Calibration and implementation of convertible bond models

Leif Andersen
Banc of America Securities LLC, 9 West 57th Street, New York, NY 10019, USA

Dan Buffum
Banc of America Securities LLC, 100 N Tryon Street, Charlotte, NC 28255, USA

While convertible bond models recently have come to rest on solid theoretical foundation, issues in model calibration and numerical implementation still remain. This paper highlights and quantifies a number of such issues, demonstrating, among other things, that naïve calibration approaches can lead to highly significant pricing biases. We suggest a number of techniques to resolve such biases. In particular, we demonstrate how applications of the Fokker–Planck PDE allows for efficient joint calibration to debt and option markets, and also discuss volatility smile effects and the derivation of forward PDEs to embed such information into model calibration. Throughout, we rely on modern finite-difference techniques, rather than the binomial or trinomial trees that so far have dominated much of the literature.

1 Introduction

While it has long been realized that a framework for pricing convertible bonds should ideally incorporate elements of both equity and debt modeling, practical efforts in this direction have long lacked a firm theoretical basis. In particular, there seems to have been considerable confusion and disagreement about how to appropriately and consistently apply a default-adjusted discount operator to cashflows generated by convertible bonds. Early papers with an ad hoc approach to discounting include McConell and Schwarz (1986), Cheung and Nelken (1994), and Ho and Pfeffer (1996). Many of these models do not explicitly model bankruptcy, and as compensation uniformly apply a somewhat arbitrary risky spread to the risk-free discount rate. More recent papers recognize that equity and debt components of convertible bonds are subject to different default risk and attempt more nuanced schemes. An often-quoted example is Tsivioritis and Fernandes (TF) (1998) (later extended by Yigitbasioglu, 2001, to multiple factors), which effectively splits the convertible bond into cash and equity components, with only the former being subject to credit risk. A related approach was promoted by Goldman Sachs (1994) and involves careful weighting of risky and risk-free

The authors wish to thank Peter Carr, Alex Lipton, Peter Forsyth, Vladimir Piterbarg, and Jesper Andreasen for insights and comments. All errors are our own. Communicating author: leif.andersen@bofasecurities.com

discounting in a binomial lattice. The TF splitting scheme is analyzed in detail in Ayache *et al* (2002) who conclude that it is inherently unsatisfactory due to its unrealistic assumption of stock prices being unaffected by bankruptcy.

With the advances of credit derivatives theory, in particular the reduced-form approach[1] of Jarrow and Turnbull (1995), the foundation for convertible bond models has recently improved significantly. A key development has been the inclusion of stock price dynamics that explicitly incorporate default events, as well as the explicit modeling of stock and bond recoveries in default. Most commonly, default is modeled as a Poisson event that drives stock prices into some low value and coupon bond prices (and convertible bonds) into a certain, fixed percentage of their notional values. Representative, and quite similar, papers include Davis and Lischka (1999) and Takahashi *et al* (2001). See Grimwood and Hodges (2002) and Olsen (2002) for comparisons of the approach in these papers against other models in the literature. In recent work, Ayache *et al* (2002) lay out a solid basis for the numerical computation of convertible bond prices, discussing in detail how modern finite-difference methods can replace the computationally sub-optimal binomial and trinomial trees that pervade most of the literature.

With theory and computing techniques now on a relatively solid basis, it remains to be determined how to best parameterize models for convertible bonds. While a number of specific parameterizations have emerged in the literature – eg, Muromachi (1999), Bloch and Miralles (2002), and Arvanitis and Gregory (2001) – these are typically based on empirical observations and do not generally result in a model that will price any particular instrument close to market. In fact, as we shall see, when applied to such simple instruments as stock options and coupon bonds, naïvely parameterized convertible bond models can yield surprisingly large price biases. In a trading setting where we might be interested in relative value plays, or perhaps want to hedge all or pieces of the convertible bond with options and straight debt (or credit derivatives), this situation is obviously not ideal.

In this paper, we will discuss the parameterization and calibration of convertible bond models to quoted prices of straight debt and equity options. That is, in the time-honored tradition of financial engineering we will attempt to "imply" parameters from market quotes on actively traded securities. The treatment of this topic will proceed as follows: in Section 2, we outline our process assumptions and discuss a number of technical issues. Section 3 discusses numerical implementation and analyzes a number of model effects in vanilla options and

[1] Briefly, the reduced-form approach directly models default as a point process focusing primarily on the properties of the default intensity process. An alternative framework, the *structural approach*, models default as the first passage time to a barrier of some process, typically either the company equity or a proxy for its assets. See Merton (1974) for a classical example. While it is in principle possible to build convertible bond models using the structural approach (see eg, Brennan and Schwarz, 1980), the reduced-form approach is, by far, the most natural for trading applications and shall be the sole focus of this paper.

straight debt. Section 4 discusses forward and backward equations for transition densities, and outlines an algorithm for joint calibration to debt and stock options markets. A number of numerical examples are provided to illustrate typical calibration results. For reference and future tests, Section 5 lists prices for a few standard convertible bonds, while Section 6 discusses certain interesting extensions and avenues for future research. In particular, we briefly illustrate how, in theory, debt markets can be tied together with equity option volatility skews using forward PDEs. Finally, Section 7 concludes the paper.

2 Model

2.1 Basics

Consider the pricing of a convertible bond issued by a company with publicly traded equity S. For most of this paper, we assume that S is the single underlying state variable of our model. (Extensions to stochastic interest rates are straightforward, albeit labor-intensive, and will be discussed in more detail in Section 6.) To incorporate the possibility of defaults on the underlying company, we make the standard assumption that default of the underlying company is governed by the first jump of a Cox process (eg, Lando, 1998) $N(t)$ with a stochastic intensity entirely captured by a functional dependence on the stock price level. Specifically, we let the time t intensity of $N(t)$ be denoted $\lambda(t, S)$, where $\lambda: \mathbb{R}_+^2 \to \mathbb{R}_+$ is some well-behaved deterministic function. Before the default time $\tau = \inf\{t: N(t) = 1\}$ we assume that S is a diffusion process driven by a single Brownian motion $W(t)$, independent of $N(t)$, and let the instantaneous diffusion volatility of S be $\sigma(t, S)$ for some smooth, bounded function $\sigma: \mathbb{R}_+^2 \to \mathbb{R}_+$. With the risk-free interest rate r and the instantaneous dividend yield q both assumed deterministic, the risk-neutral stock process can be stated as

$$\frac{\mathrm{d}S(t)}{S(t-)} = \big(r(t) - q(t) + \lambda(t, S(t-))\big)\mathrm{d}t + \sigma(t, S(t-))\mathrm{d}W(t) - \mathrm{d}N(t) \qquad (1)$$

where $t-$ is defined as the limit of $t - \varepsilon$ for $\varepsilon \downarrow 0$. A few comments to this SDE are in order. First, notice the drift term $\lambda(t, S(t-))$ which compensates for the expected downward drift of the Cox process term: $E_t(-\mathrm{d}N(t)) = -\lambda(t, S(t-)\mathrm{d}t$, where $E_t(\cdot)$ is the time t risk-neutral expectation operator. The drift compensation is required for the process to satisfy the arbitrage restriction that $S(t)\exp\big(-\int_0^t [r(u) - q(u)]\mathrm{d}u\big)$ be a martingale in the risk-neutral probability measure. Second, notice that we assume that the stock price drops to zero upon default: when N jumps from 0 to 1, $\mathrm{d}S(t) = -S(t-)$ and the stock is driven into 0, where it stays. The assumption that equity holders recover essentially nothing on default is reasonable, and consistent with much existing literature; see for instance Davis and Lischka (2001). However, note that if we instead wanted to assume, as in Ayache *et al* (2002), that some fraction R_S of the pre-default value of the stock is recovered in default, we simply multiply the terms $\mathrm{d}N(t)$ and

$\lambda(t, S(t-))\mathrm{d}t$ in (1) by $(1 - R_S)$.[2] For simplicity, however, we throughout use the approximation $R_S \approx 0$. As an aside, we notice that the assumption of a bounded $\sigma(t, S)$ ensures that the stock price cannot diffuse to 0 but only reach this value by a default jump.

Consider now the pricing of a contingent claim V with maturity T. Writing $V = V(t, S)$, the claim value is governed by the following backward PDE,[3] subject to a boundary payout condition at T:

$$\frac{\partial V}{\partial t} + \big(r(t) - q(t) + \lambda(t, S)\big) S \frac{\partial V}{\partial S} + \tfrac{1}{2}\sigma(t,S)^2 S^2 \frac{\partial^2 V}{\partial S^2} = \big(r(t) + \lambda(t, S)\big) V - \lambda(t,S) R_V(t,S) \tag{2}$$

Here, $R_V(t, S)$ is the recovery value of V in case of default at time t; the recovery value can be allowed to depend on both time and the pre-default value of the stock price. For securities, such as convertible bonds, paying intermediate coupon cashflows, additional boundary conditions are obviously needed at each cashflow date (see Section 5). Further intermediate boundary conditions are needed to capture early exercise options and put/call features, all of which are present in a typical convertible bond. The formulation of such boundary conditions is standard; see for instance Tavella and Randall (2000) for details (see also Section 5).

Derivation of the backward PDE (2) is straightforward and follows from the jump-extended Itô lemma for the stochastic differential $\mathrm{d}V(t)$, followed by an application of the standard arbitrage restriction[4] that $E_t(\mathrm{d}V(t)) = r(t)V(t)\mathrm{d}t$. Its solution generally requires the application of numerical methods, although it frequently is possible – by the Feynman–Kac theorem – to state the solution probabilistically, as an expectation. Consider, for instance, the important special case of a *risky zero-coupon bond* $B(t, T)$ which pays out \$1 at time T if no default takes place before time T, 0 otherwise. In other words, the PDE boundary condition is $B(T, T) = 1$ and the recovery rate in default is 0. From the Feynman–Kac theorem, the time 0 solution of this PDE is simply

[2] Further, we need to interpret (1) as holding only up to and including the time of default. After default the stock ceases to exist and we cannot allow the stock to continue diffusing and jumping. One option is to model the post-default stock as a cash certificate on a deposit of $R_S S(\tau -)$, ie, $\mathrm{d}S(t) = r(t)S(t)\mathrm{d}t,\ t > \tau$.

[3] For the case where there is fractional recovery on the stock in default, the term $\lambda(t, S)S \cdot \partial V/\partial S$ in equation (2) must be replaced with $(1 - R_S)\lambda(t, S)S \cdot \partial V/\partial S$, where R_S is the recovery fraction.

[4] This expresses that claim values deflated by the money-market account are martingales in the risk-neutral measure. To avoid confusion, we should point out that this is a fundamental no-arbitrage condition and does not rely on any assumptions about, say, default risk being perfectly diversifiable. As always, the parameters in the process specification are assumed to contain the necessary risk-premia. Later, calibration to market observable prices will ensure that this assumption is justified.

$$B(0,T) = E\left(\mathrm{e}^{-\int_0^T r(u)\mathrm{d}u} 1_{\tau>T}\right) = E\left(\mathrm{e}^{-\int_0^T [r(u)\mathrm{d}u+\lambda(u,S(u))]\mathrm{d}u}\right)$$
$$= P(0,T)E\left(\mathrm{e}^{-\int_0^T \lambda(u,S(u))\mathrm{d}u}\right) \tag{3}$$

where we have defined (deterministic) default-free zero-cupon bond prices as $P(t, T) = \exp\left(-\int_t^T r(u)\mathrm{d}u\right)$, and where 1_A denotes the indicator function for the event A. As shown in Appendix A, the prices of coupon bonds and credit default swaps can be stated in terms of risky zero-coupon bonds, and vice versa, making risky zero-coupon bonds an obvious and convenient target for model calibration to non-convertible debt markets.

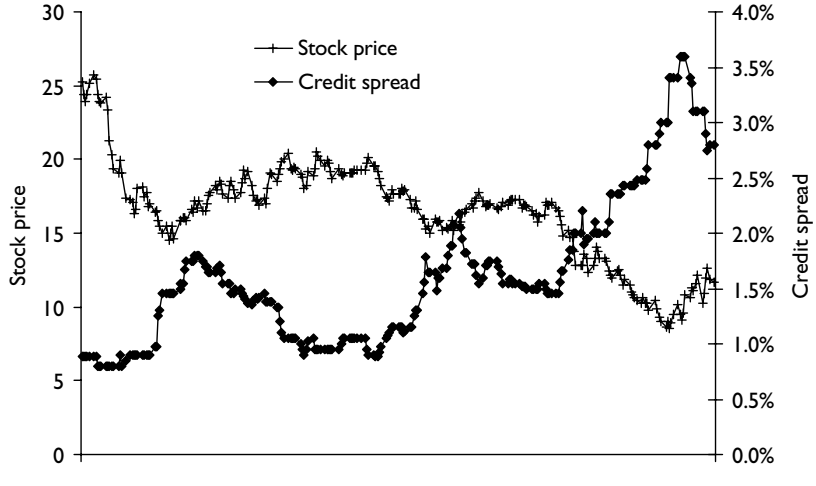
2.2 Intensity process and specification

An application of Itô's lemma results in the following pre-default ($t < \tau$) dynamics of the intensity process λ (suppressing dependency on S):

$$\mathrm{d}\lambda(t) = \frac{\partial\lambda}{\partial t}\mathrm{d}t + \frac{\partial\lambda}{\partial S}S(t)\big(r(t) - q(t) + \lambda(t)\big)\mathrm{d}t$$
$$+ \frac{1}{2}\frac{\partial^2\lambda}{\partial S^2}S(t)^2\sigma(t,S)^2\mathrm{d}t + S(t)\frac{\partial\lambda}{\partial S}\sigma(t,S)\mathrm{d}W(t), \quad t < \tau \tag{4}$$

As the real-world example in Figure 1 demonstrates, we would normally expect the term multiplying $\mathrm{d}W(t)$ to be negative (indicating perfect inverse correlation

FIGURE 1 Stock prices and credit spreads for Deutsche Telekom.



Daily data from Bloomberg covering period July 2001 to July 2002. The "Credit spread" is the five-year par credit default swap spread.

to the stock) as any reasonable model would have $\partial\lambda/\partial S \le 0$ as a reflection of the fact that companies with high stock prices are less likely to default than are those with low stock valuations. Specific parameterizations suggested in the literature include:

$$\lambda(t,S) = a + \frac{b}{S^p}$$
$$\lambda(t,S) = c - d\ln S$$
$$\lambda(t,S) = e + f\exp(-gS)$$

where $a, b, \ldots, g$ and p are constants. The first specification can be found in Takahashi *et al* (2001) and Davis and Lischka (1999), among others. The second parameterization is discussed in Bloch and Miralles (2002), and the third in Arvanitis and Gregory (2001).

While most of the methods developed in this paper are non-parametric and independent of the particular specification of $\lambda(t, S)$, for many of our numerical experiments we will use the first of these specifications with $a = 0$. Setting $a = 0$ is natural as it implies reasonable asymptotic behavior: $\lim_{S\to\infty}\lambda = 0$ and $\lim_{S\downarrow 0}\lambda = \infty$. Moreover, for this specification the dynamics of λ are particularly straightforward:

$$\frac{\mathrm{d}\lambda(t)}{\lambda(t)} = p\left[-\left(r(t) - q(t) + \lambda(t)\right) + \frac{1}{2}(p+1)\sigma\left(t,S(t)\right)^2\right]\mathrm{d}t$$
$$-p\sigma(t,S)\mathrm{d}W(t), \quad t < \tau \tag{5}$$

In other words, the log-normal volatility of λ is just the equity volatility scaled with a factor of p. The interpretation of p representing the ratio of equity and spread volatilities makes for particularly convenient estimation of this parameter. In a study on Japanese companies, Muromachi (1999) estimates p to be in the range 1.2 to 2.0; this appears reasonable and consistent with the fact that short-term credit spreads are typically more volatile than stock prices. For reference, a log-log regression of the sample data in Figure 1 returns a p-value of 1.56 (with an R-square of 0.8).

Finally, we notice that the dynamics (5) imply a certain amount of autocorrelation, with the drift of λ involving reversion at speed p around a level of $q(t) - r(t) + \frac{1}{2}(p+1)\sigma(t, S(t))^2$. Appendix C takes a closer look at the long-term properties of (5) and demonstrates that sometimes a stationary distribution exists. In particular, for constant process parameters r, q, and σ, the Appendix shows that $\lim_{t\to\infty}E(\lambda(t)) = \max(0, \frac{1}{2}\sigma^2 - r + q)$.

2.3 Hedging

A brief word on hedging in the model above. With two sources of uncertainty (W and N), hedging of contingent claims will involve taking positions in cash and two traded stock-dependent derivatives. For instance, we could take a position in a corporate bond and the stock itself. To develop the specific hedge for

this example, let V be the value of the derivative to be hedged, and let H denote the price of the bond used in the hedge. Further let w_S and w_H be the hedge positions in stock and bond, respectively, and let Π be the portfolio of V and its hedges. With the recovery rate on the bond being a constant R_H, the evolution of this portfolio is

$$\begin{aligned}\Pi(t) &= w_H H(t) + w_S S(t) + V(t) + \text{cash} \Rightarrow \\ \mathrm{d}\Pi(t) &= \ldots \mathrm{d}t + \left[w_H \frac{\partial H(t)}{\partial S} + w_S + \frac{\partial V(t)}{\partial S} \right] S(t)\sigma(t, S(t))\mathrm{d}W(t) \\ &\quad - \left[w_S S(t) + w_H \big(H(t) - R_H\big) + \big(V(t) - R_V(t)\big) \right] \mathrm{d}N(t)\end{aligned}$$

where for simplicity we have omitted the somewhat cumbersome $t-$ notation. For the hedge to work, the terms in the square brackets must be 0, leading to the following explicit expression for the hedge at time t:

$$w_H = \frac{R_V(t) - V(t) + \dfrac{\partial V(t)}{\partial S} S(t)}{H - R_H - \dfrac{\partial H(t)}{\partial S} S(t)}; \quad w_S = -\frac{\partial V(t)}{\partial S} - w_H \frac{\partial H(t)}{\partial S}$$

As is always the case, we note that the efficiency of the hedge depends on the realism of the underlying model. For instance, if the stock default jump is not truly to zero, a cashflow will be generated at the time of default. Similarly, the hedge does not protect against stochastic movements in interest rates and stock volatility. The latter issue would in practice be mitigated by adding volatility ("vega") and interest rate hedges to the portfolio.

3 Numerical implementation

3.1 Finite-difference scheme

We now turn to the solution of equation (2) by finite-difference methods. To this end, we introduce $z = \ln S$, set $v(t, z) = V(t, S)$, $R_v(t, z) = R_V(t, S)$ and rewrite (2) as

$$\frac{\partial v}{\partial t} + Lv = -\lambda(t, \mathrm{e}^z) R_V(t, z) \tag{6}$$

where L is the operator

$$\begin{aligned}L &= \left(r(t) - q(t) + \lambda(t, \mathrm{e}^z) - \frac{1}{2}\sigma(t, \mathrm{e}^z)^2 \right) \frac{\partial}{\partial z} \\ &\quad + \frac{1}{2}\sigma(t, \mathrm{e}^z)^2 \frac{\partial^2}{\partial z^2} - \left(r(t) + \lambda(t, \mathrm{e}^z) \right)\end{aligned}$$

Discretizing z-space into buckets of size Δz, we can approximate L by the finite-

difference operator (dropping t and z dependence for brevity)

$$\hat{L} = \left(r - q + \lambda - \tfrac{1}{2}\sigma^2\right)\delta_z + \tfrac{1}{2}\sigma^2\delta_{zz} - (r + \lambda)$$

where δ_z and δ_{zz} are the usual first- and second-order finite-difference operators,

$$\delta_z f(z) = \frac{1}{2\Delta z}\left[f(z + \Delta z) - f(z - \Delta z)\right],$$

$$\delta_{zz} f(z) = \frac{1}{(\Delta z)^2}\left[f(z + \Delta z) - 2f(z) + f(z - \Delta z)\right]$$

We then introduce a time grid $0 = t_0 < t_1 < \ldots < t_n$ with $\Delta t_i \equiv t_{i+1} - t_i$ and employ a modified theta discretization of the PDE (6) in the time-domain:

$$\left(\Delta t_i^{-1} - \theta\hat{L}\right)v(t_i, z) = \left(\Delta t_i^{-1} + (1-\theta)\hat{L}\right)v(t_{i+1}, z) + \theta\lambda(t_i, e^z)R_v(t_i, z) + (1-\theta)\lambda(t_{i+1}, e^z)R_v(t_{i+1}, z) \quad (7)$$

In general, (7) results in a series of tri-diagonal matrix equations and is stable for $\theta \geq ½$. For $\theta = ½$ (the *Crank–Nicolson method*), the precision of the scheme is at its maximum ($O(\Delta t^2 + \Delta z^2)$), making this the preferred choice for most smooth payoff functions. With n time-steps and m z-steps, the total computational effort is $O(mn)$ for all values of θ.

3.2 Numerical example: pricing European call options and risky bonds

For later use and as an illustration of the scheme above, consider now the pricing of a call option C with time T payout of $(S(T) - K)^+$ (recall the notation $x^+ = \max(x, 0)$). In case of default, the stock price drops to 0 and the call becomes worthless, ie, its recovery value is zero[5] and $R_C(t, S)$ for all t and S. When we state our computed call option prices, we follow the market convention of quoting European option prices in terms of their Black–Scholes *implied volatilities* $\sigma_{\text{imp}}(t; T, K)$, defined as the solution to the equation

$$C(t) = S(t)e^{-\int_t^T q(u)\,du}\Phi(d_+) - Ke^{-\int_t^T r(u)\,du}\Phi(d_-) \quad (8)$$

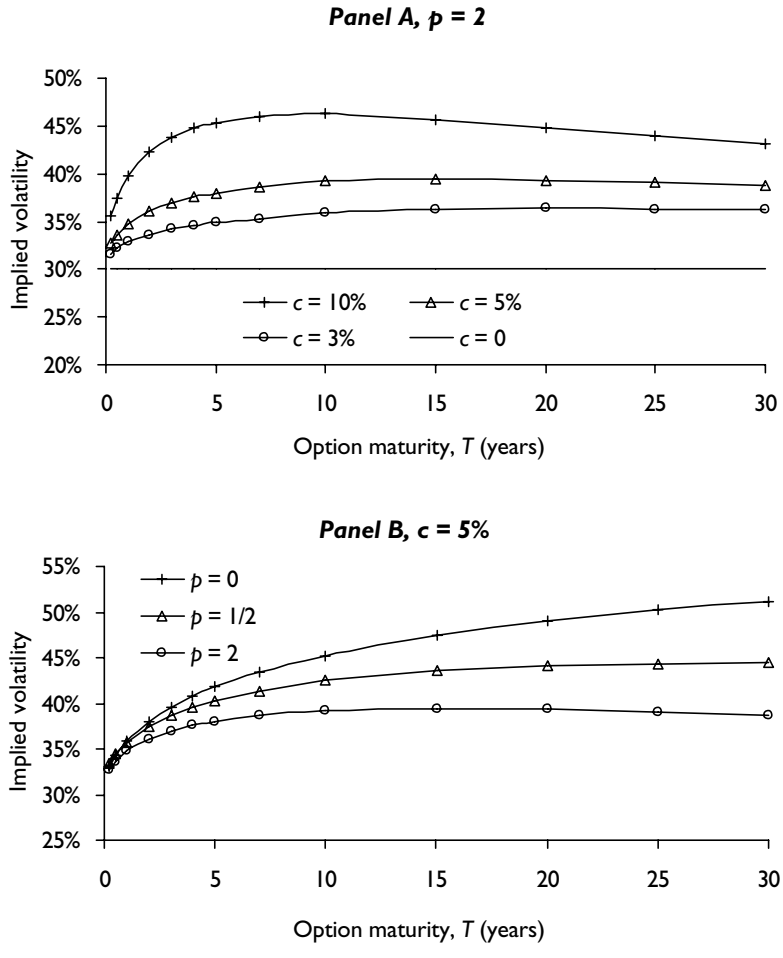
$$d_\pm = \frac{\ln\left(S(t)/K\right) + \int_t^T \left[r(u) - q(u)\right]du \pm \tfrac{1}{2}\sigma_{\text{imp}}^2(T - t)}{\sigma_{\text{imp}}\sqrt{T - t}}$$

where the left-hand side of (8) is observed in the market. Notice that implied volatilities are always quoted using default-free discounting.

Setting $\lambda(t, S) = c \cdot (S/S(0))^{-p}$ and using a constant diffusion volatility of $\sigma(t, S) = 30\%$, Figure 2 shows the term structure of at-the-money (ATM)

[5] In contrast, for a put option with time T payout $(K - S(T))^+$, the recovery value would be $R(t, S) = K\exp\left(-\int_t^T r(u)du\right)$.

FIGURE 2 Term structures of ATM implied volatilities vs intensity specification.

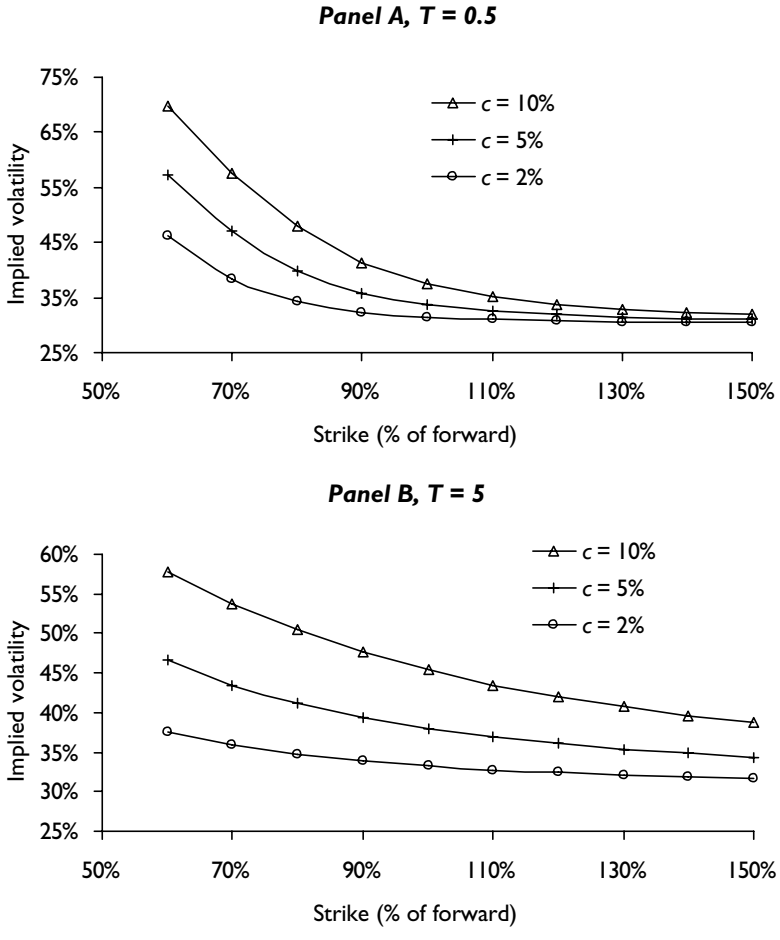


The graphs show term structures of implied volatility for options struck at the stock forward price. The model specification is as follows: $r(t) = 4\%$, $q(t) = 2\%$, $S(0) = 50$, $\sigma(t, S) = 30\%$, and $\lambda(t, S) = c \cdot (S/50)^{-p}$, with c and p varying as noted in panels A and B. All numbers in the graphs are based on a Crank–Nicolson finite-difference grid with 150 spatial steps. For $T < 1$, we used daily time-steps; for $T \in [1, 7]$ weekly time-steps; and for $T \in (7, 30]$ monthly time-steps.

implied volatilities (from (8)) for different values of c and p. A note: unless otherwise indicated, we use the term "at-the-money" for options with strikes set at the forward value ("at-the-money forward") rather than at the spot price ("at-the-money spot").

Raising c causes an increase in the variance rate of the Cox process $N(t)$ which is proportional to λ (specifically, $E([\mathrm{d}N(t)]^2) = E(\mathrm{d}N(t)) = \lambda(t, S)\mathrm{d}t$). As implied Black–Scholes volatility is an aggregate of diffusion and jump volati-

FIGURE 3 Six-month and five-year implied volatility skews vs intensity parameter *c*.

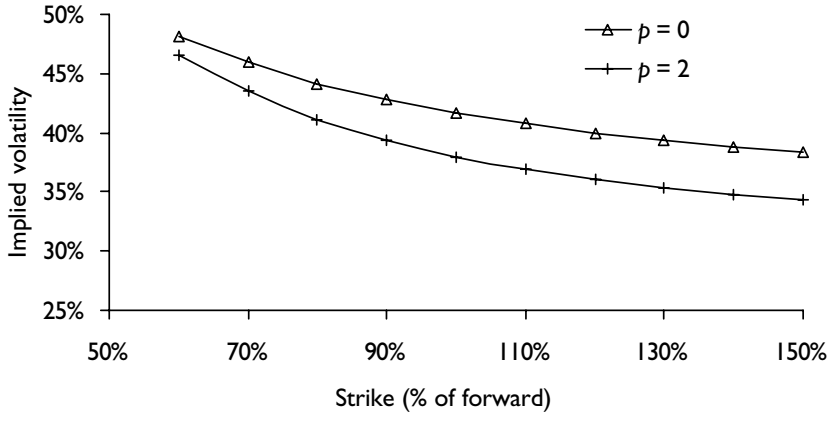


The graphs show implied volatility skews for six-month (panel A) and five-year (panel B) call options. All strikes are reported as percentages of the forward stock price. The model specification is as follows: $r(t) = 4\%$, $q(t) = 2\%$, $S(0) = 50$, $\sigma(t, S) = 30\%$, and $\lambda(t, S) = c \cdot (S/50)^{-2}$, with c varying as reported in the graphs. Finite-difference grid dimensions were as in Figure 2.

lity, an increase in c causes an increase in implied at-the-money volatility, as evidenced in Figure 2, panel A. The volatility increase is typically an increasing function of maturity for short- to medium-dated options, yet can be very substantial even for three-months options.[6] Figure 2, panel B, shows that implied at-the-money volatilities fall when the power p is increased. This effect is a

[6] The first point on the various graphs corresponds to a three-months maturity. For clarification, we point out that all graphs intersect the y-axis at a level of 30%.

FIGURE 4 Five-year implied volatility skew vs intensity parameter p.



The graph shows implied volatility skews for six-month call options. All strikes are reported as percentages of the forward stock price. The model specification is as follows: $r(t) = 4\%$, $q(t) = 2\%$, $S(0) = 50$, $\sigma(t, S) = 30\%$, and $\lambda(t, S) = 5\% \cdot (S/50)^{-p}$, with p varying as specified in the graph. All numbers were generated in a 150×150 finite-difference grid.

consequence of the fact that when p is increased, defaults at high stock prices become increasingly unlikely. As jumps associated with defaults from high stock price levels correspond to a large effective variance, implied volatility will decrease when p is increased, *ceteris paribus*.

In Figure 3, we fix the maturity and now consider the effect of default on implied volatilities at different option strikes (the so-called *volatility skew*). Adding default jumps to a diffusion process will necessarily fatten the lower tail of the distribution, raising prices of low-struck options relative to the pure diffusion setting; this effect is evident in Figure 3. Not surprisingly, the steepness of the jump-induced part of the volatility skew increases in c. Comparison of panels A and B in Figure 3 also demonstrate that the effect of default on the volatility skew decreases with option maturity, a typical characteristic of jump-induced skews.[7]

Finally, to get a feel for the impact on the skew of the stock-dependency of the jump intensity, Figure 4 graphs the one-year volatility skew for various values of the power p. While increasing the value of p here steepens the smile somewhat, the effect is comparatively mild for the parameters and strike range considered.

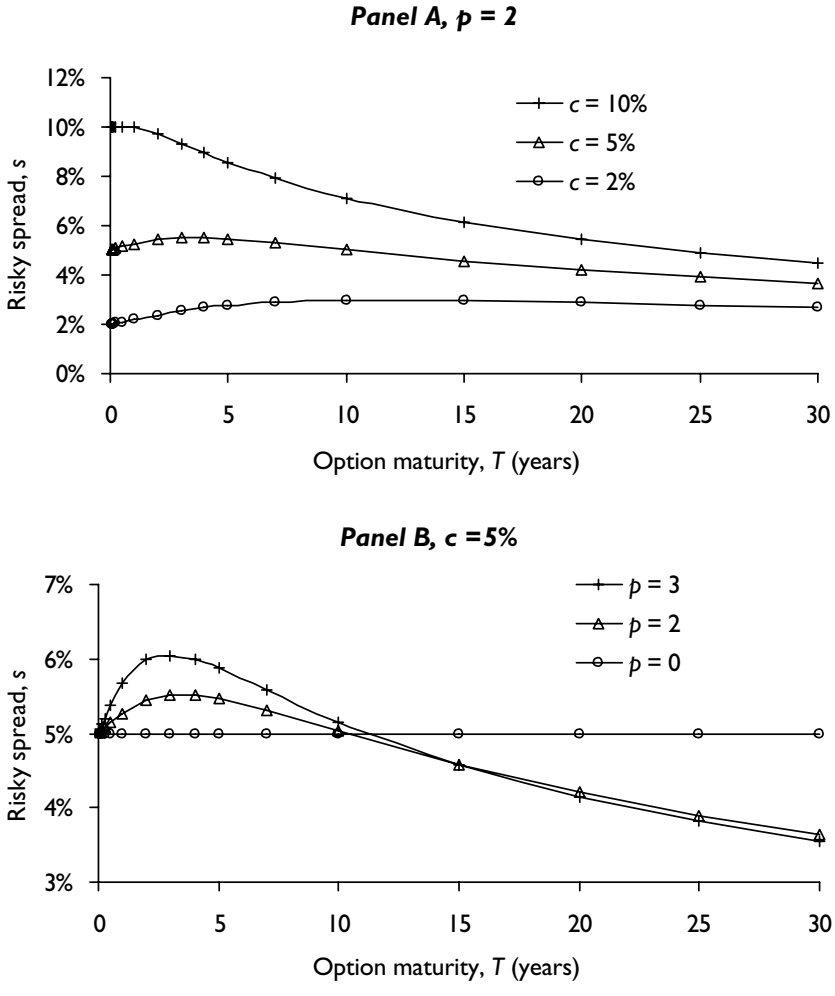
Having examined call options, we turn to the numerical pricing of risky bonds. Rather than report directly the bond prices $B(0, T)$ computed in our finite-difference grid, we instead prefer to use the concept of a *risky term spread* $s(T)$, defined through

[7] By letting the function $\sigma(t, S)$ depend explicitly on S, a diffusion can itself generate a volatility skew or smile. We will return to this issue in Section 6.

$$e^{-s(T)T} = E\left(e^{-\int_0^T \lambda(u,S(u))\,du}\right) = \frac{B(0,T)}{P(0,T)} \tag{9}$$

We note that $s(T)$ is by definition associated with an assumption of zero recovery. For bonds with a recovery rate of R, the quantity $s(T)(1-R)$ is roughly equal to the bond credit spread (see Duffie and Singleton, 1999). Figure 5 graphs the risky term spread as a function of T, for various scenarios for p and c.

FIGURE 5 Term structures of risky spreads vs intensity specification.



The graphs show term structures of spreads on risky bonds, as defined in equation (9). The model specification is as follows: $r(t) = 4\%$, $q(t) = 2\%$, $S(0) = 50$, $\sigma(t, S) = 30\%$, and $\lambda(t, S) = c \cdot (S/50)^{-p}$, with c and p varying as noted in panels A and B. Finite-difference grid dimensions were as in Figure 2.

Broadly speaking, all figures show risky spreads that initially increase in maturity, but ultimately start falling. The former effect is caused by the fact that the local drift of $\lambda(t, S)$ is typically positive for small t. Indeed, from the term multiplying dt in (5) we see that for small t the drift of $\lambda(t, S) = c(S/S(0))^{-p}$ becomes approximately $c(\frac{1}{2}p(p+1)\sigma^2 - p(r-q+c))$; this quantity is positive for all cases in Figure 5 (although barely so for the case $c = 10\%$ in Panel A). Over longer time horizons, mean reversion (see discussion at the end of Section 2.2) will slow down the growth of the expectation of $\lambda(t, S)$ and eventually pull it towards a long-term stationary level of $\max(0, \frac{1}{2}\sigma^2 - r + q)$ which here amounts to 2.5%. Convexity effects also contribute to risky spreads ultimately falling, as follows from Jensen's inequality

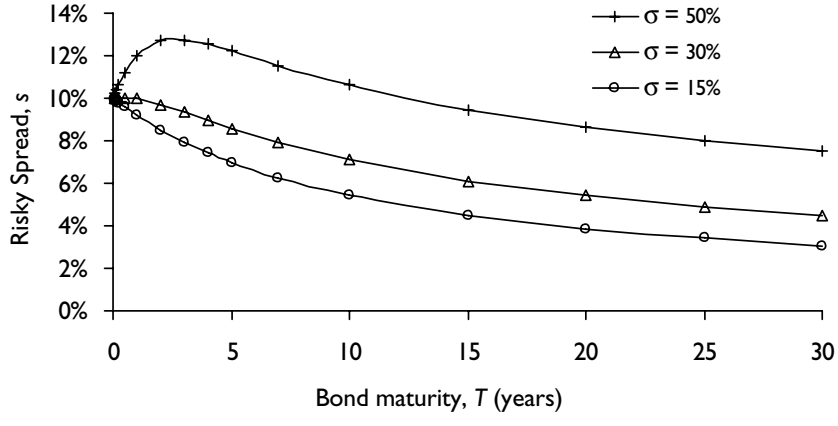
$$e^{-s(T)T} > e^{-\int_0^T E[\lambda(u,S(u))]\,du} \Rightarrow s(T) < T^{-1}\int_0^T E\big[\lambda\big(u,S(u)\big)\big]\,du$$

ie, the risky spreads will be lower than the average intensity, with the discrepancy between the two increasing in T. Increasing volatility will increase the hump in the spread curve, and lowering it will eventually remove the hump altogether; see Figure 6 for an example.

4 Calibration to risky bonds and at-the-money options

In Section 2.4, we considered, among other things, the pricing of risky zero-coupon bonds and at-the-money call options in the model (1). Among the conclusions we can draw from our numerical study is that both implied option volatilities and risky credit spreads depend in a complicated way on maturity and

FIGURE 6 Term structures of risky spreads vs stock volatility.



The graph shows term structures of spreads on risky bonds, as defined in equation (9). The model specification is as follows: $r(t) = 4\%$, $q(t) = 2\%$, $S(0) = 50$, $\lambda(t, S) = 10\% \cdot (S/50)^{-p}$, and $\sigma(t, S) = \sigma$, with σ varying as shown in the graph. Finite-difference grid dimensions were as in Figure 2.

the joint parameterization of $\lambda(t, S)$ and $\sigma(t, S)$. Indeed, even a straightforward parameterization using constant volatility and the time-homogenous intensity $\lambda(t, S) = c(S/S(0))^p$ can give rise to highly non-flat, non-monotonic term structures of implied volatility and credit spreads. In practice, it is extremely unlikely that these term structures will even remotely resemble those observed in the market. In particular, we stress that the practice of simply importing into the model (1) a volatility function $\sigma(t, S)$ maintained on a "usual", default-free equity option system is highly inappropriate as the resulting model is likely to severely overstate both at-the-money volatilities and the steepness of the volatility skew.

In this section, we will dispense with naïve time-homogenous model parameterizations and discuss schemes to explicitly bring the convertible bond model into calibration with risky zero-coupon bonds and European at-the-money call options. We will do so by introducing time-dependent functions into both $\lambda(t, S)$ and $\sigma(t, S)$. Section 6 will discuss generalizations of the procedure that allow for fitting to an entire strike-maturity surface of option prices.

4.1 Fokker–Planck equation

For numerical efficiency, we wish to base our calibration technique on a forward induction technique.[8] For this, we introduce the concept of a (log) state price density $p(t, z, s, y)$ as the time t price of delivery of a Dirac amount $\delta(\ln S(s) - y)$ at time $s > t$, given that $\ln S(t) = z$. The density solves the usual backward Kolmogorov equation (compare with (6))

$$\frac{\partial p}{\partial t} + \left(r(t) - q(t) + \lambda(t, \mathrm{e}^z) - \frac{1}{2}\sigma(t, \mathrm{e}^z)^2\right)\frac{\partial p}{\partial z} + \frac{1}{2}\sigma(t, \mathrm{e}^z)^2\,\frac{\partial^2 p}{\partial z^2}$$

$$= \left(r(t) + \lambda(t, \mathrm{e}^z)\right)p \tag{10}$$

where s and z are considered fixed. The boundary condition is $p(s, z, s, y) = \delta(y - z)$. Notice that we assume that p recovers nothing in default and consequently associate with p a defect transition density that excludes the singularity at $S = 0$. The formal adjoint to (10) is the *Fokker–Planck* (or *forward Kolmogorov*) equation

$$-\frac{\partial p}{\partial s} - \frac{\partial}{\partial y}\left(\left(r(s) - q(s) + \lambda(s, \mathrm{e}^y) - \frac{1}{2}\sigma(s, \mathrm{e}^y)^2\right)p\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(\sigma(s, \mathrm{e}^y)^2\,p\right)$$

$$= \left(r(s) + \lambda(s, \mathrm{e}^y)\right)p$$

[8] The brute-force alternative – embedding repeated applications of the backward equation in a calibration loop – is typically too slow for real-time systems. In case one wishes to calibrate to more complicated instruments such as American options or even convertible bonds, it will, however, typically be necessary to use the backward equation. In such cases employment of parallel computing platforms or other high-performance systems may still allow for acceptable computing times.

where now t and z are considered fixed. The boundary condition is here $p(t, z, t, y) = \delta(y - z)$. Assuming that σ and λ are twice and once differentiable in S, respectively, we can rewrite this equation as

$$-\frac{\partial p}{\partial s} - H(s,y)\frac{\partial p}{\partial y} + \frac{1}{2}\sigma(s,\mathrm{e}^y)^2\frac{\partial^2 p}{\partial y^2} = G(s,y)p;$$

$$H(s,y) = r(s) - q(s) + \lambda(s,\mathrm{e}^y) - \frac{1}{2}\sigma(s,\mathrm{e}^y)^2 - 2\mathrm{e}^y\sigma(s,\mathrm{e}^y)\sigma'(s,\mathrm{e}^y);$$

$$G(s,y) = r(s) + \lambda(s,\mathrm{e}^y) + \mathrm{e}^y\lambda'(s,\mathrm{e}^y)$$
$$-\mathrm{e}^y\left(\sigma(s,\mathrm{e}^y)\left[2\sigma'(s,\mathrm{e}^y) + \mathrm{e}^y\sigma''(s,\mathrm{e}^y)\right] + \mathrm{e}^y\sigma'(s,\mathrm{e}^y)^2\right) \qquad (11)$$

where $\lambda' \equiv \partial\lambda/\partial S$ and $\sigma' \equiv \partial\sigma/\partial S$, $\sigma'' \equiv \partial^2\sigma/\partial S^2$.

4.2 Calibration scheme

Now we specialize to the case where $\lambda(t, S) = \lambda(S; a(t))$ and $\sigma(t, S) = \sigma(S; b(t))$, for time-dependent scalar functions a and b to be set. Assuming that r and q are known, we wish to set the two unknown functions a and b such that we simultaneously match at all maturities (a) risky zero coupon bond prices $B(0, T)$; and (b) European at-the-money equity option prices. For the latter, we introduce the notation

$$C(T,K) = \mathrm{e}^{-\int_0^T r(u)\,\mathrm{d}u} E\big(S(T) - K\big)^+$$

That is, $C(T, K)$ is the time 0 price of a T-maturity call struck at K. In terms of the state price densities p, we can write

$$C(T,K) = \int_{\ln K}^{\infty} p(T,y)(e^y - K)\,\mathrm{d}y; \quad B(0,T) = \int_{-\infty}^{\infty} p(T,y)\,\mathrm{d}y \qquad (12)$$

where we use the abbreviated notation $p(T, y) = p(0, \ln S(0), T, y)$.

The idea is now to solve (11) forward, at each step finding $b(T)$ and $a(T)$ such that (12) is satisfied for a chosen value of the strike K. Numerically, this requires a discretization of (11), for instance by finite-difference methods.[9] Following the steps that lead to (7), we introduce a time line $0 = T_0 < T_1 < , \ldots, T_N$ and discretize y-space into buckets of size Δy. On the time interval $[T_i, T_{i+1}]$ we write

[9] As an alternative to a direct finite-difference discretization of the Fokker–Planck PDE, we can instead discretize the backward equation and then invert the resulting matrix equations to generate an alternative discrete-time version of the Fokker–Planck equation. See Andersen and Brotherton-Ratcliffe (1998) for this.

$$\left(\Delta T_i^{-1} - \theta\hat{D}\right)p\left(T_{i+1}, y\right) = \left(\Delta T_i^{-1} + (1-\theta)\hat{D}\right)p\left(T_i, y\right) \tag{13}$$

where we have introduced the finite-difference operator $\hat{D} \equiv -H\delta_y + \frac{1}{2}\sigma^2\delta_{yy} - G$. With y discretized on the grid $\{y_0 + j\Delta y\}_{j=0,1,\ldots,M}$, (12) can be written as sums:

$$C\left(T_i, K_i\right) \approx \Delta y \sum_{y_0 + j\Delta y > \ln K_i} p\left(T_i, y_0 + j\Delta y\right)\left(\mathrm{e}^{y_0 + j\Delta y} - K_i\right);$$

$$B(0, T_i) \approx \Delta y \sum_{j \geq 0} p\left(T_i, y_0 + j\Delta y\right) \tag{14}$$

where K_i is the calibration strike for time T_i, assumed here to be set equal to the forward value of the stock.

With (14) we are ready to state our iterative scheme for finding the functions a and b on the time line $\{T_i\}$:

0. Set $i = 0$ and $p(0, y) = 1_{y = \ln S(0)}/\Delta y$. Make a guess for $a(0)$ and $b(0)$.
1. Assuming that a and b are piecewise flat, solve (13) one time-step forward to find $p(T_{i+1}, y_0 + j\Delta y)$ for all j.
2. Compute the right-hand sides of the expressions for C and B in (14).
3. If the right-hand sides (14) equal the market-observable left-hand sides, store $a(T_i)$ and $b(T_i)$ and make guesses for $a(T_{i+1})$ and $b(T_{i+1})$ (for instance, we can set $a(T_{i+1}) = a(T_i)$ and $b(T_{i+1}) = b(T_i)$). Otherwise update $a(T_i)$ and $b(T_i)$ and go back to step 2.
4. Set $i = i + 1$ and go to step 1.

The joint root-search for $a(T_i)$ and $b(T_i)$ in Steps 2 and 3 can be done using any standard non-linear equation solver. While we re-emphasize that European option prices are sensitive to the default sensitivity, $C(T_{i+1}, K_{i+1})$ is typically more sensitive to $a(T_i)$ than to $b(T_i)$, and vice versa for $B(0, T_{i+1})$. As such, one can typically simplify the two-dimensional root-search problem to an inexpensive Gauss–Seidel iteration over two simple one-dimensional root-searches, as in

(a) Freeze $a(T_i)$ and find $b(T_i)$ by root-search such that $C(T_{i+1}, K_{i+1})$ is fit to market.
(b) Freeze $b(T_i)$ and find $a(T_i)$ by root-search such that $B(0, T_{i+1})$ is fit to market.
(c) Repeat until both $C(T_{i+1}, K_{i+1})$ and $B(0, T_{i+1})$ are jointly fit to market.

For each of the one-dimensional searches, we can apply a simple root-finder such as Newton–Raphson, secant search, or similar. On average, we find that we only need to repeat (a) and (b) two or three times, making the algorithm fast.

Let us comment on a few practical details of the proposed calibration routines. First, we notice that while we implicitly assumed that calibration strikes

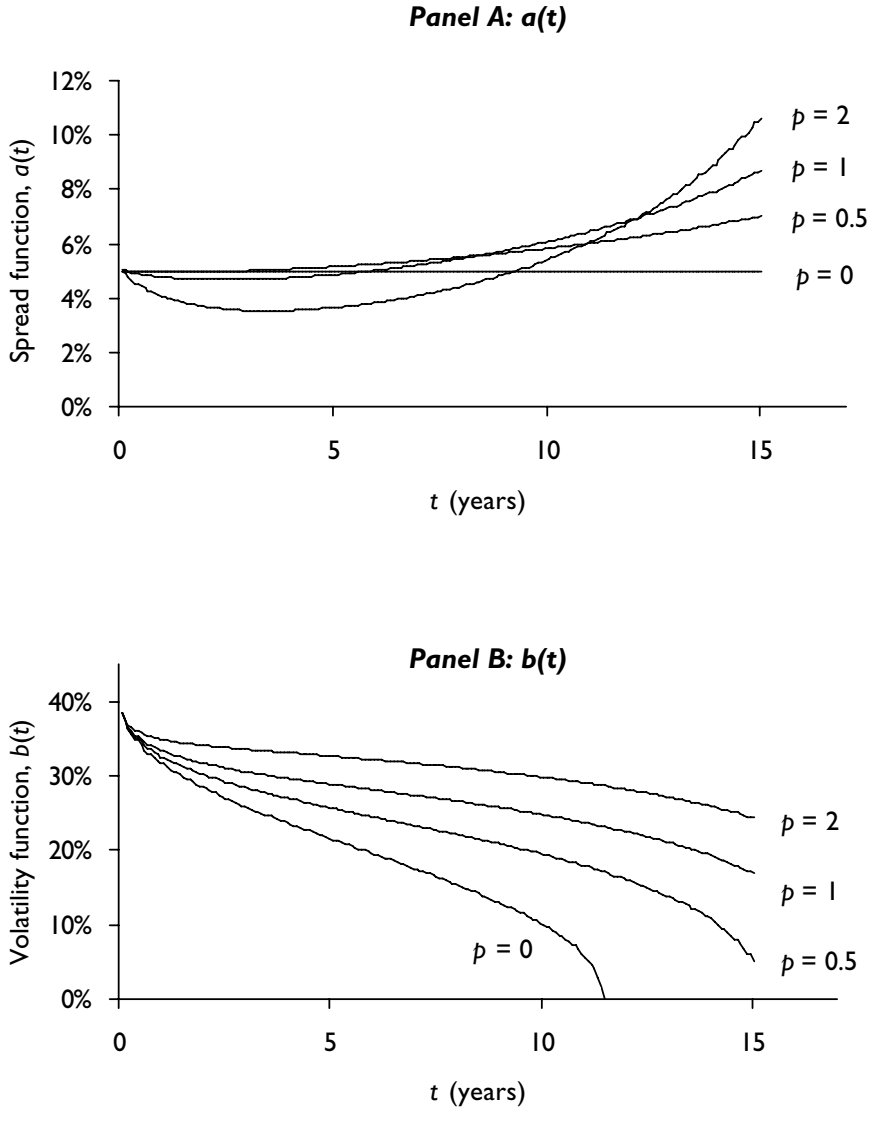
on the call options were set at the ATM point, nothing prevents us from picking other levels, as long as we restrict ourselves to a single calibration strike per time-step (see Section 6 for extensions). Second, we can work with a finer discretization in the finite-difference grid than in the fitting of C and B prices. That is, steps 2 and 3 need not be performed on all points in the finite-difference time grid, but only on a subset. Third, while not of great practical importance we can, if desired, easily relax the assumption of piecewise flat parameters through the introduction of more sophisticated interpolation schemes. Fourth, we point out that the discontinuous boundary condition of (13) in Step (1) can lead to undesirable oscillations of the solution if the discretization parameter θ is too low. Such oscillations can, however, easily be eliminated by using $\theta = 1$ (*fully implicit scheme*) for the first few time-steps, at which point we can switch to the better-converging Crank–Nicolson scheme where $\theta = \frac{1}{2}$. For details on this so-called *Rannacher stepping* technique, see Pooley *et al* (2001). Finally, we notice that it is often beneficial to make a continuity adjustment of the first term in the sum for $C(T, K)$ in (14) to match the proper area under the call payout ramp function; see eg, Tavella and Randall (2000), p. 111 for details on this. Alternatively, we can adjust the grid geometry such that the option strikes all fall precisely halfway between grid points in y-space.

4.3 Numerical example

Assume now that we have estimated from straight debt and/or credit derivatives that the risky spread term structure is flat at 500 basis points. Further, from equity option markets we observe that implied at-the-money volatilities are flat at 40% for all maturities. Assuming that both credit spreads and equity prices have deterministic volatilities, we set $\sigma(t, S) = b(t)$ and $\lambda(t, S) = a(t)(S(0)/S)^p$ where we assume that p has been estimated by time-series data and/or observations of the ratio of credit spread volatilities to equity volatilities (see Section 2.2). Estimation of the deterministic functions $a(t)$ and $b(t)$ can now be done using the algorithm from Section 4.2; the results are shown in Figure 7.

The function $a(t)$ calibrates to a U-shaped function of time that compensates for the humped credit spread term structures associated with constant $a(t)$ (see Figure 5). The larger the value of p, the more pronounced the U-shape of $a(t)$ becomes, consistent with previous results. In the same vein, the volatility function $b(t)$ becomes downward-sloping to compensate for the volatility effect of the default jump. As we would expect, the smaller the value of p, the more pronounced this effect becomes. We notice that for $p = 0$ the downward slope of b is so severe that it becomes impossible to calibrate the model for maturities beyond 12 years, a reflection of the fact that the intrinsic stock volatility induced solely by the Poisson process exceeds the target implied volatility used in our example. We note that such calibration problems are fairly uncommon in practice as they involve the somewhat unlikely combination of low long-term implied spread volatilities and high credit spreads.

FIGURE 7 Result of calibration to risky bonds and ATM options.



The graphs show the calibrated functions a, b in the model $\lambda(t, S) = a(t) \cdot (S/50)^{-p}$, $\sigma(t, S) = b(t)$. ATM call options (with strikes equal to the stock forward values) quote at 40% implied volatility for all maturities; the risky credit spread quotes at 5% for all maturities. The remaining parameters were as follows: $r(t) = 4\%$, $q(t) = 2\%$, $S(0) = 50$, with the parameter p of the λ-function varying as noted in the graphs. In the calibration algorithm, the fit was performed in increments of one month, with weekly points in the finite-difference grid time line. The spatial direction of the finite-difference grid contained 150 points.

5 Pricing of convertible bonds

For reference, this section focuses on the pricing of a few specific convertible bond contracts. We demonstrate the effect of calibration, and state the boundary conditions needed in the finite-difference solver discussed in Section 3.

5.1 Boundary conditions for convertible bonds

Fundamentally, convertible bonds are coupon-bearing instruments with an embedded option to exercise into a certain number of shares of the underlying company. Let $V(t)$ be the time t price of a T-maturity convertible bond paying an annualized coupon rate of κ on some schedule $\{t_i\}_{i=0,\ldots,K}$, $t_0 = 0$, $t_K = T$. The stream of cashflows introduce the following jump-type boundary conditions for V:

$$V(t_i-) = V(t_i+) + \kappa(t_i - t_{i-1}), \quad i = 1,2,\ldots,K-1, \quad V(T) = 1 + \kappa(T - t_{K-1}) \tag{15}$$

where we have assumed a (normalized) notional of \$1 and ignored the finer details of bond accrual conventions. On dates t where the bond can be exercised into shares, we impose the free boundary condition

$$V(t) \geq L(t)S(t) \tag{16}$$

where L is the possibly time-dependent conversion ratio (normalized to apply to a notional of \$1). Ayache *et al* (2002) show how to appropriately formulate as a linear complementary problem. In a finite-difference grid, the application of the boundary condition can be done a number of ways, the simplest of which is to treat it as a jump condition

$$V(t-) = \max\big(V(t+), L(t)S(t)\big)$$

See Forsyth and Vetzal (2002) for a more sophisticated (and better-converging) alternative to this approach.

Many convertible bonds also contain a periodic option for the issuer to buy back (or "call") the bond at some time-dependent level H, typically close to par. In case the issuer calls the bond, the investor has the option to convert the bond to shares, rather than receive the call amount H. That is,

$$V(t) \leq \max\big(H(t), L(t)S(t)\big) \tag{17}$$

on dates t for which the bond can be called. Typically, the issuer's right to call the bond early becomes active only after a certain period of time (the *non-call protection period*) has lapsed since the original issue date.[10] We point out that the free boundary condition (17) assumes rational exercise on the issuer's behalf,

[10] Like the conversion option, the issuer's call option can be exercised American-style (ie, continuously). In case exercise takes place in the middle of a coupon period, the bond owner is normally entitled to interest accrued up to the date of exercise. This can easily be incorporated into the boundary condition by modifying the call strike appropriately.

an assumption that appears to often be violated in practice (see, eg, the discussion in Ho and Pfeffer, 1996).

A third feature of many convertible bonds is an embedded right of the buyer to sell (or "put") the bond back to the issuer for some possibly time-dependent amount h. This option provides the buyer with protection against interest rate increases, as well as against deterioration in the financial health of the underlying company. The put option on a convertible bond normally involves only a limited number (two or three, say) of exercise opportunities during the life of the bond. On a date t in the put schedule, the free boundary condition is

$$V(t) \geq h(t) \tag{18}$$

For our purposes, the conditions (15)–(18) are sufficient to describe a convertible bond. In practice, it is not uncommon for convertible bonds to have additional features, some of which may be strongly path-dependent. See Olsen (2002) and Grimwood and Hodges (2002) for terminology and discussions of more advanced features in convertible bonds.

5.2 Recovery rate assumption

Consistent with market practice in bond and credit derivatives markets, we will here assume that the recovery value is a fixed fraction $0 \leq \phi \leq 1$ of the bond notional. The constant ϕ depends on the seniority of the bond and is typically somewhere around 30–50%. We point out that many papers on convertible bonds make the alternative assumption that convertible bonds and coupon bonds recover a fixed percentage of their pre-default value (*not* the notional); see, eg, Takahashi *et al* (2001), Bloch and Miralles (2001), and Ayache *et al* (2002). This recovery-of-value approach was first proposed by Duffie and Singleton (1999), but while it often allows for certain technical simplifications, it does not seem to be consistent with typical bankruptcy proceedings.

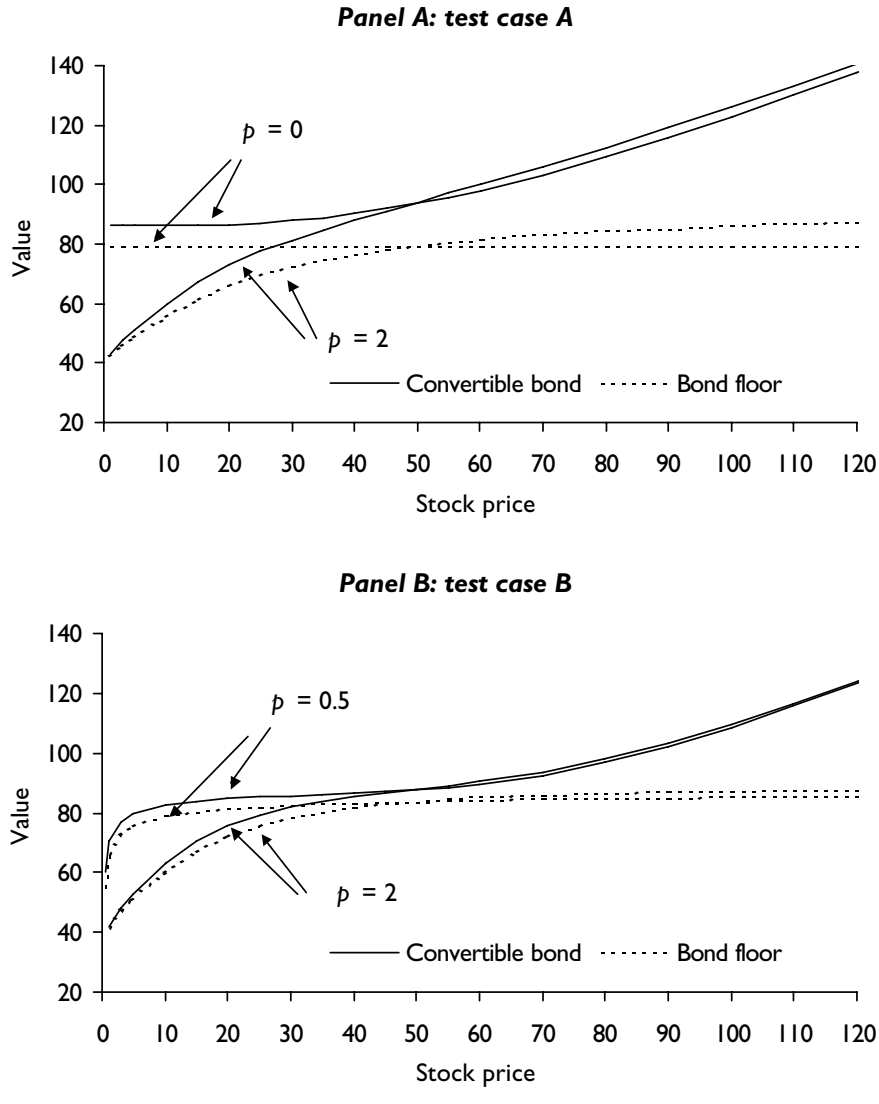
5.3 Numerical example.

We now turn to the pricing of some particular convertible bond. To stress our algorithms, we consider two cases: (a) a fairly long-dated contract with a significant non-call protection period and high equity volatility (case A); and (b) a medium-term contract with a relatively short non-call protection period and medium equity volatility (case B). Table 1 lists the specifics.

We assume that all observable yield, spread, and implied volatility term structures are flat. From the formulas in Appendix A, we compute that the pure fixed-income part of the convertible bond (the so-called *bond floor*) trades at \$79.5 at time 0 for case A, and at \$83.9 for case B. The conversion premia are thus \$29.5 and \$33.9, respectively. On the model side, we assume as in Section 4.3 that $\sigma(t, S) = b(t)$ and $\lambda(t, S) = a(t)(S(0)/S)^p$, with the deterministic parameters a and b calibrated as outlined in Section 4.

For test cases A and B, Figure 8 shows how the value of the convertible bond,

FIGURE 8 Prices of convertible bonds in test cases A and B.



The graphs show convertible bond prices for the test cases A (in panel A) and B (in panel B). See Table 1 for details about the contract and the market parameter. The parameter p of the λ-function varies as noted in the graphs. In the calibration algorithm, the fit was performed in increments of one month, with weekly points in the finite-difference grid time line. The spatial direction of the finite-difference grid contained 250 points. The same finite-difference grid was used for pricing and calibration.

as well as the bond floor, depend on the current stock price and on the parameter p in the function $\lambda(t, S)$. For $p > 0$ the model implies that both the bond floor and the convertible bond approach the recovery value of \$40 when the stock

TABLE 1 Market and contract data for convertible bond pricing example.

	Case A	Case B
Bond maturity, T	10 years	5 years
Bond coupon	3%	1.5%
Bond payments	Every 0.5 years	Every 0.5 years
Notional amount	100	100
Recovery value	40% of notional	40% of notional
Credit spread,[11] s	300 basis points	200 basis points
Conversion ratio, L	1	1
Equity price, $S(0)$	50	50
Equity implied volatility	40%	25%
Equity dividend rate, q	2%	2%
Interest rate, r	4%	4%
Call price, H	100	100
Call dates	Anytime after $t = 5$ years	Anytime after $t = 3$ years
Put price, h	100	100
Put date(s)	Years 6 and 8	Year 4

price goes to zero; this effect is sometimes referred to as the "collapsing bond floor". As a consequence, the convertible bond can locally be a concave function of stock ("negative gamma"), a situation that is not uncommon in practice for convertible bonds issued by companies in financial distress.

Let us now briefly turn to a comparison of the results above for the properly calibrated model against similar results for a naïvely calibrated model, that is, a model where the volatility function σ is assumed constant, and where $\lambda(t, S) = c(S(0)/S)^p$ for some constant c. Here, we set $\sigma = 40\%$ and $\sigma = 25\%$ for cases A and B, respectively, and set $c = s$, where s is the quoted spread[12] (300 and 200 basis points, respectively). From our earlier tests, we would expect both the effective volatility and credit spread of the "naïve" model to be too high. As these effects have opposing impact on the prices of convertible bond (the first effect increases its price, the second lowers it) we could perhaps hope that the naïve model would on average not do too poorly. To investigate this, Table 2 compares today's price of the convertible bond and the bond floor for various values of p; Figure 9 graphs the errors (as a percentage of the difference

[11] Recall that this is the spread in zero-recovery (intensity) terms. The credit spread observable in the market for coupon bonds or credit default swap would then be roughly $(1 - \phi) = 60\%$ of the numbers in the table (180 bp and 120 bp).

[12] An alternative is to set c such that bond floor is priced correctly at today's value of the stock price. However, for the cases tested in the paper this approach would result in convertible bond prices that would be significantly too high, as the increased volatility of the "naïve" model would not be countered by a decrease in the value of the bond floor. To avoid overstating our case, we do not list the pricing results of this approach (but see the case of $p = 0$ in Figure 8 to get a feel for the magnitude of the errors).

TABLE 2 Prices of convertible bonds and bond floors.

	Convertible bond				Bond floor		
p	Calibrated price	Naïve price	Price difference	Percentage difference*	Calibrated price	Naïve price	Percentage difference†
				Test case A			
2	94.1	90.5	−3.5	−24.4%	79.5	72.7	−8.5%
1	93.8	93.8	0.0	−0.2%	79.5	76.0	−4.4%
0.5	93.8	95.4	1.7	11.7%	79.5	78.1	−1.8%
0	93.7	96.6	2.9	20.5%	79.5	79.5	0.0%
				Test case B			
2	87.8	87.7	−0.1	−2.3%	83.9	83.1	−1.0%
1	87.7	88.3	0.6	14.6%	83.9	83.7	−0.2%
0.5	87.7	88.5	0.8	20.1%	83.9	83.9	0.0%
0	87.7	88.6	0.9	23.4%	83.9	83.9	0.0%

*Price difference as a percentage of price spread between convertible bond and bond floor.
†Price difference as a percentage of bond floor value.
The table shows today's ($S(0) = 50$) prices for convertible bonds and bond floors in test cases A and B. "Calibrated price" refers to prices computed by the model fully calibrated to call options and debt prices; "Naïve price" refers to prices computed by the simplified parameter estimation approach outlined in the main text.

between the convertible bond and the bond floor) of the convertible bond price as a function of the stock price.

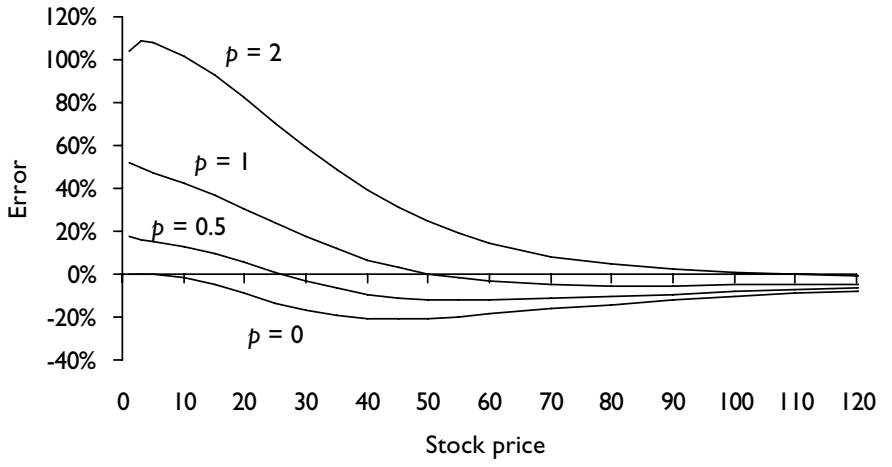
It is clear from Table 2 and Figure 9 that the convertible bond price errors introduced by a naïve calibration scheme can be highly significant, easily reaching 20–50% and more of the price spread between the bond floor and the convertible bond. The absolute price differences reach as much as \$3.9 and grow rapidly in maturity. For contracts with truly long maturities (20–30 years), we find that the absolute errors can easily reach \$5–10 on a \$100 bond.

6 Extensions

6.1 Stochastic interest rates

The model framework outlined so far assumes that risk-free interest rates are deterministic. As interest rate volatilities are generally much lower than equity volatilities, this assumption is generally defensible for the purpose of pricing plain-vanilla convertible bonds. Indeed, as stated in Grimwood and Hodges (2002), "… the stochastic modeling of the spot interest rate appears the least important model feature" in their investigations. See also Olsen (2002) and Brennan and Schwarz (1980) for similar comments. Inclusion of stochastic

FIGURE 9 Percentage price errors of "naïvely" calibrated model (case A).



The graph shows the convertible bond price error of the "naïvely" calibrated model vs the level of the stock price. Errors are represented as a percentage fraction of the excess value of the convertible bond relative to the bond floor. All market and model data are as in test case A (magnitude of the percentage error is similar for test case B); see Table 1 for details about the contract and the market parameter. The parameter p of the λ function varies as noted in the graphs. In the calibration algorithm, the fit was performed in increments of one month, with weekly points in the finite-difference grid time line. The spatial direction of the finite-difference grid contained 250 points. The same finite-difference grid was used for pricing and calibration.

interest rates into the model is, however, not conceptually difficult, as we shall now demonstrate. First, we assume that the interest rate dynamics satisfy the one-factor short-rate SDE

$$\mathrm{d}r(t) = \mu_r(t, r(t))\mathrm{d}t + \sigma_r(t, r(t))\mathrm{d}Z(t)$$

where μ_r, σ_r: $\mathbb{R}^2 \to \mathbb{R}$ are smooth well-behaved functions, and where $Z(t)$ is a Brownian motion under the risk-neutral probability measure. We assume that μ_r has been calibrated such that the model fits zero-coupon bond prices; that is,

$$P(0,T) = E\left(\mathrm{e}^{-\int_0^T r(u)\mathrm{d}u}\right)$$

Further, we assume that $Z(t)$ is correlated to the Brownian motion of the equity price, $\mathrm{d}Z(t) \cdot \mathrm{d}W(t) = \rho(t)\mathrm{d}t$, for some deterministic correlation function ρ (extensions to state-dependent correlation are trivial). The backward equation for a contingent claim $V = V(t, z, r)$, $z = \ln S$ then becomes

$$\frac{\partial V}{\partial t}+\left(r-q(t)+\lambda(t,\mathrm{e}^z)-\frac{1}{2}\sigma(t,\mathrm{e}^z)^2\right)\frac{\partial V}{\partial z}+\mu_r(t,r)\frac{\partial V}{\partial r}$$
$$+\frac{1}{2}\sigma(t,\mathrm{e}^z)^2\frac{\partial^2 V}{\partial z^2}+\frac{1}{2}\sigma_r(t,r)^2\frac{\partial^2 V}{\partial r^2}+\rho(t)\sigma_r(t,r)\sigma(t,\mathrm{e}^z)\frac{\partial^2 V}{\partial z\partial r}$$
$$=\left(r+\lambda(t,\mathrm{e}^z)\right)V-\lambda(t,\mathrm{e}^z)R_V(t,z,r) \tag{19}$$

where the recovery value default R_V can depend on both equity and interest rates. The numerical solution of this equation in a finite-difference grid is most easily accomplished by an alternating direction implicit (ADI) method; see, eg, Craig and Schneyd (1988) for a scheme that can handle mixed second-order derivatives. The computational order of such a scheme is $O(n_r n_S m)$, where m is the number of time steps, and n_r and n_S are the numbers of interest rate and equity steps, respectively. In practice n_r often needs to be around 50, making the two-factor model above significantly slower to evaluate than the one-factor model discussed previously (which as discussed earlier has computational order $O(n_S m)$).

To calibrate the model above, we again would turn to an iterative scheme relying on forward induction in a finite-difference scheme. From standard theory for adjoint operators, the necessary Fokker–Planck equation here becomes

$$-\frac{\partial p}{\partial s}+\frac{\partial}{\partial y}\left[\left(r-q(s)+\lambda(s,\mathrm{e}^y)-\frac{1}{2}\sigma(s,\mathrm{e}^y)^2\right)p\right]+\frac{\partial}{\partial r}\left(\mu_r(s,r)p\right)$$
$$-\frac{1}{2}\frac{\partial^2}{\partial y^2}\left(\sigma(s,\mathrm{e}^y)^2 p\right)-\frac{1}{2}\frac{\partial^2}{\partial r^2}\left(\sigma_r(s,r)^2 p\right)$$
$$-\rho(s)\frac{\partial^2}{\partial y\partial r}\left(\sigma_r(s,r)\sigma(s,\mathrm{e}^y)p\right)$$
$$=\left(r+\lambda(s,\mathrm{e}^y)\right)p \tag{20}$$

where $p(t, s, r_0, r, z, y)$ denotes the state price density for transition from state (z, r_0) at time t to state (r, y) at time $s > t$ (and where t, r_0, z are considered fixed). The boundary condition is $p(t, t, r_0, r, z, y) = \delta(r - t_0)\delta(y - z)$. Using the same steps that lead to (11), equation (20) above can be rewritten to only contain derivatives of p. The calibration algorithm outlined in Section 4 then holds unchanged: use a finite-difference (ADI) method to solve (20) forward in time, at each step taking care to adjust σ and λ such that the relations (14) are satisfied. With two-state variables involved, the computational effort here obviously increases significantly over that of the one-factor model discussed earlier.

6.2 Calibration to full equity volatility surface

Reverting to the one-factor model of Sections 2–5, we now wish to explore the possibility of fitting such a model to more than one equity option per maturity.

As a first observation, we notice that if option prices of sufficiently low strikes are known, it becomes unnecessary to consider risky bonds, as the prices of these are already contained in equity options. Specifically, consider a T-maturity, K-strike put option with time 0 price of $V_p(T, K)$. For sufficiently low values of K, the put option will practically only pay out if a default takes place and the stock goes to zero; that is,

$$V_P(T,K) \approx P(0,T)K\ \Pr(\tau \le T) = P(0,T)K\left(1 - \frac{B(0,T)}{P(0,T)}\right)$$
$$= K\big(P(0,T) - B(0,T)\big)$$

for low values of K. This relationship is, in fact, frequently used to trade equities against straight debt. Taking the limit, we get

$$B(0,T) = P(0,T) - \lim_{K \downarrow 0} \frac{\partial V_P(T,K)}{\partial K} = -\lim_{K \downarrow 0} \frac{\partial C(T,K)}{\partial K}$$

where the second equality follows from put–call parity. (These equations also follow from first principles – see Appendix B.)

As we have seen in Section 3, the possibility of default induces a significant volatility skew in option prices. To the extent that this skew does not fit the observed market skew, it is natural to adjust the volatility function $\sigma(t, S)$ in order to correctly price options with different strikes expiring on the same date. To perform this task, we follow the literature of forward volatility fitting (see, eg, Dupire, 1994, and Andersen and Andreasen, 2000) and work with a forward PDE for call option prices expressed as functions of strike and maturity. Appendix B derives this PDE for our model framework:

$$\frac{\partial C(T,K)}{\partial T}$$
$$= -\big(q(T) - K\lambda'(T,K)\big)C(T,K) - \big(r(T) - q(T) + \lambda(T,K)\big)K\frac{\partial C(T,K)}{\partial K}$$
$$+ \frac{1}{2}K^2\sigma(T,K)^2\frac{\partial^2 C(T,K)}{\partial K^2} + K\int_K^\infty C(T,k)\lambda''(T,k)\mathrm{d}k \qquad (21)$$

with boundary condition $C(0, K) = (S(0) - K)^+$. λ' and λ'' denote the first and second derivatives of λ with respect to K (which we assume exist; otherwise, we can use equation (B5) in Appendix B). There are a number of different ways this forward PDE can be used in a calibration exercise. First, if we assume that a continuum of option prices is known for all T and K (which in practice requires application of interpolation techniques) the PDE can be rearranged (omitting some function arguments for space considerations) to

$$\sigma(T,K)^2 = \frac{\frac{\partial C}{\partial T} + \big(q - K\lambda'(T,K)\big)C + \big(r - q + \lambda(T,K)\big)K\frac{\partial C}{\partial K} - K\int_K^\infty C(T,k)\lambda''(T,k)\mathrm{d}k}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}}$$

where we assume that $\lambda(T,K)$ is known. For the equation above to work in practice, great care needs to be taken in ensuring that the option price surface is sufficiently smooth; see Andersen and Andreasen (2000) for more on this.

Rather than assuming that a full continuum of option prices are observable, alternatively we could attempt to fit only a discrete set of options directly observable in the market. Such a fit would typically involve iterating on σ until (a) all observable option prices are reproduced with adequate accuracy; and (b) some smoothness norm is maximized. See, eg, Lagnado and Osher (1997) for a typical approach. The forward PDE (21) is crucial for this approach to be practical, as it allows for quick pricing of call options with different maturities and strikes. In particular, we notice that solving a *single* log-K finite-difference grid gives us the prices of call options at numerous strikes and maturities, with each node (T_i, y_j) in the finite-difference grid representing the time 0 price of a call option maturing at time T_i with strike e^{y_i}. This is obviously much faster than pricing the call options using the backward equation, which involves one finite-difference grid per option.

As yet, we have assumed that $\lambda(t,S)$ is known and that $\sigma(t,S)$ is fitted to the smile. In principle, we could also fix $\sigma(t,S)$ and attempt to solve for $\lambda(t,S)$ such that the option smile is replicated. Using (21) (or alternatively equation (B5) in Appendix B) this leads to a Volterra integral equation for $\lambda(t,S)$. As we have seen earlier in Figure 4, however, the effect on the volatility smile from changing $\lambda(t,S)$ can be rather limited, so in practice there may be situations where this equation will fail to have a reasonable (or even stable) solution.

Finally, we point out that the forward equation (21) is a special case of a somewhat broader theory for forward PDEs associated with general jump-diffusions with time- and state-dependent intensity. The interested reader can find some additional results along these lines at the end of Appendix B.[13]

6.3 Other extensions

In (1), the model for the pre-default stock dynamics are purely diffusive, with the stock being the single state-variable. Many recent models for stock dynamics are considerably more complicated than this and may involve discontinuities and additional state variables (such as unspanned stochastic volatility). In principle, such additional features could also form the basis for a model for convertible

[13] Some of the results in Appendix B have independently been derived in Carr and Javaheri (2002).

bonds, although it remains to be seen under what circumstances such models could be made practical. This is an interesting avenue of future research.

Finally, we note that while imposing a deterministic functional dependence between stock prices and default intensities appears to capture a number of salient characteristics of convertible bond markets well, some authors (eg, Davis and Lischka, 1999) have pursued models where the default intensity is driven by a Brownian motion different from that of the stock. While some of these models imply that credit spreads can become negative, this approach has the potential to more accurately capture the fact that the relationship between spreads and equity prices is often quite noisy. For regular convertible bonds, where such effects are likely to only have secondary effects on prices, it is doubtful that this approach justifies the additional computational effort of introducing an extra state variable.

7 Conclusion

This paper adds to the newly resurgent literature on convertible bonds modeling by addressing the important question of how to efficiently calibrate reduced-form models to observed market information. Our primary approach involves application of the Fokker–Planck equation for simultaneous calibration to at-the-money options and straight debt (or credit derivatives). Through a number of numerical experiments we demonstrated practicality of the suggested techniques and documented a number of economically significant effects associated with naïve, yet apparently widespread, model parameterizations. In situations where liquid option data are plentiful, we developed theory to guide more ambitious calibration to the full option volatility smile. Applying this theory in practice – and testing impact on prices – is a challenging and interesting avenue of future research. For the parameterized version of our algorithm, estimation of non-calibrated parameters is an important question for future research; for instance, in our example specification $\lambda(t, S) = a(t)S^{-p}$ reliable estimates of the power p are critical. Finally, more research is needed in gauging the impact of applying more sophisticated, possibly discontinuous, pre-default stock dynamics than those used in this paper.

Appendix A – Coupon bonds and credit default swaps in terms of risky zero-coupon bonds

Calibration of convertible bond models to straight debt markets is most conveniently done through term structures of risky zero-coupon bonds. While risky zero-coupon bonds normally do not trade directly, their use as calibration instruments entail no difficulties as they can be implied from credit default swap (CDS) or coupon bond quotes by bootstrap techniques. For reference, this appendix briefly derives the necessary equations for this exercise.

First, consider a regularly spaced schedule $t_i = i\delta$, $i = 1, 2, \ldots, n$ and consider

a security A that pays an annualized coupon of \$1 on this schedule (ie, a net payment of δ is made on each schedule date), up to either the time default or the final maturity t_n, whatever comes first. Assuming that no accrued interest is due on default times falling between schedule dates, the time 0 value of this security is given as a risky annuity factor:

$$A(0) = \delta \sum_{i=1}^{n} B(0, t_i)$$

Second, consider a security F that pays \$1 precisely at the time of default, under the condition that default takes place before the final schedule maturity t_n. Working, strictly for convenience, on a time-grid that coincides with the schedule dates $\{t_i\}$, we write the time 0 value of this security as follows:

$$\begin{aligned} F(0) &= \sum_{i=1}^{n} P\left(0, \tfrac{1}{2}(t_i + t_{i-1})\right) \Pr\left(\tau \in [t_{i-1}, t_i]\right) \\ &= \sum_{i=1}^{n} P\left(0, \tfrac{1}{2}(t_i + t_{i-1})\right) \left[\Pr(\tau > t_{i-1}) - \Pr(\tau > t_i)\right] \\ &= \sum_{i=1}^{n} P\left(0, \tfrac{1}{2}(t_i + t_{i-1})\right) \left[\frac{B(0, t_{i-1})}{P(0, t_{i-1})} - \frac{B(0, t_i)}{P(0, t_i)}\right] \end{aligned}$$

where $\Pr(\cdot)$ denotes risk-neutral expectation, and where $B(0, t_0) = B(0, t_0) \equiv 1$. The third equality follows directly from (3), as by definition $\Pr(t > t_i) = E(1_{t > t_i})$. Notice that we have assumed that defaults anywhere inside the bucket $[t_{i-1}, t_i]$ can on average be associated with a payment at time $\frac{1}{2}(t_i + t_{i-1})$. This is obviously an approximation, but a highly accurate one and sufficient for our purposes. (In general, we could easily make this time-discretization finer, or even take the continuous-time limit).

We now turn to the pricing of a risky coupon bond D paying an annualized coupon rate of κ on the $\{t_i\}$ schedule discussed above. Consistent with market practice, we assume all coupons are lost in case of a default with only a certain fraction ϕ of notional recovered (see also Section 5.2). For a normalized notional of \$1, the value of the coupon bond can thereby be written as

$$D(0) = B(0, T) + \kappa A(0) + \phi F(0) \tag{A1}$$

We recognize the first two terms as representing the value of a bond with zero recovery on both the back-end notional (the first term) and the coupons (the second term). The last term represents the add-back of the value associated with non-zero recovery. Consider now a CDS where we pay a premium rate θ on the $\{t_i\}$ schedule in return for a payment of $(1 - \phi)$ if a default happens before the final maturity t_n. Again assuming a normalized notional of \$1, the time 0 value of this instrument is straightforward:

$$CDS(0) = (1-\phi)F(0) - \theta A(0) \tag{A2}$$

With (A1) and (A2), the relations linking risky zero-coupon bonds to CDS and coupon bond prices are complete. There are, of course, generally less CDSs and coupon bonds trading than the total number of different dates in the various payments schedules, meaning that we normally have too few equations to directly back out all risky zero-coupon bonds in equations (A1) and (A2). As a consequence, one typically applies some type of bootstrap interpolation technique to reduce the number of independent unknown variables.

Appendix B – Derivation of forward equation for European call options

Define $\Psi(t) = (S(t) - K)^+$ for some positive strike K. By the Itô–Tanaka theorem for continuous non-differentiable functions, we get

$$\begin{aligned} d\Psi(t) = {} & 1_{S(t)>K} S(t)\big(r(t) - q(t) + \lambda(t, S(t))\big)dt \\ & + \frac{1}{2}\delta(S(t) - K)S(t)^2\sigma(t, S(t))^2 dt + \\ & 1_{S(t)>K} S(t)\sigma(t, S(t))dW(t) - \Psi(t)dN(t) \end{aligned}$$

where we have omitted the tedious usage of $t-$ in the understanding that all terms are to be evaluated to the left of any jump time. Integrating over time, taking expectations, and noting that $\delta(S(t) - K)S(t)^2\sigma(t, S(t))^2 = \delta(S(t) - K)K^2\sigma(t, K)^2$ gives

$$\begin{aligned} E(\Psi(t)) = \Psi(0) + \int_0^T [r(t) - q(t)]E\left(1_{S(t)>K} S(t)\right)dt + \\ \int_0^T E\left(1_{S(t)>K} S(t)\lambda(t, S(t))\right)dt + \frac{1}{2}\int_0^T K^2\sigma(t, K)^2 E\left(\delta(S(t) - K)\right)dt - \\ \int_0^T E\left(\Psi(t, S(t))\lambda(t, S(t))\right)dt \end{aligned} \tag{B1}$$

To relate this equation to observable quantities, first notice that $E(\delta(S(t) - K))$ is the density of $S(t)$ in K. As before, we now let $C(t, K)$ denote the time 0 price of a call option with strike K and maturity t. From Breeden and Litzenberger (1978), we then have

$$E\left(\delta(S(t) - K)\right) = \frac{\partial^2 C(t, K)}{\partial K^2} P(0, t)^{-1} \tag{B2}$$

where $\partial^2 C(t, K)/\partial K^2$ will contain a mass in $K = 0$ to reflect the probability of default in $[0, t]$:

$$P(0, t)^{-1} \int_{0-}^{0+} \frac{\partial^2 C(t, k)}{\partial k^2} dk = \Pr(\tau < t)$$

We also notice that

$$C(t,K)P(0,t)^{-1} = E\left(1_{S(t)>K}(S(t)-K)\right)$$
$$= E\left(1_{S(t)>K}S(t)\right) + K\frac{\partial C(t,K)}{\partial K}P(0,t)^{-1} \tag{B3}$$

where the last equation follows, for instance, from integration of (B2). Finally, we notice that the third and fifth integrand in (B1) combine to

$$E\left(1_{S(t)>K}S(t)\lambda(t,S(t))\right) - E\left(1_{S(t)>K}S(t)\lambda(t,S(t)) - 1_{S(t)>K}K\lambda(t,S(t))\right)$$
$$= KE\left(1_{S(t)>K}\lambda(t,S(t))\right) \tag{B4}$$

Inserting expressions (B2)–(B4) into (B1), noting that $C(T, K) = P(0, T)E(\Psi(T))$, and taking the derivative with respect to T yields

$$\frac{\partial C(T,K)}{\partial T} = -q(T)C(T,K) - \left(r(T)-q(T)\right)K\frac{\partial C(T,K)}{\partial K} +$$
$$\frac{1}{2}K^2\sigma(T,K)^2\frac{\partial^2 C(T,K)}{\partial K^2} + K\int_K^\infty \frac{\partial^2 C(T,k)}{\partial k^2}\lambda(T,k)\mathrm{d}k \tag{B5}$$

where we have used (B2) in the last term. Assuming that λ is twice differentiable, we can rewrite the equation above by integrating the last term by parts twice:

$$\int \frac{\partial^2 C(T,k)}{\partial k^2}\lambda(T,k)\mathrm{d}k$$
$$= \frac{\partial C(T,k)}{\partial k}\lambda(T,k) - \lambda'(T,k)C(T,k) + \int \lambda''(T,k)C(T,k)\mathrm{d}k$$

Under some regularity on $\lambda(T, k)$ (eg, that it and its k-derivatives are bounded for large enough K), the two first terms in the equation above approach 0 as $k \to \infty$ As such, we can write

$$\int_K^\infty \frac{\partial^2 C(T,k)}{\partial k^2}\lambda(T,k)\mathrm{d}k$$
$$= -\frac{\partial C(T,K)}{\partial K}\lambda(T,K) + \lambda'(T,K)C(T,K) + \int_K^\infty \lambda''(T,k)C(T,k)\mathrm{d}k$$

Inserting this into (B5), we get the alternative expression:

$$\frac{\partial C(T,K)}{\partial T}$$

$$= -\big(q(T) - K\lambda'(T,K)\big)C(T,K) - \big(r(T) - q(T) + \lambda(T,K)\big)K\frac{\partial C(T,K)}{\partial K} +$$

$$\frac{1}{2}K^2\sigma(T,K)^2\frac{\partial^2 C(T,K)}{\partial K^2} + K\int_K^\infty C(T,k)\lambda''(T,k)\,\mathrm{d}k \tag{B6}$$

It is of interest to note that the forward equation above is just a special case of forward equations for jump-diffusion stock processes with state-dependent jump intensity. For instance, consider the stock process

$$\frac{\mathrm{d}S(t)}{S(t-)} = \big(r(t) - q(t) + m(t)\lambda(t,S(t-))\big)\mathrm{d}t + \sigma(t,S(t-))\mathrm{d}W(t) + (J(t)-1)\mathrm{d}N(t)$$

where $J(t)$ is some positive random variable with deterministic density $\Pr(J(t) \in \mathrm{d}z)/\mathrm{d}z = \xi(t,z)$, and $m(t) \equiv 1 - E(J(t))$. (We can easily extend further to stock and time-dependent density $\xi(t,S,z)$). Following the same steps as above, we arrive at the forward PDE

$$\begin{aligned}\frac{\partial C(T,K)}{\partial T} &= -q(T)C(T,K) - \big(r(T) - q(T)\big)K\frac{\partial C(T,K)}{\partial K} + \\ &\frac{1}{2}K^2\sigma(T,K)^2\frac{\partial^2 C(T,K)}{\partial K^2} + K\int_K^\infty \frac{\partial^2 C(T,k)}{\partial k^2}\lambda(T,k)\,\mathrm{d}k \\ &+E(J(t))\left[\int_0^\infty\int_{K/z}^\infty \frac{\partial^2 C(T,k)}{\partial k^2}\left(k - \frac{K}{z}\right)\xi^*(T,z)\lambda(T,k)\,\mathrm{d}k\,\mathrm{d}z\right. \\ &\left. -\int_K^\infty \frac{\partial^2 C(T,k)}{\partial k^2}k\lambda(T,k)\,\mathrm{d}k\right]\end{aligned} \tag{B7}$$

where $\zeta^*(T,z) = z\xi(T,z)/E(J(T))$ is a density function. Let us consider some special cases. First, when $J(t) = 0$ we are left with the same expression as earlier. Second, if $J(t)$ is some deterministic function $\eta(t)$, $\xi(t,z) = \delta(z - \eta(t))$ and the equation simplifies to

$$\frac{\partial C(T,K)}{\partial T} = -q(T)C(T,K) - \big(r(T) - q(T)\big)K\frac{\partial C(T,K)}{\partial K} +$$

$$\frac{1}{2}K^2\sigma(T,K)^2\frac{\partial^2 C(T,K)}{\partial K^2} + \int_K^{K/\eta(t)} \frac{\partial^2 C(T,k)}{\partial k^2}(K - k\eta(T))\lambda(T,k)\,\mathrm{d}k$$

And third, when λ is independent of S, (B7) degenerates into the forward equation derived in Andersen and Andreasen (2000).

Appendix C – Constant-parameter stationary distribution of intensity process $\lambda(t, S) = cS^{-p}$

To understand the asymptotic behavior of $\lambda(t, S) = cS^{-p}$, $p > 0$, for large t, consider the case where all stock process parameters (r, q, σ) are constants. From equation (5), we have that

$$\frac{d\lambda(t)}{\lambda(t)} = (a - p\lambda(t))dt - p\sigma dW(t)$$

where $a \equiv p(\frac{1}{2}(p+1)\sigma^2 - r + q)$. If $a < 0$, no stationary distribution exists as the origin becomes an attractive point and $\lim_{t \to \infty} E)\lambda(t)) = 0$. For sufficiently positive values of a, however, a stationary density $\psi(\lambda)$ might exist. From standard theory (eg, Karlin and Taylor, 1981, p. 220), ψ satisfies, should it exist,

$$\frac{d}{d\lambda} = (a - p\lambda) + \frac{d^2}{d\lambda^2}\left(\psi(\lambda)\sigma^2 p^2\right) = 0$$

Integrating once (and arguing that the integration constant must be zero) and separating variables yields the result

$$\psi(\lambda) = C\lambda^{\mu-1}e^{-\lambda/\alpha} \tag{C1}$$

where C is an arbitrary constant, $\mu \equiv 2a/(p^2\sigma^2) - 1$, and $\alpha \equiv \frac{1}{2}p\sigma^2$. For the density to be integrable, we obviously must assume that $\mu \geq 0$. Recognizing (C1) as the density of a Gamma distributed variables, the constant C must equal $1/(\Gamma(\mu)\alpha^\mu)$ to ensure that ψ integrates to 1. More importantly, from the properties of the Gamma distribution, we see that

$$\int_0^\infty \lambda\psi(\lambda)d\lambda = \alpha\mu = \frac{a}{p} - \frac{1}{2}p\sigma^2 + \frac{1}{2}\sigma^2 - r + q, \quad \mu \geq 0$$

Inserting the definition of a and simplifying, we get

$$\lim_{t\to\infty} E(\lambda(t)) = \max\left(0, \frac{1}{2}\sigma^2 - r + q\right) \tag{C2}$$

REFERENCES

Andersen, L., and Andreasen, J. (2000). Jump-diffusion processes: volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research* **4**, 231–62.

Andersen, L., and Brotherton-Ratcliffe, R. (1998). The equity option volatility smile: an implicit finite difference approach. *Journal of Computational Finance* **2**, 5–38.

Arvanitis, A., and Gregory, J. *Credit: The Complete Guide to Pricing, Hedging, and Risk Management*, Risk Books, London.

Ayache, E, Forsyth, P., and Vetzal K. (2002). The valuation of convertible bonds with credit risk. Working paper, University of Waterloo.

Bloch D., and Miralles, P. (2002). Credit treatment in convertible bond model. Working paper, Dresdner Kleinwort Wasserstein

Breeden, D., and Litzenberger, R. (1978). Prices of state-contingen claims implicit in options prices. *Journal of Business* **51**, 621–51.

Brennan, M., and Schwarz, E. Analyzing convertible bonds. *Journal of Financial and Quantitative Analysis* **15**, 907–29.

Carr, P., and Javaheri, A. (2002). The forward PDE for European options on stocks with fixed fractional jumps. Working paper, Courant Institute and RBC Capital Markets.

Cheung, W., and Nelken, I. (1994). Costing the converts. *Risk*, July, 47–49.

Craig, I., and Schneyd, A. (1988). An alternating direction method for three state variables. *Numerical Mathematics* **4**, 41–63.

Davis, M., and Lischka, F. (1999). Convertible bonds with market risk and credit risk. Working paper, Tokyo-Mitsubishi International.

Duffie, D., and Singleton, K. (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies* **12**, 687–720.

Dupire, B. (1994). Pricing with a smile. *Risk*, January, 18–20.

Forsyth, P., and Vetzal, K. (2002). Quadratic convergence for valuing American options using a penalty method. *SIAM Journal on Scientific Computation* **23**, 2096–133.

Goldman Sachs (1994). Valuing convertible bonds as derivatives. Research note, Goldman Sachs.

Grimwood, R., and Hodges, S. (2002). The valuation of convertible bonds: a study of alternative pricing models. Working paper, University of Warwick

Ho, T., and Pfeffer, D. (1996). Convertible bonds: model, value attribution, and analytics. *Financial Analysts Journal* **52**, 35–44.

Jarrow, R., and Turnbull, S. (1995). Pricing derivative securities on financial securities subject to credit risk. *Journal of Finance* **50**, 53–85.

Karlin, S., and Taylor, H. (1981). *A Second Course in Stochastic Processes*. Academic Press, San Diego.

Lagnado, R., and Osher, S. (1997). Reconciling differences. *Risk*, **4**, April, 79–83.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives Research* **2**, 99–120.

McConell, J., and Schwarz, E. (1986). LYON taming. *Journal of Finance* **50**, 53–85.

Merton, R. (1974). On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance* **29**, 449–70.

Muromachi, Y. (1999). The growing recognition of credit risk in corporate and financial bond markets. Working paper, NLI Research Institute

Olsen, L. (2002). Convertible bonds: a technical introduction. Research note, Barclays Capital

Pooley, D., Vetzal, K., and Forsyth, P. Digital projection. Working paper, University of Waterloo.

Takahashi, A, Kobayashi, T., and Nakagawa, N. (2001). Pricing convertible bonds with default risk. *Journal of Fixed Income* **11**, 20–9.

Tsiveriotis, K., and Fernandes, C. (1998). Valuing convertible bonds with credit risk. *Journal of Fixed Income*, **8**, 95–102.

Yigitbasioglu, A. (2001). Pricing convertible bonds with interest rate, equity, credit, and FX risk. Working paper, University of Reading.