrins, 2018) to describe, in a schematic way, exposures of FRA and IRS. Calibration to actual exposures yields indicative values for the parameters.

In general, there is no reason to assume that the Brownian motion driving the default intensity (W) would be independent of the Brownian motion driving the exposure (W^V): it depends on the problem at hand. Usually, we consider the general case of the *wrong-way risk* (WWR) effect, obtained by introducing a correlation between Brownian drivers. For CIR++, we assume that $dW_t dW_t^V = \rho dt$, whereas for TC-CIR, we apply the synchronization procedure suggested in (Mbaye and Vrins, 2018) to preserve the correlation after time-changing the intensity process. In the special case where the default time of the counterparty is independent of the discounted exposure (i.e., $\rho = 0$, that is, no wrong-way risk), one can easily deduce from (24) the *independent* CVA formula:

$$\text{CVA}^\perp = -(1-R) \int_0^T \mathbb{E}[V_u^+] d\mathbb{E}[e^{-\Lambda_u}] = (1-R) \int_0^T f^\lambda(u) \mathbb{E}[V_u^+] P^\lambda(u) du.$$

Recall that whatever the chosen model, it is assumed to be calibrated to the survival probability curve G , extracted from CDS prices. This leads to $P^\lambda(u) = G(u)$ and to the optimal shift and clock functions, namely, φ or φ^+ in the S-CIR and PS-CIR cases, and Θ for the TC-CIR. In this case, $\text{CVA}^\perp$ does not depend on the default model:

$$\text{CVA}^\perp = -(1-R) \int_0^T \mathbb{E}[V_u^+] dG(u) = (1-R) \int_0^T h(u) \mathbb{E}[V_u^+] G(u) du.$$

However, the independent case $\rho = 0$ is unrealistic and may lead to severe over- or underestimations of CVA (Kim and Leung, 2016; Brigo and Vrins, 2018; Breton and Marzouk, 2018). Under WWR, CVA becomes model dependent. Figure 4 shows the evolution of CVA with respect to ρ for three different models: λ^φ , $\lambda^{\varphi,+}$ and λ^θ , all calibrated to Ford's survival probability curve G as before. Under no WWR, CVA is equal to $\text{CVA}^\perp$ (dotted black): it is flat, model free and can be computed using a simple integration. Under WWR, the CVAs are computed using Monte Carlo simulations (100K paths, time step of 0.01) and adaptive control variate (see (Mbaye and Vrins, 2018) for the implementation of the adaptive control variate applied to CVA computation). The TC-CIR (solid red) and S-CIR (dashed blue) models exhibit the largest WWR effects and therefore seem appropriate to address high-WWR applications. Recall that only TC-CIR is valid here because S-CIR allows for negative intensities. The PS-CIR (dash-dotted green), however, is almost flat, equal to $\text{CVA}^\perp$. This can be understood from the fact that WWR is essentially a covariance effect between V and $e^{-\Lambda}$. Hence, the models capable of generating large variance for Λ can exhibit larger WWR effects at any (non-zero) fixed Brownian correlation level ρ , suggesting the appropriateness of TC-CIR for such applications.

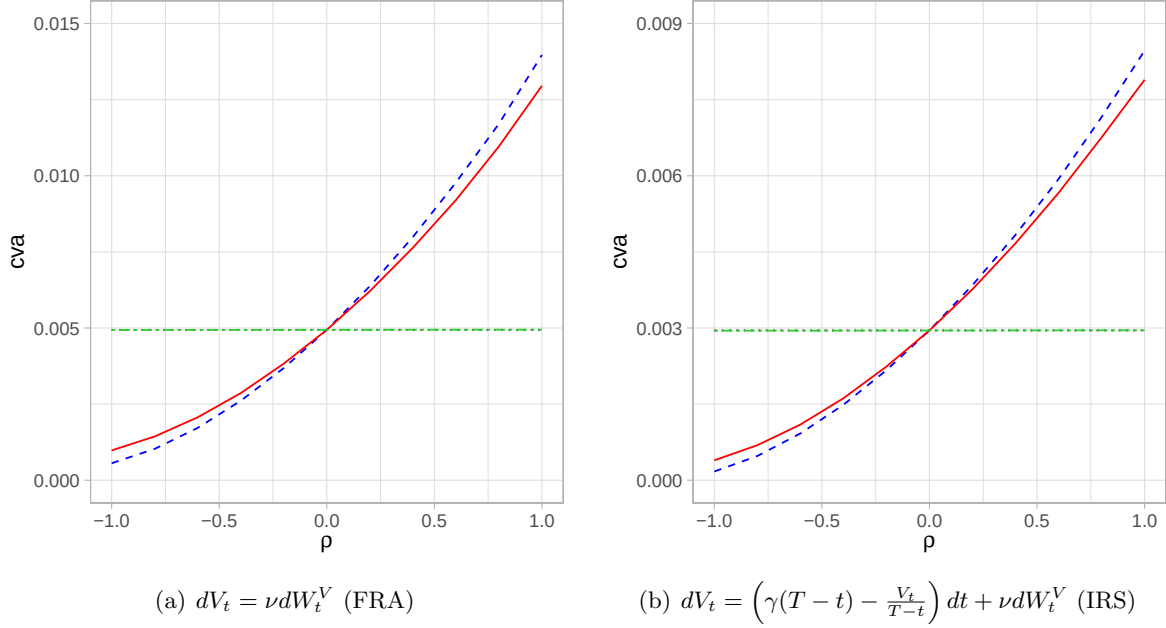


Figure 4: Impact of the exposure-credit correlation ρ on CVA levels for prototypical 5Y forward (left) and swap exposure (right) with $y_0 = h_0$ (the case in which y_0 is optimized is given in Figure 8), $\nu = 8\%$ and $\gamma = 0.1\%$. The curves correspond to different intensity models: $\lambda^{\varphi,+}$ (PS-CIR, dash-dotted green), λ^φ (S-CIR, dashed blue), and λ^θ (TC-CIR, solid red). The case without wrong-way risk corresponds to the flat (dotted black) line.

7 Conclusion

The calibration problem consists of finding the parameters of a model x to perfectly fit a given market curve. The perfect fit feature is of paramount importance in real financial applications, as it is connected to no-arbitrage opportunities and accurate valuation of trading positions. This calls for two important properties: the model must be (i) flexible enough (to be able to generate various term structures) and (ii) tractable enough (to facilitate the parameters' optimization procedure). Time-homogeneous affine models such as Vasicek, CIR or JCIR are good candidates in this respect and are widely used in interest rate and credit risk modelling. However, as such, they only feature a few constants and hence lack calibration flexibility. Moving to multi-factor models can help produce more realistic dynamics and provide additional flexibility but still does not solve the perfect fit issue. The deterministic shift extension offers an appealing solution. It consists of starting with a tractable shadow process y , which is shifted in a deterministic way with a function φ . The resulting process $x_t = y_t + \varphi(t)$ becomes fully flexible. Indeed, any discount or survival probability curve can be generated by such a model. Moreover, it has a tractability level that is very similar to that of y because φ is deterministic. Finally, for every market curve, the shift φ^* that leads to the perfect fit is known in closed form as a function of the y parameters and the market curve. However, this method is less appealing when model x needs to fulfil some range constraints. Among these, non-negativity is of primary importance when modelling interest rates (depending on the type of economy at hand), mortality rates, prepayment rates or default intensities. In the deterministic shift approach, indeed, starting

with a non-negative shadow process, y is not sufficient to guarantee that x will also be without additional constraint on φ . Furthermore, this constraint becomes increasingly severe when increasing the process volatility due to the zero lower bound.

All stochastic volatility models in force prevent negative values. However, it is currently not the case for many standard intensity models. This is quite surprising because a negative intensity is just as meaningless as a negative volatility. We believe that the reason is twofold: first, negative intensities obtained through a shift do not directly trigger numerical problems (in contrast with volatilities that often appear in square roots), so that the issue is not “blocking” in practice. Second, there is a lack of a sound alternative. The positivity constraint can be addressed by including a non-negativity constraint on φ . However, this again raises two problems. First, the parameter optimization problem becomes more difficult, but more important, the resulting process x then features a much lower variance than without the constraint, which is a strong limitation. Therefore, it seems that one prefers to disregard the “negative intensities” issue, giving priority to stochasticity and perfect fit.

In this paper, we develop a simple alternative that consists of time-changing a positive tractable shadow process. The time-change model remains analytically tractable and positive; the optimal clock is found by simple inversion and features larger implied volatility than the shift approach. Moreover, perfect fit is achievable for a broad class of discount curves, including all decreasing discount curves. The features of the model have been illustrated using topical examples taken from credit risk, but other applications could also be considered. This method thus proves to be a competitive challenger to the shift approach, at least under the (very common) positivity constraint and when large volatility levels are needed.

Finally, although specific attention has been devoted to the one-dimensional JCIR due to its popularity in credit risk, the time-change approach is not restricted to relying on homogeneous single-factor affine jump diffusions. Just as the shift approach, it can be used to calibrate any tractable term-structure framework, such as an affine transform of multi-factor affine or linear rational models.

Acknowledgment

The research of Cheikh Mbaye was funded by the *National Bank of Belgium* as well as the *Communauté Française de Belgique*. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the National Bank of Belgium. The work of F. Vrans was supported by the *Fonds de la Recherche Scientifique F.S.R.-FNRS*, grant J.0037.18, as well as by the research grant ARC 18-23/089 of the Belgian Federal Science Policy Office.

The authors are grateful to Monique Jeanblanc and Damiano Brigo for discussions and insightful comments. We also warmly thank the whole review team for relevant remarks that contributed to significantly enhancing this paper. Potential remaining errors are ours.

Data availability statement

The data used in this work is explicitly shared in the manuscript. The data source is Bloomberg.

8 Appendix

8.1 Properties of some homogeneous affine jump diffusions

Let y be an $\mathbb{F}$ -adapted jump-diffusion introduced in Definition 5 and $Y_t := \int_0^t y_u du$ its integrated version. We denote by $v^x(t) = v^x(t; \Xi) := \mathbb{V}[x_t]$ the variance of a stochastic process x parametrized by Ξ at time t . When not explicitly mentioned, all the results below that are given without proof can be found in, e.g., (Brigo and Mercurio, 2006). New results are given in lemmas for further reference.

8.1.1 Vasicek model

The dynamics of the Vasicek model read as

$$dy_t = \kappa(\beta - y_t)dt + \eta dW_t. \quad (25)$$

The A^y, B^y functions in (4) are given by:

$$\begin{aligned} A_s^{\text{VAS}}(t; \Xi) &= \left(\beta - \frac{\eta^2}{2\kappa^2} \right) (B_s^{\text{VAS}}(t; \Xi) + s - t) - \frac{\eta^2}{4\kappa} B_s^{\text{VAS}}(t; \Xi)^2, \\ B_s^{\text{VAS}}(t; \Xi) &= \frac{1}{\kappa} \left(1 - e^{-\kappa(t-s)} \right). \end{aligned}$$

If (i) $a = b = 0$ or $b \neq 0$ and (ii) $d = 0$, the dynamics of the AJD in (5) correspond to Vasicek dynamics with $(\kappa, \beta, \eta^2) = (-b, -a/b, c)$ combined with compound Poisson jumps. If $J = 0$, (i.e., $\alpha\omega = 0$), the associated functions A^y, B^y correspond to $A_s^{\text{VAS}}(t; \Xi), B_s^{\text{VAS}}(t; \Xi)$ given above with the corresponding parameters. Otherwise, if $\alpha\omega > 0$, the functions A^y, B^y coincide with $C^*(t-s, [0, 0], -1)$ and $-A^*(t-s, [0, 0], -1)$ given in (Chacko and Das, 2002, p.203) using $\theta = \beta, \sigma = \eta, \lambda_u = \omega, \eta_u = \alpha$ and $\lambda_d = \eta_d = 0$. Algebraic manipulations show that with these notations, $A_s^y(t; \Xi) = A_s^{\text{VAS}}(t; \Xi) + \omega \left(\frac{1}{\alpha + \kappa} \ln |(1 + \alpha/\kappa)e^{\kappa(t-s)} - \alpha/\kappa| + s - t \right)$ and $B_s^y(t; \Xi) = B_s^{\text{VAS}}(t; \Xi)$. The case where $a \neq 0$ and $b = d = 0$ can be treated similarly by noting that the solution y_t of (5) is then just $a \times t$ plus a Vasicek process with parameters $(\kappa, \beta, \eta^2) = (0, 0, c)$ and compound Poisson jumps, whose A^y, B^y functions can be recovered from the previous expressions, by taking the limit $\kappa \rightarrow 0$.

The forward curve associated with this model is proven to be

$$f_s^{\text{VAS}}(t) := (1 - e^{-\kappa(t-s)}) \frac{\kappa^2 \beta - \eta^2/2}{\kappa^2} + \frac{\eta^2}{2\kappa^2} e^{-\kappa(t-s)} (1 - e^{-\kappa(t-s)}) + y_t e^{-\kappa(t-s)}. \quad (26)$$

Moreover, both y and Y are normally distributed at all times, with

$$\begin{aligned} \mathbb{E}[y_t] &= y_0 e^{-\kappa t} + \beta \kappa B_0^{\text{VAS}}(t; \Xi), \\ \mathbb{E}[Y_t] &= y_0 B_0^y(t; \Xi) + \beta (t - B_0^{\text{VAS}}(t; \Xi)), \\ v^y(t) &= \frac{\eta^2}{2\kappa} (1 - e^{-2\kappa t}), \\ v^Y(t) &= \frac{\eta^2}{\kappa^2} \left[t + \frac{1 - e^{-2\kappa t}}{2\kappa} - 2B_0^{\text{VAS}}(t; \Xi) \right]. \end{aligned}$$

Lemma 5. *Let y be a Vasicek process and Y its time integral. The functions $v^y(t)$ and $v^Y(t)$ are increasing with respect to t .*

Proof. It is obvious that for $v^y(t)$, and for $v^Y(t)$, a few manipulations lead to

$$\frac{d}{dt}v^Y(t) = \left(\frac{\eta}{\kappa}(1 - e^{-\kappa t})\right)^2 \geq 0.$$

□

8.1.2 CIR model

The dynamics of the CIR model read as

$$dy_t = \kappa(\beta - y_t)dt + \delta\sqrt{y_t}dW_t \quad (27)$$

where κ, β, y_0 are positive constants. The A^y, B^y functions in eq. (4) are given by:

$$\begin{aligned} A_s^{\text{CIR}}(t; \Xi) &= \frac{2\kappa\beta}{\delta^2} \ln \frac{2\gamma \exp\{(\kappa + \gamma)(t - s)/2\}}{2\gamma + (\kappa + \gamma)(\exp\{(t - s)\gamma\} - 1)}, \\ B_s^{\text{CIR}}(t; \Xi) &= \frac{2(\exp\{(t - s)\gamma\} - 1)}{2\gamma + (\kappa + \gamma)(\exp\{(t - s)\gamma\} - 1)}. \end{aligned}$$

where $\gamma := \sqrt{\kappa^2 + 2\delta^2}$.

If (i) $a = b = 0$ or $b > 0$, (ii) $c = 0$, and (iii) $J = 0$, (i.e., $\alpha\omega = 0$), the dynamics of the AJD in (5) with $d > 0$ correspond to CIR dynamics with $(\kappa, \beta, \delta^2) = (-b, -a/b, d)$, so that the associated functions A^y, B^y correspond to $A_s^{\text{CIR}}(t; \Xi), B_s^{\text{CIR}}(t; \Xi)$ given above with the corresponding parameters. The case $c > 0$ in (5) can be treated similarly since it corresponds to a CIR shifted by $-c/d$ with parameters $(\kappa, \beta, \delta) = (-b, -a/b + c/d, \sqrt{d})$ and initial value $y_0 + c/d$. The resulting A^y, B^y functions can thus be obtained from those of the CIR, adjusting A^{CIR} along the lines given in Section 3.1 with $\varphi(t) = -c/d$.

The forward curve associated with this model is given by (Brigo and Mercurio, 2006)

$$f_s^{\text{CIR}}(t) := \frac{2\kappa\beta(e^{(t-s)\gamma} - 1)}{2\gamma + (\kappa + \gamma)(e^{(t-s)\gamma} - 1)} + y_t \frac{4\gamma^2 e^{(t-s)\gamma}}{[2\gamma + (\kappa + \gamma)(e^{(t-s)\gamma} - 1)]^2}. \quad (28)$$

Important characteristics of the CIR processes can be computed explicitly (see, e.g., (Dufresne, 2001)). For instance, y is distributed as a non-central chi-squared. The two first-order moments of y and Y are respectively given by

$$\begin{aligned} \mathbb{E}[y_t] &= y_0 e^{-\kappa t} + \beta (1 - e^{-\kappa t}), \\ \mathbb{E}[y_t^2] &= y_0^2 e^{-2\kappa t} + \left(\frac{\delta^2}{2\kappa} + \beta\right) \left[2y_0 (e^{-\kappa t} - e^{-2\kappa t}) + \beta (1 - e^{-\kappa t})^2\right], \\ \mathbb{E}[Y_t] &= t\beta + \frac{(y_0 - \beta)}{\kappa} (1 - e^{-\kappa t}), \\ \mathbb{E}[Y_t^2] &= \left(\frac{y_0 - \beta}{\kappa}\right)^2 + \delta^2 \left(\frac{y_0}{\kappa^3} - \frac{5\beta}{2\kappa^3}\right) + t\beta \left(\frac{2y_0}{\kappa} - \frac{2\beta}{\kappa} + \frac{\delta^2}{\kappa^2}\right) + t^2\beta^2 \\ &\quad + e^{-\kappa t} \left[-2\left(\frac{y_0 - \beta}{\kappa}\right)^2 + \frac{2\beta\delta^2}{\kappa^3} + t\left(-\frac{2y_0\beta}{\kappa} + \frac{2\beta^2}{\kappa} - \frac{2y_0\delta^2}{\kappa^2} + \frac{2\beta\delta^2}{\kappa^2}\right)\right] \\ &\quad + e^{-2\kappa t} \left[\left(\frac{y_0 - \beta}{\kappa}\right)^2 - \frac{y_0\delta^2}{\kappa^3} + \frac{\beta\delta^2}{2\kappa^3}\right]. \end{aligned}$$

In contrast with the Vasicek model, the variance of the CIR does not always increase monotonically with time; it depends on the parameters. However, the variance of the integrated CIR is increasing. These properties are proven in the next lemma and will be central in the proof of Theorem 2.

Lemma 6. *Let y be a CIR process and Y its time integral. Then,*

$$v^y(t) = \frac{\delta^2}{\kappa} \left[y_0(e^{-\kappa t} - e^{-2\kappa t}) + \frac{\beta}{2\kappa}(1 - e^{-\kappa t})^2 \right], \quad (29)$$

$$v^Y(t) = \frac{\delta^2}{\kappa^3} \left[(2\beta - 2\kappa t(y_0 - \beta) - (y_0 - \beta/2)e^{-\kappa t})e^{-\kappa t} + t\kappa\beta + (y_0 - 5\beta/2) \right]. \quad (30)$$

The function $v^y(t)$ is increasing if $\beta \geq y_0$. Otherwise, it is first increasing up to a time t^* and then decreasing on (t^*, ∞) . By contrast, $v^Y(t)$ is always increasing.

Proof. The computation of the variances is trivial from the first two moments recalled above. The derivative of the variance of the CIR is given by

$$\begin{aligned} \frac{d}{dt}v^y(t) &= y_0\delta^2(-e^{-\kappa t} + 2e^{-2\kappa t}) + \beta\delta^2e^{-\kappa t}(1 - e^{-\kappa t}) \\ &= \delta^2e^{-\kappa t}(\beta - y_0) - \delta^2e^{-2\kappa t}(\beta - 2y_0). \end{aligned}$$

This expression has a root on the positive half-line at

$$t^* = \frac{1}{\kappa} \ln \left(1 + \frac{y_0}{y_0 - \beta} \right)$$

only if $y_0 > \beta$. Otherwise, $v^y(t)$ is always increasing in t . The derivative of the variance of the integrated CIR with respect to time can be written, after some manipulations, as

$$\frac{d}{dt}v^Y(t) = \frac{\delta^2}{\kappa^2} \left[2y_0e^{-\kappa t}(e^{-\kappa t} - (1 - \kappa t)) + \beta(1 - (2\kappa t + e^{-\kappa t})e^{-\kappa t}) \right].$$

Because $e^{-x} \geq 1 - x$ for all $x \geq 0$ and y_0, κ are positive constants, the first term is positive for all $t \geq 0$. On the other hand, $\beta > 0$, and it is sufficient to check that $1 - g(\kappa t) \geq 0$ for all $t \geq 0$, with $g(x) := (2x + e^{-x})e^{-x}$. Clearly, $g(0) = 1$ and $g'(x) = 2e^{-x}(1 - x - e^{-x}) \leq 0$ for all $x \geq 0$. Hence, $1 - g(\kappa t) \geq 0$ for all $t \geq 0$. $\square$

8.1.3 JCIR model

The characteristics of the JCIR can be obtained by adjusting those of the corresponding CIR, i.e., with the same initial value and diffusion parameters. We denote by z the former and by y the latter and follow an analogous approach for their integrated versions (Z and Y , respectively). Hence, if the parameter set for the CIR (y, Y) is $\Xi_0 = (\kappa, \beta, \delta, 0, 0, y_0)$, the parameter set for the corresponding JCIR (z, Z) is $\Xi_0 = (\kappa, \beta, \delta, \alpha, \omega, z_0)$ with $z_0 = y_0 > 0$ and $\alpha, \omega \geq 0$. The functions associated with the discount curve are given by

$$\begin{aligned} A_s^{\text{JCIR}}(t; \Xi) &= A_s^{\text{CIR}}(t; \Xi) + \frac{\alpha\omega}{\delta^2/2 - \kappa\alpha - \alpha^2} \ln \frac{2\gamma e^{\frac{\gamma+\kappa+2\alpha}{2}(t-s)}}{2\gamma + (\kappa + \gamma + 2\alpha)(e^{(t-s)\gamma} - 1)}, \\ B_s^{\text{JCIR}}(t; \Xi) &= B_s^{\text{CIR}}(t; \Xi). \end{aligned}$$

As for the CIR case, if (i) $a = b = 0$ or $b > 0$, the dynamics of the AJD in (5) correspond to those of a JCIR shifted by $-c/d$, with parameters $(\kappa, \beta, \delta) = (-b, -a/b + c/d, \sqrt{d})$ and initial value $y_0 + c/d$. The resulting A, B functions can be obtained from those of the JCIR, as detailed for the CIR case. From the above functions, it is easy to see that the forward curve associated with the JCIR model reads as

$$f_s^{\text{JCIR}}(t) := f_s^{\text{CIR}}(t) + \frac{2\omega\alpha(e^{(t-s)\gamma} - 1)}{2\gamma + (\kappa + \gamma + 2\alpha)(e^{(t-s)\gamma} - 1)}, \quad (31)$$

where $f_s^{\text{CIR}}(t)$ is given in (28). For every valid parameter, $f_s^{\text{JCIR}}(t) \geq f_s^{\text{CIR}}(t)$ for all $t \geq s$. Regarding the moments, we have the following result.

Lemma 7. *Let y (resp. Y) be a CIR (resp. integrated CIR) and z (resp. Z) be a JCIR (resp. integrated JCIR) with the same initial value and the same diffusion parameters but with jumps governed by (ω, α) . Then,*

$$\begin{aligned} \mathbb{E}[z_t] &= \mathbb{E}[y_t] + \frac{\omega\alpha}{\kappa}(1 - e^{-\kappa t}), \\ \mathbb{E}[Z_t] &= \mathbb{E}[Y_t] + \frac{\omega\alpha}{\kappa^2}(\kappa t - (1 - e^{-\kappa t})). \\ v^z(t) &= v^y(t) + \omega\alpha \left[\frac{\delta^2}{2} \left(\frac{1 - e^{-\kappa t}}{\kappa} \right)^2 + \alpha \frac{1 - e^{-2\kappa t}}{\kappa} \right], \\ v^Z(t) &= v^Y(t) + \frac{\alpha\omega}{\kappa^3} \left[\frac{1 - e^{-\kappa t}}{\kappa} (\xi(3 - e^{-\kappa t}) - 4\delta^2) + 2\delta^2 t e^{-\kappa t} + t(2\alpha\kappa + \delta^2) \right], \end{aligned}$$

where $\xi := \delta^2/2 - \alpha\kappa$. The function $v^z(t)$ is increasing with respect to t unless $y_0 > \beta + \omega\alpha/\kappa$, in which case it is first increasing up to a time t_1 and then decreasing on (t_1, ∞) . Moreover, $v^z(t) \geq v^y(t)$, $v^Z(t) \geq v^Y(t)$ and $v^Z(t)$ are always increasing.

Proof. Applying Ito's lemma, we can solve the JCIR SDE (15) by

$$z_t = z_0 e^{-\kappa t} + \beta(1 - e^{-\kappa t}) + \delta \int_0^t e^{-\kappa(t-s)} \sqrt{z_s} dW_s + \int_0^t e^{-\kappa(t-s)} dJ_s, \quad (32)$$

and find the SDE governing the integrated JCIR process

$$Z_t = t\beta + \frac{(z_0 - \beta)}{\kappa}(1 - e^{-\kappa t}) + \delta \int_0^t \int_0^s e^{-\kappa(s-u)} \sqrt{z_u} dW_u ds + \int_0^t \int_0^s e^{-\kappa(s-u)} dJ_u ds. \quad (33)$$

From (32), we can write

$$\begin{aligned} \mathbb{E}[z_t] &= z_0 e^{-\kappa t} + \beta(1 - e^{-\kappa t}) + \mathbb{E} \left[\int_0^t e^{-\kappa(t-s)} dJ_s \right] \\ &= \mathbb{E}[y_t] + \omega\alpha \int_0^t e^{-\kappa(t-s)} ds \\ &= \mathbb{E}[y_t] + \frac{\omega\alpha}{\kappa}(1 - e^{-\kappa t}). \end{aligned}$$

Using Ito isometry,

$$\begin{aligned}
v^z(t) &= \mathbb{E}[(z_t - \mathbb{E}[z_t])^2] \\
&= \mathbb{E} \left[\left(\delta \int_0^t e^{-\kappa(t-s)} \sqrt{z_s} dW_s + \int_0^t e^{-\kappa(t-s)} dJ_s - \omega \alpha \int_0^t e^{-\kappa(t-s)} ds \right)^2 \right] \\
&= \mathbb{E} \left[\left(\delta \int_0^t e^{-\kappa(t-s)} \sqrt{z_s} dW_s \right)^2 \right] + \mathbb{E} \left[\left(\int_0^t e^{-\kappa(t-s)} (dJ_s - \omega \alpha ds) \right)^2 \right] \\
&= v^y(t) + \frac{\delta^2 \omega \alpha}{\kappa} \int_0^t e^{-2\kappa(t-s)} (1 - e^{-\kappa s}) ds + 2\omega \alpha^2 \int_0^t e^{-2\kappa(t-s)} ds \\
&= v^y(t) + \omega \alpha \left[\frac{\delta^2}{2} \left(\frac{1 - e^{-\kappa t}}{\kappa} \right)^2 + \alpha \frac{1 - e^{-2\kappa t}}{\kappa} \right].
\end{aligned}$$

Using a similar procedure applied to (33) combined with Fubini's theorem, one can derive the expectation and variance of the integrated JCIR:

$$\begin{aligned}
\mathbb{E}[Z_t] &= t\beta + \frac{(z_0 - \beta)}{\kappa} (1 - e^{-\kappa t}) + \mathbb{E} \left[\int_0^t \int_0^s e^{-\kappa(s-u)} dJ_u ds \right] \\
&= \mathbb{E}[Y_t] + \mathbb{E} \left[\int_0^t \int_s^t e^{-\kappa u} du e^{\kappa s} dJ_s \right] \\
&= \mathbb{E}[Y_t] + \frac{\omega \alpha}{\kappa} \int_0^t (1 - e^{-\kappa(t-s)}) ds \\
&= \mathbb{E}[Y_t] + \frac{\omega \alpha}{\kappa^2} (\kappa t - (1 - e^{-\kappa t}))
\end{aligned}$$

and

$$\begin{aligned}
v^Z(t) &= \mathbb{E}[(Z_t - \mathbb{E}[Z_t])^2] \\
&= \mathbb{E} \left[\left(\delta \int_0^t \int_0^s e^{-\kappa(s-u)} \sqrt{z_u} dW_u ds + \int_0^t \int_0^s e^{-\kappa(s-u)} dJ_u ds - \mathbb{E} \left[\int_0^t \int_0^s e^{-\kappa(s-u)} dJ_u ds \right] \right)^2 \right] \\
&= \mathbb{E} \left[\left(\frac{\delta}{\kappa} \int_0^t (1 - e^{-\kappa(t-s)}) \sqrt{z_s} dW_s + \frac{1}{\kappa} \int_0^t (1 - e^{-\kappa(t-s)}) dJ_s - \frac{\omega \alpha}{\kappa} \int_0^t (1 - e^{-\kappa(t-s)}) ds \right)^2 \right] \\
&= \mathbb{E} \left[\left(\frac{\delta}{\kappa} \int_0^t (1 - e^{-\kappa(t-s)}) \sqrt{z_s} dW_s \right)^2 \right] + \mathbb{E} \left[\left(\frac{1}{\kappa} \int_0^t (1 - e^{-\kappa(t-s)}) (dJ_s - \omega \alpha ds) \right)^2 \right] \\
&= v^Y(t) + \frac{\delta^2 \omega \alpha}{\kappa^2} \int_0^t (1 - e^{-\kappa(t-s)})^2 (1 - e^{-\kappa s}) ds + \frac{2\omega \alpha^2}{\kappa^2} \int_0^t (1 - e^{-\kappa(t-s)})^2 ds \\
&= v^Y(t) + \frac{\alpha \omega}{\kappa^3} \left[\frac{1 - e^{-\kappa t}}{\kappa} ((\delta^2/2 - \alpha \kappa)(3 - e^{-\kappa t}) - 4\delta^2) + 2\delta^2 t e^{-\kappa t} + t(2\alpha \kappa + \delta^2) \right].
\end{aligned}$$

Note that the above results can be obtained using another procedure, namely, by deriving once or twice the characteristic function $\Psi_t(u, v) = \mathbb{E}[e^{uz_t + vZ_t}]$ of (z_t, Z_t) , which can be recovered from eq. (A.1) in (Duffie and Gârleanu, 2001). However, this procedure is considerably more burdensome. ¹⁵

¹⁵Be aware that there are typos in this formula. The correct expression can be found in eq. (B.9) in the draft version of Duffie and Gârleanu's paper, which is available for download on the authors' webpage.

The derivative of the variance of the JCIR is given by

$$\begin{aligned}\frac{d}{dt}v^z(t) &= y_0\delta^2(-e^{-\kappa t} + 2e^{-2\kappa t}) + \beta\delta^2e^{-\kappa t}(1 - e^{-\kappa t}) + \frac{\delta^2\omega\alpha}{\kappa}e^{-\kappa t}(1 - e^{-\kappa t}) + 2\omega\alpha^2e^{-2\kappa t} \\ &= \delta^2e^{-\kappa t}(\beta - y_0 + \omega\alpha/\kappa) - \delta^2e^{-2\kappa t}(\beta - 2y_0 + \omega\alpha/\kappa - 2\omega\alpha^2/\delta^2) .\end{aligned}$$

This expression has a root on the positive half-line at

$$t_1 := \frac{1}{\kappa} \ln \left(1 + \frac{y_0 + 2\omega\alpha^2/\delta^2}{y_0 - \beta - \omega\alpha/\kappa} \right) \quad (34)$$

only if $y_0 > \beta + \omega\alpha/\kappa$. Otherwise, $v^y(t)$ is always increasing in t .

It seems intuitive that the variance of the JCIR cannot be smaller than that of the CIR, and the case for the integrated versions is analogous. Because of the mean reverting effect, however, this needs to be confirmed. It is obvious that $v^z(t) - v^y(t) \geq 0$. The term associated with $v^Z(t) - v^Y(t)$ starts at zero (since obviously $v^Z(0) = v^Y(0) = 0$). This difference is increasing:

$$\frac{d}{dt}(v^Z(t) - v^Y(t)) = \frac{\delta^2\omega\alpha}{\kappa^3} (1 - (2\kappa t + e^{-\kappa t})e^{-\kappa t}) + \frac{2\omega\alpha^2}{\kappa^2} (1 - e^{-\kappa t})^2 \geq 0 .$$

Indeed, the second term is obviously positive, and the first term takes the form $\frac{\delta^2\omega\alpha}{\kappa^3}(1 - g(\kappa t))$, where the function $g(x) = (2x + e^{-x})e^{-x}$ is shown to be bounded by 1 for $x \geq 0$ in the proof of Lemma 6. This shows that $v^Z(t) \geq v^Y(t)$. Because both $v^Y(t)$ (from Lemma 6) and $v^Z(t) - v^Y(t)$ are increasing, $v^Z(t)$ is itself increasing. $\square$

8.2 Some special cases where P^{x+y} is a discount curve

Observe first that P^{x+y} is a discount curve whenever P^x, P^y are in the case in which x, y are independent since then $P_s^{x+y} = P_s^x P_s^y$, and the product of two time- s discount curves is itself a time- s discount curve. The next lemma provides sufficient conditions on y for P^y to be a discount curve in the general case.

Lemma 8. *Let T be a fixed time horizon. Then, P^y is a discount curve whenever y is positive and $\sup_{t \in [0, T]} y_t$ is integrable.*

Proof. We start with the lemma providing sufficient conditions to swap the expectation and derivative operators, which can be found in, e.g., (Pagès, 2018).

Lemma 9. *Let I be a nontrivial interval of $\mathbb{R}$, $\mathcal{B}(I)$ be the Borel set of I and $\Psi : I \times \Omega \rightarrow \mathbb{R}$, $(x, \omega) \mapsto \Psi(x, \omega)$ be a $\mathcal{B}(I) \otimes \mathcal{G}$ -measurable function. If the function Ψ satisfies:*

- (i) *For every $x \in I$, the random variable $\Psi(x, \omega) \in L^1$,*
- (ii) *$\Psi_x(x, \omega) := \frac{\partial \Psi(x, \omega)}{\partial x}$ exists for all $x \in I$ a.s.,*
- (iii) *There exists $Z \in L^1$ such that for every $x \in I$,*

$$|\Psi_x(x, \omega)| \leq Z(\omega) \text{ a.s. .}$$

Then, the function $\psi(x) := \mathbb{E}[\Psi(x, \omega)]$ is defined and differentiable at every $x \in I$ with derivative

$$\frac{d\psi(x)}{dx} = \mathbb{E}[\Psi_x(x, \omega)] .$$

We now proceed with the proof of Lemma 8.

Let us fix $t \leq T$. Hence,

$$\left| \frac{d}{dt} e^{-\int_0^t y_u du} \right| = |y_t e^{-\int_0^t y_u du}| \leq y_t \leq \sup_{t \in [0, T]} y_t$$

for all $t \in [0, T]$. Noting that $\sup_{t \in [0, T]} y_t$ is integrable, one can use Lemma 9 with $\Psi(t, w) \leftarrow e^{-\int_0^t y_u(w) du}$ and $Z(\omega) \leftarrow S_T^y := \sup_{t \in [0, T]} y_t$, justifying the swap between the derivative and expectation operators:

$$\frac{d}{dt} P^y(t) = \frac{d}{dt} \mathbb{E} \left[e^{-\int_0^t y_u du} \right] = \mathbb{E} \left[\frac{d}{dt} e^{-\int_0^t y_u du} \right] = -\mathbb{E} \left[y_t e^{-\int_0^t y_u du} \right],$$

where the right-hand side is bounded by the expectation of Z , which is integrable. This concludes the proof. $\square$

For the assumption regarding the integrability of the running supremum of y to be useful in practice, it needs to be “checkable”. Hence, we need to provide simpler sufficient conditions (e.g., based on the coefficients of the SDE of y) that would guarantee that $S_T^y := \sup_{t \in [0, T]} |y_t|$ satisfies $\mathbb{E}[S_T^y] < \infty$.

Lemma 10. *Let W be a Brownian motion and J be a compound Poisson process with constant jump intensity ω and the jump sizes are exponentially distributed with mean α and y solving*

$$dy_t = \mu(t, y_t)dt + \sigma(t, y_t)dW_t + dJ_t$$

where $\mathbb{E}[\int_0^T |\mu(t, y_t)|dt] < \infty$ and $\mathbb{E}[\int_0^T \sigma^2(t, y_t)dt] < \infty$. Then,

$$\mathbb{E}[|S_T^y|] = \mathbb{E}[S_T^y] < \infty \quad \text{where } S_T^y := \sup_{t \in [0, T]} |y_t|.$$

Proof. The solution of the SDE is

$$y_t = y_0 + \underbrace{\int_0^t \mu(s, y_s)ds}_{A_t} + \underbrace{\int_0^t \sigma(s, y_s)dW_s}_{M_t} + J_t \Rightarrow |y_t| \leq |A_t| + |M_t| + J_t,$$

showing that

$$S_T^y \leq \underbrace{\sup_{t \in [0, T]} |A_t|}_{S_T^A} + \underbrace{\sup_{t \in [0, T]} |M_t|}_{S_T^M} + \underbrace{\sup_{t \in [0, T]} J_t}_{S_T^J}.$$

We show in the sequel that S_T^A , S_T^M and S_T^J are integrable. This would conclude the proof since it would lead to

$$\mathbb{E}[|S_T^y|] = \mathbb{E}[S_T^y] \leq \mathbb{E}[S_T^A] + \mathbb{E}[S_T^M] + \mathbb{E}[S_T^J] < \infty.$$

Suppose that $\mathbb{E}[\int_0^T |\mu(s, y_s)|ds] < \infty$. Then,

$$S_T^A = \sup_{t \in [0, T]} |y_0 + \int_0^t \mu(s, y_s)ds| \leq |y_0| + \sup_{t \in [0, T]} \int_0^t |\mu(s, y_s)|ds \leq |y_0| + \int_0^T |\mu(s, y_s)|ds.$$

showing that $\mathbb{E}[|S_T^A|] = \mathbb{E}[S_T^A] \leq |y_0| + \mathbb{E}[\int_0^T |\mu(s, y_s)|ds] < \infty$.

On the other hand, M is a martingale, so that $|M|$ is a submartingale:

$$\mathbb{E}[|M_t| | \mathcal{F}_s] \geq |\mathbb{E}[M_t | \mathcal{F}_s]| = |M_s| .$$

We can then apply Doob's inequality,

$$\mathbb{E}[S_T^M] = \mathbb{E}[\sup_{t \in [0, T]} |M_t|] \leq \frac{e}{e-1} (1 + \mathbb{E}[|M_T| \log^+ |M_T|]) .$$

Using $-\frac{1}{e} \leq x \log x \leq x^2$ for $x \geq 0$, $|x \log^+ x| = x \log^+ x \leq |x \log x| \leq \max(e^{-1}, x^2)$:

$$\begin{aligned} \mathbb{E}[|M_T| \log^+ |M_T|] &\leq \mathbb{E}[\max(e^{-1}, M_T^2)] \\ &= \mathbb{E}[e^{-1} 1_{\{M_T^2 \leq e^{-1}\}} + M_T^2 1_{\{M_T^2 > e^{-1}\}}] \\ &\leq e^{-1} + \mathbb{E}[M_T^2] . \end{aligned}$$

Hence,

$$\mathbb{E}[S_T^M] \leq \frac{e}{e-1} (1 + e^{-1} + \mathbb{E}[M_T^2]) .$$

Using Ito isometry, $\mathbb{E}[M_T^2] = \mathbb{E} \left[\int_0^T \sigma^2(t, y_t) dt \right]$ which is bounded, by assumption. Similarly, one can prove that $\mathbb{E}[S_T^J]$ is finite by applying Doob's inequality to the martingale $(J_t - \omega \alpha t), t \leq T$. Indeed,

$$J_t = |J_t - \omega \alpha t + \omega \alpha t| \leq |J_t - \omega \alpha t| + \omega \alpha t, \quad \forall t \leq T ,$$

which implies that

$$\begin{aligned} \mathbb{E}[S_T^J] &\leq \mathbb{E}[\sup_{t \in [0, T]} |J_t - \omega \alpha t|] + \omega \alpha T \\ &\leq \frac{e}{e-1} (1 + e^{-1} + \mathbb{E}[(J_T - \omega \alpha T)^2]) + \omega \alpha T \\ &= \frac{e}{e-1} (1 + e^{-1} + 2\omega \alpha^2 T) + \omega \alpha T < \infty . \end{aligned}$$

□

8.3 Proof of Theorem 1

Observe first that for every θ and every t , one obtains

$$\int_0^t x_u^\theta du = \int_0^t \theta(u) y_{\theta(u)} du = \int_0^{\Theta(t)} y_u du .$$

Hence, the expectations of their negative exponentials also agree:

$$P^{x^\theta}(t) = P^y(\Theta(t)) .$$

The specific clock rate θ^* given by the calibration equation thus satisfies, for all t ,

$$P(t) = P^{x^{\theta^*}}(t) = P^y(\Theta^*(t)) . \tag{35}$$

Formulating this equality in terms of instantaneous forward rates yields

$$\int_0^t f(u) du = \int_0^{\Theta^*(t)} f^y(u) du .$$

Eq. (13) is simply the differential form of the latter.

It is not clear, in general, when this ODE admits a solution. However, a simple case is when P^y is a strictly decreasing discount curve. In this case, indeed, P^y admits an inverse on the positive half-line, denoted Q^y . Applying Q^y to (35) yields $\Theta^* = Q^y(P(t))$. Furthermore, the inverse of a decreasing function is decreasing, and the combination of two decreasing functions is itself increasing. Hence, if P is decreasing, $\Theta^*(t)$ is continuous and strictly increasing. Moreover, $\Theta^*(0) = Q^y(P(0)) = Q^y(1) = 0$. Hence, Θ^* exists and is a clock.

8.4 Proof of Theorem 2

It is known from Corollary 1 that for any (non-trivial) (J)CIR process y with parameter Ξ , there exists a clock $\Theta^*(t) = \Theta^*(t; \Xi)$ that yields a perfect fit between the curves P^x generated by the TC-JCIR $x_t^{\theta^*} := \theta^*(t)y_{\Theta^*(t)}$. The same holds true for the S-JCIR, $x_t^{\varphi^*}$. This means:

$$P^{x^{\varphi^*}(\Xi)}(t) = P(t) = P^{x^{\theta^*}(\Xi)}(t) ,$$

or equivalently,

$$e^{-\int_0^t \varphi^*(u) du} P^{y(\Xi)}(t) = P(t) = P^{y(\Xi)}(\Theta^*(t)) .$$

Because y is a JCIR, it can be arbitrarily close to 0 at any time t ; hence, the calibration constraint amounts to forcing $\varphi^*(t) \geq 0$ (or equivalently, $f^{\text{JCIR}}(t) \leq f(t)$) $\forall t \geq 0$. This implies that $P^{y(\Xi)}(\Theta^*(t)) \leq P^{y(\Xi)}(t)$. Because $P^{y(\Xi)}$ is a decreasing function, the last inequality is equivalent to $\Theta^*(t) \geq t$.

To prove 1), we simply note that $v^Y(t)$ is increasing (Lemma 7). Combined with the condition $\Theta^*(t) \geq t$ resulting from the positivity constraint, one gets $\mathbb{V}[X_t^{\theta^*}] = \mathbb{V}[Y_{\Theta^*(t)}] = v^Y(\Theta^*(t)) \geq v^Y(t) = \mathbb{V}[X_t^{\varphi^*}]$.

Dealing with v^y is more difficult, because v^y may not be increasing and because the expression of the variance is different, $\mathbb{V}[x_t^{\theta^*}] := \theta^*(t)^2 v^y(\Theta^*(t))$. Because we know that under positivity constraint $\Theta^*(t) \geq t$, a sufficient condition for the inequality $\mathbb{V}[x_t^{\theta^*}] \geq \mathbb{V}[x_t^{\varphi^*}] = v^y(t)$ to hold is thus to have both $\theta^*(\cdot) \geq 1$ and v^y non-decreasing. From the proof of Lemma 7, we know that v^y (i) is increasing if $y_0 \leq \beta + (\omega\alpha)/\kappa$ and (ii) is increasing up to t_1 otherwise. On the other hand, the positivity constraint implies $f(t) \geq f^{\text{JCIR}}(t)$, hence $\theta^*(t) \geq \frac{f^{\text{JCIR}}(t)}{f^{\text{JCIR}}(\Theta^*(t))}$. Therefore, because $\Theta^*(t) \geq t$, a sufficient condition for $\theta^*(\cdot) \geq 1$ is to have f^{JCIR} decreasing. This is the case when (i) $y_0 \geq \beta + (\omega\alpha)/\kappa$ or (ii) after t_2 if $\frac{\beta + \omega\alpha}{\gamma} < y_0 < \beta + \frac{\omega\alpha}{\kappa}$. To see this, we use (31) and compute the derivative of the instantaneous forward curve in the JCIR model,

$$\begin{aligned} \frac{d}{dt} f^{\text{JCIR}}(t) &= \frac{4\gamma^2 e^{t\gamma} [\kappa\beta(\gamma - \kappa + (\kappa + \gamma)e^{t\gamma}) + y_0\gamma(\gamma - \kappa - (\kappa + \gamma)e^{t\gamma}) + \omega\alpha(2\gamma + (\kappa + \gamma)(e^{t\gamma} - 1))]}{[2\gamma + (\kappa + \gamma)(e^{t\gamma} - 1)]^3} \\ &= \frac{4\gamma^2 e^{t\gamma} [(\gamma - \kappa)(\kappa\beta + y_0\gamma + \omega\alpha) - 2\omega\alpha^2] + 4\gamma^2 e^{2t\gamma} [(\gamma + \kappa)(\kappa\beta - y_0\gamma + \omega\alpha) + 2\omega\alpha^2]}{[2\gamma + (\kappa + \gamma)(e^{t\gamma} - 1)]^3} . \end{aligned}$$

From this expression, one can check that f^{JCIR} is strictly increasing if $y_0 < \beta + \omega\alpha/\kappa$ and $y_0\gamma \leq \kappa\beta + \omega\alpha$. It is strictly decreasing if $y_0 \geq \beta + \omega\alpha/\kappa$. Otherwise, i.e., if $y_0 < \beta + \omega\alpha/\kappa$ and $y_0\gamma > \kappa\beta + \omega\alpha$, the derivative has a root at

$$t_2 := \frac{1}{\gamma} \ln \frac{(\gamma - \kappa)(\kappa\beta + y_0\gamma + \omega\alpha) - 2\omega\alpha^2}{(\kappa + \gamma)(y_0\gamma - \kappa\beta - \omega\alpha) - 2\omega\alpha^2} .$$

i.e., f^{JCIR} is first increasing, then decreasing.

Observe that the condition $y_0 = \beta + \omega\alpha/\kappa$ corresponds to the case in which $v^y(t)$ is increasing and $f^{\text{JCIR}}(t)$ is decreasing; hence, (i) holds.

If f is constant, we have $f(t) \geq f^{\text{JCIR}}(t)$, which implies that $f(t) \geq f^{\text{JCIR}}(\Theta^*(t))$. Clearly, if $f(t)$ is constant or f^{JCIR} is decreasing, then $\theta^*(t) \geq 1$. If $v^y(t)$ is increasing, $v^y(\Theta^*(t)) \geq v^y(t)$ since $\Theta^*(t) \geq t$. From the fact that $\mathbb{V}[x_t^{\theta^*}] := \theta^*(t)^2 v^y(\Theta^*(t))$ and the variation of $v^y(t)$ (Lemma 7), (ii), (iii) and (iv) follow.

In particular, taking $\omega\alpha = 0$, we recover the CIR case that corresponds to the TC-CIR model.

8.5 Variance analysis of the TC-CIR

In this section, we illustrate the variance relationships given at the end of Section 5.2 from the calibrated S-CIR, PS-CIR and TC-CIR intensity processes involved in Section 6.

First, inequality (20) is illustrated in Fig. 5. Figure 5(a) is essentially the same as Fig. 1(c) but using $\Xi^{*,+}$ instead of Ξ^* (Fig. 5(b) is the same but where y_0 is a calibrated parameter instead of taken equal to h_0).

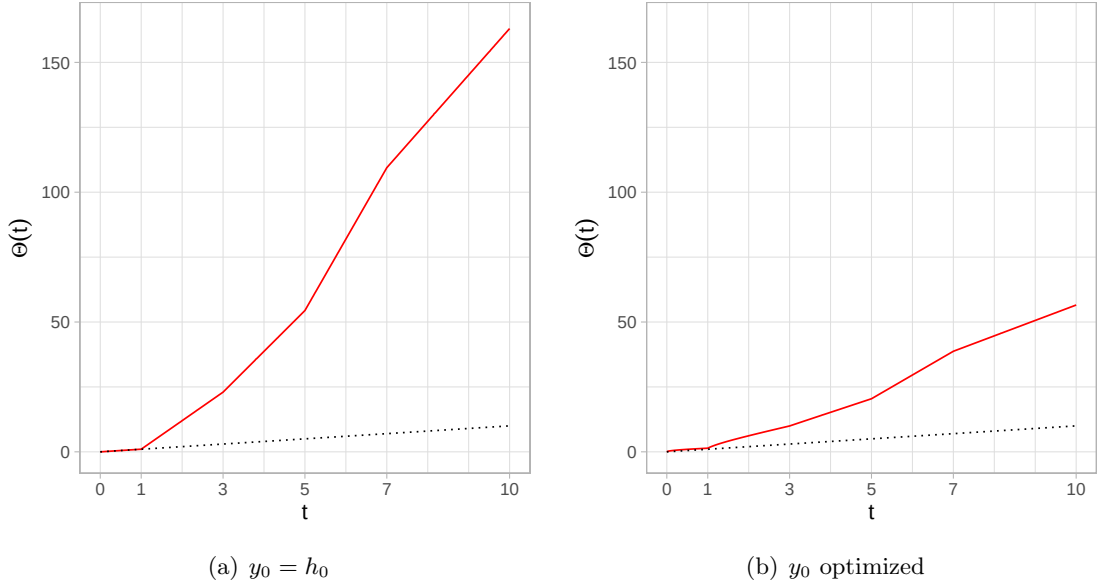


Figure 5: Illustration of ineq. (20) applied to our Ford example. The clock $\Theta^{*,+}(t) = \Theta(t; \Xi^{*,+})$ (solid red) is always larger than t (dotted black).

It is clear from Fig. 3(a) and Fig. 3(b) that imposing the non-negative shift constraint yields processes that feature little randomness: $y(\Xi^{*,+})$ and its shifted version $\lambda^{\varphi,+}$ are essentially deterministic, in contrast to $y(\Xi^*)$ and λ^{φ} , which exhibit more randomness (Fig. 3(c) and Fig. 3(d)). Inequality (19) can be checked in Figure 6: the variance of the integrated S-CIR (dashed blue) dominates that of the PS-CIR (dash-dotted green), which remains almost flat at zero. In our context, this reads as

$$\mathbb{V}[\Lambda_t^{\varphi}(\Xi)] \geq \mathbb{V}[\Lambda_t^{\varphi,+}(\Xi^{*,+})] .$$

Approximation (21) can be observed from the same figure by comparing the variance of the integrated intensity modelled as a TC-CIR with Ξ^* (solid red) to the variance of the integrated

intensity modelled as S-CIR (dashed blue):

$$\mathbb{V}[\Lambda_t^\theta(\Xi^*)] \approx \mathbb{V}[\Lambda_t^\varphi(\Xi^*)] .$$

In summary, we observe that when dealing with CIR++ under a positivity constraint, one has to choose between a valid (but low-volatility) PS-CIR process $\lambda^{\varphi,+}$ or a flawed (but high-volatility) S-CIR process λ^φ . By contrast, the TC-CIR model λ^θ is always valid (Corollary 1) and always features a variance that is larger than the PS-CIR counterpart (Theorem 2), and its variance is actually comparable to the large levels generated by the S-CIR. The TC-CIR thus proves to be a solid challenger to CIR++ models.

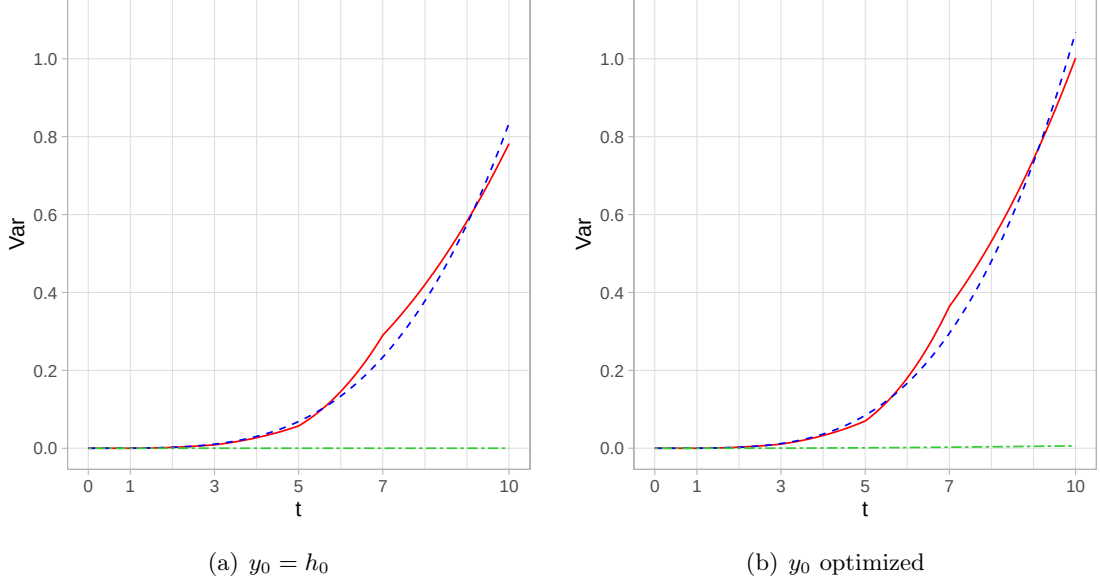


Figure 6: Illustration of inequality (19) and approximation (21) applied to our Ford example. Variances of the integrated versions of $\lambda^{\varphi,+}$ (PS-CIR $\Xi = \Xi^{*,+}$ (left) and $\Xi = \Xi_0^{*,+}$ (right), dash-dotted green), λ^φ (S-CIR $\Xi = \Xi^*$ (left) and $\Xi = \Xi_0^*$ (right), dashed blue), and λ^θ (TC-CIR with $\Xi = \Xi^*$ (left) and $\Xi = \Xi_0^*$ (right), solid red).

8.6 Time-changed Vasicek

Our time-change approach can be easily used to solve Problem 1 in the most common case where the forward curve f is arbitrary (monotonic, humped, etc.), provided that it is positive. As there is no constraint on the range of the process x^θ , let us postulate Vasicek dynamics for the shadow process: $(a, b, c, d) = (\kappa\beta, -\beta, \eta, 0)$ and no jump:

$$dy_t = \kappa(\beta - y_t)dt + \eta dW_t, \quad y_0 \in \mathbb{R} .$$

The forward curve associated with the Vasicek model is

$$f^y(t) = (1 - e^{-\kappa t}) \frac{\kappa^2 \beta - \eta^2/2}{\kappa^2} + \frac{\eta^2}{2\kappa^2} e^{-\kappa t} (1 - e^{-\kappa t}) + y_0 e^{-\kappa t} . \quad (36)$$

It can thus be used to select an appropriate Vasicek model. The next corollary provides guidelines to generate a decreasing discount curve associated with the most common case of positive instantaneous forwards.

Corollary 2. *Let P be a strictly decreasing market curve. Then, for every Vasicek model with parameters satisfying $y_0 \geq 0$ and $2\kappa^2\beta > \eta^2$, there exists a clock Θ^* such that $P^{x^\theta} \equiv P$.*

Proof. Because y is a Vasicek process, P^y is a discount curve. Moreover, the conditions $y_0 \geq 0$ and $2\kappa^2\beta > \eta^2$ guarantee that the forward curve (36) is strictly positive; hence, P^y is strictly decreasing. From Theorem 1, the clock Θ^* exists and is given by (13) with f^y given in (36). $\square$

The dynamics of our process are as follows:

$$dx_t^\theta = x_t^\theta d \log \theta(t) + \kappa(t) \left(\beta(t) - x_t^\theta \right) dt + \eta(t) dB_t ,$$

where $\beta(t) := \beta\theta(t)$, $\kappa(t) := \kappa\theta(t)$ and $\eta(t) := \eta\sqrt{\theta^3(t)}$. The dynamics of x^θ thus contain two parts. The second part corresponds to inhomogeneous Vasicek dynamics. The first part is an extra proportional correction term impacting the drift. Observe that because the calibration process is such that $\Theta(t) \approx t$, one expects to have $\theta \approx 1$. Hence, the resulting dynamics are expected to be close to those of a plain Vasicek where the coefficients evolve slightly over time. Fitting perfectly a strictly decreasing discount curve (without further constraints on the process) is a special case of Problem 1, which can also be solved using the shift approach (6) by taking $x \leftarrow x^\varphi$, where y is a Vasicek with arbitrary parameters Ξ and $\varphi(t) \leftarrow \varphi^*(t; \Xi) = f(t) - f^y(t)$. The main interest of the time-changed approach is actually when considering Problem 2 using a non-negative shadow model.

8.7 Appendix to Section 6

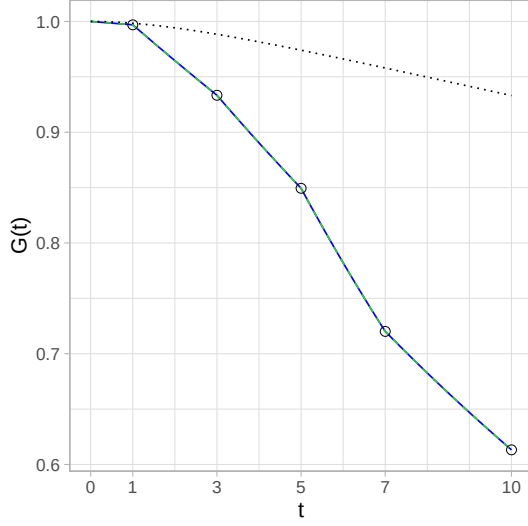
8.7.1 Extension of the numerical results when y_0 is optimized

Ξ	κ	β	δ	y_0
Ξ_0^*	0.0624	0.2975	0.3343	0.0000
$\Xi_0^{*,+}$	$3.8252 \cdot 10^{-01}$	$9.6881 \cdot 10^{-03}$	$1.5195 \cdot 10^{-01}$	$3.2093 \cdot 10^{-10}$

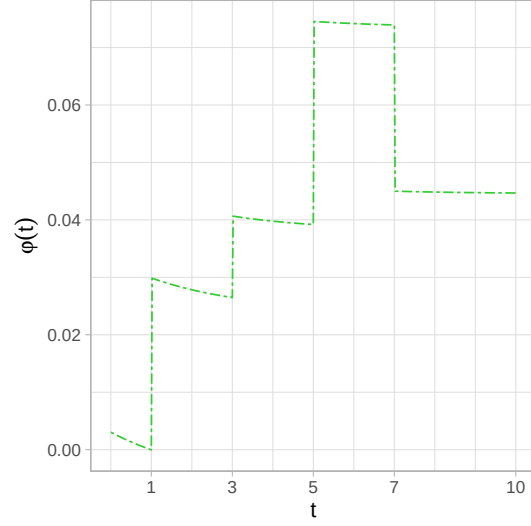
Table 5: This is the same as Table 3 except that y_0 is a parameter that enters the optimization procedure.

T_a	T_b	S-CIR	PS-CIR	TC-CIR
1	3	65.21%	9.48%	44.00%
1	5	42.35%	5.88%	26.56%
1	7	25.30%	4.87%	16.30%
1	10	19.37%	2.94%	12.27%
3	5	63.87%	8.02%	58.67%
3	7	35.20%	4.50%	36.10%
3	10	27.02%	3.44%	26.62%
5	7	35.77%	4.59%	43.32%
5	10	30.35%	3.70%	30.61%
7	10	45.05%	5.16%	39.40%

Table 6: This is the same as Table 4 except that y_0 is a parameter that enters the optimization procedure.

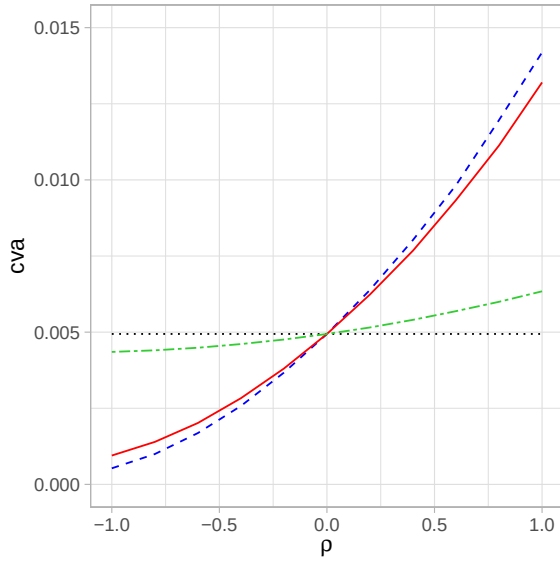


(a) Survival probability ($P^{y(\Xi_0^{*,+})}, G = P^{\lambda^{\varphi,+}}$)

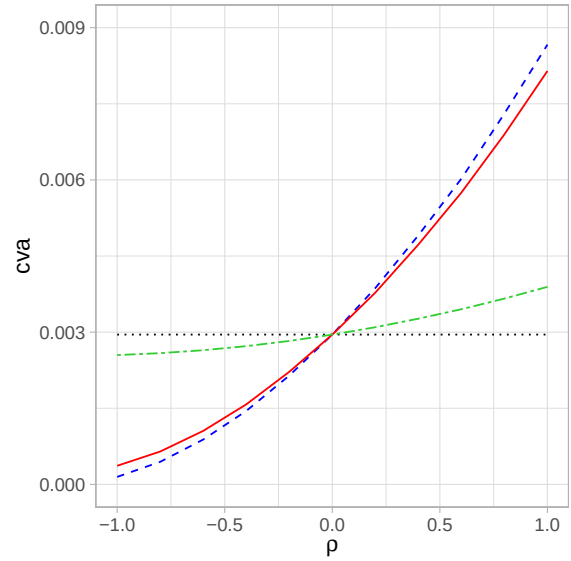


(b) Shift function ($\varphi^{*,+} = \varphi(\Xi_0^{*,+})$)

Figure 7: Same as Figure 2 except that y_0 is a parameter that enters the optimization procedure ($\Xi_0^{*,+}$ instead of $\Xi^{*,+}$).



(a) $dV_t = \nu dW_t^V$ (FRA)



(b) $dV_t = \left(\gamma(T-t) - \frac{V_t}{T-t} \right) dt + \nu dW_t^V$ (IRS)

Figure 8: Same as Figure 4 except that y_0 is a parameter that enters the optimization procedure. This provides slightly more room for randomness (and hence covariance impact) in PS-CIR (dash-dotted green).

8.7.2 Black model for PSO

In this context, the Black-Scholes model works as follows. We start by noting that the forward start CDS can be written in terms of the difference between the fair and the agreed premium

cashflows. Indeed, the former corresponds to the protection leg. Inserting (23) into (22) yields

$$CDS_t(a, b, k) = \mathbb{1}_{\{\tau > t\}}(s_t(a, b) - k) C_t(a, b) .$$

Using the results based on Brigo and Morini (2005), the payer default swaption becomes

$$PSO(a, b, k) = \mathbb{E} [(s_{T_a}(a, b) - k)^+ C_{T_a}(a, b) D(T_a)] = C_0(a, b) \mathbb{E}^{(a, b)} [(s_{T_a}(a, b) - k)^+] \quad (37)$$

where $\mathbb{E}^{(a, b)}$ stands for the expectation under the equivalent measure $\mathbb{Q}^{(a, b)}$, associated with the numéraire $C(a, b)$. Interestingly, it is clear from (23) that the par spread $s(a, b)$ is a $\mathbb{Q}^{(a, b)}$ -martingale on $[0, T_a]$. Hence, the Black-Scholes model for CDSO naturally postulates $\mathbb{Q}^{(a, b)}$ -martingale dynamics for the par spread

$$ds_t(a, b) = \bar{\sigma} s_t(a, b) dW_t^s, \quad t \leq T_a$$

where W^s is a $\mathbb{Q}^{(a, b)}$ -Brownian motion. Finally, the expectation in (37) is given by the standard Black-Scholes formula by setting $r \leftarrow 0$. Hence, the Black-Scholes price of the PSO is given by

$$PSO^{Black}(a, b, k, \bar{\sigma}) = C_0(a, b) [s_0(a, b) \Phi(d_1) - k \Phi(d_2)]$$

where

$$d_1 = \frac{\ln \frac{s_0(a, b)}{k} + \frac{1}{2} \bar{\sigma}^2 T_a}{\bar{\sigma} \sqrt{T_a}}, \quad d_2 = d_1 - \bar{\sigma} \sqrt{T_a}$$

and Φ is the distribution function of a standard normal random variable.

References

- T. Bielecki and M. Rutkowski. *Credit risk : modeling, valuation and hedging*. Springer finance. Springer, 2002.
- T. Bielecki, M. Jeanblanc, and M. Rutkowski. Credit risk modeling. Technical report, Center for the Study of Finance and Insurance, Osaka University, Osaka (Japan), 2011.
- F. Black. Interest rates as options. *The Journal of Finance*, 50(5):1371–1376, 1995.
- M. Breton and O. Marzouk. Evaluation of counterparty risk for derivatives with early-exercise features. *Journal of Economic Dynamics and Control*, 88:1–20, 2018.
- D. Brigo and A. Alfonsi. Credit default swaps calibration and option pricing with the SSRD stochastic intensity and interest rate model. *Finance and Stochastics*, 9:29–42, 2005.
- D. Brigo and L. Cousot. A comparison between the SSRD model and the market model for cds options pricing. *International Journal of Theoretical and Applied Finance*, 9(3), 2006.
- D. Brigo and N. El-Bachir. An exact formula for default swaptions pricing in the SSRJD stochastic intensity model. *Mathematical Finance*, 20(3):365–382, 2010.
- D. Brigo and F. Mercurio. A deterministic-shift extension to analytically-tractable and time-homogeneous short-rate models. *Finance and Stochastics*, 5:369–388, 2001.
- D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice*. Springer, 2006.

- D. Brigo and M. Morini. CDS Market Formulas and Models. *Working paper, Banca IMI*, 2005.
- D. Brigo and F. Vrans. Disentangling Wrong-Way Risk: Pricing CVA via change of measures and drift adjustment. *European Journal of Operational Research*, 269:254–264, 2018.
- D. Brigo, A. Capponi, and A. Pallavicini. Arbitrage-free bilateral counterparty risk valuation under collateralization and application to credit default swaps. *Mathematical Finance*, 24(1):125–146, 2014.
- P. Carr and D. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2(4):61–73, 1999.
- P. Carr, H. Geman, D. Madan, and M. Yor. Stochastic volatility for Lévy processes. *Mathematical Finance*, 13:345–382, 2003.
- G. Chacko and S. Das. Pricing interest rate derivatives: A general approach. *The Review of Financial Studies*, 15(1):195–241, 2002.
- P. Collin-Dufresne, R. Goldstein, and J. Hugonnier. A general formula for valuing defaultable securities. *Econometrica*, 72(5):1377–1407, 2004.
- J.C. Cox, J.E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.
- S. Crépey, M. Jeanblanc, and D. Wu. Informationally dynamized gaussian copula. Technical report, 2012. URL <https://www.maths.univ-evry.fr/prepubli/363.pdf>.
- K.E. Dambis. On the decomposition of continuous submartingales. *Theory of Probability and its Applications*, 10:40914100, 1965.
- B. Dellacherie and P.-A. Meyer. *Probabilités et Potentiel - Théorie des martingales*. Hermann, 1980.
- L. Dubins and G. Schwartz. On continuous martingales. *Proceedings of the National Academy of Sciences*, 53:913916, 1965.
- D. Duffie and N. Gârleanu. Risk and valuation of collateralized debt obligations. *Financial Analyst Journal*, 57:41–59, 2001.
- D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379406, 1996.
- D. Duffie and K. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12:687–720, 1999.
- D. Duffie and K. Singleton. Credit risk: pricing, measurement, and management. *Princeton University Press*, 2003.
- D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions, 1999.

- D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1056, 2003.
- D. Dufresne. The integrated square-root process. Technical report, Research Paper 90, Center for Actuarial Studies, Department of Economics, University of Melbourne, 2001.
- P.H. Dybvig. Bond and bond option pricing based on the current term structure. In M. Dempster and S. Pliska, editors, *Mathematics of Derivatives Securities (pp. 271-293)*. Cambridge University Press, 1997.
- D. Filipovic and Larsson. Polynomial jump-diffusion models. *Stochastic Systems*, 10(1):71–97, 2020. URL <http://arxiv.org/pdf/17711.08043.pdf>.
- D. Filipovic, M. Larsson, and A.B. Trolle. Linear-rational term structure models. *Journal of Finance*, 72(2):655–704, 2017.
- H. Geman, D. Madan, and M. Yor. Time changes for Lévy processes. *Mathematical Finance*, 11:79–96, 2001.
- V. Gorovoy and V. Linetsky. Intensity-based valuation of residential mortgages: an analytical tractable model. *Mathematical Finance*, 9(4):541–573, 2007.
- J. Gregory. *Counterparty Credit Risk*. Wiley Finance, 2010.
- S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.
- T. Ho and S.-B. Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41(5):1011–1029, 1986.
- M. Jeanblanc and F. Vriens. Conic martingales from stochastic integrals. *Mathematical Finance*, 28(2):516–535, 2018.
- M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Methods for Financial Markets*. Springer Verlag, Berlin, 2009.
- M. Joshi. *The Concepts and Practice of Mathematical Finance*. Cambridge University Press, 2003.
- J. Kim and T. Leung. Pricing derivatives with counterparty risk and collateralization: A fixed point approach. *European Journal of Operational Research*, 249(2):525–539, 2016.
- D. Lando. *Credit Risk Modeling: Theory and Applications*. Princeton Series in Finance. Princeton University Press, 2004.
- Markit. ISDA CDS Standard Model. Technical report, Markit, 2004. <http://www.cdsmodel.com/cdsmodel/>.
- Markit. The CDS Big-Bang: understanding the changes to the Global CDS Contract and North American conventions. Technical report, Markit, March 13, 2009. http://www.markit.com/cds/announcements/resource/cds_big_bang.pdf.

- C. Mbaye and F. Vrina. A subordinated CIR intensity model with application to wrong-way risk CVA. *International Journal of Theoretical and Applied Finance*, 21(7):1850045, 2018.
- R. Mendoza-Arriaga and V. Linetsky. Time-changed cir default intensities with two-sided mean-reverting jumps. *Annals of Applied Probability*, 24(2):811–586, 2014.
- R. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29:449–470, 1974.
- G. Pagès. *Numerical Probability: An Introduction with Applications to Finance*. Springer, Cham, 2018.
- L.C.G. Rogers. Which model for the term-structure of interest rates should one use? *Mathematical Finance, IMA*, 65:93–116, 1995.
- H. Stein and K. Pong. Counterparty valuation adjustments. In Brigo Bielecki and Patras, editors, *Credit Risk Frontiers: Subprime crisis, Pricing and Hedging, CVA, MBS, Ratings and Liquidity*, chapter 15. Wiley/Bloomberg Press, 2011.
- A. Swishchuk. *Change of Time Methods in Quantitative Finance*. SpringerBriefs in Mathematics. Springer, 2016.
- O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.
- P. Veronesi. *Fixed Income Analysis: Valuation, Risk, and Risk Management*. Wiley, 2010.
- F. Vrina. Wrong-way risk CVA models with analytical EPE profiles under Gaussian exposure dynamics. *International Journal of Theoretical and Applied Finance*, 20(7):1750045, 2017.