Self-Exciting Corporate Defaults: Contagion or Frailty?

Kay Giesecke
CreditLab
Stanford University
giesecke@stanford.edu
www.stanford.edu/~giesecke

Joint work with Shahriar Azizpour, Credit Suisse
Defaults cluster
Sources of clustering

• Firms’ exposure to observable factors: **doubly-stochastic models**
  – Defaults are conditionally independent

• Firms’ exposure to unobservable factors: **frailty models**
  – Bayesian updating of frailty distribution at events

• Firms’s exposure to default events: **contagion models**
  – Failure of a firm tends to weaken the others
  – Complex web of contractual relationships in the economy
Feedback from events

GMAC 7.75 01/19/2010 - CDS Bond Px LEH

Source: LehmanLive.com

Delphi Chapter 11
Contagion or frailty?

- Both generate similar statistical effects in conditional default rates
  - Jumps at events

- Yet they have distinct economic foundations
  - Contagion: contractual linkages among firms
  - Frailty: asymmetric information

- This paper:
  - Develop, estimate and test a model of correlated event timing that incorporates contagion and frailty
  - Understand the relative empirical importance for U.S. corporate defaults of these phenomena
Preview of empirical results

- A default is estimated to have a significant influence on the conditional default rates of surviving firms
  - Rejection of doubly-stochastic hypothesis

- Contagion and frailty are roughly equally important sources for the feedback from events

- Can “explain” the dramatic time-variation of U.S. corporate default rates during 1970–2006
Default data: 1374 events on 909 dates
Events per default date
Self-exciting model of event timing

- \((Ω, \mathcal{F}, P)\) a complete probability space and \(\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}\) a complete information filtration satisfying the usual conditions

- Events arrive at ordered stopping times \(T_n\) that generate a non-explosive counting process \(N\) with \(\mathbb{F}\)-intensity \(\lambda\)
  
  \[ N - \int_0^\cdot \lambda_s ds \text{ is an } \mathbb{F}\text{-local martingale} \]

- \(\lambda\) evolves through time according to the self-exciting model
  
  \[ d\lambda_t = \kappa(c - \lambda_t)dt + \sigma \sqrt{\lambda_t} dW_t + \delta dL_t \]

where \(W\) is an \(\mathbb{F}\)-standard Brownian motion and

\[ L = \sum_{n=0}^N \ell(D_n) \]

where \(\ell\) is a positive and bounded weight function and \(D_n \in \mathcal{F}_{T_n}\) is the number of defaults at \(T_n\)
Sample path of \((\lambda, L)\)
Default correlation channels

1. Impact of an event on the surviving firms through $\delta L$

2. Exposure to a Feller diffusion risk factor $W$

3. Uncertainty about the current value of $W$
   - This occurs when $W$ is not adapted to the observation filtration $\mathcal{G} = (\mathcal{G}_t)_{t \geq 0}$, which may be much coarser than $\mathcal{F}$
   - In this case $W$ is a frailty whose values must be filtered from the information in $\mathcal{G}$
Filtered intensity

- The econometrician’s filtered intensity $h$ of $N$ is a $\mathcal{G}$-adapted process $h$ such that $N. - \int_0^t h_s ds$ is a $\mathcal{G}$-local martingale.

- It is given by the optional projection of the complete information intensity $\lambda$ onto $\mathcal{G}$, assuming $\mathcal{G}$ is fine enough to distinguish $N$.
  - $h_t = E[\lambda_t | \mathcal{G}_t]$, almost surely.

- The filtered intensity $h$ is revised at events.
  - Contagious impact of the event (inherited from $\lambda$).
  - Bayesian updating of the $\mathcal{G}$-conditional distribution of $W$. 
Filtered likelihood

- $G$ is generated by the $(T_n, D_n)$ and a covariate process $X$
- The parameter vector to be estimated is $(\theta, \gamma, \nu)$, where $\theta = (\kappa, c, \sigma, \delta, \lambda_0, w)$ represents the intensity parameters
- For the sample period $[0, \tau]$, the likelihood function is
  \[ f_\tau(N_\tau, T; \theta \mid D, X) \cdot g(D; \gamma) \cdot p_\tau(X; \nu) \]
  - $f_\tau(\cdot, \cdot; \theta \mid D, X)$ is the conditional “density” of $N_\tau$ and $T = (T_1, \ldots, T_{N_\tau})$ given $D = (D_1, \ldots, D_{N_\tau})$ and the covariate path over $[0, \tau]$
  - $g(\cdot; \gamma)$ is the probability function of $D$
  - $p_\tau(\cdot; \nu)$ is the density of the covariate path over $[0, \tau]$
- The three terms can be maximized separately to give the full likelihood estimates
Filtered likelihood

- To evaluate the density \( f_\tau(N_\tau, T; \theta \mid D, X) \), we transform the point process \((L, \mathbb{F})\) into a standard \(\mathbb{F}\)-compound Poisson process by an equivalent change of measure.

- With the standard abuse of notation,

\[
f_\tau(N_\tau, T; \theta \mid D, X) = \hat{E}[Z_\tau^{-1} e^{-\tau} \mid N_\tau, T, D, X]
\]

where \(\hat{E}\) denotes expectation with respect to the measure \(\hat{P}\) on \(\mathcal{F}_\tau\) defined by the density \(Z_\tau\), where

\[
Z_t = \exp \left( - \int_0^t \log(\lambda_s) dN_s - \int_0^t (1 - \lambda_s) ds \right)
\]

- \(Z_\tau\) is a function of \((T_n, D_n)_{n=1,\ldots,N_\tau}\) and \(\{W_t : 0 \leq t \leq \tau\}\)
  - If \(W\) can be identified from \(X\) or \(\sigma = 0\), then

\[
f_\tau(N_\tau, T; \theta \mid D, X) = Z_\tau^{-1} e^{-\tau}
\]
Filtered likelihood

• If $W$ is a frailty that cannot be identified from $X$, then the conditional expectation is a nontrivial filter

• Since $W$ is a $(\hat{P}, \mathcal{F})$-standard Brownian motion that is $\hat{P}$-independent of $L$, we can show that

$$
\hat{E} \left[ Z_\tau^{-1} e^{-\tau} \mid N_\tau, T, D, X \right]
$$

$$
= \hat{E} \left[ \prod_{k=1}^{N_\tau} \lambda_{T_k} - \phi_{T_{k-1}, T_k} (\lambda_{T_{k-1}}, \lambda_{T_k}) \phi_{T_n, \tau} (\lambda_{T_n}, \lambda_\tau) \mid N_\tau, T, D \right]
$$

• Here, for constants $0 \leq a \leq b \leq \tau$ and positive $v$ and $w$,

$$
\phi_{a,b} (v, w) = \hat{E} \left[ \exp \left( - \int_a^b \lambda_s ds \right) \mid \lambda_a = v, \lambda_b = w \right]
$$

which can be expressed explicitly (Broadie & Kaya (2006)) since $\lambda$ follows an $\mathcal{F}$-Feller diffusion between events
Goodness-of-fit tests via time change

• We wish to assess the goodness-of-fit of a specification $(N, \lambda, G)$

• Meyer’s (1971) theorem implies that the $G$-counting process $(N, h)$ can be transformed into a standard Poisson process by a change of time that is given by the $G$-compensator $A = \int_0^t h_s ds$

• If $\lambda$ and $G$ are correctly specified, then the $(AT_n)$ form a standard Poisson process in the time-changed filtration generated by $(A_t^{-1})$

• We test the Poisson property using two tests
  – Kolmogorov-Smirnov test
  – Prahl’s (1999) test

• The tests are applied in-sample and out-of-sample
Goodness-of-fit tests via time change
Zero-factor model: $\sigma = 0$

- $\lambda$ is $\mathbb{G}$-adapted so $h = \lambda$ and likelihood is in closed form
- We solve $\sup_{\theta} \log f_{\tau}(N_\tau, T; \theta \mid D, X)$ by grid search over discretized parameter space
  - Quadratic weight function $\ell(n) = n + wn^2$ fits best
- MLEs: $\hat{\kappa} = 1.84$, $\hat{c} = \hat{\lambda}_0 = 5.48$, $\hat{\delta} = 0.43$, $\hat{w} = 0.45$
- Observations
  - An event has a significant impact on fitted default rates
  - The fitted $\lambda$ responds quickly to event bursts
  - The simple 4-parameter model captures the substantial time-series variation of default rates during 1970–2006
**Zero-factor model:** $\sigma = 0$

Fitted intensity $\lambda$ vs. events per year
Zero-factor model: $\sigma = 0$

Empirical distribution of re-scaled inter-event times
Zero-factor model: $\sigma = 0$

QQ plot of the re-scaled inter-event times vs. standard exponential
Zero-factor model: $\sigma = 0$

1Y Forecast conditional portfolio loss distribution (out-of-sample)
Zero-factor model: $\sigma = 0$

1Y Forecast conditional portfolio loss distribution vs. actual events
Zero-factor model: $\sigma = 0$

Forecast portfolio value at risk (out-of-sample)
Zero-factor model: $\sigma = 0$

$\mathcal{G}_\tau$-conditional portfolio loss surface (LGD uniform on $\{0.4, 0.6, 0.8, 1\}$)
**One-factor model:** $\sigma > 0$

- $\lambda$ is not always $\mathbb{G}$-adapted
  - **Non-informative** covariate $X$:
    $W$ is independent of $X$, and therefore not $\mathbb{G}$-adapted
    $\rightarrow$ (1) Contagion, (2) Factor exposure to $W$, (3) Frailty
  - **Informative** covariate $X$:
    $W$ is $\mathbb{G}$-adapted since it can be recovered from $X$
    $\rightarrow$ (1) Contagion, (2) Factor exposure

- We treat these cases separately to understand the relative empirical importance of contagion and frailty
One-factor model with non-informative $X$

- $\lambda$ is not $\mathcal{G}$-adapted and the likelihood must be filtered
- MLEs: $\hat{\kappa} = 1.0$, $\hat{c} = \hat{\lambda}_0 = 6.2$, $\hat{\sigma} = 3.5$, $\hat{\delta} = 0.2$, $\hat{w} = 0.5$
  - Compare with zero-factor model estimate $\hat{\delta} = 0.43$
- The filtered intensity

$$h_t = E[\lambda_t | \mathcal{G}_t] = \frac{\hat{E}[Z_t^{-1}\lambda_t | \mathcal{G}_t]}{\hat{E}[Z_t^{-1} | \mathcal{G}_t]}, \quad t \leq \tau$$

  - Jumps at $T_n$ due to contagion and Bayesian updating of the $\mathcal{G}$-conditional distribution of $W$
  - Deterministic between events
One-factor model with non-informative $X$

Filtered intensity $h_t = E[\lambda_t \mid G_t]$ vs. zero-factor intensity
One-factor model with non-informative $X$

Smoothed intensity $H_t = E[\lambda_t | G_t]$ vs. zero-factor intensity
One-factor model with non-informative $X$

Empirical distribution of re-scaled inter-event times: fit deteriorated
One-factor model with informative $X$

- $\lambda$ is $\mathcal{G}$-adapted so $h = \lambda$ (no frailty)
- We explore two covariates: S&P 500 index value, 1Y Treasury yield
  - Modeled as $\mathcal{G}$-Feller diffusions driven by $W$
  - Can recover $W$ from $X$ using covariate MLE
  - Treating the estimated $W$ as though error-free, we then estimate $\lambda$ as in the complete information case
- MLEs are similar to that of zero-factor model
- S&P 500 covariate performs slightly better than yield
One-factor model with informative $X$

Fitted intensity $\lambda$ for two covariate choices
Discussion and conclusion

- An event is estimated to have a significant impact on fitted U.S. default rates, in all model variants
  - Implications for modeling of correlated default risk
- We found that contagion and frailty are roughly equally important sources for this impact
- Feedback through contagion or frailty is necessary to fit the dramatic time variation of U.S. default rates during 1970–2006, in-sample and out-of-sample
- Feedback “explains” the excess event clustering found by Das, Duffie, Kapadia and Saita (2007) for doubly-stochastic models
Discussion and conclusion

- The simple zero-factor model ($\sigma = 0$) is hard to beat
  - The one-factor model without frailty does about as well
  - Indicates the information content of event times

- Do we need frailty in our self-exciting intensity model $\lambda$?
  - Fit is worse, the estimation is challenging
  - No new statistical features relative to complete information
  - self-exciting model $\lambda$

- Re-interpret our self-exciting $(\lambda, F)$ as filtered intensity in a frailty model in a super-filtration $\mathcal{H} \supseteq F$
  - $F$ then takes the role of the observation filtration
  - Feedback jumps of $\lambda$ can be interpreted in terms of contagion,
    or Bayesian updating of the frailty distributions
References


Prahl, Jürgen (1999), A fast unbinned test of event clustering in