Statistical Inference for Time-changed Brownian Motion Credit Risk Models

T. R. Hurd* and Zhuowei Zhou
Dept. of Mathematics and Statistics
McMaster University
Hamilton ON L8S 4K1
Canada

February 11, 2011

Abstract

We consider structural credit modeling in the important special case where the log-leverage ratio of the firm is a time-changed Brownian motion (TCBM) with the time-change taken to be an independent increasing process. Following the approach of Black and Cox, one defines the time of default to be the first passage time for the log-leverage ratio to cross the level zero. Rather than adopt the classical notion of first passage, with its associated numerical challenges, we accept an alternative notion applicable for TCBMs called "first passage of the second kind". We demonstrate how statistical inference can be efficiently implemented in this new class of models. This allows us to compare the performance of two versions of TCBMs, the variance gamma (VG) model and the exponential jump model (EXP), to the Black-Cox model. When applied to a 4.5 year long data set of weekly credit default swap (CDS) quotes for Ford Motor Co, the conclusion is that the two TCBM models, with essentially one extra parameter, can significantly outperform the classic Black-Cox model.

Key words: Credit risk, structural model, first passage problem, Lévy process, fast Fourier transform, credit default spread, maximum likelihood estimation.

AMS Subject Classification: 91G40, 91G70, 91G20, 60G35, 60G51

---

1 Introduction

Next to the Merton credit model of 1974 [12], the Black-Cox (BC) model [2] is perhaps the best known structural credit model. It models the time of a firm's default as the first passage time for the firm's log-leverage process, treated as an arithmetic Brownian motion, to cross zero. The BC model is conceptually appealing, but its shortcomings, such as the rigidity of credit spread curves, the counterfactual behaviour of the short end of the credit spread curve and the difficulty of computing correlated multifirm defaults, have been amply discussed elsewhere, see e.g. [8]. Indeed remediation of these different flaws has been the impetus for many of the subsequent developments in credit risk.

One core mathematical difficulty that has hampered widespread implementation of Black-Cox style first passage models has been the computation of first passage distributions for a richer class of processes one might want to use in modeling the log-leverage process. This difficulty was circumvented in [5], enabling us to explore the consequences of using processes that lead to a variety of desirable features: more realistic credit spreads, the possibility of strong contagion effects, and "volatility clustering" effects. [5] proposed a structural credit modeling framework where the log-leverage ratio $X_t := \log(V_t/K(t))$, where V_t denotes the firm asset value process and $K(t)$ is a deterministic default threshold, is a time-changed Brownian motion (TCBM). The time of default is the first passage time of the log-leverage ratio across zero. In that paper, the time change was quite general: our goal in the present paper is to make a thorough investigation of two simple specifications in which the time change is of Lévy type that lead to models that incorporate specific desirable characteristics. We focus here on a single company, Ford Motor Co., and show that with careful parameter estimation, TCBM models can do a very good job of explaining the observed dynamics of credit spreads. TCBMs have been used in other credit risk models, for example [13], [4], [1] and [11].

One model we study is an adaptation of the variance gamma (VG) model introduced by [10] in the study of equity derivatives, and remaining very popular since then. We will see that this infinite activity pure jump exponential Lévy model adapts easily to the structural credit context, and that the extra degrees of freedom it allows over and above the rigid structure of geometric Brownian motion correspond to desirable features of observed credit spread curves. The other model, the exponential (EXP) model, is a variation of the Kou-Wang double exponential jump model [7]. Like the VG model it is an exponential Lévy model, but now with a finite activity exponential jump distribution. We find that the EXP model performs remarkably similarly to the VG model when fit to our dataset.

We apply these two prototypical structural credit models to a dataset, divided into 3 successive 18 month periods, that consists of weekly quotes of credit default swap spreads (CDS) on Ford Motor Company. On each date, seven maturities are quoted: 1, 2, 3, 4, 5, 7, and 10 years. The main advantages of CDS data over more traditional debt instruments such as coupon bonds are their greater price transparency, greater liquidity, their standardized structure, and the fact that they are usually quoted for more maturities.

Our paper presents a complete and consistent statistical inference methodology applied to this time series of credit data, one that takes full advantage of the fast Fourier transform

to speed up the large number of pricing formula evaluations. In our method, the model parameters are taken as constants to be estimated for each 18 month time period: in contrast to "daily calibration" methods, only the natural dynamical variables, not the parameters, are allowed to be time varying.

Section 2 of this paper summarizes the financial case history of Ford Motor Co. over the global credit crisis period. Section 3 reviews the TCBM credit modeling framework introduced in [5]. There we include the main formulas for default probability distributions, defaultable bond prices and CDS spreads. Each such formula is an explicit Fourier transform representation that will be important for achieving a fast algorithm. Section 4 gives the detailed specification of the two TCBM models under study. Section 5 outlines how numerical integration of the default probability formula can be cast in terms of the fast Fourier transform. The main theoretical innovation of the paper is the statistical inference method unveiled in section 6. In this section, we argue that the naive measurement equation is problematic due to nonlinearities in the pricing formula, and that an alternative measurement equation is more appropriate. We claim that the resultant inference scheme exhibits more stable and faster performance than the naive method. In Section 7, we outline an approximate numerical scheme that implements the ideal filter of Section 6. The detailed results of the estimation to the Ford dataset are summarized in Section 8.

2   Ford: The Test Dataset

We chose to study the credit history of Ford Motor Co. over the 4.5 year period from January 2006 to June 2010. The case history of Ford over this period spanning the global credit crisis represents the story of a major firm and its near default, and is thus full of financial interest. We have also studied the credit data for a variety of other types of firm over this period, and achieved quite similar parameter estimation results. Thus our study of Ford truly exemplifies the capabilities of our modeling and estimation framework.

We divided the period of interest into three nonoverlapping successive 78 week intervals, one immediately prior to the 2007-2008 credit crisis, another starting at the outset of the crisis, the third connecting the crisis and its early recovery. We used Ford CDS and US Treasury yield data, taking only Wednesday quotes in order to remove weekday effects.

1. Dataset 1 consisted of Wednesday midquote CDS swap spreads $\widehat{\mathrm{CDS}}_{m,T}$ and their bid-ask spreads $w_{m,T}$ on dates $t_m = m/52, m = 1, \ldots, M$ for maturities $T \in \mathcal{T} := \{1, 2, 3, 4, 5, 7, 10\}$ years for Ford Motor Co., for the $M = 78$ consecutive Wednesdays from January 4th, 2006 to June 27, 2007, made available from Bloomberg.

2. Dataset 2 consisted of Wednesday midquote CDS swap spreads $\widehat{\mathrm{CDS}}_{m,T}$ and their bid-ask spreads $w_{m,T}$ on dates $t_m = m/52, m = M + 1, \ldots, 2M$ for maturities $T \in \mathcal{T} := \{1, 2, 3, 4, 5, 7, 10\}$ years for Ford Motor Co., for the $M = 78$ consecutive Wednesdays from July 11, 2007 to December 31, 2008, made available from Bloomberg.

3. Dataset 3 consisted of Wednesday midquote CDS swap spreads $\widehat{\text{CDS}}_{m,T}$ and their bid-ask spreads $w_{m,T}$ on dates $t_m = m/52, m = 2M + 1, \ldots, 3M$ for maturities $T \in \mathcal{T} := \{1, 2, 3, 4, 5, 7, 10\}$ years for Ford Motor Co., for the $M = 78$ consecutive Wednesdays from January 7th, 2009 to June 30, 2010, made available from Bloomberg.

4. The US treasury dataset[1] consisted of Wednesday yield curves (the "zero curve") on dates $t_m = m/52, m = 1, \ldots, 3M$, for maturities

$$T \in \tilde{\mathcal{T}} := \{1m, 3m, 6m, 1y, 2y, 3y, 5y, 7y, 10y, 20y, 30y\}$$

for the period January 4th, 2006 to June 30, 2010.

We note that Ford Motor Company experienced a large number of credit rating changes during this four-and-a-half year period. The history of Standard & Poors (S & P) ratings is as follows: BB+ to BB- on January 5, 2006; BB- to B+ on June 28, 2006; B+ to B on September 19, 2006; B to B- on July 31, 2008; B- to CCC+ on November 20, 2008. The downgrades continued into 2009, with a move from CCC+ to CC on March 4, 2009 and to SD ("structural default") on April 6, 2009. The latest news was good: on April 13, 2009, S & P raised Ford's rating back to CCC, on November 3, 2009 to B-, and on August 2, 2010 to B+, the highest since the onset of the credit crisis.

In hindsight we see that Ford never actually defaulted, although it came close. In the following estimation methodology, we consider the non-observation of default as an additional piece of information about the firm.

3 The TCBM Credit Setup

The time-changed Brownian motion credit framework of [5] starts with a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, which is assumed to support a Brownian motion W and an independent increasing process G where the natural filtration $\mathcal{F}_t$ contains $\sigma\{G_u, W_v : u \leq t, v \leq G_t\}$ and satisfies the "usual conditions". $\mathbb{P}$ is taken to be the physical probability measure.

Assumptions 1. 1. The log-leverage ratio of the firm, $X_t := \log(V_t/K(t)) := x + \sigma W_{G_t} + \beta\sigma^2 G_t$ is a TCBM with parameters $x > 0, \sigma > 0$ and β. The time change G_t is characterized by its Laplace exponent $\psi(u, t) := -\log \mathbb{E}[e^{-uG_t}]$ which is assumed to be known *explicitly* and has average speed normalized to 1 by the condition

$$\lim_{t \to \infty} t^{-1}\partial\psi(0, t)/\partial u = 1.$$

2. The time of default of the firm is the *first passage time of the second kind* for the log-leverage ratio to hit zero (see the definition that follows). The recovery at default is modelled by the "recovery of treasury" mechanism[2] with constant recovery fraction $R \in [0, 1)$.

---

[1]Obtained from US Federal Reserve Bank, www.federalreserve.gov/datadownload
[2]See [8].

3. The family of default-free zero-coupon bond price processes $\{B_t(T), 0 \leq t \leq T < \infty\}$ is free of arbitrage and independent of the processes W and G.

4. There is a probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$ and called the risk-neutral measure, under which all discounted asset price processes are assumed to be martingales. Under $\mathbb{Q}$, the distribution of the time change G is unchanged while the Brownian motion W has constant drift.[3] We may write $X_t = x + \sigma W_{G_t}^Q + \beta_Q \sigma^2 G_t$ for some constant β_Q where $W_u^Q = W_u + \sigma(\beta - \beta_Q)u$ is driftless Brownian motion under $\mathbb{Q}$.

We recall the definitions from [5] of first passage times for a TCBM X_t starting at a point $X_0 = x \geq 0$ to hit zero.

Definition 2.   • The standard definition of first passage time is the $\mathcal{F}$ stopping time

$$t^{(1)} = \inf\{t | X_t \leq 0\} . \tag{1}$$

The corresponding stopped TCBM is $X_t^{(1)} = X_{t \wedge t^{(1)}}$. Note that in general $X_{t^{(1)}}^{(1)} \leq 0$.

• The first passage time of the second kind is the $\mathcal{F}$ stopping time

$$t^{(2)} = \inf\{t | G_t \geq t^*\} \tag{2}$$

where $t^* = \inf\{t | x + \sigma W_t + \beta \sigma^2 t \leq 0\}$. The corresponding stopped TCBM is

$$X_t^{(2)} = x + \sigma W_{G_t \wedge t^*} + \beta \sigma^2 (G_t \wedge t^*) \tag{3}$$

and we note that $X_{t^{(2)}}^{(2)} = 0$.

The general relation between $t^{(1)}$ and $t^{(2)}$ is studied in detail in [6] where it is shown how the probability distribution of $t^{(2)}$ can approximate that of $t^{(1)}$. For the remainder of this paper, however, we consider $t^{(2)}$ to be the definition of the time of default.

The following proposition[4], proved in [5], is the basis for computing credit derivatives in the TCBM modeling framework.

Proposition 3. *Suppose the firm's log-leverage ratio X_t is a TCBM with $\sigma > 0$ and that Assumptions 1 hold.*

1. *For any $t > 0, x \geq 0$ the risk-neutral survival probability $P^{(2)}(t, x) := \mathbb{E}_x[\mathbf{1}_{\{t^{(2)} > t\}}]$ is given by*

$$\frac{e^{-\beta x}}{\pi} \int_{-\infty}^{\infty} \frac{u \sin(ux)}{u^2 + \beta^2} e^{-\psi(\sigma^2(u^2 + \beta^2)/2, t)} du + (1 - e^{-2\beta x}) \mathbf{1}_{\{\beta > 0\}}, \tag{4}$$

---

[3]This assumption can be justified by a particular version of the Girsanov theorem. It would be natural to allow the distribution of G to be different under $\mathbb{Q}$, but for simplicity we do not consider this possibility further here.

[4]Equation (6) given in [5] only deals with the case $\beta < 0$. The proof of the extension for all β is available by contacting the authors.

The density for X_t conditioned on no default is

$$\rho(y; t, x) \quad := \quad \frac{d}{dy} \mathbb{E}_x[\mathbf{1}_{\{X_t \leq y\}} | t^{(2)} > t] \tag{5}$$

$$= \quad P^{(2)}(t,x)^{-1} \mathbf{1}_{\{y>0\}} \frac{e^{\beta(y-x)}}{2\pi} \int_{\mathbb{R}} \left[e^{iu(y-x)} - e^{-iu(y+x)} \right] e^{-\psi(\sigma^2(u^2+\beta^2)/2,t)} du$$

The characteristic function for X_t conditioned on no default is

$$\mathbb{E}_x[e^{ikX_t} | t^{(2)} > t] \quad = \quad P^{(2)}(t,x)^{-1} \mathbb{E}_x[e^{ikX_t} \cdot \mathbf{1}_{\{t^{(2)}>t\}}] \tag{6}$$

$$= P^{(2)}(t,x)^{-1} \frac{e^{-\beta x}}{\pi} \int_{\mathbb{R}} \frac{u \sin(ux)}{(\beta+ik)^2 + u^2} e^{-\psi(\sigma^2(u^2+\beta^2)/2,t)} du$$

$$+ \left(e^{ikx} - e^{-ikx-2\beta x} \right) e^{-\psi(\sigma^2(k^2-2i\beta k)/2,t)} \left(\frac{1}{2} \mathbf{1}_{\{\beta=0\}} + \mathbf{1}_{\{\beta>0\}} \right)$$

2. The time 0 price $\bar{B}^{RT}(T)$ of a defaultable zero coupon bond with maturity T and recovery of treasury with a fixed fraction R is

$$\bar{B}^{RT}(T) = B(T)[P^{(2)}(T,x)) + R(1 - P^{(2)}(T,x))] \tag{7}$$

3. The fair swap rate for a CDS contract with maturity $T = N\Delta t$, with premiums paid in arrears on dates $t_k = k\Delta t, k = 1, \ldots, N$, and the default payment of $(1-R)$ paid at the end of the period when default occurs, is given by

$$CDS(x,T) = \frac{(1-R)\left[\sum_{k=1}^{N-1}[1 - P^{(2)}(t_k,x)][B(t_k) - B(t_{k+1})] + B(T)[1 - P^{(2)}(T,x)]\right]}{\Delta t \sum_{k=1}^{N} P^{(2)}(t_k,x)B(t_k)} \tag{8}$$

Remarks 4. • We shall be using the above formulas in both measures $\mathbb{P}$ and $\mathbb{Q}$, as appropriate.

• We observe in (4) that the survival and default probabilities are invariant under the following joint rescaling of parameters

$$(x, \sigma, \beta) \to (\lambda z, \lambda \sigma, \lambda^{-1}\beta), \text{ for any } \lambda > 0. \tag{9}$$

It follows that all pure credit derivative prices are invariant under this rescaling.

4 Two TCBM Credit Models

The two credit models we introduce here generalize the standard Black-Cox model that takes $X_t = x + \sigma W_t + \beta\sigma^2 t$. They are chosen to illustrate the flexibility inherent in our modeling approach. Many other specifications of the time change are certainly possible and remain to be studied in more detail. The following models are specified under the measure $\mathbb{P}$: by Assumption 1 they have the same form under the risk-neutral measure $\mathbb{Q}$, but with β replaced by β_Q.

4.1 The Variance Gamma Model

The VG credit model with its parameters $\theta = (\sigma, \beta, b, c, R)$ arises by taking G to be a gamma process with drift defined by the characteristic triple $(b, 0, \nu)_0$ with $b \in (0, 1)$ and jump measure $\nu(z) = ce^{-z/a}/z, a > 0$ on $(0, \infty)$. The Laplace exponent of G_t is

$$\psi^{VG}(u, t) := -\log E[e^{-uG_t}] = t[bu + c\log(1 + au)]. \tag{10}$$

and by choosing $a = \frac{1-b}{c}$ the average speed of the time change is $t^{-1}\partial\psi^{VG}(0, t)/\partial u = 1$. This model and the next both lead to a log-leverage process of Lévy type, that is, a process with identical independent increments that are infinitely divisible.

4.2 The Exponential Model

The EXP credit model with its parameters $\theta = (\sigma, \beta, b, c, R)$ arises taking by G to be a Lévy process with a characteristic triple $(b, 0, \nu)_0$ with $b \in (0, 1)$ and jump measure $\nu(z) = ce^{-z/a}/a, a > 0$ on $(0, \infty)$. The Laplace exponent of G_t is

$$\psi^{Exp}(u, t) := -\log E[e^{-uG_t}] = t\left[bu + \frac{acu}{1 + au}\right].$$

and by choosing $a = \frac{1-b}{c}$ the average speed of the time change is $t^{-1}\partial\psi^{VG}(0, t)/\partial u = 1$.

5 Numerical Integration

Statistical inference in these models requires a large number of evaluations of the integral formula (4) that must be done carefully to avoid dangerous errors and excessive costs. To this end, we approximate the integral by a discrete Fourier transform over the lattice

$$\Gamma = \{u(k) = -\bar{u} + k\eta | k = 0, 1, \ldots, N - 1\}$$

for appropriate choices of $N, \eta, \bar{u} := N\eta/2$. It is convenient to take N to be a power of 2 and lattice spacing η such that truncation of the u-integrals to $[-\bar{u}, \bar{u}]$ and discretization leads to an acceptable error. If we choose initial values x_0 to lie on the reciprocal lattice with spacing $\eta^* = 2\pi/N\eta = \pi/\bar{u}$

$$\Gamma^* = \{x(\ell) = \ell\eta^* | \ell = 0, 1, \ldots, N - 1\}$$

then the approximation is implementable as a fast Fourier transform (FFT):

$$P^{(2)}(t, x(\ell)) \sim \frac{-i\eta e^{-\beta x(\ell)}}{\pi} \sum_{k=0}^{N-1} \frac{u(k)e^{iu(k)x(\ell)}}{u(k)^2 + \beta^2} \exp[-\psi(\sigma^2(u(k)^2 + \beta^2)/2, t)] \tag{11}$$

$$= -i(-1)^n \eta e^{-\beta x(\ell)} \sum_{k=0}^{N-1} \frac{u(k)e^{2\pi i k\ell/N}}{u(k)^2 + \beta^2} \exp[-\psi(\sigma^2(u(k)^2 + \beta^2)/2, t)] \tag{12}$$

7

Note that we have used the fact that $e^{-iN\eta x(\ell)/2} = (-1)^n$ for all $\ell \in \mathbb{Z}$.

The selection of suitable values for N and η in the above FFT approximation of (8) is determined via general error bounds proved in [9]. In rough terms, the pure truncation error, defined by taking $\eta \to 0, N \to \infty$ keeping $\bar{u} = N\eta/2$ fixed, can be made small if the integrand of (4) is small and decaying outside the square $[-\bar{u}, \bar{u}]$. Similarly, the pure discretization error, defined by taking $\bar{u} \to \infty, N \to \infty$ while keeping η fixed, can be made small if $e^{-|\beta|\bar{x}}P^{(2)}(\bar{x}, t)$, or more simply $e^{-|\beta|\bar{x}}$, is small, where $\bar{x} := \pi/\eta$. One expects that the combined truncation and discretization error will be small if $\bar{u}$ and $\eta = \pi/\bar{x}$ are each chosen as above. These error bounds for the FFT are more powerful than bounds one finds for generic integration by the trapezoid rule, and constitute one big advantage of the FFT. A second important advantage to the FFT is its $O(N \log N)$ computational efficiency that yields $P^{(2)}$ on a lattice of x values with spacing $\eta^* = 2\pi/N\eta = \pi/\bar{u}$: this aspect will be very useful in estimation. These two advantages are offset by the problem that the FFT computes values for x only on a grid.

We now discuss choices for N and η in our two TCBM models. For $\beta < 0$, the survival function of the VG model is

$$P^{(2)}(0, t, x, \beta) = \frac{e^{-\beta x}}{\pi} \int_{-\infty}^{\infty} \exp[-tb\sigma^2(u^2 + \beta^2)/2] \left(1 + \frac{a\sigma^2(u^2 + \beta^2)}{2}\right)^{-ct} \frac{u \sin ux}{u^2 + \beta^2} du$$

while for the EXP model

$$P^{(2)}(0, t, x, \beta) = \frac{e^{-\beta x}}{\pi} \int_{-\infty}^{\infty} \exp\left[-t\left(b\sigma^2(u^2 + \beta^2)/2 + \frac{ac\sigma^2(u^2 + \beta^2)}{2 + a\sigma^2(u^2 + \beta^2)}\right)\right] \frac{u \sin ux}{u^2 + \beta^2} du$$

In both models, the truncation error has an upper bound ϵ when $\bar{u} > C|\Phi^{-1}(\epsilon C')|$, where Φ^{-1} is the inverse normal CDF and C, C' are constants depending on t. On the other hand, provided $\beta < 0$, the discretization error will be small (of order ϵ or smaller) if $N > \frac{\bar{u}}{2\pi|\beta|} \log\left(\epsilon^{-1}(1 + \exp(-2\beta x))\right)$. Errors for (6) can be controlled similarly.

6   The Statistical Method

The primary aim of this exercise is to demonstrate that our two TCBM credit models can be successfully and efficiently implemented to fit market CDS data on a single firm, in this case Ford Motor Company, and to compare these models' performance to the original Black-Cox structural model.

We were able to reduce the complexity of our models with negligible loss in accuracy by removing what appear to be two "nuisance parameters". First, we expect, and it was observed, that parameter estimations were not very sensitive to β near $\beta = 0$, so we arbitrarily set $\beta = -0.5$. Secondly, we observed insensitivity to the parameter b and a tendency for it to drift slowly to zero under maximum likelihood iteration: since $b = 0$ is a singular limit, we set $b = 0.2$. Finally, in view of the rescaling invariance (9), and the interpretation of σ as the volatility of X, without loss of generality we set $\sigma = 0.3$ in all models. So specified, the two TCBM models have three free parameters $\Theta = (c, \beta_Q, R)$ as well as three frozen parameters $\sigma = 0.3, \beta = -0.5, b = 0.2$. The Black-Cox model with

its free parameters $\Theta = (\beta_Q, R)$ and frozen parameters $\sigma = 0.3, \beta = -0.5$ then nests as the $c = 0$ limit inside both the VG and EXP models.

We summarize the modeling ingredients:

- an unobserved Markov process $X_t \in \mathbb{R}^d$;

- model parameters $\Theta \in D \subset \mathbb{R}^n$. We augment the vector $\Theta \to (\Theta, \eta)$ to include an additional measurement error parameter η;

- model formulas $F^k(X, \Theta)$ for $k = 1, \ldots, K$, which in our case are theoretical CDS spreads given by (8) for $K = 7$ different tenors;

- a dataset consisting of spreads $Y := \{Y_t\}$ observed at times $t = 1, \ldots, M$ where $Y_t = \{Y_t^k\}$ for a term structure of $k = 1, \ldots, K$, plus their associated quoted bid/ask spreads w_t^k. We use notation $Y_{\leq t} := \{Y_1, \ldots, Y_t\}$ and $Y_{<t} := \{Y_1, \ldots, Y_{t-1}\}$ etc.

Since we do not attempt to estimate an underlying interest rate model, we treat the US Treasury dataset as giving us exact information about the term structure of interest rates, and hence the discount factors entering into (8). We treat the quoted bid/ask spreads w_t^k as a proxy for measurement error: these will simplify our treatment of the measurement equation. We also treat the non-default status of Ford on each date as an additional observation.

To complete the framework, an arbitrary Bayesian prior density of Θ is taken

$$\rho_0(\Theta) := e^{\mathcal{L}_0(\Theta)}.$$

with support on $D \subset \mathbb{R}^{n+1}$. The statistical method appropriate to a problem like this is thus some variant of a nonlinear Kalman filter, combined with maximum likelihood parameter estimation.

Based on these assumptions, it is rather natural to assume that observed credit spreads provide measurements of the hidden state vector X with independent gaussian errors. Moreover the measurement errors may be taken proportional to the observed bid/ask spread. Thus a natural measurement equation is

$$Y_t^k = F^k(X_t, \Theta) + \eta w_t^k \zeta_t^k \tag{13}$$

where ζ_t^k are independent standard gaussian random variables and η is constant. In this case the full measurement density of Y would be

$$\mathcal{F}(Y|X, \Theta) = \prod_{t=1,\ldots,M} \prod_{k=1,\ldots,K} \left[\frac{1}{\sqrt{2\pi}\eta w_t^k} \exp\left(-\frac{(Y_t^k - F^k(X_t, \Theta))^2}{2\eta^2 (w_t^k)^2} \right) \right] \tag{14}$$

However, we observed an important deficiency that seems to arise in any scheme like this where the measurement equation involves a nonlinear function of an unobserved process X. This nonlinearity leads to nonconvexity in the log-likelihood function for X, which in turn can destabilize the parameter estimation procedure. For such reasons, we instead follow an alternative scheme that in our problem, and perhaps many others of this

type, gives a great improvement in estimation efficiency. It works in our case because the model formula (8) for $F^k(x, \Theta)$, although nonlinear in x, is monotonic and approximately linear in x. We will call our scheme the "linearized measurement" scheme and it is justified as follows.

We define $G^k(Y, \Theta)$ to be the solution x of $Y = F^k(x, \Theta)$, and note that $f^k := \partial_x F^k > 0$. Then, provided ηw^k are small enough, we may linearize the x dependence of the measurement equation using the Taylor expansion

$$
\begin{aligned}
Y^k - F^k(x) &= Y^k - F^k(G^k(Y^k) + x - G^k(Y^k)) \\
&\approx Y^k - F^k(G^k(Y^k)) + f^k(G^k(Y^k))(G^k(Y^k) - x) \\
&= f^k(G^k(Y^k))(G^k(Y^k) - x)
\end{aligned}
$$

This equation above justifies the following alternative to the measurement equation (13):

$$
\tilde{X}_t^k = X_t + \eta \tilde{w}_t^k \xi_t^k \tag{15}
$$

Now $\xi_t^k, k = 1, 2, \ldots, K, t = 1, 2, \ldots, M$ are iid $N(0,1)$ random variables and the transformed measurements are

$$
\tilde{X}_t^k = \tilde{X}^k(Y_t^k, \Theta) := G^k(Y_t^k, \Theta).
$$

Furthermore,

$$
\tilde{w}_t^k = \tilde{w}_t^k(\tilde{X}_t^k, \Theta) = f^k(\tilde{X}_t^k, \Theta)^{-1} w_t^k.
$$

Note that $\tilde{X}_t^k, k = 1, \ldots, K$ have the interpretation as independent direct measurements of the unobserved state value X_t.

The full measurement density of Y in our linearized measurement scheme is thus:

$$
\mathcal{F}(Y|X, \Theta) := \prod_{t=1,\ldots,M} f(Y_t|X_t, \Theta) \tag{16}
$$

$$
f(Y_t|X_t, \Theta) := \prod_{k=1,\ldots,K} \left[\frac{1}{\sqrt{2\pi}\eta w_t^k} \exp\left(-\frac{(\tilde{X}_t^k(Y_t^k, \Theta) - X_t)^2}{2\eta^2 \tilde{w}_t^k(\Theta)^2} \right) \right] \tag{17}
$$

where we have recombined denominator factors of $\tilde{w}^k$ with Jacobian factors $(f^k)^{-1}$. The multiperiod transition density conditioned on nondefault is

$$
\mathcal{P}(X|\Theta, \text{no default}) = \prod_{t=2,\ldots,M} p(X_t|X_{t-1}, \Theta) \tag{18}
$$

where $p(y|x, \Theta)$ is the one period conditional transition density given by (5) with $t = \Delta t$. Finally the full joint density for (X, Y, Θ) is

$$
\rho(X, Y, \Theta) := \mathcal{F}(Y|X, \Theta)\mathcal{P}(X|\Theta)\rho_0(\Theta) \tag{19}
$$

10

Integration over the hidden state variables X leads to the partial likelihood function, which can be defined through an iteration scheme:

$$\rho(Y,\Theta) = \int f(Y_M|X_M,\Theta)\rho(X_M,Y_{<M},\Theta)dX_M \qquad (20)$$

where for $t < M$

$$\rho(X_{t+1},Y_{\leq t},\Theta) = \begin{cases} \int p(X_{t+1}|X_t,\Theta)f(Y_t|X_t,\Theta)\rho(X_t,Y_{<t},\Theta)dX_t, & t > 0 \\ \\ \rho_0(\Theta) & t = 0 \end{cases} \qquad (21)$$

The following summarizes statistical inference within the linearized measurement scheme.

Statistical Inference using the Linearized Measurement Scheme: Let $(\hat{Y},w) := \{\hat{Y}_t^k, w_t^k\}$ be the time series of CDS observations.

1. Maximum Likelihood Inference: The maximum likelihood parameter estimates $\widehat{\Theta}$ are the solutions of

$$\widehat{\Theta} = \text{argmax}_{\Theta \in D} \log\left(\rho(\hat{Y},\Theta)/\rho_0(\Theta)\right) \qquad (22)$$

where $\rho(\hat{Y},\Theta)$ is given by (20). The log-likelihood achieved by this solution is

$$\widehat{\mathcal{L}} := \log\left(\rho(\hat{Y},\widehat{\Theta})/\rho_0(\widehat{\Theta})\right),$$

and the Fisher information matrix is

$$\widehat{\mathcal{I}} := -\left[\partial_\Theta^2 \log\left(\rho(\hat{Y},\widehat{\Theta})/\rho_0(\widehat{\Theta})\right)\right];$$

2. Filtered State Inference: The time series of filtered estimates of the state variables $X_1,\ldots,X_M$ are the solutions $\hat{X}_1,\ldots,\hat{X}_M$ of

$$\hat{X}_t = \text{argmax}_{x \in \mathbb{R}_+} \log\left(f(\hat{Y}_t|x,\widehat{\Theta})\rho(x,\hat{Y}_{\leq t-1},\widehat{\Theta})\right) \qquad (23)$$

7 Approximate Inference

The previous discussion on inference was exact, but computationally infeasible. Our aim now is to give a natural and simple approximation scheme that will be effective for the problem at hand. Our scheme is to inductively approximate the likelihood function $\rho(X_{t+1},Y_{\leq t},\Theta)$ defined by (21) by a truncated normal distribution through matching of the first two moments. The truncation point of 0 is determined by the no default condition. The rationale is that the non-gaussian nature of the transition density p will have only a small effect when combined with the gaussian measurement density f. We expect our approximation to be appropriate for a firm like Ford that spent a substantial period near default. As we discuss at the end of this section, a simpler approximation is available

11

that is applicable to a firm of high credit quality. The more complicated method we now describe is intended to be more robust when applied to firms of a range of credit qualities.

We describe a single step of the inductive computation of ρ given by (21). We fix t, denote the time t state variable as x and the time $t+1$ state variable as capital X. The length between t and $t+1$ is denoted as Δt. We also suppress $Y_{\leq t}$ and Θ. In this context, we are looking for $\bar{\mu}$ and $\bar{\sigma}$ that satisfy

$$\rho(X) \approx \frac{m_0 \phi\left(\frac{X-\bar{\mu}}{\bar{\sigma}}\right)}{\Phi\left(\frac{\bar{\mu}}{\bar{\sigma}}\right)}, X > 0 \tag{24}$$

where

$$m_0 = \int_0^\infty f(x)\tilde{\rho}(x)dx. \tag{25}$$

Here ϕ and Φ are probability density and cumulative distribution functions of the standard normal distribution and $\tilde{\rho}$ is carried over from the previous time step. The first two moments of the truncated normal distribution are straightforward to derive and are given here for completeness:

$$\begin{aligned} m_1^{trunc} &= \bar{\mu} + \bar{\sigma}\lambda(\alpha) \\ m_2^{trunc} &= \bar{\sigma}^2[1 - \delta(\alpha)] + (m_1^{trunc})^2 \end{aligned}$$

where $\alpha = -\frac{\bar{\mu}}{\bar{\sigma}}, \lambda(\alpha) = \frac{\phi(\alpha)}{1-\Phi(\alpha)}, \delta(\alpha) = \lambda(\alpha)[\lambda(\alpha) - \alpha]$. Note that the truncated normal distribution has a larger mean and smaller variance than the original normal distribution.

Using the Fubini theorem, the first two moments of the distribution $\rho(X)$ are:

$$\begin{aligned} m_1 &= m_0^{-1} \int_0^\infty g_1(x)f(x)\tilde{\rho}(x)dx \tag{26} \\ m_2 &= m_0^{-1} \int_0^\infty g_2(x)f(x)\tilde{\rho}(x)dx \end{aligned}$$

Here $g_1(x)$ and $g_2(x)$ are the first and second moments of X with respect to the transition density $p(X|x)$ and are given using (6) by

$$\begin{aligned} g_1(x) &= \frac{1}{i}\partial_k|_{k=0}\mathbb{E}_x^P[e^{ikX}|t^{(2)} > \Delta t] \tag{27} \\ g_2(x) &= -\partial_k^2|_{k=0}\mathbb{E}_x^P[e^{ikX}|t^{(2)} > \Delta t] \end{aligned}$$

Note that $\tilde{\rho}(x)$ has a gaussian kernel approximation by induction and the measurement density $f(x)$ is also gaussian. Their product gaussian kernel is then simply a scaled normal probability density function:

$$f(x)\tilde{\rho}(x) = \frac{m_0}{\sqrt{\bar{v}}\Phi(\bar{m}/\sqrt{\bar{v}})}\phi\left(\frac{x-\bar{m}}{\sqrt{\bar{v}}}\right) \tag{28}$$

We also notice that the transition density $p(X|x)$ with a short period Δt resembles a Dirac δ function of X and fitting it to a polynomial would require very high order to

12

guarantee accuracy in a local domain. In our method, by contrast, the moment functions $g_1(x)$ and $g_2(x)$ that appear in the integrals in (26) are much smoother functions of x and usually low order polynomials can approximate them quite accurately in a local domain. Take a normal transition density for example: $g_1(x)$ is linear in x and $g_2(x)$ is quadratic in x. Their counterparts for time changed Brownian motion conditional on no default can also be well approximated by low order polynomials in a local domain. We stress the word "local" because the product gaussian kernel $f\rho$ typically has a moderate variance $\bar{v}$ and relatively large mean $\bar{m}$: therefore the integrals in Equation (26) are dominated by a local domain $[\bar{m} - a\sqrt{\bar{v}}, \bar{m} + a\sqrt{\bar{v}}]$ with a safely taken to be 4. Thus we need to fit $g_1(x)$ and $g_2(x)$ over the interval $[\bar{m} - a\sqrt{\bar{v}}, \bar{m} + a\sqrt{\bar{v}}]$ which can be done quite accurately with quartic polynomials:

$$
\begin{aligned}
g_1(x) &= \Sigma_{k=0}^4 c_{1k}(x - \bar{m})^k \\
g_2(x) &= \Sigma_{k=0}^4 c_{2k}(x - \bar{m})^k.
\end{aligned}
\tag{29}
$$

Equation (26) is now approximated by

$$
\begin{aligned}
m_1 &= \frac{1}{m_0\sqrt{\bar{v}}\Phi(\bar{m}/\sqrt{\bar{v}})} \int_0^\infty \Sigma_{k=0}^4 c_{1k}(x - \bar{m})^k \phi\left(\frac{x - \bar{m}}{\sqrt{\bar{v}}}\right) dx \\
m_2 &= \frac{1}{m_0\sqrt{\bar{v}}\Phi(\bar{m}/\sqrt{\bar{v}})} \int_0^\infty \Sigma_{k=0}^4 c_{2k}(x - \bar{m})^k \phi\left(\frac{x - \bar{m}}{\sqrt{\bar{v}}}\right) dx
\end{aligned}
$$

which can be evaluated analytically in terms of the error function. Matching m_1 and m_2 with m_1^{trunc} and m_2^{trunc} determines $\bar{\mu}$ and $\bar{\sigma}$ and completes the iteration scheme for (20).

Remarks 5.

- In our numerical examples, we enlarge the integral domain in Equation (30) from $\mathbb{R}^+$ to $\mathbb{R}$ if $\bar{m} > 4\sqrt{\bar{v}}$, which leads to a simpler implementation. It turns out in our study that this condition is satisfied for all sampling periods.

- An alternative moment matching approximation is possible which approximates $\rho(X)$ by a regular normal distribution, rather than a truncated normal. Then the truncated density in Equation (24) should be replaced by the regular density $\phi\left(\frac{X-\bar{\mu}}{\bar{\sigma}}\right), X \in \mathbb{R}$. Although this approximation conflicts with the default barrier, for a firm that is far from default this does not introduce a serious numerical error. Moreover, this approximation leads to linear gaussian transition density and is thus a Kalman filter.

Here we summarize the computation of $\rho(Y_{\leq M}, \Theta)$ for a fixed value of Θ:

1. Set $\rho_1 = \rho_0(\Theta)$;

2. Compute the measurement density $f(Y_1|X_1)$ (i.e. compute its mean and variance: this step requires efficient use of the FFT to invert the CDS spread formula);

3. For $t = 1 : M - 1$

(a) Approximate $\rho(X_{t+1}, Y_{\leq t}, \Theta)$ given by (21) by a truncated normal density with mean and variance computed by matching moments. For this one uses the exact formula for the first two moments of the conditional transition density (5), and the assumed normal form of $f(Y_t|X_t)$ and $\rho(X_t, Y_{<t}, \Theta)$;

(b) Compute the measurement density $f(Y_{t+1}|X_{t+1})$ (ie. compute its mean and variance, again with efficient use of FFT);

(c) End loop;

4. Finally compute $\rho(Y_{\leq M}, \Theta)$ by integrating X_M as in (20).

8 Numerical Implementation

From the considerations described in section 6 we fix $\beta = -0.5, \sigma = 0.3, b = 0.2$. We choose $\bar{u} = 300$ which controls the truncation error within 10^{-10}. Depending on Θ, we allowed the size of the FFT lattice, N, to vary from 2^8 to 2^{10}, keeping the discretization error within 10^{-10}. We use the Matlab function *fmincon* to implement the quasi-Newton method to maximize the likelihood function. Since *fmincon* also calculates the gradient and Hessian of the objective function, we also obtain standard errors of the parameter estimates.

Table 1 summarizes the estimation results for each of the three models, for the three datasets in 2006-2010, using our time series approximate inference. Estimated parameter values are given with standard errors, as well as summary statistics for the resulting filtered time series of X_t. We also present the root mean square error (RMSE) defined as the average error of the CDS spreads quoted in units of the bid/ask spread.

$$\text{RMSE} = \sqrt{\frac{1}{M \cdot K} \sum_{t=1}^{M} \sum_{k=1}^{K} \frac{\left(F^k(X_t, \Theta) - Y_t^k\right)^2}{(w_t^k)^2}}$$

Overall, the finite activity EXP model shares quite a few similarities with the infinite activity VG model, both in behavior and performance. For these two TCBM models, their model parameters are quite similar between dataset 1 and dataset 3 respectively. It is consistent with Ford's history of credit ratings that dataset 3 has lower, more volatile log-leverage ratios and lower recovery rate than dataset 1. We can also see that during the peak of the credit crisis in dataset 2, the estimated parameters show noticeable signs of stress. The mean time change jump size is up by approximately 50%, driven mainly by the increased short term default probability. The recovery rate is significantly lower. In the very stressed financial environment at that time, a firm's value would be greatly discounted and its capacity to liquidate assets would be limited. On the other hand the risk neutral drift β_Q is significantly higher, reflecting a certain positive expectation on the firm. At the peak of the credit crisis, Ford's annualized credit spreads exceeded 100%. The log-leverage ratios are much suppressed to a level of about 65% of that of dataset 1.

By definition, RMSE measures the deviation of the observed CDS spreads from the model CDS spreads while η measures the deviation of the "observed" log-leverage ratios

$\tilde{X}_t$ from the "true" log-leverage ratios X_t. We can see that RMSE and η are very close in all cases, which implies that the objective functions based on the naive CDS measurement density (14) and the linearized measurement density (16) are fundamentally very similar.

In terms of RMSE and η, both TCBM models performed much better than the Black-Cox model. The TCBM fitting is typically within two times the bid/ask spread across 3 datasets, while the errors of the Black-Cox model are about 30% higher on average. Figure 1 shows that on three typical days, the TCBM models can fit the market CDS term structure curves reasonably well while the Black-Cox model, with its restrictive hump-shaped term structures, has difficulties for some tenors. To fit high short spreads, the log-leverage ratio is forced to unreasonably low levels. The TCBM models, with only one extra parameter than the Black-Cox model, generate more flexible shapes, and do a better job of fitting the data.

Figure 2 displays histograms of the signed relative error $(w_t^k)^{-1}\big(F^k(X_t, \Theta) - Y_t^k\big)$ for the three models, for the short and long end of the term structure. For both TCBM models we can see that most errors are bounded by ± 2 and are without obvious bias. By comparison, the errors of the Black-Cox model are highly biased downward in the both the short and long terms. For 1-year spreads the majority of errors stay near -2 and for 10-year spreads there is a concentration of errors near -4. Surprisingly, all the three models perform better and more closely to one another during the crisis period of dataset 2. For the TCBM models, the great majority of errors are near 0 and without obvious bias. The Black-Cox model does not have obvious bias either, but there are more errors beyond the ± 2 range. The performance of all three models is better for intermediate tenors between 1 and 10 years, with the mid-range 5-year and 7-year tenors having the best fit. The histograms for these tenors (not shown) do still indicate that the TCBM models perform better than the Black-Cox model, in regard to both bias and absolute error.

The estimation results (not shown here) using the Kalman filter method described in Remarks 5 are very close to the results shown in Table 1, indicating that the transition density can be safely approximated by a gaussian density. The Kalman filter is convenient for calculating the weekly likelihood function, which is needed in the Vuong test [15], a test to compare the relative performance of nested models. If $\bar{X}_t$ and $\bar{P}_t$ denote the ex-ante forecast and variance of time t values of the measurement series obtained from Kalman filtering, the weekly log-likelihood function can be written as

$$l_t = -\frac{1}{2}\log|\bar{P}_t| - \frac{1}{2}(\tilde{X}_t - \bar{X}_t)^\top(\bar{P}_t)^{-1}(\tilde{X}_t - \bar{X}_t) - \sum_k f^k(\tilde{X}_t^k, \Theta). \tag{30}$$

The log-likelihood ratio between two models i and j is

$$\lambda_{ij} = \sum_{t=1}^{M}(l_{it} - l_{jt})$$

and the Vuong test statistic is

$$\mathcal{T}_{ij} = \frac{\lambda_{ij}}{\hat{s}_{ij}\sqrt{M}},$$

15

where $\hat{s}_{ij}^2$ is the variance of $\{l_{it} - l_{jt}\}_{t=1,...,M}$. Vuong proved that $\mathcal{T}_{ij}$ is asymptotic to a standard normal under the null hypothesis that models i and j are equivalent in terms of likelihood function. Due to the serial correlation within the log-likelihood functions, Newey and West's estimator [14] is used for $\hat{s}$. The Vuong test results are shown in Table 2 and confirm that the Black-Cox model is consistently outperformed by the two TCBM models. Moreover, by this test, the EXP model shows an appreciable improvement over the VG model that could not be easily observed in the previous comparison.

It is interesting to compare the time series of Ford stock prices to the filtered log-leverage ratios X_t. Fig 3 shows there is a strong correlation between these two quantities, indicating that the equity market and credit market are intrinsically connected. The empirical observations supporting this connection and thereafter financial modeling interpreting this connection can be found in [11], [3] and their references.

Finally, we mention that a stable model estimation over a 78 week period typically involved about 120 evaluations of the function $\rho(Y, \Theta)$, and took around one minute on a standard laptop.

9 Conclusions

In this paper, we demonstrated that the Black-Cox first passage model can be efficiently extended to a very broad class of firm value processes that includes exponential Lévy processes. We tested the fit of two realizations of Lévy subordinated Brownian motion models to observed CDS spreads for Ford Motor Co., a representative firm with an interesting credit history in recent years. We found that the two Lévy process models can be implemented very easily, and give similarly good performance in spite of the very different characteristics of their jump measures. With one extra parameter, both models outperform the Black-Cox model in fitting the time series of CDS term structures over 1.5 year periods. However, they still have limitations in fitting all tenors of the CDS term structure, suggesting that further study is needed into models with more flexible time changes.

We also proposed a new method for filtered statistical inference, based on what we call the linearized measurement equation. This new method inductively creates "quasi-gaussian" likelihood functions that can be approximated either as truncated gaussians, or as true gaussians in which case we are lead to a Kalman filter. By their strategic use of the fast Fourier transform, both of our two approximation methods turn out to be very efficient: parameter estimation for a time series of term structures for 78 weeks can be computed in about a minute. Finally, we observe a strong correlation between Ford's stock price and the filtered values of its unobserved log-leverage ratios. This final observation provides the motivation for our future research that will extend these TCBM credit models to TCBM models for the joint dynamics of credit and equity.

References

[1] M. Baxter. Dynamic modelling of single-name credits and CDO tranches. working paper downloadable at `www.defaultrisk.com/pp_crdrv_07.htm`, 2006.

[2] F. Black and J. C. Cox. Valuing corporate securities. *J. Finance*, 31:351–367, 1976.

[3] P. Carr and L. Wu. Stock options and credit default swaps: A joint framework for valuation and estimation. *J. Fin. Economet.*, doi:10.1093/jjfinec/nbp010:1–41, 2009.

[4] Fang Fang, Henrik Jönsson, Cornelis W. Oosterlee, and Wim Schoutens. Fast Valuation and Calibration of Credit Default Swaps Under Levy Dynamics. *Journal of Computational Finance*, 14:??, 2010.

[5] T. R. Hurd. Credit risk modeling using time-changed Brownian motion. *Int. J. Theor. App. Fin.*, 12:1213–1230, 2009.

[6] T. R. Hurd and A. Kuznetsov. On the first passage time for Brownian motion subordinated by a Lévy process. *J. Appl. Probab.*, 46:181–198, 2009.

[7] S. G. Kou and H. Wang. First passage times of a jump diffusion process. *Adv. in Appl. Probab.*, 35(2):504–531, 2003.

[8] D. Lando. *Credit Risk Modeling: Theory and Applications*. Princeton University Press, 2004.

[9] R. Lee. Option pricing by transform methods: Extensions, unification, and error control. *Journal of Computational Finance*, 7:51–86, 2004.

[10] D. Madan and E. Seneta. The VG model for share market returns. *Journal of Business*, 63:511–524, 1990.

[11] Rafael Mendoza-Arriaga, Peter Carr, and Vadim Linetsky. Time-changed markov processes in unified credit-equity modeling. *Mathematical Finance*, 20(4):527–569, 2010.

[12] R. C. Merton. On the pricing of corporate debt: the risk structure of interest rates. *J. Finance*, 29:449–470, 1974.

[13] T. Moosbrucker. Pricing CDOs with correlated Variance Gamma distributions. Working paper `http://www.defaultrisk.com/pp_crdrv103.htm`, 2006.

[14] Whitney K Newey and Kenneth D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–08, 1987.

[15] Q. H. Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57:307–333, 1989.

		Dataset 1	Dataset 2	Dataset 3
	number of weeks	78	78	78
	$\hat{\sigma}$	0.3	0.3	0.3
	$\hat{b}$	0.2	0.2	0.2
	$\hat{c}$	1.039(0.060)	0.451(0.034)	1.08(0.11)
	$\hat{\beta}_Q$	-1.50(0.12)	-0.879(0.061)	-1.368(0.066)
VG Model	$\hat{R}$	0.626(0.026)	0.450(0.029)	0.611(0.018)
	$\hat{\eta}$	1.53	0.897	1.797
	$\hat{x}_{av}$	0.693	0.457	0.480
	$\hat{x}_{std}$	0.200	0.239	0.267
	RMSE	1.43	0.837	1.792
	$\hat{\sigma}$	0.3	0.3	0.3
	$\hat{b}$	0.2	0.2	0.2
	$\hat{c}$	2.23(0.12)	1.17(0.07)	2.33(0.20)
	$\hat{\beta}_Q$	-1.44(0.12)	-0.780(0.060)	-1.286(0.067)
Exponential Model	$\hat{R}$	0.609(0.028)	0.395(0.033)	0.588(0.022)
	$\hat{\eta}$	1.503	0.882	1.775
	$\hat{x}_{av}$	0.702	0.479	0.486
	$\hat{x}_{std}$	0.199	0.242	0.266
	RMSE	1.41	0.821	1.763
	$\hat{\sigma}$	0.3	0.3	0.3
	$\hat{\beta}_Q$	-2.02(0.10)	-1.793(0.067)	-1.78(0.12)
	$\hat{R}$	0.773(0.011)	0.757(0.009)	0.760(0.013)
Black-Cox Model	$\hat{\eta}$	2.38	1.29	2.18
	$\hat{x}_{av}$	0.624	0.406	0.422
	$\hat{x}_{std}$	0.187	0.214	0.237
	RMSE	2.19	1.19	2.14

Table 1: Parameter estimation results and related statistics for the VG, EXP and Black-Cox models. $\hat{X}_t$ derived from (23) provide the estimate of the hidden state variables. The numbers in the brackets are standard errors. The estimation uses weekly (Wednesday) CDS data from January 4th 2006 to June 30 2010. $\hat{x}_{std}$ is the square root of the annualized quadratic variation of $\hat{X}_t$.

	VG	EXP	B-C
VG	0	-2.21/-1.41/-2.33	5.42/5.10/2.03
EXP	2.21/1.41/2.33	0	5.46/5.22/2.19
B-C	-5.42/-5.10/-2.03	-5.46/-5.22/-2.19	0

Table 2: Results of the Vuong test for the three models, for dataset 1, dataset 2 and dataset 3. A positive value larger than 1.65 indicates that the row model is more accurate than the column model with 95% confidence level.
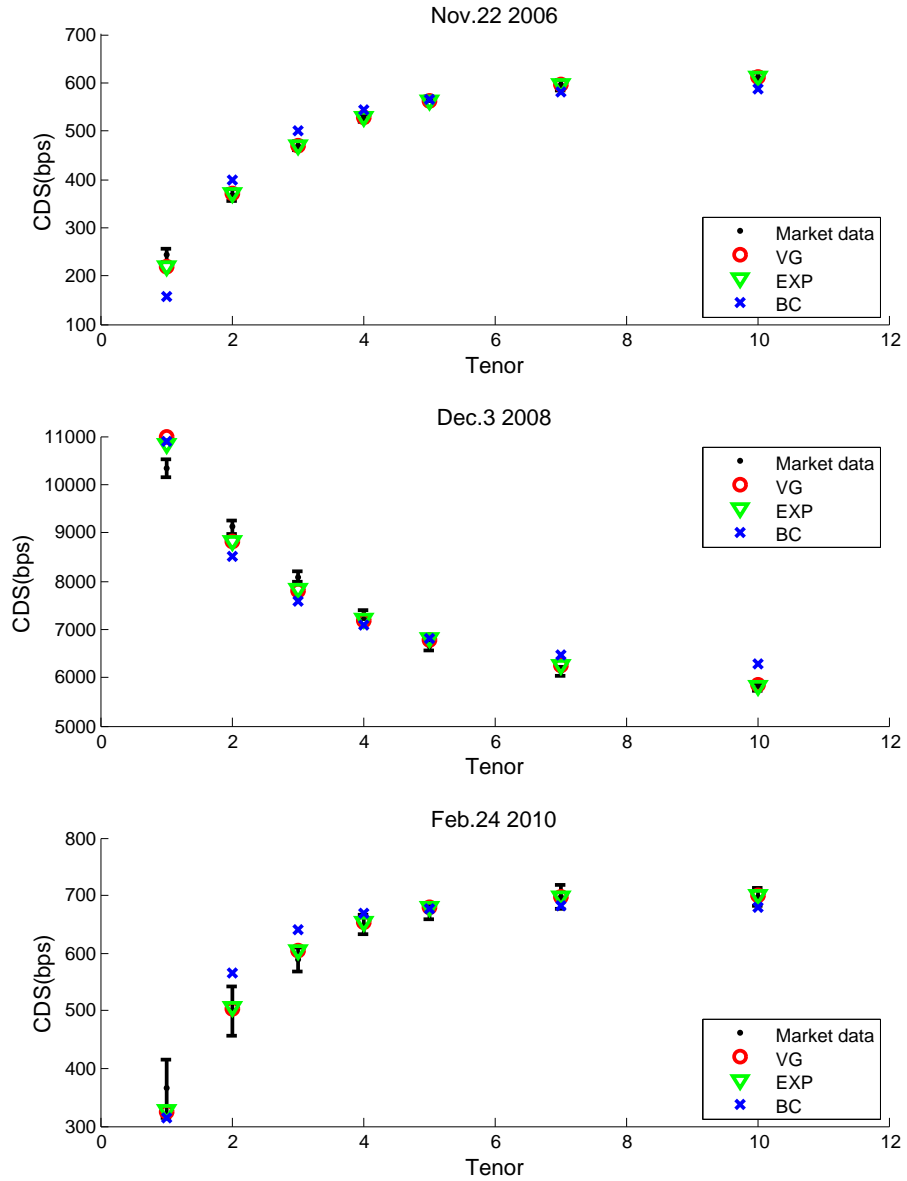
Figure 1: The in-sample fit of the two TCBM models and Black-Cox model to the observed Ford CDS term structure for November 22, 2006 (top), December 3, 2008 (middle) and February 24, 2010 (bottom). The error bars are centered at the mid-quote and indicate the size of the bid-ask spread.
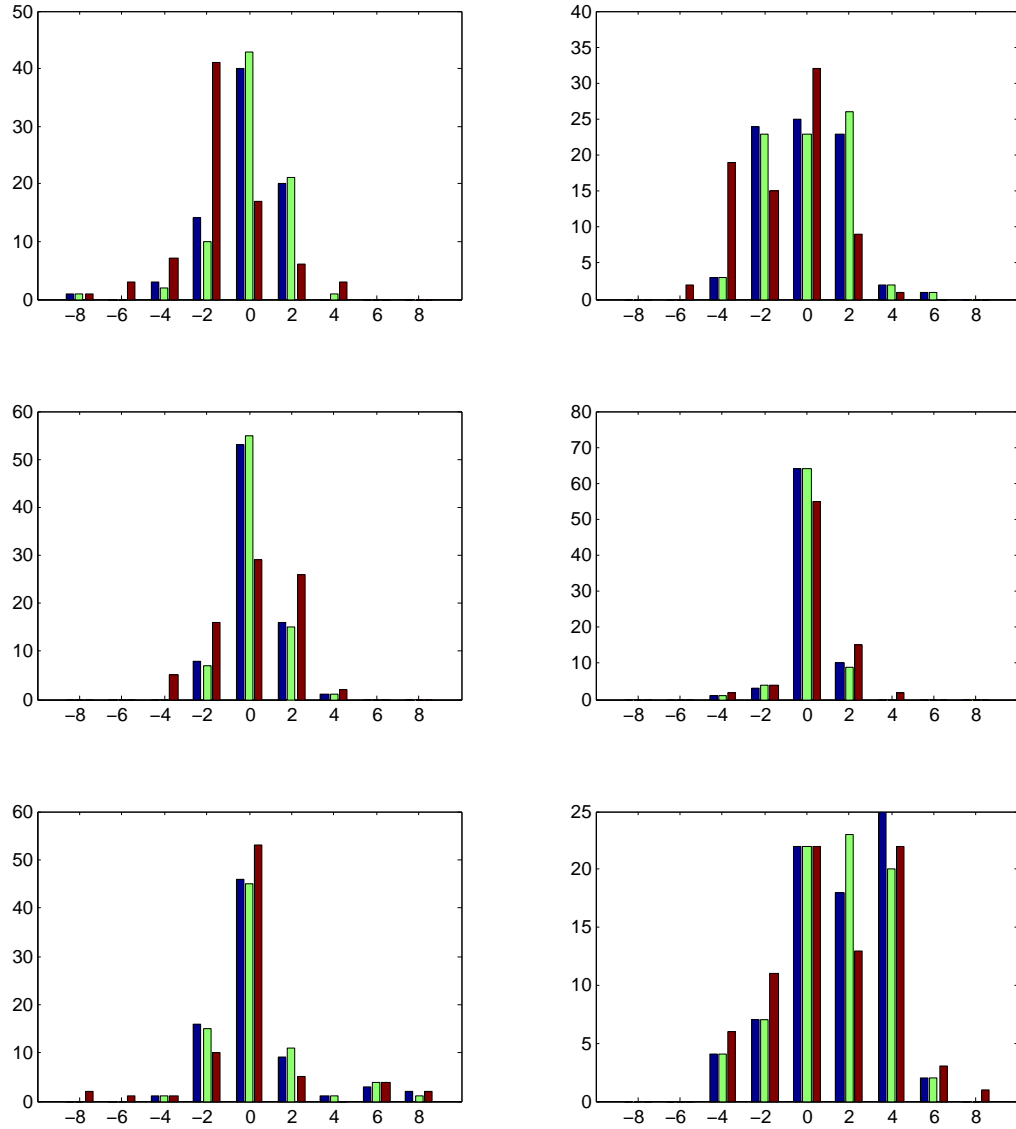
Figure 2: Histograms of the relative errors, in units of bid-ask spread, of the in-sample fit for the VG model (blue bars), EXP model (green bars) and Black-Cox model (red bars) for dataset 1 (top), dataset 2 (middle) and dataset 3 (bottom). The tenor on the left is 1-year and on the right, 10-year.
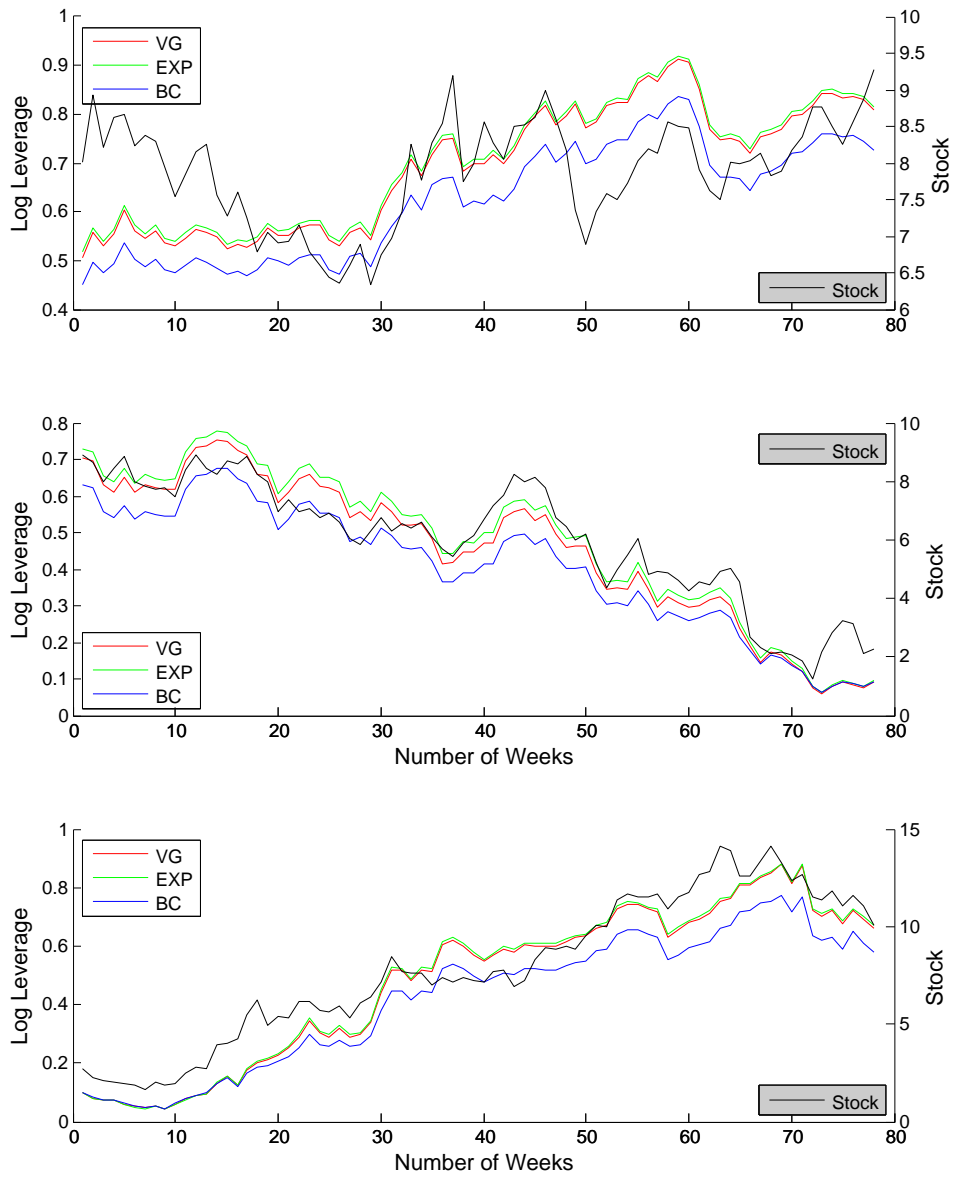
Figure 3: Filtered values of the unobserved log-leverage ratios X_t versus stock price for Ford for dataset 1(top), 2 (middle) and 3 (bottom).