SIAM J. FINANCIAL MATH.
Vol. 1, pp. 642–665

Affine Point Processes and Portfolio Credit Risk*

Eymen Errais†, Kay Giesecke‡, and Lisa R. Goldberg§

Abstract. This paper analyzes a family of multivariate point process models of correlated event timing whose arrival intensity is driven by an affine jump diffusion. The components of an affine point process are self- and cross-exciting and facilitate the description of complex event dependence structures. ODEs characterize the transform of an affine point process and the probability distribution of an integer-valued affine point process. The moments of an affine point process take a closed form. This guarantees a high degree of computational tractability in applications. We illustrate this in the context of portfolio credit risk, where the correlation of corporate defaults is the main issue. We consider the valuation of securities exposed to correlated default risk and demonstrate the significance of our results through market calibration experiments. We show that a simple model variant can capture the default clustering implied by index and tranche market prices during September 2008, a month that witnessed significant volatility.

Key words. self-exciting point process, affine jump diffusion, Hawkes process, transform, portfolio credit derivative, correlated default, index and tranche swap

AMS subject classifications. 60-08, 60J99, 60G55, 90-08, 90B25

DOI. 10.1137/090771272

1. Introduction.

The collapse of Lehman Brothers brought the financial system to the brink of a breakdown. The dramatic repercussions point to the existence of feedback phenomena that are channeled through the complex web of informational and contractual relationships in the economy. Lehman was an important node in a network of derivative traders. It had bought and sold default insurance on a large number of firms and was itself a reference entity in countless other insurance contracts. Its downfall triggered payments that forced some insurance sellers into default, leaving the corresponding protection buyers with losses. It also exposed the counterparties to the contracts Lehman itself had written.

This and related episodes motivate the design of models of correlated default timing that incorporate the feedback phenomena that plague credit markets. This paper analyzes a family of computationally tractable self-exciting point processes that can capture event feedback. The future evolution of a self-exciting point process is influenced by the timing of past events and their marks, for example, the financial loss they caused. This feature takes account of the direct impact of events. It also generates a dependence structure between arrival rates and losses, a property that is empirically well documented.

*Received by the editors September 16, 2009; accepted for publication (in revised form) June 12, 2010; published electronically September 16, 2010. This paper was previously circulated under the title "Pricing Credit From the Top-Down With Affine Point Processes."

http://www.siam.org/journals/sifin/1/77127.html

†CreditFlow Partners, 345 86th Street, Suite 12D, New York, NY 10028 (Eymen.Errais@thecreditflow.com).

‡Corresponding author. Department of Management Science & Engineering, Stanford University, Stanford, CA 94305-4026 (giesecke@stanford.edu, http://www.stanford.edu/∼giesecke).

§MSCI Barra, 2100 Milvia St., Berkeley, CA 94704 (Lisa.Goldberg@mscibarra.com).

Our stepping stone is the Hawkes process, perhaps the most parsimonious self-exciting point process. The conditional event arrival rate or intensity of a Hawkes process jumps in response to events and tends toward a target level in the absence of an event. While the Hawkes process is widely used in a range of disciplines, its distributional properties are poorly understood. We develop these properties, exploiting the fact that the two-dimensional process consisting of a Hawkes process and its intensity is Markov. The structure of the associated infinitesimal generator leads to a Dynkin formula and closed expressions for the moments of the Hawkes intensity. We show that a transform of the Hawkes process satisfies a certain partial integral differential equation (PIDE). The solution to that equation turns out to be an exponentially affine function of the initial value of the two-dimensional process, whose coefficients satisfy a system of ODEs. Analysis of the transform leads to ODEs that characterize the probability distribution of a Hawkes process. We obtain closed formulae for the moments of the process.

The transform, distribution, and moment formulae generate computational tractability for a range of applications in portfolio credit risk. To illustrate, we use a Hawkes process to model the cumulative loss due to default in a portfolio of firms. The jump times represent default times, and the jump magnitudes represent the random losses at default. This formulation captures the impact of a default on the surviving names. It also incorporates the negative correlation between default and recovery rates. The transform formulae facilitate the valuation, hedging, and calibration of a portfolio credit derivative, which is a security whose payoff is a specified function of the portfolio loss, and which provides insurance against default losses in the portfolio. An index swap, for example, pays any portfolio losses before the contract maturity. A tranche swap pays a slice of the portfolio loss specified by a lower and an upper attachment point. Our market calibration experiments, which are based on index and tranche price data observed before and after Lehman's collapse, indicate the empirical significance of the self-exciting property of the loss process. An empirical analysis demonstrates that the parsimonious Hawkes process can capture the default correlation implied by credit market rates on each trading day in September 2008, a month that witnessed dramatic volatility. This is a significant improvement over the industry standard copula model.

The insight into the probabilistic structure of the Hawkes process leads us to consider an extension to a multivariate point process whose intensity is driven by an affine jump diffusion process that represents a vector of stochastic risk factors. The variation of the factors generates diffusive and jump volatility of conditional event arrival rates. The point process itself can be a risk factor so that the point process components are self- and cross-exciting. The basic transform, distribution, and moment characterization arguments developed for the Hawkes process extend to this family of affine point processes.

While this article highlights derivatives valuation applications, the self-exciting point processes considered here are also potential tools for the risk management of corporate debt portfolios, for which event feedback and the dependence between default and recovery rates are significant issues. For example, Das et al. [14] test a doubly stochastic model whose features are similar to those of the models endorsed by the regulatory authorities to estimate portfolio credit risk. In a doubly stochastic model event times are conditionally independent. Das et al. [14] find evidence of historical default clustering in excess of that implied by the model they tested. This suggests that doubly stochastic models underestimate risk capital.

A self-exciting point process model implies a more realistic degree of default clustering. It also accounts for the dependence between default and recovery rates, whose importance is emphasized in Basel II's pillar I guidelines; see [5].

The point processes examined in this paper facilitate a top-down approach to portfolio credit risk, in which the portfolio loss process is specified without reference to the constituent names. There are other examples of this approach in the literature. Arnsdorf and Halperin [2] describe arrivals by a nonlinear death process and its generalizations. Brigo, Pallavicini, and Torresetti [9] model events by a mixed Poisson process and its extensions. Cont and Minca [12] describe arrivals by a nonhomogeneous Markov chain. Davis and Lo [15] model events by a piecewise deterministic Markov process. Ding, Giesecke, and Tomecek [16] propose a time-changed linear birth process as a model of arrivals. In [28], defaults are driven by independent Poisson processes that model idiosyncratic, sector specific, and economy-wide events. Lopatin and Misirpashaev [29] introduce a point process model with an intensity whose drift is modulated by the portfolio loss. The family of loss point processes proposed in this paper is distinct from, or in some special cases encompasses, the models introduced in these contributions. This paper features multivariate, interacting point processes with stochastic recoveries, and correlated arrival and recovery rates. Calibration experiments establish the fit of a basic model variant to the market data in September 2008.

The top-down specification of the portfolio loss process generates computational advantages for portfolio derivatives valuation. Constituent name hedging requires the sensitivities of the portfolio derivative price with respect to changes in the prices of the single-name derivatives referenced on the constituents. The sensitivities determine the amount of single-name protection on each portfolio constituent to be bought or sold in order to neutralize portfolio derivative price fluctuations due to small changes in the constituent risks. Giesecke, Goldberg, and Ding [20] develop a random thinning approach to estimate these hedges for a given portfolio loss process. For September 2008, they demonstrate the effectiveness of this approach for the Hawkes process. The empirical results indicate that a top-down model enables better hedging than the industry standard copula model.

Section 2 develops the distributional properties of the Hawkes process. Section 3 applies these results to the valuation of credit derivatives. Market calibration experiments demonstrate the significance of the self-exciting feature and the fit of the Hawkes model. Section 4 provides further parametric examples of multivariate self- and cross-exciting point processes. Section 5 concludes. The appendix contains the proofs.

2. Hawkes process.

Consider a sequence of default stopping times $0 < T_1 < T_2 < \dots$ that are defined on a complete probability space $(\Omega, \mathcal{F}, P)$ with right-continuous and complete information filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. The nature of the probability measure P depends on the application. In risk management applications, P is the actual or statistical measure. In valuation applications, P is a risk-neutral pricing measure, relative to which the discounted price of a traded security is a martingale. The financial loss at T_n is given by a random variable $\ell_n \in \mathcal{F}_{T_n}$. The sequence (T_n, ℓ_n) generates a nonexplosive default counting process N given by $N_t = \sum_{n\geq 1} 1_{\{T_n \leq t\}}$ and a loss point process L defined by $L_t = \sum_{n\geq 1} \ell_n 1_{\{T_n\leq t\}}$.

We propose to specify the processes N and L directly through a conditional arrival rate or intensity λ and a distribution ν on $(0, \infty)$ for the loss ℓ_n at an event. We assume that the

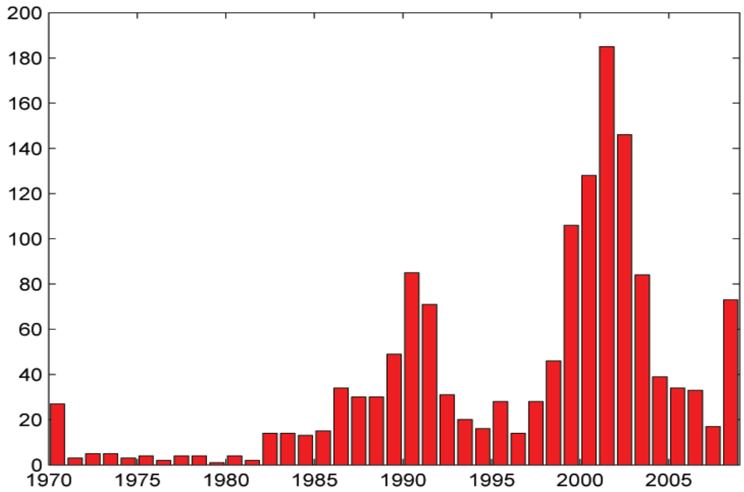

Figure 1. *Annual defaults of Moody's rated U.S. firms between January* 1970 *and November* 17, 2008*. Source: Moody's Default Risk Service. The peak in* 1970 *represents a cluster of* 24 *railway defaults triggered by the collapse of Penn Central Railway on June* 21, 1970*. The fallout of the* 1987 *crash is indicated by the peak in the early* 1990*s. The burst of the internet bubble caused many defaults during* 2001–2002*. From a trough in* 2007*, default rates increased significantly in* 2008*.*

jump transform $\int e^{\omega z} d\nu(z)$ exists and is finite for complex ω and admits a finite derivative $\int z e^{\omega z} d\nu(z)$. The intensity follows a strictly positive stochastic process that describes the conditional mean default rate in the sense that $E(N_{t+\Delta} - N_t \,|\, \mathcal{F}_t) \approx \lambda_t \Delta$ for small $\Delta > 0$. This means that $N - \int_0^{\cdot} \lambda_s ds$ is a local martingale relative to P and $\mathbb{F}$. The process followed by λ completely determines the conditional distribution of N. We analyze a family of models for λ whose features are empirically motivated.

2.1. Empirical motivation. Empirical observation dictates the properties of the stochastic process followed by λ. Most importantly, λ must replicate the *clustering of defaults* seen in Figure 1, which is due to the dependence of the default times T_n. The dependence is a result of the sensitivity of firms to common economic risk factors that vary stochastically through time. It is also caused by the informational and contractual linkages between firms, which provide a channel for the propagation of financial distress from one firm to another. The existence of these feedback phenomena is indicated by the ripple effects associated with the default of Lehman Brothers on September 15, 2008 and is further empirically documented in [4], [11], [17], [24], [27], and others. We propose to model the impact of a default on the other firms by ramping up the intensity at an event. This amounts to including the default process itself as a risk factor.

Another important empirical feature is the negative correlation between default and recovery rates. During periods with elevated default rates, creditors tend to recover less at a default than during periods with relatively few defaults. While documented by Altman et al. [1] and many others, this property has been largely ignored in a theoretical literature that is focused on modeling default timing. We propose to capture this property by modeling the magnitude of the response of the intensity to a default as a linear function of the realized loss

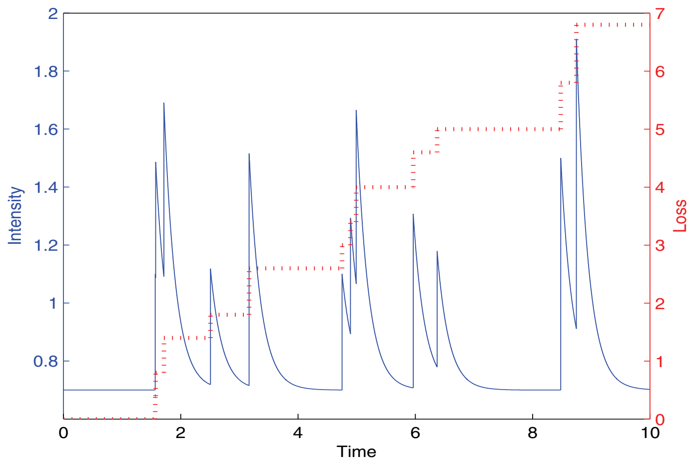

Figure 2. *Sample paths of the intensity* (1) *and the associated loss process L. A jump in the intensity represents the impact of a default. The jump size is the product of the loss at default and the sensitivity parameter $\delta = 1$. The loss at default is drawn from a uniform distribution on $\{0.4, 0.6, 0.8, 1\}$. The rate $\kappa = 5$ controls the exponential decay after an event. The reversion level $c = \lambda_0 = 0.7$.*

at an event. The larger the loss, the larger the impact of an event on the other firms, and the bigger the increase of the intensity at an event.

2.2. Specification. We examine a basic intensity model that incorporates the features described above. We suppose events arrive with intensity λ given by

$$\lambda_t = u(t) + \int_0^t h(t-s)\, dL_s. \tag{1}$$

The first to default intensity $u(t) = c + e^{-\kappa t}(\lambda_0 - c)$ is a deterministic function of time, $c > 0$, $\lambda_0 > 0$, and the impact of a loss on the intensity is governed by the function

$$h(v) = \delta e^{-\kappa v}, \qquad v \geq 0, \tag{2}$$

with $\kappa \geq 0$ and $\delta \geq 0$. As illustrated in the sample path of (λ, L) shown in Figure 2, the randomness in the default rate is driven by two sources of uncertainty: the timing of events and the recovery at these events. At a default, the intensity ramps up by the realized loss scaled with the sensitivity parameter δ. The lower the recovery, the higher the jump of the default rate. The impact of an event decays exponentially over time with rate κ. The reversion level is c. It follows that defaults are positively self-affecting, or *self-exciting*. Further, default and recovery rates are *negatively correlated*. Despite its parsimony, in section 3.3 we show that the basic model (1) is very effective at replicating the clustering of defaults implied by market prices.

While parsimonious and empirically motivated, the intensity model (1) and its extensions discussed in section 4 may not be appropriate for relatively small portfolios. This is because (1) generates a point process (N, L) that does not terminate when all firms in the portfolio are

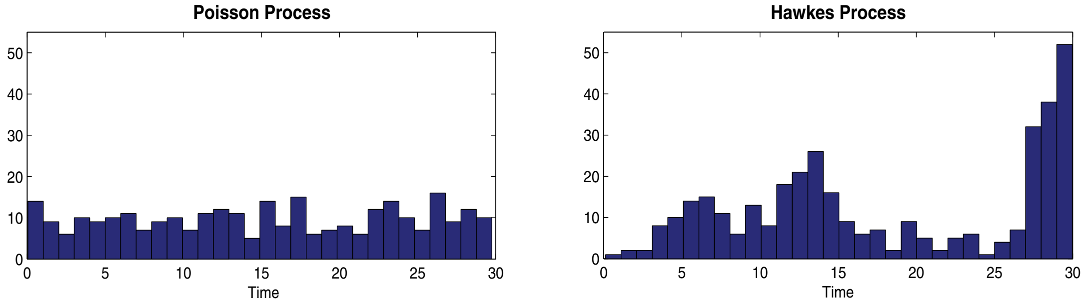

Figure 3. *Sample paths of the Poisson and Hawkes processes. A bar indicates the number of arrivals in a given year. While the Poisson arrivals are evenly distributed over time due to the order statistics property, the Hawkes arrivals are clustered thanks to the self-exciting property. For the Hawkes process,* $c = \lambda_0 = 1$, $\kappa = 1.5$, $\delta = 2$, *and the loss at default has a uniform distribution* ν *on* $\{0.4, 0.6, 0.8, 1\}$. *The Poisson process is the special Hawkes process for which* $\delta = 0$ *so that* $\lambda_t = u(t)$. *We choose* $c = \lambda_0 = 10.57$ *to match the expected number of Hawkes events over* 30 *years.*

in default. This feature tends to be innocuous for the large diversified portfolios in practice, which can have as many as several thousand names, and for which the likelihood of total default is typically negligible. It is fully appropriate only for portfolios in which defaulted names are replaced with new names, as may be the case for collateralized debt obligations with actively managed collateral pools.

The intensity λ governs the common event times of N and L. While the jumps of the default process N are unit-sized, the jumps of the loss process L are drawn from the distribution ν on $(0, \infty)$ of loss given default. The two-dimensional process $J = (L, N)^\top$ is a Hawkes process; see [22] and [23]. Variants of the Hawkes process have been applied to a range of problems in science and engineering; see [13] for a sample of these. They have only recently been used in financial economics. For example, Bowsher [7] fits different Hawkes models to security trade data. Surprisingly, despite these applications, the distributional properties of Hawkes-type processes are poorly understood. We establish these properties below and extend the Hawkes process to a broad family of self-exciting point processes.

We note two special cases. If $\kappa = 0$ in the specification (1), then λ is constant between events, and so J is a birth process. If $\delta = 0$, then J is a Poisson process whose intensity $u(t)$ evolves deterministically through time and whose interarrival times are independent. Figure 3 contrasts the sample paths of the Poisson and Hawkes processes. It shows how the self-exciting property of the Hawkes process generates arrivals that are overdispersed relative to Poisson arrivals. The resulting event clusters are strikingly similar to those of the empirical default process in Figure 1. The parameter δ in (1) allows us to directly control the frequency of these clusters. The parameter κ governs their magnitude.

While the Hawkes process induces clustered arrivals, it does not explode, and therefore the integral $A_t = \int_0^t \lambda_s ds$ is finite for all t. To see this, it suffices to consider the case $\kappa = 0$, when λ has nondecreasing sample paths. Then N is a birth process, which is nonexplosive and integrable. From this we can conclude that N_t is integrable also in the case $\kappa > 0$. This, in turn, implies that A_t is integrable, and that the compensated local jump martingale $M = N - A$ is a martingale. To see this, we can appeal to Corollary 3 in Chapter II.6 of [30], which implies that M is a martingale if $[M,M]$ is integrable. But $[M,M]=N$. Further, we obtain that $E(M_t^2)=E(N_t)=E(A_t)$.

2.3. Dynkin formula. While the process J itself does not have the Markov property, the process $Y=(\lambda,J)^\top$ is a Markov process in a state space $D=\mathbb{R}_+\times(\mathbb{R}_+\times\mathbb{N})$. Y has an infinitesimal generator $\mathcal{D}$, defined at a function $g:D\to\mathbb{R}$ with continuous partial derivative $g_\lambda(\lambda,x)$, by

$$(\mathcal{D}g)(\lambda,x)=\kappa(c-\lambda)g_\lambda(\lambda,x)+\lambda\int[g(\lambda+\delta z,x+(z,1)^\top)-g(\lambda,x)]d\nu(z). \tag{3}$$

Proposition 2.1. *If the expectation*

$$E\left(\int_0^t|(\mathcal{D}g)(\lambda_s,J_s)-\kappa(c-\lambda_s)g_\lambda(\lambda_s,J_s)|ds\right) \tag{4}$$

is finite for each t, then for each $t\le T$ we have a Dynkin formula

$$E(g(\lambda_T,J_T)\,|\,\mathcal{F}_t)=g(\lambda_t,J_t)+E\left(\int_t^T(\mathcal{D}g)(\lambda_s,J_s)ds\,\middle|\,\mathcal{F}_t\right). \tag{5}$$

Proposition 2.1 provides a formula expressing the conditional expectation of a function g of the Markov process $Y=(\lambda,J)^\top$ in terms of the infinitesimal generator (3). This formula is useful for functions g for which the expectation of $(\mathcal{D}g)(\lambda_s,J_s)$ can be calculated. We exploit it to obtain explicit formulae for the first and second moments of the intensity. Consider the function $g(\lambda,x)=\lambda$. We calculate

$$\begin{aligned}E((\mathcal{D}g)(\lambda_s,J_s))&=E(\kappa(c-\lambda_s))+E\left(\lambda_s\int\delta z d\nu(z)\right)\\&=\kappa c+\mu E(\lambda_s),\end{aligned}$$

where $\mu=\delta\ell-\kappa$ and $\ell=\int zd\nu(z)$ is the expected loss at default. Letting $m(t)=E(\lambda_t)$, equation (5) implies that

$$m(t)=E(g(\lambda_t,J_t))=\lambda_0+\kappa ct+\mu\int_0^t m(s)ds. \tag{6}$$

In other words, the mean intensity solves the ODE $m'(t)=\kappa c+\mu m(t)$ with initial condition $m(0)=\lambda_0$. The solution is

$$m(t)=E(\lambda_t)=\left(\frac{\kappa c}{\mu}+\lambda_0\right)e^{\mu t}-\frac{\kappa c}{\mu} \tag{7}$$

provided that $\mu\neq 0$. Equation (7) shows that for $\mu>0$, the mean intensity grows exponentially (without bound) in time. For $\mu\to 0$, the mean intensity grows linearly with time as $\lim_{\mu\to 0}E(\lambda_t)=\lambda_0+\kappa ct$. For $\mu<0$, the mean intensity decays exponentially with time to the long-run mean

$$\lim_{t\to\infty}E(\lambda_t)=-\frac{\kappa c}{\mu}>0. \tag{8}$$

We apply a similar argument to the function $g(\lambda, x) = \lambda^2$ to show that the second moment $v(t) = E(\lambda_t^2)$ solves the ODE

$$v'(t) = \rho m(t) + 2\mu v(t) \tag{9}$$

with initial condition $v(0) = \lambda_0^2$, where $\rho = 2\kappa c + \delta^2 \int z^2 d\nu(z)$. Here, we assume that ν has finite second moment. For $\mu \neq 0$, the solution to (9) is given by

$$v(t) = \frac{1}{2\mu^2}\left[\kappa c\rho(e^{\mu t} - 1)^2 + 2\mu\lambda_0 e^{\mu t}(\rho(e^{\mu t} - 1) + \mu\lambda_0 e^{\mu t})\right]. \tag{10}$$

We conclude that for $\mu < 0$, the long run variance of the intensity satisfies

$$\lim_{t\to\infty} \mathrm{Var}(\lambda_t) = \lim_{t\to\infty} (v(t) - m^2(t)) = \frac{\kappa c\delta^2}{2\mu^2} \int z^2 d\nu(z). \tag{11}$$

2.4. Transform. Along with the distribution ν of the loss at default, the intensity specification (1) determines the distribution of the processes N and L. We develop a formula for a transform of the point process $J = (L, N)^\top$ that includes as special cases the Laplace and Fourier transforms. The transform can be inverted to obtain the corresponding distribution.

Let $u \in \mathbb{C}_-^2$, the set of pairs of complex numbers with nonpositive real part. Express the conditional transform $E(\exp(u \cdot J_T) \,|\, \mathcal{F}_t)$ as $f(t, \lambda_t, J_t)$ for some complex-valued function f on $[0, T] \times D$. For $f(t, \lambda_t, J_t)$ to be a martingale, its drift must vanish. This means that f must satisfy the PIDE

$$0 = f_t(t, \lambda, x) + (\mathcal{D}f)(t, \lambda, x)$$

with boundary condition $f(T, \lambda, x) = \exp(u \cdot x)$, where $(\mathcal{D}f)(t, \lambda, x)$ is obtained by applying the generator $\mathcal{D}$ to the real and imaginary parts of $f(t, \cdot, \cdot)$. This PIDE reduces to a system of ODEs when f is taken to be an exponentially affine function of λ and x. This argument is made precise in the next result.

Proposition 2.2. *The transform of the point process $J = (L, N)^\top$ is given by*

$$E(\exp(u \cdot J_T) \,|\, \mathcal{F}_t) = \exp(a(t) + b(t)\lambda_t + u \cdot J_t), \tag{12}$$

where $t \leq T$, $u \in \mathbb{C}_-^2$, and the coefficient functions $a(t) = a(u, t, T)$ and $b(t) = b(u, t, T)$ satisfy the following ODEs:

$$\partial_t b(t) = \kappa b(t) + 1 - \theta(\delta b(t) + u \cdot (1, 0)^\top) \exp(u \cdot (0, 1)^\top), \tag{13}$$

$$\partial_t a(t) = -\kappa c b(t) \tag{14}$$

with boundary conditions $a(T) = b(T) = 0$, where θ is the jump transform

$$\theta(\omega) = \int e^{\omega z} d\nu(z), \quad \omega \in \mathbb{C}. \tag{15}$$

The argument behind Proposition 2.2 immediately extends to more general functionals of $Y = (\lambda, J)^\top$. For example, the argument leads to the transform of the vector Y and not just its component J. Section 4 covers this extension in a more general setting with multiple stochastic risk factors driving the intensity λ of J. We then obtain the joint transform of the point process and the risk factors. This yields, in particular, the transform of the intensity of the point process.

In some cases the ODEs (13) and (14) can be solved analytically—for example, when $\kappa = 0$ and J is a birth process. In the general case, the ODEs are quickly solved numerically, using the Runge–Kutta algorithm, for example.

2.5. Moments. By differentiating the conditional transform of J with respect to u and evaluating the derivative at $u = 0$, we find the conditional expectation

$$E(v \cdot J_T \,|\, \mathcal{F}_t) = \mathcal{A}(0,t,T) + \mathcal{B}(0,t,T)\lambda_t + v \cdot J_t \tag{16}$$

for $v \in \mathbb{R}^2$ and $t \leq T$, where the functions $\mathcal{A}(t) = \mathcal{A}(u,t,T)$ and $\mathcal{B}(t) = \mathcal{B}(u,t,T)$ satisfy

$$\partial_t \mathcal{B}(t) = \kappa \mathcal{B}(t) - (\delta \mathcal{B}(t) + v \cdot (1,0)^\top)\theta'(\delta b(t)) - v \cdot (0,1)^\top \theta(\delta b(t)), \tag{17}$$

$$\partial_t \mathcal{A}(t) = -\kappa c \mathcal{B}(t) \tag{18}$$

with boundary conditions $\mathcal{A}(T) = \mathcal{B}(T) = 0$. The derivative of the jump transform is

$$\theta'(\omega) = \int z e^{\omega z} d\nu(z), \quad \omega \in \mathbb{C},$$

and $a(t)$ and $b(t)$ satisfy (13) and (14) and the relevant boundary conditions. Since $u = 0$, we can choose $a(t) = b(t) = 0$. Under mild assumptions on the distribution ν (continuity of $\theta'(\omega)$ suffices), this solution is unique and (17) simplifies to

$$\partial_t \mathcal{B}(t) = (\kappa - \delta\ell)\mathcal{B}(t) - v \cdot (\ell,1)^\top, \tag{19}$$

where $\ell = \int z d\nu(z)$ is the expected loss at default. Provided that $\mu = \delta\ell - \kappa \neq 0$, we obtain the following explicit solutions:

$$\mathcal{B}(t) = \frac{1}{\mu} v \cdot (\ell,1)^\top (e^{\mu(T-t)} - 1), \tag{20}$$

$$\mathcal{A}(t) = \frac{1}{\mu} v \cdot (\ell,1)^\top \kappa c \left(\frac{1}{\mu}(e^{\mu(T-t)} - 1) - T + t \right). \tag{21}$$

For $\mu \neq 0$, the mean number of events takes the form

$$E(N_t) = c_1(e^{\mu t} - 1) + c_2 t,$$

where $c_1 = (\kappa c + \mu\lambda_0)/\mu^2$ and $c_2 = -\kappa c/\mu$. If $\kappa = 0$ and N is a birth process, then $E(N_t)$ grows exponentially in t at rate $\delta\ell > 0$. If $\kappa > 0$, then the growth of $E(N_t)$ is determined by the sign of μ. If $\mu > 0$, then we have exponential growth at rate μ that is counteracted by linear decay at rate c_2. If $\mu < 0$, then we have linear growth at rate $c_2 = \lim_{t\to\infty} E(\lambda_t)$ that is counteracted by exponential decay at rate μ.

Higher order conditional moments of J_T can be calculated similarly, by successively differentiating the transform. Thus, we obtain a closed form expression for conditional expectations of the form $E(f(J_T) \,|\, \mathcal{F}_t)$, where $f : \mathbb{R}_+ \times \mathbb{N} \to \mathbb{R}$ is an integrable function that is polynomial. We can estimate the conditional expectation $E(h(J_T) \,|\, \mathcal{F}_t)$ for any continuous, integrable function h on $\mathbb{R}_+ \times \mathbb{N}$ by approximating h with f.

2.6. Transform inversion. The transform of J_T can be inverted using fast Fourier transform techniques to obtain the conditional distribution given $\mathcal{F}_t$ of N_T and L_T for all future dates T. The distribution of N_T can also be obtained directly. Proposition 2.2 implies that the conditional probability generating function of the Hawkes process takes the form

$$E(v^{N_T - N_t} \mid \mathcal{F}_t) = \exp(c(t) + d(t)\lambda_t) \tag{22}$$

for $v \in (0, 1)$ and $t \leq T$, where the coefficient functions

$$\begin{aligned} c(t) &= c(v, t, T) = a((0, \log v)^\top, t, T), \\ d(t) &= d(v, t, T) = b((0, \log v)^\top, t, T) \end{aligned}$$

satisfy the ODEs

$$\partial_t d(t) = \kappa d(t) + 1 - v\theta(\delta d(t)), \tag{23}$$

$$\partial_t c(t) = -\kappa c d(t) \tag{24}$$

with boundary conditions $c(T) = d(T) = 0$. Equations (23)–(24) determine the distribution of the Hawkes process. Expansion of the left-hand side of (22) into a power series shows that for $n = 0, 1, 2, \ldots$

$$\varphi_t(n, T) = P(N_T - N_t = n \mid \mathcal{F}_t) = \frac{1}{n!}\partial_v^n \exp(c(v, t, T) + d(v, t, T)\lambda_t)|_{v=0}. \tag{25}$$

In practice, the calculation of these probabilities calls for the solution of a system of ODEs derived from (23)–(24). In case $n = 0$ we get

$$\varphi_t(0, T) = \exp(C(t) + D(t)\lambda_t),$$

where the functions $D(t) = d(0, t, T)$ and $C(t) = c(0, t, T)$ satisfy the ODEs

$$\begin{aligned} \partial_t D(t) &= \kappa D(t) + 1, \\ \partial_t C(t) &= -\kappa c D(t) \end{aligned} \tag{26}$$

with boundary conditions $C(T) = D(T) = 0$. We obtain the formula

$$\varphi_t(0, T) = \exp\left(\frac{c - \lambda_t}{\kappa}(1 - e^{-\kappa(T-t)}) - c(T - t)\right) \tag{27}$$

provided that $\kappa > 0$. Note that the probability (27) of no events during $(t, T]$ does not explicitly depend on the clustering parameter δ (this parameter influences only the current intensity value λ_t). It agrees with the probability of no events during $(t, T]$ for an inhomogeneous Poisson process with intensity $u(t) = c + (\lambda_0 - c)\exp(-\kappa t)$. This is because between events, the Hawkes intensity (1) is deterministic and governed by $u(t)$.

For $n \geq 1$ the probability (25) depends on the clustering parameter δ. Using Faà di Bruno's formula (chain rule for higher derivatives), we get

$$\varphi_t(n, T) = \varphi_t(0, T) \sum \frac{1}{m_1! \cdots m_n!} \prod_{k=1}^{n} \left(\frac{C_k(t) + D_k(t)\lambda_t}{k!}\right)^{m_k}, \tag{28}$$

where the sum is over all n-tuples $(m_1, \ldots, m_n)$ of nonnegative integers that satisfy the constraint $m_1 + 2m_2 + \cdots + nm_n = n$ and the functions

$$D_k(t) = \partial_v^k d(v,t,T)|_{v=0},$$
$$C_k(t) = \partial_v^k c(v,t,T)|_{v=0}$$

satisfy the ordinary differential equations

$$\partial_t D_k(t) = \kappa D_k(t) - kE_{k-1}(t), \tag{29}$$
$$\partial_t C_k(t) = -\kappa c D_k(t) \tag{30}$$

with boundary conditions $C_k(T) = D_k(T) = 0$, where $E_k(t) = \partial_v^k \theta(\delta d(v,t,T))|_{v=0}$ can be calculated by another application of Faà di Bruno's formula,

$$E_k(t) = \sum \frac{k!}{m_1! \cdots m_k!} \theta^{(m_1+\cdots+m_k)}(\delta D(t)) \prod_{i=1}^{k} \left(\frac{\delta D_i(t)}{i!} \right)^{m_i}, \tag{31}$$

and the sum is over all k-tuples $(m_1, \ldots, m_k)$ of nonnegative integers that satisfy $m_1 + 2m_2 + \cdots + km_k = k$. Equations (29) and (30) are solved numerically unless $\delta = 0$, in which case N is a time-inhomogeneous Poisson process with intensity $u(t)$. Klimko [25] provides a fast algorithm for the calculation of the summands in Faà di Bruno's formula, which can be applied to treat the computation of (28) and (31) efficiently.

3. Portfolio credit derivatives.

We use the results developed above to analyze portfolio credit derivatives, which expose an investor to the risk of correlated default in a reference portfolio of credit-sensitive instruments such as loans or bonds. We consider index and tranche swaps, which are the most popular portfolio derivatives. These are bilateral contracts, in which one party provides default protection on the reference portfolio, and the other party pays a premium for this protection. The basic valuation problem, addressed below, is to determine the fair premium, which is governed by the correlated default timing in the reference portfolio. Through a market calibration, we show that the basic intensity model (1) captures the correlated default risk implied by the premia offered in the market.

3.1. Valuation.

Index and tranche swaps are based on a portfolio whose n constituent securities have a notional 1, maturity date T, and premium payment dates (t_m). The loss at the kth default is $\ell_k \in [0,1]$. The swap is specified by a lower attachment point $\underline{K} \in [0,1]$ and an upper attachment point $\overline{K} \in (\underline{K}, 1]$. An index swap has attachment points $\underline{K} = 0$ and $\overline{K} = 1$. The swap notional $K = n(\overline{K} - \underline{K})$. The protection seller covers portfolio losses as they occur, given that the cumulative losses are larger than $\underline{K}$ but do not exceed $\overline{K}$. The cumulative payments at time t, denoted U_t, are given by the "call spread"

$$U_t = (L_t - \underline{K}n)^+ - (L_t - \overline{K}n)^+. \tag{32}$$

The value at time $t \leq T$ of these payments is given by

$$D_t = E\left(\int_t^T e^{-r(s-t)} dU_s \,\middle|\, \mathcal{F}_t \right), \tag{33}$$

where, here and below, the reference measure P is a risk-neutral pricing measure with respect to an interest rate $r > 0$. By Stieltjes integration by parts, we can conveniently express the value D_t in terms of conditional expectations of U:

$$D_t = e^{-r(T-t)} E(U_T \,|\, \mathcal{F}_t) - U_t + r \int_t^T e^{-r(s-t)} E(U_s \,|\, \mathcal{F}_t)\, ds. \tag{34}$$

The protection buyer receives the loss payments and, in return, makes premium payments to the protection seller. Each premium payment has two parts. The first part is an upfront payment, which is expressed as a fraction F of the tranche notional K. For an index swap, $F = 0$. The second part consists of payments that are proportional to the premium notional I_t, which is given by $n - (N_t \wedge n) = n + (N_t - n)^+ - N_t$ for an index swap and $K - U_t$ for a tranche swap with $\overline{K} < 1$. Let c_m be the day count fraction for the period m, roughly $1/4$ for quarterly payments. Then, with S denoting the running premium rate, the value at time $t \leq T$ of the premium payments is given by

$$P_t(F, S) = FK + S \sum_{t_m \geq t} e^{-r(t_m - t)} c_m E(I_{t_m} \,|\, \mathcal{F}_t). \tag{35}$$

For a fixed upfront rate F, the running spread S_t at time t is the solution $S = S_t$ to the equation $D_t = P_t(F, S)$. Setting $D_t = P_t(F, S)$ for a fixed S gives a value $F = F_t$ for the time t upfront rate F_t. Formulae (34) and (35) indicate that these rates depend only on call options $E((v \cdot J_s - c)^+ \,|\, \mathcal{F}_t)$ with various strikes c, maturities $s \in (t, T]$, and values $v \in \{(0,1)^\top, (1,0)^\top\}$. Thus, to value a swap we need only calculate the values of options on N and L. One approach to calculating these values is to integrate the option payoff function $(x - c)^+$ against the point process distribution. The latter is obtained by inverting the point process transform in Proposition 2.2. In the case of N, the inverse transform is given by formula (25). Alternatively, we can invert the transform of the option price, which is a function of the transform of the point process and that of the payoff function; see [26]. We may also apply the saddlepoint approximations developed by Glasserman and Kyoung-kuk [21] for affine jump diffusion models. These are applicable here since the transform of J_s is an exponentially affine function of the intensity.

The valuation formulae developed above lead to approximate tranche and index rates to the extent that the distribution of the loss process L, which does not terminate at the nth default in the portfolio, has nonnegligible mass beyond the total portfolio notional n. This is mainly relevant for super senior tranches, which depend on the tail of the loss distribution. As pointed out in section 2.2 above, the approximation tends to be accurate for the large diversified reference portfolios typical in practice.[1] Note also that in this case, the index swap spread can be expressed in terms of $E(v \cdot J_s \,|\, \mathcal{F}_t)$; from formula (16), we then obtain a closed formula for the index rate.

[1]Exact valuation results can always be obtained at an additional computational expense. To this end, we replace N by $N^n = N \wedge n$ and L by $L^n = \sum_{k=1}^{N^n} \ell_k$ in the formulae. The distribution of the process N^n, which has intensity $\lambda 1_{\{N<n\}}$, is easily calculated from the distribution of N: we have $P(N_s^n = k) = P(N_s = k)$ for all $k < n$ and $P(N_s^n = n) = P(N_s \geq n)$. The calculation of the distribution of L^n requires further steps. If, for example, the losses ℓ_k are independent and identically distributed (i.i.d.) and independent of N^n, then this amounts to a convolution operation.

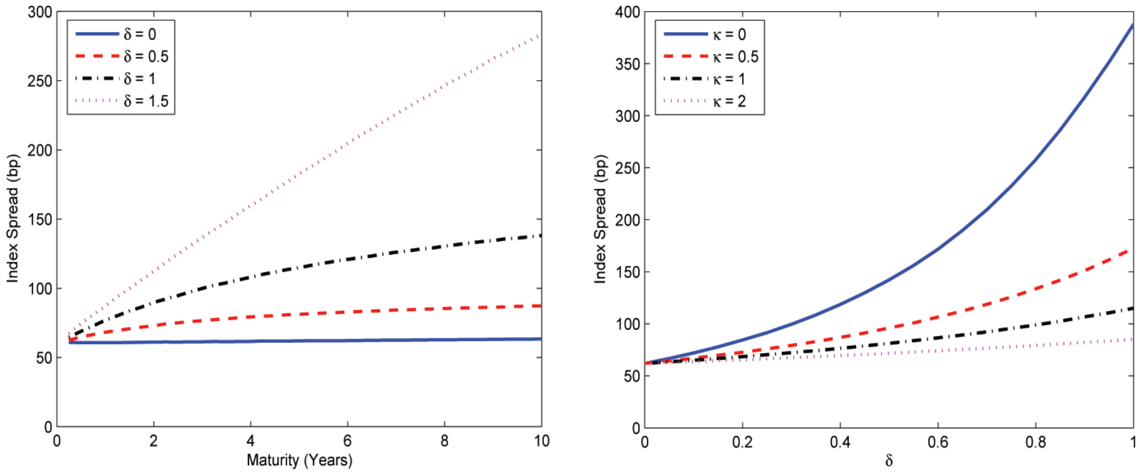

Figure 4. *Annualized premium for protection against default losses in a portfolio of* 100 *equally weighted securities as implied by the model specification* (1)*, stated in basis points* (1 *basis point* $= 10^{-4}$)*. We assume quarterly premium payments and set* $\lambda_0 = c = 1$*. The risk-free rate* $r = 5\%$*. The loss at default* ℓ_n *is uniformly distributed on* $\{0.24, 0.96\}$ *with expectation* $\ell = 0.6$*. Left panel: Index swap spread as a function of the maturity date* T *for each of several values of the clustering parameter* δ*. The decay rate* $\kappa = 1$*. Right panel:* 5*-year index spread as a function of* δ *for each of several values of* κ*.*

3.2. Numerical examples. To develop some intuition for the model parameters, we provide numerical examples of index and tranche rates. Based on the approximate index swap spread formula, the left panel of Figure 4 shows the index swap spread S_0 as a function of the maturity date T for each of several values of the clustering parameter δ. For $\delta = 0$, the portfolio loss process is a compound Poisson process, which generates a flat spread term structure. All else fixed, the higher δ, the more frequent the event clusters, the higher the spread, and the steeper the term structure. The right panel of Figure 4 shows the $T = 5$ year index swap spread as a function of δ for each of several values of the decay rate κ. For $\kappa = 0$, the portfolio loss process is a compound birth process. All else equal, the higher κ, the faster the impact of an event decays, and the smaller the chance of large losses.

The tranche swap rate increases with maturity and decreases with seniority. An increase in the lower attachment point, $\underline{K}$, leads to greater subordination (that is, buffer capital) available to the protection seller, who covers only the losses in excess of $\underline{K}$. The left panel of Figure 5 shows the upfront tranche swap rate[2] F_0 as a function of the clustering parameter δ for each of several sets of standard attachment points, assuming the running spread $S_0 = 0$. The sensitivity of the tranche rate to δ increases with the seniority of the tranche. This is because senior tranches are most exposed to default clustering, the frequency of which is controlled by δ. The bigger δ, the fatter the tail of the portfolio loss distribution. Since relatively few defaults suffice to wipe out the subordinated tranches, the frequency of clusters is less relevant for the pricing of these tranches. The right panel of Figure 5 shows the upfront tranche rate as a function of the volatility (standard deviation) of the loss at default ℓ_n. The sensitivity to the volatility of the loss increases with seniority.

[2]See section 3.3 for details on the method of computation.

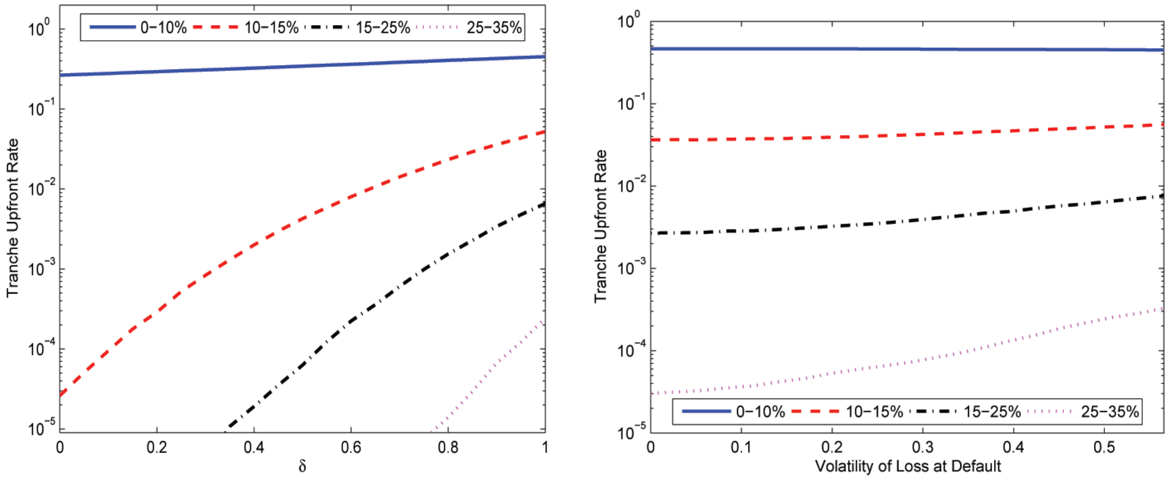

Figure 5. *Upfront premium rate for protection against default losses in a tranche of a portfolio of* 100 *equally weighted securities as implied by the model specification* (1). *We assume the running spread is equal to zero. The maturity $T = 5$ years. We set $\lambda_0 = c = 1$ and $\kappa = 1$. The risk-free rate $r = 5\%$. Left panel: Tranche upfront rate as a function of the clustering parameter δ for each of several sets of standard attachment points. The loss at default ℓ_n is uniformly distributed on $\{0.24, 0.96\}$ with expectation $\ell = 0.6$. Right panel: Tranche upfront rate as a function of the volatility of the loss at default ℓ_n for each of several sets of standard attachment points. The expected loss at default is* 0.6. *The clustering parameter $\delta = 1$.*

3.3. Market calibration. We perform market calibration experiments to illustrate the empirical importance of the self-exciting feature. To this end, we fit the parameters of the specification (1) from index and tranche swap market rates for each of the 21 trading days of September 2008. This month witnessed significant volatility due to the demise of Fannie Mae and Freddy Mac on the 8th, the default of Lehman Brothers on the 15th, the collapse of American International Group on the 16th, the problems appearing at Morgan Stanley on the 18th, and the default of Washington Mutual on the 25th. The swap market rates are for the $T = 5$ year maturity, and are obtained from UBS. They reference the CDX High Yield portfolio, which consists of 100 equally weighted names of relatively low credit quality. The tranches have attachment points $(0, 10\%)$, $(10\%, 15\%)$, $(15\%, 25\%)$, and $(25\%, 35\%)$. Thus, together with the quote for the index contract, we have five quotes per calibration date, from which we fit the parameter vector $\theta = (c, \kappa, \delta, \lambda_0)$. In accordance with market practice, we assume that the distribution of the loss at default $\nu = \delta_{0.6}$. The risk-free rate of interest r is set to the five year rate quoted on a calibration date (from Bloomberg). Swap premium payments are made quarterly. Using a gradient-based method, we numerically solve the nonlinear optimization problem

$$\min_{\theta\in\Theta} \sum_m \frac{(\mathrm{Mid}(m) - \mathrm{Model}(m,\theta))^2}{\mathrm{Mid}(m)}, \tag{36}$$

where $\Theta = (0, 8] \times [0, 10] \times [0, 20] \times (0, 30]$ and the sum ranges over the contracts. The market midquote $\mathrm{Mid}(m)$ is the arithmetic average of the market bid and ask quotes for index or tranche m. The model rate for index or tranche m is denoted $\mathrm{Model}(m,\theta)$ and is given by the formulae developed in section 3.1. The optimization is initialized at a set of parameter values drawn from a uniform distribution over the parameter space Θ and is repeated for each of 100 independent draws. The optimal parameter vector θ^* is the solution to (36) with the minimum objective function value among all 100 runs.

The index and tranche swap pricer and the optimization are implemented in MATLAB. The ODEs (13)–(14) determining the point process transform (12) are solved numerically using the Runge–Kutta method. The transform (12) is inverted numerically using the fast Fourier transform. We use 1100 sample points, a value we found to be a good tradeoff between speed and accuracy. To get a sense of the numerical accuracy of this procedure, we contrast the distribution of N generated by the fast Fourier transform with that based on formula (28). We find that the fast Fourier transform produces very accurate results. It is also quicker than the evaluation of formula (25), which requires high-precision algorithms. The computations are performed on a PC with a 2.66 GHz Intel Processor and 4 GB of RAM.

For each of the 21 calibration dates, we calculate the average absolute percentage pricing error (AAPE), given by $(1/5)\sum_{m=1}^{5}|\text{Model}(m,\theta^*) - \text{Mid}(m)|/\text{Mid}(m)$. The left panel of Figure 6 shows the time series of these errors during September 2008. The model fits the market on each date, with an AAPE over all calibration dates of 4.4% and a variance of 0.086. For comparison, the market standard copula model completely failed calibration on several days during September 2008.[3] The calibration errors are relatively low during the most volatile period, which started on September 15th with the default of Lehman. This indicates that the model captures the default correlation implied by the market prices, especially in periods of extreme stress. The calibrated values of the clustering parameter δ, shown in the right panel of Figure 6, suggest that this may be related to the self-exciting property of the default process. The time series behavior of the calibrated values of δ clearly reflects the events in mid-September. It indicates the fear of investors of a cluster of events triggered by the default of Lehman on the 15th (first peak) and the near collapse of Morgan Stanley on the 18th (second peak).

4. Affine point processes. We extend the basic Hawkes model. Our primary objective is to permit a richer structure for the intensity without reducing the computational tractability of the basic specification. The extension is required for applications in which the intensity is influenced by a set of risk factors that follow stochastic processes on their own. In the basic model (1), the point process itself is the only such risk factor. The extension facilitates the inclusion of exogenous (jump) diffusion risk factors that are relevant to arrivals. This is particularly important for empirical applications, in which the intensity model is estimated from a time series of portfolio derivative market prices, as in [3] and [28]. In these applications, the model must replicate the diffusive fluctuation of market prices, and this requires the presence of diffusive risk factors.

The idea behind the extension is to replace the intensity component λ of the Markov process $(\lambda, J)^\top$ analyzed in Proposition 2.1 by a Markov process X that represents a vector of stochastic risk factors. The process X drives the intensity of the jumps of J. The transform of $(X, J)^\top$ is computationally tractable if X is taken to be an affine jump diffusion. This

[3]That is, calibration required parameter values outside the set of admissible model parameter values. The copula model called for a correlation coefficient larger than 1, which is not meaningful.

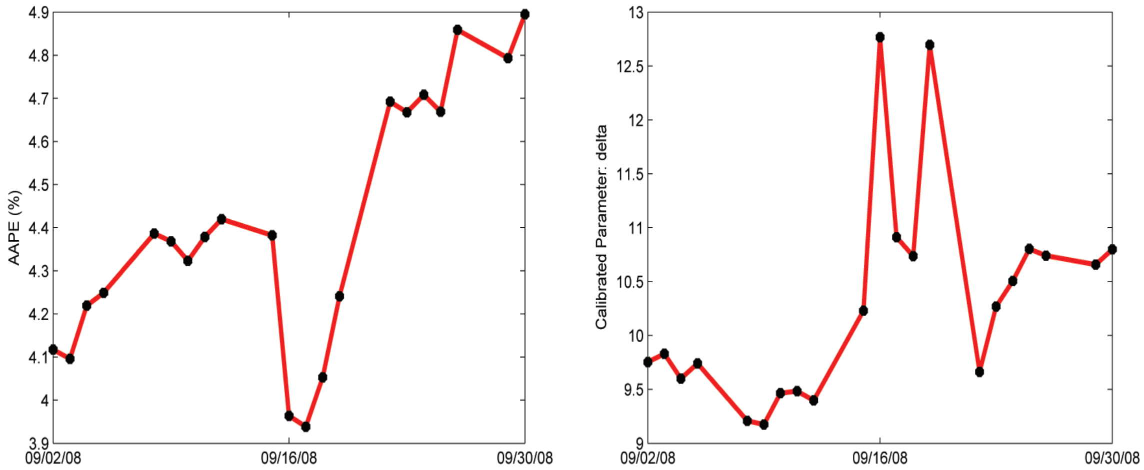

Figure 6. *Calibration results for the basic intensity model* (1) *for the* 21 *trading days of September* 2008*, for the CDX High Yield portfolio. Left panel: Average absolute percentage error (AAPE) for each calibration date. Right panel: Value of the clustering parameter* δ *for each calibration date.*

formulation leads to an affine point process.

4.1. Specification. We call a point process *affine* if its event arrival intensity is an affine function of an affine jump diffusion and its jump sizes are drawn from a fixed distribution. A Markov process X in a state space $D \subset \mathbb{R}^d \times \mathbb{R}_+$ is an affine jump diffusion in the sense of [19] if X is a strong solution to the SDE

$$dX_t = \mu(X_t, t)\,dt + \sigma(X_t, t)\,dW_t + \sum_{i=1}^{m} \zeta^i\,dZ_t^i, \qquad X_0 \in \mathbb{R}^d, \tag{37}$$

where W is an $\mathbb{R}^d$-valued standard Brownian motion, $\mu : D \to \mathbb{R}^d$ is the drift, $\sigma : D \to \mathbb{R}^{d\times d}$ is the volatility, and each Z^i is a *temporally consistent* $\mathbb{R}^d_+$-valued point process. In other words, the component processes of each vector Z^i share event times and differ only in jump sizes, and we denote their common intensity by $\lambda^i(X_t, t)$ for some $\lambda^i : D \to \mathbb{R}_+$. The jump sizes are drawn from a distribution ν^i on $\mathbb{R}^d_+$ that has no mass zero at 0. Each parameter ζ^i is a d-dimensional diagonal matrix. We assume that for each t, $\{x : (x,t) \in D\}$ contains an open subset of $\mathbb{R}^d$. For time-dependent coefficient functions that are bounded and continuous on $\mathbb{R}_+$, we assume that

$$\begin{aligned}
\mu(x,t) &= K_0(t) + K_1(t)x, \quad K_0(t) \in \mathbb{R}^d, \quad K_1(t) \in \mathbb{R}^{d\times d},\\
(\sigma(x,t)\sigma(x,t)^\top)_{jk} &= (H_0)_{jk}(t) + (H_1)_{jk}(t)\cdot x, \quad H_0(t) \in \mathbb{R}^{d\times d}, \quad H_1(t) \in \mathbb{R}^{d\times d\times d},\\
\lambda^i(x,t) &= \Lambda_0^i(t) + \Lambda_1^i(t)\cdot x, \quad \Lambda_0^i(t) \in \mathbb{R}, \quad \Lambda_1^i(t) \in \mathbb{R}^d, \quad i = 1, 2, \ldots, m.
\end{aligned}$$

Whether an affine point process J driven by X has the self-exciting property depends on the relation between J and its intensity $\lambda_t = \lambda(X_t, t)$, where $\lambda : D \to \mathbb{R}_+$. The self-exciting property holds if X, and hence λ, depends on J itself. A sufficient condition is that at least one of the component processes of J is temporally consistent with one of the component processes of one of the jump terms Z^i of X. Since the intensity is a function of the realized loss at default, we also generate a dependence structure between default and recovery rates in the self-exciting case.

4.2. Self-exciting examples. To illustrate the specification of an affine point process, we first consider the two-dimensional case $J = (L, N)^\top$. J is driven by a one-dimensional risk factor X with a single jump term Z that we identify with the component process L. The remaining component process N is temporally consistent with Z. Thus, L and N are one-dimensional affine point processes that share common event times that arrive with intensity $\lambda(X_t, t)$. By construction, the jump sizes of L are governed by the distribution ν. The jumps of N are unit-sized since N counts the arrivals. The specification of J as an affine point process is completed with the specification of the coefficient functions of the risk factor X.

Example 4.1. Suppose $K_0(t) = \kappa c$ for $\kappa \geq 0$ and $c > 0$, $K_1(t) = -\kappa$, $H_0(t)$ is a matrix of zeros, $H_1(t)$ is a tensor of zeros, $X_0 = c$, and $\zeta = \delta \geq 0$. Let $\Lambda_0(t) = 0$ and $\Lambda_1(t) = 1$. Then the intensity $\lambda = X$ of J satisfies the basic model (1):

$$d\lambda_t = \kappa(c - \lambda_t)\, dt + \delta\, dL_t.$$

A significant generalization of the basic specification is made by introducing Brownian terms and independent jump terms in the intensity. The Brownian terms model the diffusive fluctuation in the default rate and can be driven by a stochastic volatility. The jump terms model the sensitivity of the intensity to market events, such as macroeconomic shocks or defaults to names that are outside the portfolio.

Example 4.2. Suppose the coefficient functions of X are chosen as in Example 4.1, with the exception of the tensor $H_1(t)$, for which $(H_1)_{111}(t) = \sigma^2$ for $\sigma \geq 0$; all other elements are zero. Then the intensity $\lambda = X$ of J satisfies the SDE

$$d\lambda_t = \kappa(c - \lambda_t)\, dt + \sigma\sqrt{\lambda_t}\, dW_t + \delta\, dL_t, \tag{38}$$

showing that between events, the intensity drifts stochastically toward c with diffusive fluctuations driven by W. A sample path of the intensity is in Figure 7. Compare with the sample path of the basic model (1) in Figure 2.

In practice, the impact of a default on the other firms may depend on the characteristics of the defaulter. The next example allows us to distinguish between different types of firms represented in the portfolio. We consider a two-dimensional affine point process $J = (L^1, L^2)^\top$ whose components record portfolio losses triggered by defaults of two firm types. The portfolio loss process L is the sum $L^1 + L^2$. The process J is driven by a two-dimensional risk factor $X = (X^1, X^2)^\top = (\lambda^1, \lambda^2)^\top$ with two jump terms Z^1 and Z^2 whose components are identified, respectively, with L^1 and L^2. This means that both component processes of Z^i are indistinguishable from L^i.

Example 4.3. Suppose

$$K_0(t) = (\kappa^1 c^1, \kappa^2 c^2)^\top, \quad K_1(t) = \mathrm{diag}(-\kappa^1, -\kappa^2),$$

$H_0(t)$ is a matrix of zeros, $H_1(t)$ is a tensor of zeros, and

$$\zeta^1 = \mathrm{diag}(\delta^{1,1}, \delta^{2,1}), \quad \zeta^2 = \mathrm{diag}(\delta^{1,2}, \delta^{2,2}).$$

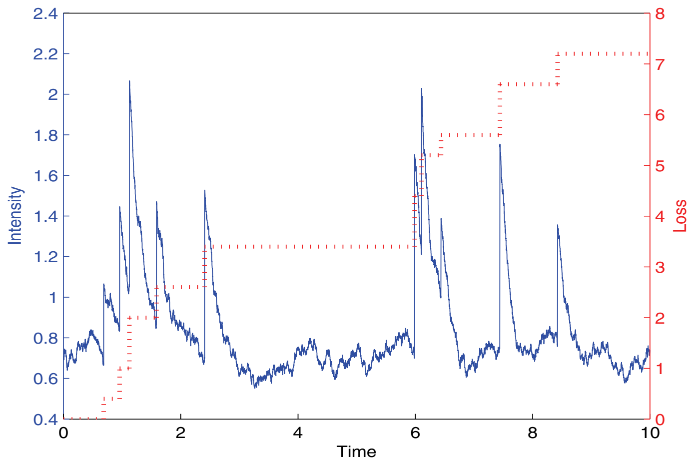

Figure 7. *Sample paths of the intensity (38) and the associated loss process L. A jump in the intensity represents the impact of a default. The jump size is the product of the loss at default and the sensitivity parameter $\delta = 1$. The loss at default is drawn from an independent uniform distribution on $\{0.4, 0.6, 0.8, 1\}$. The reversion rate $\kappa = 5$ and the reversion level $c = \lambda_0 = 0.7$. The volatility $\sigma = 0.2$ controls the diffusive fluctuation of the intensity between events.*

Let $\Lambda_0^i(t) = 0$ for $i = 1, 2$, $\Lambda_1^1(t) = (1, 0)^\top$, and $\Lambda_1^2(t) = (0, 1)^\top$. Then the intensities $\lambda^1 = X^1$ and $\lambda^2 = X^2$ satisfy the SDE

$$d\lambda_t^i = \kappa^i(c^i - \lambda_t^i)\,dt + \delta^{i,1}\,dL_t^1 + \delta^{i,2}\,dL_t^2.$$

While the parameters $\delta^{1,1}$ and $\delta^{2,2}$ control the self-excitation, the parameters $\delta^{1,2}$ and $\delta^{2,1}$ control the cross-excitation of the processes L^1 and L^2.

The preceding example illustrates the significance of allowing J to be vector-valued. The multidimensional setting facilitates the specification of (univariate) point processes whose arrivals are correlated. The dependence between the components of J can be induced by a diffusion risk factor that is common to the component intensities. It can also be generated by direct interaction terms, as in Example 4.3. In the latter case, an arrival of one component process has an impact on the intensities of other component processes. This facilitates the modeling of cross-excitation phenomena. For example, the component processes can model distinct portfolios. A default causes a loss in the respective portfolio and also has an impact on firms in the other portfolios.

4.3. Transform.

A general affine point process J is as computationally tractable as the basic Hawkes model (1). To derive a formula for the transform of J, we can apply the PIDE arguments developed in Propositions 2.1 and 2.2, which are based on the generator of the Markov process $Y = (X, J)^\top$. An alternative approach is to construct Y as an affine jump diffusion process in an enlarged state space and then to apply to this process the transform characterization for affine jump diffusions developed by [19]. This approach, pursued below, allows us to embed our point process model formulation into the affine jump diffusion setting, which is standard in many areas.

Proposition 4.4. *Suppose J is a dimension k affine point process driven by a dimension d affine jump diffusion X with state space $D^X \times \mathbb{R}_+ \subset \mathbb{R}^d \times \mathbb{R}_+$. Then $Y = (X, J)^\top$ is a dimension $(k+d)$ affine jump diffusion with state space $D^Y \times \mathbb{R}_+ \subset \mathbb{R}^{k+d} \times \mathbb{R}_+$ whose coefficients $(K_0, K_1, H_0, H_1, \Lambda_0^i, \Lambda_1^i, \zeta^i)$ and jump distributions ν^i are canonically determined by the corresponding items for X.*

In Proposition 4.4 and below, we use the same symbols for the coefficients of X and Y to avoid further complicating notation. We illustrate the idea in Example 4.3.

Example 4.5. In the setting of Example 4.3, the four-dimensional affine jump diffusion $Y = (X, J)^\top = (\lambda^1, \lambda^2, L^1, L^2)^\top$ has two (four-dimensional) jump terms Z^1 and Z^2 whose components are indistinguishable from L^1 and L^2, respectively. Since the third and fourth components of Y are driftless, the third and fourth rows of the drift coefficients are populated with zeros. These coefficients are given by

$$K_0(t) = (\kappa^1 c^1, \kappa^2 c^2, 0, 0)^\top, \quad K_1(t) = \operatorname{diag}(-\kappa^1, -\kappa^2, 0, 0).$$

The volatility coefficients $H_0(t)$ and $H_1(t)$ are zero since X has no Brownian term. The sensitivity matrices are

$$\zeta^1 = \operatorname{diag}(\delta^{1,1}, \delta^{2,1}, 1, 0), \quad \zeta^2 = \operatorname{diag}(\delta^{1,2}, \delta^{2,2}, 0, 1).$$

Finally, we give coefficients of the arrival intensities $\lambda_t^i = \Lambda_0^i(t) + \Lambda_1^i(t) \cdot Y_t$. We have $\Lambda_0^1(t) = \Lambda_0^2(t) = 0$ and

$$\Lambda_1^1(t) = (1, 0, 0, 0)^\top, \quad \Lambda_1^2(t) = (0, 1, 0, 0)^\top.$$

Proposition 4.4 facilitates the application of Proposition 1 in [19], which expresses the conditional transform of an affine jump diffusion as an exponentially affine function of its current value. Let $\mathbb{C}^n$ denote the set of n-tuples of complex numbers. When well defined at $t \le T$ and $u \in \mathbb{C}^{d+k}$, the conditional transform of the $(d+k)$-dimensional affine jump diffusion $Y = (X, J)^\top$ constructed in Proposition 4.4 is given by

$$\psi(u, Y_t, t, T) = E(\exp(u \cdot Y_T) \mid \mathcal{F}_t). \tag{39}$$

Under technical conditions that are stated in the appendix for completeness,

$$\psi(u, Y_t, t, T) = \exp(\alpha(u, t, T) + \beta(u, t, T) \cdot Y_t), \tag{40}$$

where the coefficient functions $\beta(t) = \beta(u, t, T)$ and $\alpha(t) = \alpha(u, t, T)$ satisfy the ODEs

$$\partial_t \beta(t) = -K_1(t)^\top \beta(t) - \frac{1}{2}\beta(t)^\top H_1(t)\beta(t) - \sum_{i=1}^m \Lambda_1^i(t)(\theta^i(\zeta^i \beta(t)) - 1), \tag{41}$$

$$\partial_t \alpha(t) = -K_0(t) \cdot \beta(t) - \frac{1}{2}\beta(t)^\top H_0(t)\beta(t) - \sum_{i=1}^m \Lambda_0^i(t)(\theta^i(\zeta^i \beta(t)) - 1) \tag{42}$$

with boundary conditions $\alpha(T) = 0$ and $\beta(T) = u$ and jump transforms

$$\theta^i(\omega) = \int_{\mathbb{R}_+^{d+k}} e^{\omega \cdot z} d\nu^i(z). \tag{43}$$

Under additional technical conditions stated in the appendix for completeness, the transform (40) can be differentiated with respect to u to get the formula

$$E(e^{u\cdot Y_T}\, v\cdot Y_T \,|\, \mathcal{F}_t) = \psi(u, Y_t, t, T)(A(u, v, t, T) + B(u, v, t, T)\cdot Y_t), \tag{44}$$

where $t \leq T$, $u \in \mathbb{C}^{d+k}$, and $v \in \mathbb{R}^{d+k}$, and where the coefficient functions $B(t) = B(u, v, t, T)$ and $A(t) = A(u, v, t, T)$ satisfy the ODEs

$$\partial_t B(t) = -\, K_1(t)^\top B(t) - \beta(t)^\top H_1(t) B(t) - \sum_{i=1}^m \Lambda_1^i(t) \nabla\theta^i(\zeta^i\beta(t))\cdot\zeta^i B(t), \tag{45}$$

$$\partial_t A(t) = -\, K_0(t)\cdot B(t) - \beta(t)^\top H_0(t) B(t) - \sum_{i=1}^m \Lambda_0^i(t) \nabla\theta^i(\zeta^i\beta(t))\cdot\zeta^i B(t) \tag{46}$$

with boundary conditions $A(T) = 0$ and $B(T) = v$.

As in the case of the Hawkes process, we get a closed formula for $E(v\cdot Y_T \,|\, \mathcal{F}_t)$. To see this, note that the assumption $u = 0$ implies that $\alpha = \beta = 0$. Therefore, since $\nabla\theta^i(0)$ is equal to the $(d+k)$-vector ℓ^i of expected losses of the ith jump term, (45) and (46) simplify to[4]

$$\partial_t B(t) = -K_1(t)^\top B(t) - \sum_{i=1}^m \Lambda_1^i(t) \otimes \ell^i\zeta^i B(t),$$

$$\partial_t A(t) = -K_0(t)\cdot B(t) - \sum_{i=1}^m \Lambda_0^i(t) {\ell^i}^\top \zeta^i \cdot B(t).$$

We conclude that

$$B(0, v, t, T) = v\, \exp\left(\int_t^T \left(K_1(s)^\top + \sum_{i=1}^m \Lambda_1^i(s) \otimes \ell^i\zeta^i \right) ds \right), \tag{47}$$

$$A(0, v, t, T) = \int_t^T \left(K_0(s) + \sum_{i=1}^m \Lambda_0^i(s) {\ell^i}^\top \zeta^i \right) \cdot B(s)\, ds. \tag{48}$$

The transform (40) encodes the *joint* distribution of the risk factor process X and the point process J. It facilitates applications that require the calculation of functionals of the form $E(h(X_T, J_T) \,|\, \mathcal{F}_t)$ for suitable functions h. An example application is the valuation of forward or option contracts on index and tranche swaps; see [6] and [16]. Other applications include empirical time-series estimation problems as in [3] and the calculation of the future mark-to-market exposure generated by a portfolio derivative position.

We can follow the argument in section 2.6 to characterize in terms of ODEs derived from (41)–(42) the probability distribution of an integer-valued affine point process that may be driven by a multidimensional risk factor process X. Alternatively, we can apply Fourier inversion to the transform, which would also cover multidimensional and real-valued affine point processes.

[4]The tensor product $u \otimes v$ of vectors u and v is a matrix whose ijth element is $u_i v_j$.

4.4. Further extensions. The specification can be generalized to include a stochastic discount rate that is driven by the affine jump diffusion Y. Since Y includes the affine point process J, this specification would generate a dependence structure among default, recovery, and risk-free rates. Another potential extension is the generalization of the driving risk factor process X to a more general Markov process, such as the general affine process analyzed in [18] or the linear-quadratic jump diffusion process analyzed in [10].

5. Conclusion.

We study a family of self-exciting point processes for applications in portfolio credit risk. These processes can capture the feedback from default events and the dependence structure between default and recovery rates, both features that are emphasized in the empirical literature on portfolio credit risk. The processes are also computationally tractable, because ODEs characterize their probability distribution. We illustrate this with an application to the valuation of portfolio credit derivatives, which are securities with payoffs that depend on the cumulative loss due to default in a portfolio of corporate bonds or loans. Market calibration experiments demonstrate the fit of a basic specification and highlight the empirical importance of the self-exciting feature.

Appendix. Proofs.

Proof of Proposition 2.1. Note that Y is a process with right-continuous paths of finite variation. By a change of variables for Stieltjes integrals,

$$g(\lambda_t, J_t) - g(\lambda_0, J_0) = \int_0^t g_\lambda(\lambda_s, J_s)\kappa(c-\lambda_s)ds + \sum_{0<s\le t}[g(\lambda_s, J_s) - g(\lambda_{s-}, J_{s-})].$$

We can write

$$\sum_{0<s\le t}[g(\lambda_s, J_s) - g(\lambda_{s-}, J_{s-})] = U_t^g + \int_0^t \int [g(\lambda_s + \delta z, J_s + (z,1)^\top) - g(\lambda_s, J_s)]d\nu(z)\lambda_s ds,$$

where U^g is a process defined by

$$U_t^g = \int_0^t \int [g(\lambda_{s-} + \delta z, J_{s-} + (z,1)^\top) - g(\lambda_{s-}, J_{s-})]d\nu(z)dM_s$$

and $M = N - \int_0^\cdot \lambda_s ds$ is a martingale. The integrability condition on the predictable integrand guarantees that U^g is a martingale; see Theorem 8 in Chapter II of [8]. We conclude that $g(\lambda, J)$ is a special semimartingale with unique decomposition into a sum of a predictable finite variation process and a martingale:

$$g(\lambda_t, J_t) = g(\lambda_0, J_0) + \int_0^t (\mathcal{D}g)(\lambda_s, J_s)ds + U_t^g.$$

Since U^g is a martingale, so is the process defined by $g(\lambda_t, J_t) - g(\lambda_0, J_0) - \int_0^t(\mathcal{D}g)(\lambda_s, J_s)ds$, and this yields formula (5). ■

Proof of Proposition 2.2. For fixed T and $u \in \mathbb{C}_-^2$, define the function $f : [0,T] \times D \to \mathbb{C}$ by the conditional expectation

$$f(t,\lambda,x) = E(\exp(u \cdot J_T) \,|\, \lambda_t = \lambda, J_t = x).$$

Then $E(\exp(u\cdot J_T)\,|\,\mathcal{F}_t) = f(t,\lambda_t,J_t)$. Since for $u\in\mathbb{C}^2_-$ the real and imaginary parts of the function $g(\lambda,x)=\exp(u\cdot x)$ are in the domain of the generator $\mathcal{D}$, so are the real and imaginary parts of the function $f(t,\lambda,x)$ for each time t. The function $f(t,\lambda,x)$ is also differentiable with respect to t, and the derivative is continuous; see Proposition VII.1.2 in [31]. For this, we note that $Y=(\lambda,J)^\top$ is a Feller process.

Now, for $(f(t,\lambda_t,J_t))_{t\le T}$ to be a martingale, we must have that

$$\begin{cases} f_t(t,\lambda,x)+(\mathcal{D}f)(t,\lambda,x)=0,\\ f(T,\lambda,x)=\exp(u\cdot x),\end{cases} \tag{49}$$

where $(\mathcal{D}f)(t,\lambda,x)$ is obtained by applying $\mathcal{D}$ to the real and imaginary parts of $f(t,\cdot,\cdot)$. We propose the following change of variables: $w(t,\lambda,u)=f\,(t,\lambda,x)\,e^{-u\cdot x}$. Then, after rearranging terms and denoting $\bar z=(z,1)^\top$, (49) takes the form

$$\begin{cases} w_t(t,\lambda,u)+\lambda\int\left[w\,(t,\lambda+\delta z,u)-w\,(t,\lambda,u)\right]e^{u\cdot\bar z}d\nu\,(z)\\ +\lambda\int\left(e^{u\cdot\bar z}-1\right)w\,(t,\lambda,u)\,d\nu\,(z)+\kappa\,(c-\lambda)\,w_\lambda(t,\lambda,u)=0,\\ w(T,\lambda,u)=1.\end{cases}$$

We take $w(t,\lambda,u)=e^{a(t)+b(t)\lambda}$. Then,

$$\lambda\partial_t b+\partial_t a+\lambda\int(e^{\delta bz}-1)e^{u\cdot\bar z}d\nu\,(z)+\lambda\int\left(e^{u\cdot\bar z}-1\right)d\nu\,(z)+\kappa\,(c-\lambda)\,b=0$$

so that

$$\begin{cases}\partial_t b=\kappa b-\int(e^{\delta bz+u\cdot\bar z}-1)d\nu(z),\\ \partial_t a=-\kappa cb\end{cases}$$

with boundary conditions $a(T)=b(T)=0$. ■

Proof of Proposition 4.4. The process $Y=(X,J)^\top$ is Markov since the intensity of J depends only on the current state of X, which is Markov, and the jump distribution of J is fixed. We can append the components of J one at a time. Therefore, it is sufficient to consider the case where J is one-dimensional. By assumption, its arrival intensity λ can be expressed as a time-dependent affine function $\lambda(X_t,t)=\Lambda_0(t)+\Lambda_1(t)\cdot X_t$ of X_t.

The first step is to extend the d-dimensional jump processes Z^i by one dimension. Suppose the affine point process J is self-affecting so that it is temporally consistent with some Z^i. Renumbering if necessary, assume that $i=1$ and extend Z^1 by appending the component process J. We retain the symbol ν^1 for the distribution on $\mathbb{R}^{d+1}$ that governs the jump sizes of Z^1. The extensions of the remaining Z^is are made by appending temporally consistent counting processes. In other words, the jumps in the $(d+1)$st component of $Z^2,Z^3,\dots,Z^m$ are of unit size. It follows that the (degenerate) jump distributions ν^i on $\mathbb{R}^{d+1}$ are canonically determined for $i\ge 2$.

Next, we specify the affine jump diffusion coefficients of Y. Since the intensity λ^i of Z^i is an affine function of X, it can be expressed as an affine function of Y. The one-dimensional coefficients Λ_0^i remain the same, and the d-dimensional vectors Λ_1^i are extended to dimension $(d+1)$ with a zero. The dimension $(d+1)$ sensitivity matrix of Y to Z^1 is given by $\zeta^1=\mathrm{diag}(0,0,\dots,0,1)$. We extend the $d\times d$ sensitivity matrices $\zeta^2,\zeta^3,\dots,\zeta^m$ by adding one additional row and column with all entries set to zero.

If the affine point process J is not temporally consistent with any Z^i, then all of the X^is are extended by counting processes and all the (degenerate) ν^is are canonically determined. Here, we add an $(m+1)$st jump term given by $(d+1)$ copies of J. The intensity $\lambda^{m+1} = \lambda$, and so $\Lambda_0^{m+1}(t) = \Lambda_0(t)$, and the $(d+1)$ vector $\Lambda_1^{m+1}(t)$ is obtained from the d vector $\Lambda_1(t)$ by appending a zero. Each of the m sensitivity matrices ζ^i is extended from dimension d to dimension one with a row and column of zeros. The $(d+1)$-dimensional sensitivity matrix $\zeta^{m+1} = \mathrm{diag}(0, 0, \dots, 0, 1)$.

Since J is a pure jump process, the drift coefficients K_0 and K_1 and the covariance matrix coefficients H_0 and H_1 are obtained from the corresponding coefficients for X by extending with rows and columns of zeros. ■

Based on Proposition 4.4, we apply [19, Proposition 1], which implies that the conditional transform of $Y = (X, J)^\top$ is given by (40) if the coefficient functions β and α uniquely solve the differential equations (41) and (42) and if for $(u, T) \in \mathbb{C}^{d+k} \times \mathbb{R}_+$ the following expectations are all finite:

(1) $E(|\Psi_T|)$, where $\Psi_t = \exp(\alpha(t) + \beta(t) \cdot Y_t)$.
(2) $E((\int_0^T \eta_t \cdot \eta_t \, dt)^{1/2})$, where $\eta_t = \Psi_t \beta(t)^\top \sigma(Y_t, t)$.
(3) $E(\int_0^T |\gamma_t| \, dt)$, where $\gamma_t = \Psi_t \sum_{i=1}^m (\theta^i(\zeta^i \beta(t)) - 1)\lambda^i(Y_t, t)$.

Similarly, based on Proposition 4.4, we apply [19, Proposition 3], which implies that the conditional expectation of $e^{u \cdot Y_T}(v \cdot Y_T)$ is given by formula (44) if the coefficient functions β and α uniquely solve the differential equations (41) and (42), the coefficient functions B and A uniquely solve the differential equations (45) and (46), and if for $(u, v, T) \in \mathbb{C}^{d+k} \times \mathbb{R}^{d+k} \times \mathbb{R}_+$ the following expectations are all finite:

(1) $E(|\Phi_T|)$, where $\Phi_t = \Psi_t \, (A(t) + B(t) \cdot Y_t)$.
(2) $E((\int_0^T \bar{\eta}_t \cdot \bar{\eta}_t \, dt)^{1/2})$, where $\bar{\eta}_t = \Phi_t(\beta(t)^\top + B(t)^\top)\sigma(Y_t, t)$.
(3) $E(\int_0^T |\bar{\gamma}_t| \, dt)$, where $\bar{\gamma}_t = \sum_{i=1}^m \lambda^i(Y_t, t)(\Phi_t(\theta^i(\zeta^i \beta(t)) - 1) + \Psi_t \nabla\theta^i(\zeta^i \beta(t)) \cdot \zeta^i B(t))$.

Acknowledgments. We thank Andrew Abrahams, Benjamin Armbruster, Damiano Brigo, Peter Carr, Valdo Durrleman, Andreas Eckner, Bjorn Flesaker, Igor Halperin, Steven Hutt, Peter Jäckel, Rajnish Kamat, Andrei Lopatin, Woujiang Lou, Timur Misirpashaev, George Papanicolaou, Vijay Poduri, David Soronow, Tom Stone, Pascal Tomecek, Nicolas Victoir, Stefan Weber, and Ulrich Zier for helpful discussions and comments. We also thank two anonymous referees for their helpful comments. We are extremely grateful for excellent research assistance from Xiaowei Ding, Baeho Kim, Jack Kim, and Kao-chih Syao.

REFERENCES

[1] E. Altman, B. Brady, A. Resti, and A. Sironi, *The link between default and recovery rates: Theory, empirical evidence and implications*, J. Business, 78 (2005), pp. 2203–2227.
[2] M. Arnsdorf and I. Halperin, *BSLP: Markovian bivariate spread-loss model for portfolio credit derivatives*, J. Comput. Finance, 12 (2008), pp. 77–100.
[3] S. Azizpour, K. Giesecke, and B. Kim, *Premia for correlated default risk*, Working paper, Stanford University, Stanford, CA, 2008.
[4] S. Azizpour, K. Giesecke, and G. Schwenkler, *Exploring the sources of default clustering*, Working paper, Stanford University, Stanford, CA, 2008.

[5] Basel Commission on Bank Regulation, *International convergence on capital measurement and capital standards*, Bank for International Settlements, 2004.
[6] E. Bayraktar and B. Yang, *Multi-scale time-changed birth processes for pricing multi-name credit derivatives*, Appl. Math. Finance, 16 (2009), pp. 429–449.
[7] C. Bowsher *Modelling security market events in continuous time: Intensity based, multivariate point process models*, J. Econometrics, 141 (2007), pp. 876–912.
[8] P. Brémaud, *Point Processes and Queues – Martingale Dynamics*, Springer-Verlag, New York 1981.
[9] D. Brigo, A. Pallavicini, and R. Torresetti, *Calibration of CDO tranches with the dynamical generalized Poisson loss model*, Risk, 20 (2007), pp. 70–75.
[10] P. Cheng and O. Scaillet, *Linear-quadratic jump-diffusion modeling*, Math. Finance, 17 (2007), pp. 575–598.
[11] P. Collin-Dufresne, R. Goldstein, and J. Helwege, *How large can jump-to-default risk premia be? Modeling contagion via the updating of beliefs*, Working paper, Columbia University, New York, NY, 2009.
[12] R. Cont and A. Minca, *Extracting portfolio default rates from CDO spreads*, Working paper, Columbia University, New York, NY, 2008.
[13] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes, Volume* I, Springer-Verlag, New York, 2003.
[14] S. Das, D. Duffie, N. Kapadia, and L. Saita, *Common failings: How corporate defaults are correlated*, J. Finance, 62 (2007), pp. 93–117.
[15] M. Davis and V. Lo, *Modeling default correlation in bond portfolios*, in Mastering Risk Volume 2: Applications, C. Alexander, ed., Prentice Hall, Englewood Cliffs, NJ, 2001, pp. 141–151.
[16] X. Ding, K. Giesecke, and P. Tomecek, *Time-changed birth processes and multi-name credit derivatives*, Oper. Res., 57 (2009), pp. 990–1005.
[17] D. Duffie, A. Eckner, G. Horel, and L. Saita, *Frailty correlated default*, J. Finance, 64 (2009), pp. 2089–2123.
[18] D. Duffie, D. Filipovic, and W. Schachermayer, *Affine processes and applications in finance*, Ann. Appl. Probab., 13 (2003), pp. 984–1053.
[19] D. Duffie, J. Pan, and K. Singleton, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica, 68 (2000), pp. 1343–1376.
[20] K. Giesecke, L. Goldberg, and X. Ding, *A top down approach to multi-name credit*, Oper. Res., to apear.
[21] P. Glasserman and K.-K. Kim, *Saddlepoint approximations for affine jump diffusion models*, J. Economic Dynamics and Control, 33 (2009), pp. 37–52.
[22] A. G. Hawkes *Spectra of some self-exciting and mutually exciting point processes*, Biometrika, 58 (1971), pp. 83–90.
[23] A. G. Hawkes and D. Oakes, *A cluster process representation of a self-exciting process*, J. Appl. Probab., 11 (1974), pp. 493–503.
[24] P. Jorion and G. Zhang, *Good and bad credit contagion: Evidence from credit default swaps*, J. Financial Economics, 84 (2007), pp. 860–883.
[25] E. M. Klimko, *An algorithm for calculating indices in Faa di Bruno's formula*, BIT, 13 (1973), pp. 38–49.
[26] R. W. Lee, *Option pricing by transform methods: Extensions, unification, and error control*, J. Comput. Finance, 7 (2004), pp. 51–86.
[27] F. Longstaff, *The subprime credit crisis and contagion in financial markets*, J. Finance, to appear.
[28] F. Longstaff and A. Rajan, *An empirical analysis of collateralized debt obligations*, J. Finance, 63 (2008), pp. 529–563.
[29] A. Lopatin and T. Misirpashaev, *Two-dimensional Markovian model for dynamics of aggregate credit loss*, Adv. Econometrics, 22 (2008), pp. 243–274.
[30] P. Protter, *Stochastic Integration and Differential Equations*, Springer-Verlag, New York, 2004.
[31] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer-Verlag, Heidelberg, 2005.