XAIA INVESTMENT

Negative basis measurement for callable bonds

Jan-Frederik Mai
XAIA Investment GmbH
Sonnenstraße 19, 80331 München, Germany
jan-frederik.mai@xaia.com

Mohammadali Montazeri
XAIA Investment GmbH
Sonnenstraße 19, 80331 München, Germany
mohammadali.bs@gmail.com

Date: October 17, 2016

Abstract The negative basis is an annualized earnings figure. It is supposed to measure the annualized excess return of a bond investment after having hedged away issuer-specific default risk completely via credit default swap (CDS) protection. For callable bonds a mathematically rigorous definition of the negative basis is particularly challenging because the future cash flows of the bond depend on the issuer's call decision and are random. We propose a definition based on a specific default intensity model. It is demonstrated by means of an example that the resulting negative basis measurement is smaller than it would be under the assumption of knowledge about the call time point. This is a desirable property because eliminating the credit risk associated with a bond issuer via a CDS hedge is more expensive in the presence of call rights for the issuer.

1 Introduction

For a bond that is deliverable into a credit default swap (CDS), the *negative basis* is an annualized rate that can be earned without exposure to credit risk when CDS protection on the bond issuer is bought. Intuitively, the negative basis is a compensation for taking the residual risks associated with the bond investment after issuer-specific credit risk has been eliminated. There are different economic explanations for these residual risks like counterparty credit risk, funding issues, or legal gaps, see Bernhart, Mai (2016) and the references therein. Despite its clear intuition, the negative basis is surprisingly difficult to define in a mathematically rigorous manner. For non-callable bonds the article Bernhart, Mai (2016) surveys different measurements and demonstrates why the so-called *hidden yield methodology* is the most appropriate one. The idea is to shift a reference discounting short rate parallely until both observed bond and CDS prices can be replicated jointly with one and the same piecewise-constant default intensity model. It is shown in Bernhart, Mai (2016) that there exists a portfolio consisting of the bond and several CDS with different maturities such that the internal rate of return of the respective investment equals the reference rate plus the negative basis defined in that way, irrespectively of the timing of the default. This argument provides a sound theoretical justification for the use of the hidden yield method for negative basis measurement.

The present article shows that the hidden yield definition of the negative basis can be transferred in a reasonable way also to callable bonds. The general idea of the definition is the same: the negative basis is defined as a parallel shift of a reference discounting short rate that makes it possible to find a mathematical pricing model explaining both bond price and CDS prices jointly. The major challenge concerning the extension to callable bonds is of a rather practical nature. Whereas in the non-callable case the mathematical pricing model can be based on a piecewise-constant default intensity, this is no longer sufficient for callable bonds. Pricing a callable bond with a piecewise constant default intensity model is known as the *"worst-ansatz"*. Intuitively, it implies that the market already knows when the bond is called, leading to a systematic overestimation of the bond's yield. In reality, however, the bond issuer need not decide immediately but can make his or her call decision in the future, based on then possibly changed fundamentals. In other words, it is necessary to link the call decision to a randomly evolving bond price process. Generally speaking, the main drivers influencing the volatility of the bond price process are the discounting interest rate and the default intensity, which consequently both require to be modeled stochastically. However, in the present article we focus on the stochastic modeling of only the default intensity and treat the discounting interest rate non-stochastic. Why? On the one hand, this makes the whole implementation simpler and the presentation a lot clearer. On the other hand, a negative basis can typically only be found for issuers in the high-yield sector, where the default intensity is much higher than the discounting interest rate in absolute terms, especially in a low-interest rate environment such as in the current Draghi regime.

Since stochastic default intensity models are more complex than non-stochastic default intensity models, the numerical task of efficiently computing the negative basis in practice requires much more engineering finesse. The particular challenge is to be able to fit the model to the market CDS quotes within fractions of a second (like possible for the piecewise constant intensity model), but at the same time use a model that implies realistic paths for the callable bond, based on randomly evolving expectations about the issuer's call decision (like it is present, e.g., in deliberately chosen stochastic intensity models). The crucial idea in the present paper is to use a stochastic intensity model which implies the same survival function as a given piecewise constant intensity model that is bootstrapped from market-observed CDS quotes. This guarantees that the required fit to market CDS prices is separated from the stochasticity of the model.

The remaining article is organized as follows. Section 2 introduces the methodology, Section 3 provides an example, and Section 4 concludes.

2 Negative basis for callable bonds

We observe a market price for the bond with maturity T, denoted $B(T)$, and market prices for CDS with different maturities T_i, denoted by $C(T_i)$, $i = 1, \dots, n$, where $T_1 < T_2 \dots < T_n$. Typically,

$T_n \geq T$, which is assumed henceforth, so that our model horizon equals T_n. Throughout we focus on a default intensity pricing model. This means the default time τ of the bond in concern is modeled as

$$\tau := \inf\Big\{t > 0 \,:\, \int_0^t \lambda_s \,\mathrm{d}s > \epsilon\Big\}, \tag{1}$$

with a non-negative[1] default intensity process $\lambda = \{\lambda_t\}_{t\in[0,T_n]}$ and an independent exponential random variable ϵ with unit mean. The market's filtration $(\mathcal{F}_t)_{t\geq 0}$ is defined as the natural filtration of λ, enhanced by the natural filtration of the default indicator process $\{1_{\{\tau\leq t\}}\}_{t\geq 0}$. Intuitively, this means that λ, as well as the default event, are observed by market participants. The stochastic behavior of the default intensity is controlled by a parameter (vector) θ, i.e. we choose λ from a parametric family of positive (possibly stochastic) processes, indexed by parameters θ. We denote the model price of the bond in dependence of the discounting short rate and the model parameters as $B\big(r(.), \theta, T\big)$. Similarly, the model prices of the CDS are denoted $C\big(r(.), \theta, T_i\big)$, $i = 1, \ldots, n$. The discounting short rate $r(.)$ is modeled non-stochastically throughout, which means that each future cash flow at time t is multiplied by the discounting factor $\exp(-\int_0^t r(s)\,\mathrm{d}s)$ to compute its present value.

2.1 Definition of the negative basis

The crucial idea of the hidden yield negative basis is to question the risk-free rate $r(.)$. Actually, it is postulated that another, parallelly shifted rate $r(.) + x$ can be earned without exposure to credit risk, but the parallel shift x, which is defined as the negative basis, is initially unknown. Based on the introduced notations, the hidden yield negative basis is defined as follows.

Definition 2.1 (The hidden yield negative basis)
The negative basis is defined as solution x of the following equation system:

$$\begin{aligned} C\big(r(.) + x, \theta(x), T_i\big) &= C(T_i), \quad i = 1, \ldots, n,\\ B\big(r(.) + x, \theta(x), T\big) &= B(T). \end{aligned}$$

In words, it is defined as a parallel shift of the reference discounting rate $r(.)$ that makes a successful calibration of the model parameters θ to all observed market prices possible. In general, the successfully calibrated model parameters depend on x and are hence denoted by $\theta(x)$.

On the one hand, defining the negative basis as in Definition 2.1 is motivated by arbitrage pricing theory. In the arbitrage-free model world of Definition 2.1 any self-financing trading strategy in bond and CDS necessarily must earn the rate $r(.) + x$ on average, which provides x with an intuitive meaning. For non-callable bonds, Bernhart, Mai (2016) even show the existence of a static trading strategy that is invested in the bond and several CDS with different maturities such that the rate $r(.)+x$ is earned

[1] We need to postulate non-negativity for technical reasons, because it implies that $\mathbb{P}(\tau > t)$ equals the expected value of $\exp(-\int_0^t \lambda_s\,\mathrm{d}s)$, see, e.g., Bielecki et al. (2006), which is a very convenient property.

almost surely until the minimum of either a credit event or bond maturity. This shows how the rate $r(.) + x$ can essentially be earned risk-free. For callable bonds, a static trading strategy is not recommended in general, because changing opinions about the call time point should obviously affect the maturities in the CDS hedge portfolio dynamically.
On the other hand, it is unfortunate that from a practical standpoint Definition 2.1 causes a lot of challenges regarding its implementation. It only becomes unambiguous in case the solution x to the involved equation system is unique. In general, this requires a well-deliberated choice of default intensity model. For non-callable bonds Bernhart, Mai (2016) show how to efficiently and unambiguously implement Definition 2.1 based on a piecewise constant default intensity model, in which θ equals the vector of (piecewise constant) values of the default intensity. The by far more challenging implementation for callable bonds is the content of the present article.

2.2 What is the challenge?

Unlike in the case of a non-callable bond it is not recommended to use a piecewise constant default intensity model for callable bonds, as it ignores the uncertainty about future cash flows and call decisions. However, the use of a truly stochastic default intensity makes the model fit to observed data computationally much more expensive than in the case of a piecewise constant default intensity. Even worse, it is not an easy task to find a reasonable default intensity model that perfectly matches observed market quotes at all. Mathematically, the callable bond's model price $B\big(r(.), \theta, T\big)$ is an infimum over a set of stopping times in the case of a truly stochastic default intensity. Denoting by $\mathcal{T}$ the set of all possible call time points[2] and by $\mathcal{C}$ the set of all $(\mathcal{F}_t)$-stopping times with values in $\mathcal{T}$, the bond's model price is given by

$$B\big(r(.), \theta, T\big) = \inf_{\eta \in \mathcal{C}} \Big\{ \mathbb{E}\Big[DCF\big(r(.), \theta, \eta\big) \Big] \Big\}, \tag{2}$$

where $DCF\big(r(.), \theta, \eta\big)$ denotes the discounted value of all bond cash flows that arise when the bond is called according to the stopping rule η. The computation of this infimum over a set of stopping times is achieved by means of dynamic programming, resulting in a so-called tree pricing algorithm.
The model CDS prices depend on the model (1) for the default time only via its induced survival function

$$\bar{F}(t) := \mathbb{P}(\tau > t) = \mathbb{E}\Big[e^{-\int_0^t \lambda_s \, \mathrm{d}s} \Big], \quad t \geq 0. \tag{3}$$

In contrast – and this is a fundamental difference compared with the pricing of CDS, and also compared with the pricing of non-callable bonds – the price $B\big(r(.), \theta, T\big)$ does not only depend on $\bar{F}$, but additionally depends on the level of dispersion of the default intensity process. To this end, it is important to recall that the function $\bar{F}$ does not uniquely determine the model for λ. This means we face an identifiability issue, because in general there

[2] Without loss of generality, we assume that the maturity date of the bond is in $\mathcal{T}$, i.e. "no call" is rephrased as a "call at maturity".

exist several different default intensity processes yielding the precisely same survival function $\bar{F}$, and hence CDS prices (for any fixed deterministic discounting short rate, on which the latter also depend, of course). Intuitively, the level of dispersion of λ is not determined by the observed market CDS data, only its “average behavior” in the sense that (3) must be satisfied by the model for λ for a fixed function $\bar{F}$, which is partially observed via CDS market prices. Figure 1 schematically visualizes these dependencies of model prices on model constituents.

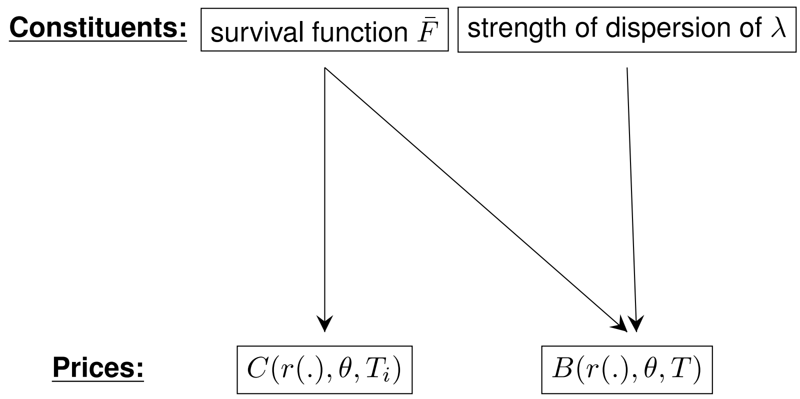

Fig. 1: Which model constituents explain which model prices?

According to this viewpoint, it is desirable to split the model parameters θ into two parts, say $\theta = (\theta_1, \theta_2)$, where θ_1 determines the survival function $\bar{F}$ and θ_2 only controls the level of dispersion of λ, and hence only affects the bond price but not the CDS prices. This is convenient with regards to Definition 2.1, since θ_1 can be fitted to CDS prices without having to worry about θ_2 at all. It is obvious that there is a lot of modeling freedom, not to say model uncertainty, regarding the strength of dispersion of λ. Since the resulting negative basis depends critically on this model choice, our approach to tackle this critical issue is to work with two extreme models: one with minimal and one with maximal dispersion (in a certain sense). In practice, this provides a continuous range for the negative basis from which the preferred measurement can be picked, based on the specific application and/or on one’s personal level of risk aversion.

2.3 Our proposed implementation

As just explained, our strategy to compute a negative basis for a callable bond is to interpolate between the respective measurements for two extreme default intensity models:

(a) **Zero dispersion of λ:** A piecewise constant intensity model, in which there is zero uncertainty about the call time point.

(b) **Maximal dispersion of λ:** A stochastic default intensity model, in which there is “maximal” uncertainty about the call time point.

The negative basis measurement according to (a) is greater than the negative basis measurement according to (b), simply because the bond price in model (a) is by definition always greater or equal to the bond price in model (b), provided both models imply the same survival function $\bar{F}$. Intuitively, the negative basis is a decreasing function in the strength of dispersion of λ (which is a measurement of the uncertainty about the call time point). In our opinion, a reasonable negative basis measurement should then be chosen somewhere from the continuous range between the two measurements in the extreme cases (a) and (b), depending on the level of uncertainty about the call time point and/or one's personal risk aversion. The following remark briefly indicates a possible method to quantify the uncertainty about the call time point. This can be helpful in order to provide a feel for how different the two models (a) and (b) are.

Remark 2.2 (Entropy of the call time point)
Denote by η an optimizer in (2), i.e. the random variable $\eta \in \mathcal{C}$ is the future time point at which the issuer calls the bond. Discretizing the probability distribution of η to only finitely many potential call time points $t_1, \ldots, t_m \in \mathcal{T}$, one particular measurement of the uncertainty about η may be accomplished via the Shannon entropy of the discrete probability distribution

$$p_i := \mathbb{P}(\eta = t_i), \quad i = 1, \ldots, m.$$

These probabilities can be obtained as a byproduct of the tree pricing algorithm which is used in order to compute the callable bond's model price. In general, the maximal possible entropy is obtained for a uniform distribution on $\{t_1, \ldots, t_m\}$ and equals $\log_2(m)$. Consequently, the normalized Shannon entropy

$$H := \frac{-\sum_{i=1}^{m} p_i \, \log_2(p_i)}{\log_2(m)} \in [0, 1]$$

provides a feel for the amount of uncertainty about the call time point. Obviously, using the piecewise constant default intensity model (a), the induced entropy equals zero. It is our intuition that a minimizer η in (2) induces a high entropy, since the entropy of η is a measurement that is thought to be highly correlated with any reasonable dispersion measure of λ. Consequently, the model (b) implies an entropy H_b that is expected to be quite close to the maximally possible entropy consistent with observed market prices. If a market participant can quantify his or her subjective opinion on the uncertainty about the issuer's call time point in terms of an entropy $H \in [0, H_b]$, the value $H/H_b \in [0, 1]$ is one possible choice for the interpolation value between the two extreme negative basis figures of methods (a) and (b).

The computation (a) is straightforward, since the default intensity is non-stochastic. In this case, the seemingly difficult bond pricing formula (2) boils down to the easy-to-implement "worst-ansatz"

$$B\big(r(.), \theta, T\big) = \inf_{t \in \mathcal{T}} \Big\{ BB\big(r(.), \theta, t\big) \Big\}, \tag{4}$$

where $BB\big(r(.), \theta, t\big)$ denotes the value of the non-callable (bullet) bond that arises when the callable bond is called at time t for sure. Approximating the appearing infimum in (4) by a minimum over a finite number of time points in $\mathcal{T}$, this "worst price" of the callable bond provides an easy-to-compute upper bound for the callable bond price in general, which becomes sharp in case the default intensity is non-stochastic. The practical computation of the negative basis with the piecewise constant intensity model then works precisely along the lines described in Bernhart, Mai (2016), with the sole exception that the bond price formula in this reference has to be replaced by the "worst price" formula. This task is numerically efficient and the resulting negative basis is unique, i.e. there are no problems with an ambiguous definition, as already demonstrated[3] in Bernhart, Mai (2016).

The numerically challenging part is the second computation (b). Recall that the market CDS prices only depend on the survival function $\bar{F}$ of τ. We denote by $\theta_1(x)$ the parameters (i.e. values) of a non-stochastic, piecewise constant default intensity function that perfectly explains all observed CDS prices under the assumption that the risk-free discounting rate equals $r(.) + x$. The associated piecewise constant default intensity function is denoted by $\lambda_x^{(\theta_1(x))}(t)$, $t \geq 0$, and the associated survival function by $\bar{F}_x^{(\theta_1(x))}$, i.e.

$$\bar{F}_x^{(\theta_1(x))}(t) = e^{-\int_0^t \lambda_x^{(\theta_1(x))}(s)\,\mathrm{d}s}, \quad t \geq 0. \tag{5}$$

The calibration of $\theta_1(x)$, resp. $\bar{F}_x^{(\theta_1(x))}$, to observed market CDS prices is a standard algorithm that is quick, unambiguous, and known by every market participant in the CDS marketplace. Borrowing an idea of Brigo, Mercurio (2001), we model the default intensity as the sum of a truly stochastic part and a non-stochastic, compensating part which is used in order to match the resulting survival function perfectly to the one in (5). More clearly, we define the default intensity $\lambda_t := \xi_t + \varphi_x(t)$ as the sum of a non-negative Itô diffusion $\xi = \{\xi_t\}_{t \in [0,T_n]}$, whose parameters we denote by θ_2 in the sequel and which do not depend at all on x, and a non-stochastic part $\varphi_x = \{\varphi_x(t)\}_{t \in [0,T_n]}$, which does depend on x. Following Brigo, Mercurio (2001), the compensating part φ_x is uniquely determined from the model for ξ and from $\theta_1(x)$ as

$$\varphi_x(t) := \lambda_x^{(\theta_1(x))}(t) - \Bigg(\underbrace{-\frac{\partial}{\partial t}\log\Big(\mathbb{E}\Big[e^{-\int_0^t \xi_s\,\mathrm{d}s}\Big]\Big)}_{=:f_{\theta_2}(t)}\Bigg), \quad t \geq 0.$$

By construction, under the assumption that φ_x is a non-negative function (which one needs to ensure), the survival function of τ in this model is given by

$$\mathbb{P}(\tau > t) = \mathbb{E}\Big[e^{-\int_0^t \xi_s\,\mathrm{d}s}\Big]\, e^{-\int_0^t \varphi_x(s)\,\mathrm{d}s} = \bar{F}_x^{(\theta_1(x))}(t).$$

The crucial idea is that this model can be matched to the observed CDS market prices very quickly, because the calibration is fully independent of (the parameters θ_2 of) ξ, it only requires to bootstrap the parameters $\theta_1(x)$ along the lines described in O'Kane, Turnbull (2003). The parameters θ_2 of the diffusion ξ are

[3] For the piecewise constant default intensity model it is shown in (Bernhart, Mai, 2016, Lemma 2) that the function $x \mapsto BB(r(.) + x, \theta(x), t)$ is decreasing and continuous for each fixed t, which carries over to $x \mapsto B(r(.) + x, \theta(x), T)$ as the infimum over continuous and decreasing functions is continuous and decreasing.

independent of x and control solely the level of dispersion of the default intensity, as desired according to Figure 1. The choice of ξ must be well-deliberated and constitutes the most challenging part of the whole modeling process. On the one hand, it should exhibit a great level of dispersion in order to be consistent with our underlying idea of constituting an extremely dispersed default intensity model. On the other hand, its parameters θ_2 must be chosen according to the restriction

$$f_{\theta_2}(t) \leq \inf_x \left\{\lambda_x^{(\theta_1(x))}(t)\right\}, \quad \text{for all } t \geq 0, \tag{6}$$

in order to guarantee non-negativity of φ_x, and hence almost sure non-negativity of the default intensity λ. The infimum in (6) is taken over all x in a reasonable range of potential negative basis measurements. In particular, since the negative basis is decreasing in the level of dispersion of λ, the negative basis measurement according to model (a) establishes an upper bound, say $\bar{x}$, for the solution x in Definition 2.1 according to model (b). A lower bound $\underline{x}$ may be set according to the specific business case (e.g. $\underline{x} = 0$ in case one is not interested in bonds trading expensive relative to CDS anyway), and the infimum in (6) is taken over $x \in [\underline{x}, \bar{x}]$. In this regard, it is very helpful to notice that this infimum can be computed in closed form very efficiently, as the following lemma shows.

Lemma 2.3 (Effect of x on $\lambda_x^{(\theta_1(x))}(.)$)
For every $t \geq 0$ we have that

$$\inf_{x\in[\underline{x},\bar{x}]} \left\{\lambda_x^{(\theta_1(x))}(t)\right\} = \min\left\{\lambda_{\underline{x}}^{(\theta_1(\underline{x}))}(t), \lambda_{\bar{x}}^{(\theta_1(\bar{x}))}(t)\right\} =: \underline{\lambda}(t;\underline{x},\bar{x}),$$
$$\sup_{x\in[\underline{x},\bar{x}]} \left\{\lambda_x^{(\theta_1(x))}(t)\right\} = \max\left\{\lambda_{\underline{x}}^{(\theta_1(\underline{x}))}(t), \lambda_{\bar{x}}^{(\theta_1(\bar{x}))}(t)\right\} =: \overline{\lambda}(t;\underline{x},\bar{x}).$$

Proof
See Appendix. □

Lemma 2.3 is based on the observation that the function $x \mapsto \lambda_x^{(\theta_1(x))}(t)$ is monotone for each fixed t, which implies that

$$\underline{\lambda}(t;\underline{x},\bar{x}) \leq \lambda_x^{(\theta_1(x))}(t) \leq \overline{\lambda}(t;\underline{x},\bar{x})$$

for any parallel shift $x \in [\underline{x}, \bar{x}]$ of the reference discounting rate $r(.)$, and the lower and upper bound intensities are easy to compute. Moreover, the interval $[\underline{\lambda}(t;\underline{x},\bar{x}), \overline{\lambda}(t;\underline{x},\bar{x})]$ is typically quite small, which follows from the fact that the sensitivity of model CDS prices $C(r(.), \theta_1, T_i)$ with respect to the short rate $r(.)$ is relatively small compared to its sensitivity with respect to the default intensity parameters θ_1. This phenomenon results from the fact that a change in $r(.)$ affects the discount factors in both legs of the CDS pricing formula in a similar way, thus netting themselves out.

2.4 The model for ξ

We define the Itô diffussion ξ as the unique solution to the SDE

$$\mathrm{d}\xi_t = \xi_t \left(\sqrt{\xi_t}\,\frac{2\,\beta\,\sigma}{\sqrt{\xi_0}}\,\mathrm{d}W_t + \xi_t\,\beta\left(2 + \frac{\sigma^2\,(2\,\beta - 1)}{\xi_0}\right)\mathrm{d}t\right)$$

for model parameters $\xi_0, \sigma > 0$ and $\beta < 0$, i.e. $\theta_2 = (\xi_0, \beta, \sigma)$. The parameter ξ_0 is also chosen as the starting value for the diffusion, justifying its notation. The choice of this particular Itô diffusion model is a result of the following deliberations:

- **Analytical tractability:** The model is a three-parametric subfamily of a default intensity model implicitely used in Carr, Linetsky (2006). In the latter reference the default intensity is modeled as a function of a stock price process, which is itself modeled as a diffusion. Using relations with Bessel processes, the results of Carr, Linetsky (2006) imply closed formulas for the associated survival function in that model, from which we obtain f_{θ_2} efficiently by finite differencing. Moreover, the connection to Bessel processes can be used in order to define efficient trinomial tree approximations for ξ, hence λ, resulting in efficiently implementable dynamic programming pricing algorithms for the computation of the callable bond's model price $B(r(.), \theta, T)$. The availability of such a pricing engine is a necessary requirement for the model of ξ.

- **Flexible shapes of f_{θ_2}:** Even though the model is only three-parametric, the set of possible shapes for the function f_{θ_2} is quite rich. It turns out that our goal of choosing θ_2 such that ξ is maximally dispersed without violating the restriction (6) corresponds to fitting θ_2 in such a way that f_{θ_2} is very close to its upper bound $\underline{\lambda}(.; \underline{x}, \bar{x})$. This is intuitively clear, because the further it stays away from its upper bound, the larger is the non-stochastic compensating part φ_x of the default intensity, and consequently the smaller is the level of dispersion of λ, which can only be induced by the stochastic part ξ. Summarizing, the ability to replicate with f_{θ_2} as many different shapes as possible is important, because the given upper bound $\underline{\lambda}(.; \underline{x}, \bar{x})$ can exhibit any shape (flat, inverted, upward sloping).

- **Explosive property possible:** One can immediately check using Feller's test for explosions[4] that the diffusion ξ is strictly positive almost surely, but may explode to infinity if its parameters are chosen according to the criterion $2\,\xi_0 < \sigma^2$. On the one hand, this is a very desirable property when credit risk is modeled via a default intensity, as nicely explained in Andreasen (2001). In particular, many popular stochastic short rate models, like the CIR process, do not exhibit this desirable property. On the other hand, it is intuitively clear that a possibly explosive behavior of ξ, leading to an explosive behavior of λ, results in a very high level of dispersion, as desired.

How do we find the parameters $\theta_2 = (\xi_0, \beta, \sigma)$ in practice? We define a grid for both ξ_0 and β. For ξ_0 we choose 10 equidistant values, ranging from 0.001 to $\underline{\lambda}(0; \underline{x}, \bar{x}) - 0.001$. This corresponds to an equidistant covering of the admissible range $\big(0, \underline{\lambda}(0; \underline{x}, \bar{x})\big)$. For β, we choose the grid $\{-1, -0.9, \ldots, -0.2\}$, which is an equidistant covering of the range $[-1, -0.2]$, which we consider reasonable based on our empirical experience. The choice of ξ_0

[4]See, e.g., Karatzas, Ruf (2015).

determines the value $f_{\theta_2}(0) = \xi_0$. Keeping (ξ_0, β) fixed, one can show that $f_{\theta_2}(t)$ is increasing in σ for all $t > 0$. Since it is intuitive that the level of dispersion of ξ is increased by an increase of f_{θ_2}, hence by an incrase in σ, for each pair (ξ_0, β) we choose $\sigma = \sigma(\xi_0, \beta)$ as large as possible before (6) is violated. Then, for each triplet $\theta_2 = (\xi_0, \beta, \sigma)$ we have found in this way, we find the unique $x = x(\theta_2)$ according to Definition 2.1 via a bisection root search applied to the continuous and decreasing[5] function

$$x \mapsto B\big(r(.) + x, (\theta_1(x), \theta_2), T\big) - B(T). \tag{7}$$

Finally, the parameters θ_2 are chosen as the minimizer of $x(\theta_2)$ among all considered triplets $\theta_2 = (\xi_0, \beta, \sigma)$. A concrete example visualizing the process is provided in the following section.

3 Example

On 7th April 2015 we consider the ABC[6] bond with maturity at 1st June 2020, which pays the coupon rate 8.875% semi-annually. The bond trades above par value, namely at 101.3%. Tables 1 and 2 depict the bond information regarding the call option and the observed CDS prices.

from	until	call right
now	May 29th 2017	not callable
May 30th 2017	May 29th 2018	at strike 104.438%
May 30th 2018	May 29th 2019	at strike 102.219%
May 30th 2019	maturity	at strike 100%

Table 1: Call information for the ABC bond.

Maturity	Par CDS spread (in bps)	upfront (in $\%$)
1y	144	-4.29
2y	260	-5.17
3y	376	-3.73
4y	436	-2.44
5y	472	-1.28

Table 2: CDS market prices for the ABC case. Upfronts refer to a running coupon of 500 bps, which is the standardized strike level for ABC.

Figure 2 depicts the negative basis measurement in the extreme model (a) based on a piecewise constant default intensity. Each bar in the plot corresponds to a potential call time point and illustrates the negative basis in case of a certain call at that respective time point. The minimum of all these measurements equals precisely the negative basis based on model (a), which in this example means that no call (=call at maturity) is assumed. It equals 220.3385 bps. Intuitively, this means that if the issuer had to fix the call time point immediately (on 7th April 2015), then he or she would decide not to call at all.
Figure 3 illustrates the model for the stochastic default intensity in the extreme model (b) that minimizes the negative basis.

[5] The continuity and decreasingness of this function is proved in the Appendix.
[6] All numbers are inspired by a real-world example, which for discretionary reasons we call ABC in the present article.

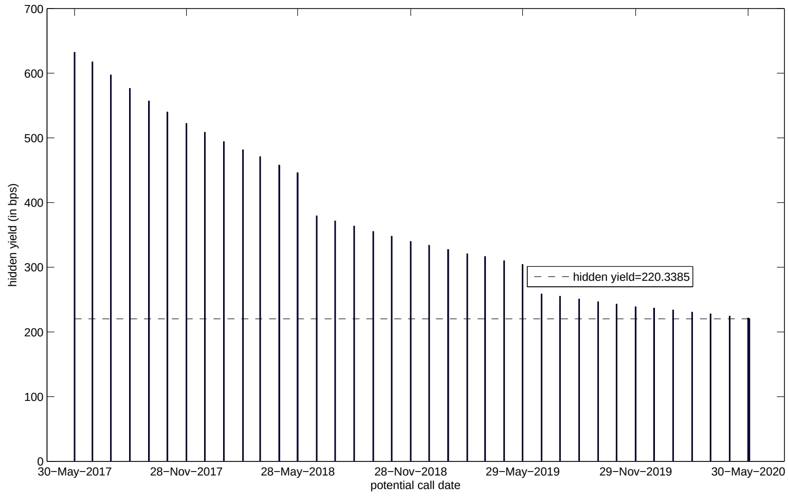

Fig. 2: Negative basis using the extreme model (a) based on a non-stochastic default intensity.

The minimal and maximal functions $\underline{\lambda}(.;\underline{x},\bar{x})$ and $\bar{\lambda}(.;\underline{x},\bar{x})$ are computed for $\underline{x} = 0$ and $\bar{x} = 0.02203385$. One observes that the sensitivity of the bootstrapped piecewise constant default intensity $\lambda_x^{(\theta_1(x))}$ on the interest rate shift x is very small, as mentioned after Lemma 2.3. The function f_{θ_2} that results from the model found for ξ lies below the minimal function $\underline{\lambda}(.;\underline{x},\bar{x})$ in order to guarantee non-negativity of the default intensity, as desired. The parameters of ξ in this example are $\theta_2 = (\xi_0, \beta, \sigma) = (0.001, -0.5, 0.5034)$.

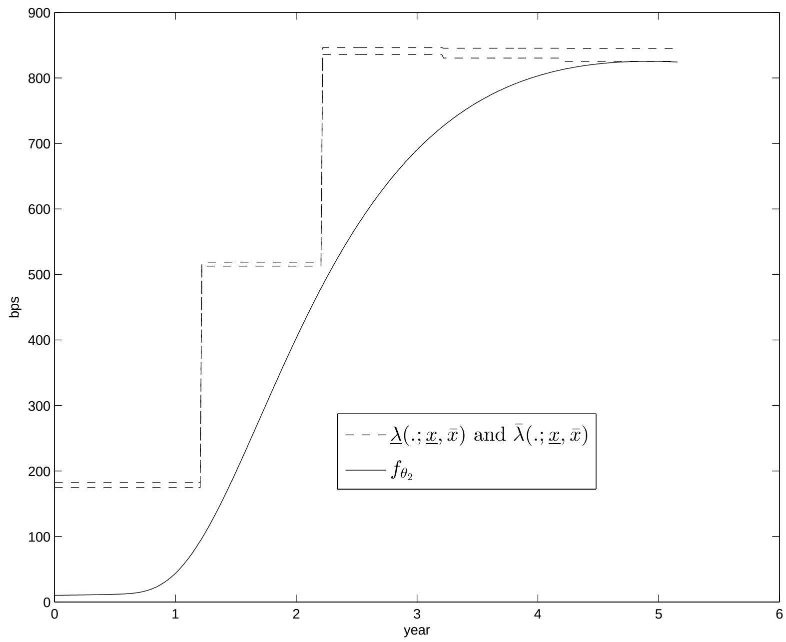

Fig. 3: Visualization of the functions f_{θ_2}, $\underline{\lambda}(.;\underline{x},\bar{x})$, and $\bar{\lambda}(.;\underline{x},\bar{x})$.

Figures 4 and 5 illustrate three simulations of the default intensity and of the resulting bond price process in the extreme model (b) for the stochastic default intensity. The bond price simulations have been performed precisely along the depicted paths of the default intensity. In two of the three simulations the default intensity $\lambda = \xi + \varphi_x$ behaves with almost no volatility like the smooth non-stochastic component φ_x, which can clearly be observed in the plot.The stochastic component ξ remains almost constantly on its low starting value $\xi_0 = 0.001$. In the third simulation, however, the default intensity suddenly explodes to infinity, leading to a default of the bond. This explosive property makes clear that the distribution of ξ_t at any fixed $t > 0$ is highly dispersed, as desired. Moreover, the model obviously induces a highly volatile bond price process, confirming our intuition that this is required in order to increase the uncertainty about the call time point, and thus decreasing the negative basis.

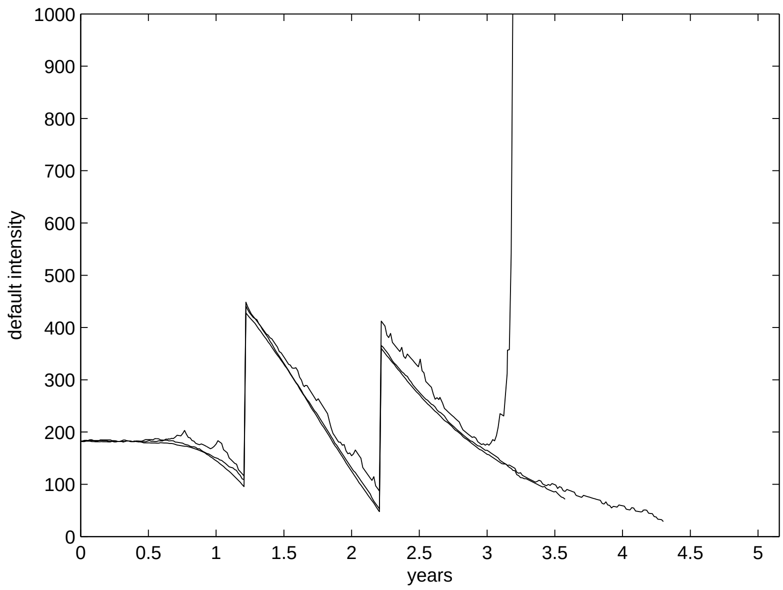

Fig. 4: Three sample paths of the default intensity λ in the extreme model (b) leading to the minimal negative basis.

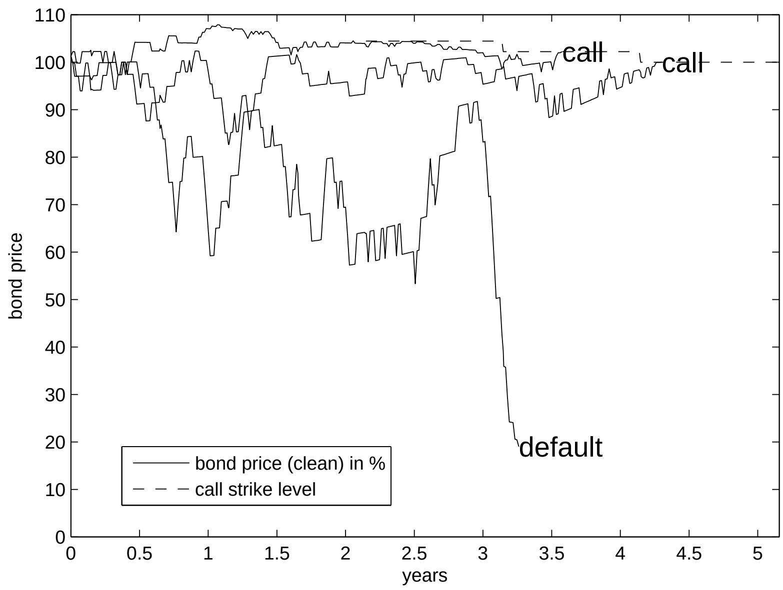

Fig. 5: Three sample paths of the bond price process in the extreme model (b) leading to the minimal negative basis.

The discounting short rate $r(.)$ that we use is retrieved from market-observed prices for interest rate swaps with different maturities, based on a quarterly payment schedule. The optimal parameter choice for θ_2 took about 30 minutes on a standard PC. Recall that we search the latter on a two-dimensional grid for (ξ_0, β) with $10 \cdot 10 = 100$ grid points. For each grid point we first have to find σ, and then compute the resulting negative basis by means of a bisection with a function that runs a dynamic program in order to determine the callable bond's price. Concluding, the negative basis computed according to the extreme stochastic default intensity model (b) is found to be 124 bps, which means that the use of a deterministic discounting short rate $r(.)+0.0124$ makes the observed prices of the CDS and the bond arbitrage-consistent in the respective model with the extremely dispersed stochastic default intensity of model (b). It is noticeable that this negative basis figure is far below the value 220 bps of the non-stochastic default intensity model according to the approach (a). This shows that the given market prices allow for a lot of modeling freedom regarding the call time point uncertainty, respectively the strength of dispersion of λ. The normalized Shannon entropy of the optimal call time point, according to Remark 2.2, is given by 43.87% in the extreme model (b).

4 Conclusion

We have presented a method to compute a negative basis figure for callable bonds. We discussed the challenges of computing such measurement in practice and proposed a specific implementation. The latter was based on a pricing model in which the default intensity process was modeled as the sum of a stochastic and a deterministic part. The key to an efficient implementation was to separate the stochasticity of the default intensity model from its ability to match observed market CDS prices, which has been achieved by a deterministic shift-extension in the sense of Brigo, Mercurio (2001).

References

J. Andreasen, Credit explosives, *Working paper, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=262682* (2001).

G. Bernhart, J.-F. Mai, Negative basis measurement: finding the holy scale, in *Innovations in Derivatives Markets - Fixed Income modeling, valuation adjustments, risk management, and regulation, edited by K. Glau et al., Springer-Verlag* (2016) pp. 385–403.

Bielecki T., Jeanblanc M., and Rutkowski M., Credit risk, *Lecture notes of M. Jeanblanc, available at http://www.maths.univ-evry.fr/pages_perso/jeanblanc/conferences/lisbon.pdf* (2006).

D. Brigo, F. Mercurio, A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models, *Finance and Stochastics* **5** pp. 369–387 (2001).

P. Carr, V. Linetsky, A jump to default extended CEV model: an application of Bessel processes, *Finance and Stochastics* **10:3** (2006) pp. 303–330.

D. O'Kane, S. Turnbull, valuation of credit default swaps, *Fixed Income Quantitative Research Lehman Brothers* (2003).

I. Karatzas, J. Ruf, Distribution of the time to explosion for one-dimensional diffusions, *Probability theory and related fields, forthcoming* (2015).

Appendix

Proof of Lemma 2.3 It suffices to prove that the function $x \mapsto \lambda_x^{(\theta_1(x))}(t)$ is monotone, for any fixed $t \geq 0$. For the sake of a rigorous proof, it is convenient to first introduce some further notation. Recall that we observe n CDS prices with maturities $T_1 < T_2 < \ldots < T_n$, and the parameter $\theta_1(x)$ denotes the values of the piecewise constant default intensity function $\lambda_x^{(\theta_1(x))}(.)$ that is bootstrapped from these market prices when the applied discounting short rate is $r(.) + x$. This bootstrap procedure determines the values of the piecewise constant default intensity uniquely and is described in O'Kane, Turnbull (2003). In total, $\theta_1(x)$ is a vector of n numbers, namely the n values of the default intensity on the intervals $[T_0, T_1)$, $[T_1, T_2)$, $\ldots$, and $[T_{n-1}, T_n)$, where $T_0 := 0$. We denote the value on $[T_{i-1}, T_i)$ by $y_i(x)$ for the remainder of the proof, i.e. $\theta_1(x) = \big(y_1(x), \ldots, y_n(x)\big)$. We prove by induction over n that each function $x \mapsto y_i(x)$ is monotone, which proves the claim. Finally, we denote by $upf(T_i)$ the observed upfront (i.e. market) price for the CDS with maturity T_i, corresponding to the CDS running coupon of s, which is assumed not to depend on the CDS maturity[7].

In the induction start we prove that $x \mapsto y_1(x)$ is a monotone function. We denote by $R \in [0, 1)$ the constant recovery assumption for the CDS contracts and assume that the CDS running spread is paid continuously for the sake of a simplified notation. We define the functions

$$f(x, y) := upf(T_1) + s \int_0^{T_1} e^{-\int_0^t r(s)+x+y \,\mathrm{d}s} \,\mathrm{d}t,$$

$$g(x, y) := (1 - R)\, y \int_0^{T_1} e^{-\int_0^t r(s)+x+y \,\mathrm{d}s} \,\mathrm{d}t.$$

For given x, the value $y_1(x)$ is defined as the unique root of the equation $f(x, y) = g(x, y)$ in the variable y, i.e. $y_1(x)$ is implicitly defined by the relation $f\big(x, y_1(x)\big) = g\big(x, y_1(x)\big)$. We distinguish three cases:

(i) $upf(T_1) = 0$: In this case, obviously $y_1(x) = s/(1 - R)$ is fully independent of x. In particular, $x \mapsto y_1(x)$ is a constant, hence monotone.

(ii) $upf(T_1) > 0$: Fix $\epsilon > 0$ arbitrary. Introducing the help variable

$$v(x) := \int_0^{T_1} e^{-\int_0^t r(s)+x+y_1(x) \,\mathrm{d}s} \,\mathrm{d}t,$$

[7] The CDS running coupon is standardized in the marketplace, typically to either 100 bps or 500 bps across all maturities for one reference entity. Changes in creditworthiness are not reflected in a change of the standardized CDS running spread, but in a change of the CDS upfront price, which equals its market value.

the defining equation $f\big(x, y_1(x)\big) = g\big(x, y_1(x)\big)$ reads

$$upf(T_1) + s\, v(x) = (1-R)\, y_1(x)\, v(x). \tag{8}$$

Considering Equation (8) as an equation in v, it corresponds to two lines with non-negative slope intersecting in $v = v(x)$. Since $upf(T_1) > 0$, the y-intercept of the left-hand line is greater than the one of the right-hand line (which is zero), so that the left-hand line lies strictly above the right-hand line for all $v < v(x)$. It is shown in the proof of (Bernhart, Mai, 2016, Lemma 2(b)) that the function $x \mapsto x + y_1(x)$ is non-decreasing, implying that $v(x+\epsilon) < v(x)$. It follows that $y_1(x+\epsilon) > y_1(x)$ in order to achieve the required equality

$$upf(T_1) + s\, v(x+\epsilon) = (1-R)\, y_1(x+\epsilon)\, v(x+\epsilon).$$

Since $\epsilon > 0$ was arbitrary, the function $x \mapsto y_1(x)$ is increasing, hence montone.

(iii) $upf(T_1) < 0$: An analogous argument as in case (ii) implies that the function $x \mapsto y_1(x)$ is decreasing, hence montone.

We now proceed with the induction step, assuming that monotonicity of the functions $x \mapsto y_j(x)$ is shown for $j = 1, \ldots, i-1$. We seek to prove monotonicity of $x \mapsto y_i(x)$. We define the functions

$$\begin{aligned} f(x,y) :&= upf(T_i) - upf(T_{i-1}) \\ &\quad + s \int_{T_{i-1}}^{T_i} e^{-\int_0^t r(s) + x + \lambda_x^{(y_1(x),\ldots,y_{i-1}(x),y)}\,\mathrm{d}s}\,\mathrm{d}t, \\ g(x,y) :&= (1-R)\, y \int_{T_{i-1}}^{T_i} e^{-\int_0^t r(s) + x + \lambda_x^{(y_1(x),\ldots,y_{i-1}(x),y)}\,\mathrm{d}s}\,\mathrm{d}t, \end{aligned}$$

highlighting the dependence of the piecewise constant default intensity on $[0, T_i)$ on its respective values $(y_1(x), \ldots, y_{i-1}(x), y)$, denoting the last value on the piece $[T_{i-1}, T_i)$ as the free variable y. The value $y_i(x)$ equals the unique solution of the equation $f(x,y) = g(x,y)$ in the variable y, i.e. is implicitly defined by the equation $f\big(x, y_i(x)\big) = g\big(x, y_i(x)\big)$. The precisely same argument as in the induction start now proves that

$$x \mapsto y_i(x) \text{ is } \left\{ \begin{array}{ll} \text{constant} & \text{, if } upf(T_i) - upf(T_{i-1}) = 0, \\ \text{increasing} & \text{, if } upf(T_i) - upf(T_{i-1}) > 0, \\ \text{decreasing} & \text{, if } upf(T_i) - upf(T_{i-1}) < 0, \end{array} \right\}.$$

Consequently, $x \mapsto y_i(x)$ is monotone.

Proof of decreasingness of the function (7)

According to (2) the model bond price function equals

$$B\big(r(.) + x, (\theta_1(x), \theta_2), T\big) = \inf_{\eta \in \mathcal{C}} \Big\{ \mathbb{E}\Big[DCF\big(r(.) + x, (\theta_1(x), \theta_2), \eta\big) \Big] \Big\}.$$

Now the continuity and decreasingness of the function

$$x \mapsto DCF\big(r(.) + x, (\theta_1(x), \theta_2), \eta(\omega)\big)$$

for each possible realization $\eta(\omega)$ of the call decision is shown precisely along the same lines as this is shown for the non-callable case in (Bernhart, Mai, 2016, Lemma 2). This obviously implies that the function

$$x \mapsto \mathbb{E}\Big[DCF\big(r(.) + x, (\theta_1(x), \theta_2), \eta\big)\Big]$$

is continuous and decreasing as well, and since the infimum of continuous and decreasing functions is also of such kind, the claimed continuity and decreasingness follows.