Bayesian Estimation of Time-Changed Default Intensity Models

Michael B. Gordy and Pawel J. Szerszen

Bayesian Estimation of
Time-Changed Default Intensity Models[*]

Michael B. Gordy and Pawel J. Szerszen
Federal Reserve Board, Washington, DC, USA

January 6, 2015

Abstract

We estimate a reduced-form model of credit risk that incorporates stochastic volatility in default intensity via stochastic time-change. Our Bayesian MCMC estimation method overcomes nonlinearity in the measurement equation and state-dependent volatility in the state equation. We implement on firm-level time-series of CDS spreads, and find strong in-sample evidence of stochastic volatility in this market. Relative to the widely-used CIR model for the default intensity, we find that stochastic time-change offers modest benefit in fitting the cross-section of CDS spreads at each point in time, but very large improvements in fitting the time-series, i.e., in bringing agreement between the moments of the default intensity and the model-implied moments. Finally, we obtain model-implied out-of-sample density forecasts via auxiliary particle filter, and find that the time-changed model strongly outperforms the baseline CIR model.

JEL classification: C11, C15, C58, G12, G17
Keywords: Bayesian estimation; MCMC; Particle filter; CDS; Credit derivatives; CIR process; Stochastic time change

1   Introduction

Credit spreads widened dramatically during the financial crisis of 2007–09. The CDX index of five year North American investment grade corporate credit default swaps (CDS) rose from under 30 basis points (bp) in February 2007 to 280bp in November 2008. Perhaps less widely appreciated, the financial crisis also witnessed bursts of extreme *volatility* in spreads. This can be seen in Figure 1, where we plot daily changes in the log of the par spread on the on-the-run CDX index. The figure suggests that stochastic volatility is not a phenomenon peculiar to the financial crisis, as bursts of volatility can be seen, for example, around March 2005 and September 2011. More formally, Gordy and Willemann (2012) find that a hypothesis of constant volatility in a Vasicek model of log-spreads is strongly rejected even in the pre-crisis data, and Alexander and Kaeck (2008) provide evidence of regime-dependent volatility in a Markov switching model of CDS spreads. Arguably, financial institutions might have been better prepared for the financial crisis had risk management and rating models incorporated stochastic volatility. As a case study in model risk, Gordy and Willemann (2012) show that the high investment grade ratings assigned pre-crisis to constant proportion debt obligations depended crucially on the rating agencies' assumption of constant volatility in spreads.

In practical application and in the empirical literature, the most widely-used pricing models for CDS take the so-called reduced form approach, pioneered by Jarrow and Turnbull (1995) and Duffie and Singleton (1999), in which a firm's default occurs at the first event of a non-explosive counting process with stochastic intensity. Broadly speaking, there are three ways in which to accommodate the patterns of Figure 1 in this class of models. First, if we begin with the assumption that the log of the intensity follows a Vasicek process, as in Pan and Singleton (2008), then we can simply augment the process with positive and negative jumps. The diffusion component would drive the low volatility periods, and the jumps would accommodate the high volatility periods. Observe here that the log transform of the intensity is what allows us to accommodate negative jumps without violating the zero lower bound on the intensity process, but it carries the price of computational intractability. A second approach builds on the widely-used and analytically tractable single-factor CIR specification for the intensity. Jacobs and Li (2008) employ a two-factor specification in which a second CIR process controls the volatility of the intensity process. They estimate the model on time-series of corporate bond spreads and find strong evidence of positive volatility of volatility. The model retains tractability in pricing, but at the expense of the zero lower bound. Except when volatility of volatility is zero, there is no region of the parameter space for which the default intensity is bounded nonnegative. The third approach is to induce stochastic volatility via stochastic time-change. The constant volatility model is assumed to apply in a latent "business time." The speed of business time with respect to

1

calendar time is stochastic, and captures the intuition of time-variation in the rate of arrival of news to the market. When applied to default intensity models, analytic tractability is sacrificed to a degree, but Mendoza-Arriaga and Linetsky (2014) and Costin, Gordy, Huang, and Szerszen (forthcoming) derive computationally efficient series solutions to the time-changed model. An advantage to this approach is that the calendar-time default intensity is bounded nonnegative so long as the business-time intensity is bounded nonnegative.

In this paper, we take the third approach. Although CDS pricing is tractable, estimation of the model presents significant challenges. The model has two latent state variables (one for the default intensity in business time, and one for the time-change process). Due to nonlinearities in the pricing function and state-dependent variance in the state evolution, the model is not well-suited to maximum likelihood estimation by Kalman filter or its extensions. Bayesian Markov chain Monte Carlo (MCMC) estimation is viable, but the conventional MCMC algorithm will be prone to very poor convergence rates in our setting primarily because the latent default intensity has high serial dependence. Variables that are highly correlated are best sampled in blocks, but the model nonlinearity and state-dependent variance make this infeasible. To overcome these obstacles, we build on the approach introduced by Stroud, Müller, and Polson (2003) in which an auxiliary linearized model facilitates sampling of the latent state variables in blocks. As a by-product of the MCMC estimation, we obtain smoothed estimates of the path of the default intensity and time-change increments. Moreover, missing data is easily addressed in the MCMC algorithm.

We estimate the model on firm-level time-series of CDS spreads and find strong evidence of material stochastic volatility. Relative to the model without time-change, we find that stochastic time-change offers modest benefit in fitting the cross-section of CDS spreads at each point in time. However, we find that stochastic time-change offers very large improvements in fitting the time-series, i.e., in bringing agreement between the in-sample moments of the default intensity and the model-implied moments.

Finally, we assess model performance in out-of-sample density forecasts. We hold parameters fixed at the mean of the in-sample posterior distribution, and filter state variables through the out-of-sample period. Although non-adapted particle filters are straightforward to implement in our model setting, they are prone to particle degeneration whenever abrupt changes of CDS spreads occur. We develop a partially-adapted auxiliary particle filter which effectively mitigates particle degeneration while preserving tractability. For nine of the ten firms in our sample, we reject the baseline CIR model in favor of the model with stochastic time-change.

Our model specification and CDS pricing are set forth in Section 2. Our econometric methodology is described in Section 3. Data and estimation results are presented in Section 4. Out-of-sample performance is assessed in Section 5. Section 6 concludes.

2 Model specification and pricing

Mendoza-Arriaga and Linetsky (2014) and Costin et al. (forthcoming, hereafter CGHS) introduce stochastic time change to default intensity models for pricing credit default swaps and other credit sensitive instruments. A firm's default occurs at the first event of a non-explosive counting process. Under the business-time clock, the intensity of the counting process is λ_t. The intuition driving default intensity models is that $\lambda_t dt$ is the probability of default before business time $t + dt$, conditional on survival to business time t.

We assume complete markets, no arbitrage, and the existence of an equivalent martingale measure $\mathbb{Q}$. Unless stated otherwise, all specifications of stochastic processes and expectations are taken with respect to this measure. We assume that λ_t is a CIR process under the business clock, i.e., that it follows the stochastic differential equation

$$d\lambda_t = (\mu - \kappa\lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t \tag{1}$$

where W_t is a Brownian motion. We assume that $\mu > 0$ and initial condition $\lambda_0 \geq 0$, but do not restrict κ. Our approach can easily be generalized to allow for independent positive Poisson jumps in the default intensity process, but we do not pursue this extension here.

Define $\Lambda_t(s)$ as the time-integral (or "compensator") of the default intensity from time t to time $t + s$, i.e., $\Lambda_t(s) = \int_t^{t+s} \lambda_u du$. Conditional on the survival to time t and the state λ_t, the probability of survival to date $t + s$ is

$$S_t(s; \ell) = \Pr(\tau > t + s | \tau > t, \lambda_t = \ell) = \mathrm{E}\left[\exp(-\Lambda_t(s))|\lambda_t = \ell\right]. \tag{2}$$

where τ denotes the default time of the firm. This function has an exponential affine form as detailed in Duffie and Singleton (2003, Appendix A.5).

We now introduce stochastic time-change. Let T_t be the stochastic business time associated with calendar time t. For a given firm, let $\tilde{\tau}$ denote the calendar default time, so that $\tau = T_{\tilde{\tau}}$ is the corresponding time under the business clock. We assume that processes λ and T are independent, i.e., that there is no leverage effect. In the empirical literature on stochastic volatility in stock returns, there is strong evidence for dependence between the volatility factor and stock returns (e.g., Andersen et al., 2002; Jones, 2003; Jacquier et al., 2004). In the credit risk literature, the evidence is rather less compelling. Across the firms in their sample, Jacobs and Li (2008) find a median correlation of around 1% between the default intensity diffusion and the volatility factor.

We assume that the time-change T_t is an inverse Gaussian (IG) process with mean parameter t and shape parameter αt^2. This is an almost surely increasing Lévy process such that $\mathrm{E}\left[\exp(u(T_{t+s} - T_t))\right] = \exp(s\Psi(u))$ where the Laplace exponent is $\Psi(u) = \alpha\left(1 - \sqrt{1 - 2u/\alpha}\right)$. As $\mathrm{E}\left[T_{t+s} - T_t\right] = s$ for all s and t, we say that the business clock is

unbiased, i.e., business time moves at the same speed on average as calendar time. The parameter α can be interpreted as a precision parameter. As $\alpha \to \infty$, business time converges in probability to calendar time (hence, no time-change). As a Lévy subordinator, the process has increments $T_{t+s} - T_t, T_{t+2s} - T_{t+s}, T_{t+3s} - T_{t+2s}, \ldots$ that are independent and identically distributed, so the specification rules out the possibility of volatility clustering in time. The methodology of Mendoza-Arriaga and Linetsky (2014) allows T_t to be any Lévy subordinator, and CGHS allow for a still broader class of time-change processes. Nonetheless, we have found the IG specification to be flexible and numerically well-behaved in our econometric application, so impose this specification throughout this paper.

Independence of the default intensity and business clock implies that time-changing the default time is equivalent to time-changing Λ_t; that is, that the compensator in calendar time is $\tilde{\Lambda}_t(s) = \Lambda_{T(t)}(T_{t+s} - T_t)$. Since λ_t and T_t are Markov processes, the triplet $(T_t, \lambda(T_t), \tilde{\tau} > t)$ is a sufficient statistic for the information at time t. The conditional calendar-time survival probability function is

$$
\tilde{S}_t(s; \ell) = \Pr(\tilde{\tau} > t + s | \tilde{\tau} > t, T_t, \lambda(T_t) = \ell) = \Pr(\tau > T_{t+s} | \tau > T_t, T_t, \lambda(T_t) = \ell)
$$
$$
= \mathrm{E}\left[\mathrm{E}\left[\exp(-\Lambda_{T(t)}(T_{t+s} - T_t)) | T_{t+s}, T_t, \lambda(T_t) = \ell \right] \Big| T_t, \lambda(T_t) = \ell \right]
$$
$$
= \mathrm{E}_t\left[S_{T(t)}(T_{t+s} - T_t; \ell) \right].
$$

Mendoza-Arriaga and Linetsky (2014) provide a series solution to $\tilde{S}_t(s; \ell)$ via Bochner subordination of the eigenfunction expansion. The expansion is uniformly convergent when the intensity process is stationary ($\mu > 0$ and $\kappa > 0$), but cannot be employed in the case of non-stationarity ($\kappa \leq 0$). Both Duffee (1999) and Jacobs and Li (2008) find that the default intensity process is indeed non-stationary under the risk-neutral measure for the typical firm. In our empirical results below, we find that the mean of the posterior distribution for κ under the risk-neutral measure is negative for every firm in our sample. Therefore, we employ the "expansion in derivatives" method of CGHS, which does not impose any restriction on κ. Calendar-time survival probabilities are given by

$$
\tilde{S}_t(s; \ell) \approx \sum_{m=0}^{M} \alpha^{-m} \sum_{j=0}^{m} c_{m,j} s^j D_s^{m+j} S_{T(t)}(s; \ell) \tag{3}
$$

where D_s is the differential operator $\frac{d}{ds}$, M is the order of approximation, and

$$
c_{m,j} = \frac{1}{2^m} \frac{1}{m} \frac{1}{(j-1)!} \binom{2m}{m-j}
$$

for $m \geq j \geq 1$. For $j = 0$, $c_{0,0} = 1$ and $c_{m,0} = 0$ for $m > 0$. Derivatives of the business-time

4

survival function are easily computed using a recurrence rule in §4 of CGHS. CGHS show that expansion (3) diverges as $M \to \infty$ but yields an accurate approximation when the series is truncated close to the numerically least term. In our empirical application, we fix $M = 2$, which CGHS find generally accurate to within the bid-ask spread.

Let r_t denote the riskfree short rate, which we assume to be independent of the default intensity and time-change. Our pricing and estimation methods each can be generalized to allow for dependence, but at significant cost in complexity and computational resources.[1] Duffee (1999) finds that the correlation between riskfree rates and default intensity is typically negative but has second-order importance, so we do not pursue the extension here. The discount function is given by

$$P_t(s) = \mathrm{E}_t \left[\exp \left(- \int_t^{t+s} r_u du \right) \right]$$

CDS pricing follows the treatment of Leeming et al. (2010, §2). By independence of the short rate and default risk, we can write the value of the remaining quarterly *premium leg* payments at $t + s_1, \ldots, t + s_n$ at date t as

$$\mathrm{PremiumLeg} = c \sum_{i=1}^{n} P_t(s_i) \tilde{S}_t(s_i; \lambda_{T(t)}) B(s_i)$$

where $B(s) = \min\{0.25, s\}$ is the coupon time interval. The value of *protection leg* payments is

$$\mathrm{ProtectLeg} = (1 - \rho) \int_0^{s(n)} P_t(s) \tilde{q}_t(s; \lambda_{T(t)}) \, ds$$

where ρ is the expected recovery rate as a fraction of face value, and $\tilde{q}_t(s; \ell)$ is the density of the default time under calendar time conditional on business-time intensity $\lambda_{T(t)} = \ell$. This density is simply $\tilde{q}_t(s; \ell) = -\tilde{S}_t'(s; \ell)$. The model-implied par CDS spread equates the value of the premium leg to that of the protection leg.[2]

In Section 4, we will compare the smoothed estimates of the default intensity in a model without time-change to the calendar-time intensity $\tilde{\lambda}_t$ in the full model with time-change. The default intensity coincides with the instantaneous forward default rate, so $\tilde{\lambda}_t = \tilde{q}_t(0; \lambda_{T(t)})$ is readily computed.[3]

As emphasized by Jarrow et al. (2005), market prices can reveal only the $\mathbb{Q}$-intensity,

---

[1] On pricing in a multifactor framework with dependence, see Remark 4.1 in Mendoza-Arriaga and Linetsky (2014) and §7 in CGHS.

[2] Since April 2009, CDS have been traded on an "upfront" basis with annual coupons fixed to 1% or 5%. For observations since this change in market convention, we convert upfront to par spread using the standard formula published by ISDA. Leeming et al. (2010, §4).

[3] Formal treatment of the existence of the calendar-time default intensity can be found in Mendoza-Arriaga and Linetsky (2014, Theorems 3.2(iii), 3.3(iii)).

which may or may not contain a risk-premium for the event risk associated with default. While the intensity under the physical measure $\mathbb{P}$ is not identified in our setting, our smoothed estimate of λ_t reveals the $\mathbb{P}$-dynamics of the $\mathbb{Q}$-intensity.[4] We specify the dynamics of λ_t under the two measures jointly in the standard fashion. The drift and volatility parameters in (1) are restricted to be the same across the two measures, but the risk-neutral mean-reversion parameter $\kappa^{\mathbb{Q}}$ differs from the physical $\kappa^{\mathbb{P}}$ by an unrestricted risk-premium. This change of measure is called "drift change in the intensity" by Jarrow et al. (2005), and has been adopted by Duffee (1999) and Jacobs and Li (2008), among others. In the empirical implementation below, we will restrict $\kappa^{\mathbb{P}} > 0$ to guarantee stationarity under the physical measure, but will not restrict $\kappa^{\mathbb{Q}}$.

In principle, we can introduce a second risk-premium on the uncertainty due to time-change. Say we assume that $T_{t+s} - T_t \sim \mathrm{I}G(s, \alpha s^2)$ under the physical measure $\mathbb{P}$. For the time-change process T_t to remain within the IG family under $\mathbb{Q}$, the increments to the business clock must be distributed $T_{t+s} - T_t \sim \mathrm{I}G(s\psi, \alpha s^2)$ under $\mathbb{Q}$ where the risk-premium parameter ψ is nonnegative.[5] The issue is whether ψ can be estimated. For large α, $T_{t+s} - T_t$ converges in probability to $s\psi$. Since λ_t follows a scale-invariant CIR process, the parameters of the time-changed default intensity could not be identified in the asymptotic case of $\alpha \to \infty$. In numerical experiments based on simulated time-series of length comparable to our data and values of α from the range of estimates reported below in Section 4, we find that ψ is only very weakly identified. Therefore, we make the simplifying assumption that uncertainty due to time-change is unpriced, i.e., that $T_{t+s} - T_t \sim \mathrm{I}G(s, \alpha s^2)$ holds under both $\mathbb{P}$ and $\mathbb{Q}$.

Stochastic time-change can have a dramatic effect on the kurtosis of changes in the time-series of CDS spreads. In Figure 2, we plot kurtosis of the stationary distribution of changes in the CDS spread as a function of α on a log-log scale. Let $\mathrm{CDS}(\lambda; \alpha)$ be the five-year par spread as a function of the current business-time intensity. For each value of α and horizon δ, we first draw a sample $\lambda_0^{(i)}$, $i = 1, \ldots, \mathcal{I}$, from the stationary distribution for λ_t. Next, we draw the elapsed business-time $T_\delta^{(i)}$; and then draw $\lambda^{(i)}(T_\delta^{(i)})$ from the CIR transition density. From this sample, we calculate the kurtosis of the stationary distribution of $\mathrm{CDS}(\lambda^{(i)}(T_\delta^{(i)}); \alpha) - \mathrm{CDS}(\lambda_0^{(i)}; \alpha)$. We plot separate curves for a one day horizon ($\delta = 1/250$, assuming 250 trading days per year), a one month horizon ($\delta = 1/12$), and an annual horizon ($\delta = 1$). Parameters for the business-time CIR process are fixed to $\kappa^{\mathbb{P}} = \kappa^{\mathbb{Q}} = 0.2, \mu = 0.004, \sigma_0 = .1$, and recovery is fixed to $\rho = 0.4$. As we expect, kurtosis at all horizons tends to its asymptotic CIR limit as $\alpha \to \infty$. For fixed α, kurtosis also tends to

---

[4]This limitation applies as well to Duffee (1999), Jacobs and Li (2008), and Pan and Singleton (2008). Driessen (2005) brings additional ratings-based information to bear to identify $\lambda_t^{\mathbb{P}}$ and $\lambda_t^{\mathbb{Q}}$ separately.

[5]This is the only form that satisfies the necessary integrability conditions on the Radon-Nikodym derivative of the Lévy measures of jumps under $\mathbb{P}$ and $\mathbb{Q}$ set forth in Theorem 7.3 of Barndorff-Nielsen and Shiryaev (2010).

its CIR limit as $\delta \to \infty$. This is because an unbiased trend stationary time-change has no effect on the distribution of a stationary process far into the future. For intermediate values of α (say, between 1 and 10), we see that time-change has a modest impact on kurtosis beyond one year, but a material impact at a one month horizon, and a very large impact at a daily horizon.

3 Estimation method

Our choice of estimation method is guided by two characteristics of the model of Section 2. First, the mapping from default intensity to CDS spreads is nonlinear and sensitive to parameters. Second, the latent default intensity is persistent under the physical measure and has state-dependent variance. Bayesian Markov chain Monte Carlo (MCMC) estimation is a natural choice in this setting, but the conventional MCMC algorithm will suffer from very poor convergence rates in our setting primarily because the latent default intensity has high serial dependence. Variables that are highly correlated are best sampled in blocks, but the model nonlinearity and state-dependent variance make this infeasible. To overcome this obstacle, we build on the approach introduced by Stroud, Müller, and Polson (2003, hereafter SMP) in which an auxiliary linear model facilitates sampling of the latent state variables in blocks.

MCMC estimation has additional advantages. As a by-product of the estimation, we obtain smoothed estimates of default intensity and time-change increments. Moreover, missing data is easily addressed in the MCMC algorithm by augmenting the states. Relative to particle filtering, a limitation of MCMC is that online estimation of states and parameters is computationally infeasible. We turn to particle filtering for out-of-sample forecasting exercises in Section 5.

Section 3.1 offers an introduction to Bayesian MCMC methods in general and the SMP algorithm in particular. The presentation is in a simplified setting and may be skipped over by readers familiar with the SMP method. In Section 3.2, we adapt the SMP method to the model of Section 2.

3.1 MCMC estimation with a linear auxiliary model

Our introduction to Bayesian MCMC methods is modeled after Chib and Greenberg (1996) and Jones (2003). Let $Y = (y_1, \ldots, y_{\mathcal{T}})$ denote the vector of observations, $X = (x_1, \ldots, x_{\mathcal{T}})$ be the vector of latent state variables and Θ be the vector of model parameters.[6] In Bayesian inference the joint posterior of states and parameters is of interest and can be derived given

---

[6]Throughout the paper, we employ caligraphic font to distinguish dataset dimensions from model quantities, e.g., $\mathcal{T}$ as a count of time-series observations vs. T_t as the stochastic time-change process.

the prior distribution on the parameters:

$$p(\Theta, X|Y) \; \propto \; p(Y|X,\Theta) \cdot p(X|\Theta) \cdot p(\Theta),$$

where $p(Y|X,\Theta)$ is the likelihood function of the model, $p(X|\Theta)$ is the probability distribution of state variables conditional on the parameters and $p(\Theta)$ is the prior probability distribution on the parameters of the model. The joint posterior distribution $p(\Theta, X|Y)$ is in general analytically and computationally intractable. MCMC methods overcome this problem by breaking the high-dimensional vectors X and Θ into low-dimensional subvectors with complete conditional posterior distributions that are more easily sampled. The joint posterior is, by construction, the invariant distribution of the chain.

The Gibbs sampler of Geman and Geman (1984) is typically the preferred method when the conditional posterior can be sampled directly. We partition X and Θ into, respectively, $\mathcal{J}^X$ and $\mathcal{J}^\Theta$ subvectors $X^{(1)}, X^{(2)}, ..., X^{(\mathcal{J}^X)}$ and $\Theta^{(1)}, \Theta^{(2)}, ..., \Theta^{(\mathcal{J}^\Theta)}$. The Markov chain is initialized at chosen values X_0 and Θ_0. The chain is formed by drawing iteratively from transition densities $p(X_n^{(i)}|X_n^{(-i)}, \Theta_{n-1}, Y)$, $i = 1, 2, \ldots, \mathcal{J}^X$, and $p(\Theta_n^{(j)}|\Theta_n^{(-j)}, X_n, Y)$ $j = 1, 2, \ldots, \mathcal{J}^\Theta$, where $X_n^{(-i)} \equiv (X_n^{(k)})_{k<i} \cup (X_{n-1}^{(k)})_{k>i}$ and $\Theta_n^{(-j)} \equiv (\Theta_n^{(k)})_{k<j} \cup (\Theta_{n-1}^{(k)})_{k>j}$. Under mild regularity conditions, the chain (X_n, Θ_n) converges to its invariant distribution $p(\Theta, X|Y)$.

When it is difficult or impossible to draw directly from a complete conditional posterior for a subvector $X^{(i)}$ or $\Theta^{(j)}$, that step is replaced by the Metropolis-Hastings (MH) algorithm (Metropolis et al., 1953). Say $X^{(i)}$ is the variable to be sampled in iteration n. We sample from a tractable *proposal density* $\breve{p}(\breve{X}; X_{n-1}^{(i)}|X_n^{(-i)}, \Theta_{n-1}, Y)$. The draw $\breve{X}$ is accepted with probability

$$\min\left\{1, \frac{p(\breve{X}|X_n^{(-i)}, \Theta_{n-1}, Y) \cdot \breve{p}(X_{n-1}^{(i)}; \breve{X}|X_n^{(-i)}, \Theta_{n-1}, Y)}{p(X_{n-1}^{(i)}|X_n^{(-i)}, \Theta_{n-1}, Y) \cdot \breve{p}(\breve{X}; X_{n-1}^{(i)}|X_n^{(-i)}, \Theta_{n-1}, Y)}\right\}$$

If accepted, we set $X_n^{(i)} = \breve{X}$. If rejected, then the value $X_{n-1}^{(i)}$ is retained, i.e., we set $X_n^{(i)} = X_{n-1}^{(i)}$. A similar procedure is used when $\Theta^{(j)}$ is the variable to be sampled.

Observe that if the true conditional density is taken as the proposal density, the draw is accepted with probability one, and the MH step is simply a Gibbs sampler. In application, it is often the case that the true conditional density is too complicated to approximate with any reliability, in which case a standard approach is to sample as a normal random walk, i.e., draw $\breve{X}$ as a normal random variable centered at $X_{n-1}^{(i)}$. While simple to implement, this approach may suffer from slow convergence.

Correlation across parameters or state variables poses a computational challenge in MCMC estimation. Joint sampling of correlated variables yields faster convergence rates, but is often much more difficult to implement. In our application, this dilemma is especially

acute for sampling of the default intensity. Strong persistence in the default intensity implies that we ought to sample jointly the default intensity states over extended blocks of time. Were the model linear and Gaussian, we could sample jointly via the forward filtering backward sampling (FFBS) method of Carter and Kohn (1994) and Frühwirth-Schnatter (1994). In our model setting, nonlinearity and state-dependent variance make this infeasible. The single-move MCMC sampler (Carlin et al., 1992) can be applied, but we have found the cost in rate of convergence to be insurmountable.

SMP propose an MCMC algorithm that is well-suited to our problem. Their approach is quite general, but for illustration we adopt a simplified setting with nonlinearity in the measurement equation, constant variance in the state equation, and both y_t and x_t uni-dimensional. The essential idea is to extend the state vector to include auxiliary mixing variables $\vec{k} = (k_1, \ldots, k_{\mathcal{T}})$ with k_t having discrete support $\{1, \ldots, \mathcal{K}\}$. Conditional on the mixing variable, the measurement equation is approximated as Gaussian and linear in x_t, which allows FFBS to be applied to the proposal distribution. Finally, this FFBS draw is accepted or rejected, as in MH, to account for the gap between the true measurement equation and the linearized approximation that provides the proposal density. There is a trade-off in choosing the number of nodes in the approximation: The larger is $\mathcal{K}$, the smaller is the gap between true and linearized models, so the higher the acceptance rate. However, the computational cost per MCMC iteration increases with $\mathcal{K}$.

SMP specify the auxiliary model for y_t as Gaussian with $p^a(y_t|x_t, k_t = k, \Theta) \sim N(\gamma_{t,k} + \beta_{t,k}x_t, \omega_{t,k}^2)$ for (possibly) time-varying coefficients $\gamma_{t,k}, \beta_{t,k}, \omega_{t,k}$. The mixture weights are standardized Gaussian weights

$$p^a(k_t = k|x_t, \Theta) = \frac{\phi(x_t; \nu_k, \xi_k)}{w(x_t)}$$

where $\phi(x; \nu, \xi)$ denotes the normal density with mean ν and variance ξ^2 and the scaling factor $w(x) = \sum_{k=1}^{\mathcal{K}} \phi(x; \nu_k, \xi_k)$ guarantees that the weights sum to one. For tractability, the mixture variables are assumed to be serially independent, conditional on x_t. The auxiliary mixture model is then

$$p^a(y_t|x_t, \Theta) = \sum_{k=1}^{\mathcal{K}} p^a(y_t|x_t, k_t = k, \Theta) \cdot p^a(k_t = k|x_t, \Theta) \tag{4}$$

The node constants (ν_k, ξ_k) and identifying restrictions for $(\gamma_{t,k}, \beta_{t,k}, \omega_{t,k})$ are chosen so that the auxiliary density $p^a(y_t|x_t, \Theta)$ is close to the target density $p(y_t|x_t, \Theta)$.

In each iteration of the MCMC procedure, we draw Θ in the usual way, and draw the state variables as follows. For notational compactness, we drop the subscript n for indexing MCMC iterations.

1. Generate mixture indicators k_t from the complete conditional posterior $p^a(k_t|y_t, x_t, \Theta)$, which has a multinomial distribution. By Bayes' rule,

$$p^a(k_t|y_t, x_t, \Theta) \propto p^a(y_t|x_t, k_t, \Theta) \cdot p^a(k_t|x_t, \Theta)$$

This is a Gibbs step.

2. Conditional on the mixture indicators, draw states jointly via FFBS from the MH proposal density for the mixture model. The proposal density is given by:

$$\check{p}(X|\vec{k}, Y, \Theta) = p(x_0) \prod_{t=1}^{\mathcal{T}} p(x_t|x_{t-1}, \Theta) \cdot p^a(y_t|x_t, k_t, \Theta) \cdot \phi(x_t; \nu_{k(t)}, \xi_{k(t)})$$

This is similar to the MH proposal density in a linear model without auxiliary variables, but incorporates an extra Gaussian kernel $\phi(x_t; \nu_{k(t)}, \xi_{k(t)})$ to capture the additional conditioning information from the auxiliary variables. Denote the proposal draw $\check{X} = (\check{x}_1, \ldots, \check{x}_{\mathcal{T}})$.

3. Accept or reject $\check{X}$. The acceptance probability is

$$\min\left\{1, \prod_{t=1}^{\mathcal{T}} \frac{p(y_t|\check{x}_t, \Theta)}{w(\check{x}_t)p^a(y_t|\check{x}_t, \Theta)} \frac{w(x_t)p^a(y_t|x_t, \Theta)}{p(y_t|x_t, \Theta)}\right\} \tag{5}$$

where the $\{x_t\}$ denote the previous draw of X in the Markov chain.

Observe that the acceptance probability is maximized when the auxiliary model likelihood $p^a(y_t|x_t, \Theta)$ is close to the true likelihood $p(y_t|x_t, \Theta)$. When the acceptance probability is low, performance may be improved by dividing the time-series into blocks, $b = 1, \ldots, \mathcal{B}$, and tuning the node constants (ν_k^b, ξ_k^b) on a block-by-block basis.

3.2 Application to estimation of the time-changed default intensity model

In this section, we cast the model of Section 2 as a discrete-time state-space model, and then adapt the SMP algorithm to accommodate a multidimensional measurement equation and state-dependent variance in the state equation for the default intensity.

Our notation needs to track three time-scales, one for observations, one for calendar time, and one for business time. Henceforth, we let $t = 1, \ldots, \mathcal{T}$ index a set of daily observations. The associated (deterministic) calendar time is denoted $\tilde{T}_t$. For simplicity, we take trading days as equally spaced in calendar time at an interval of $\Delta = \tilde{T}_{t+1} - \tilde{T}_t = 1/250$ of a year. The stochastic business time associated with observation time t is $T(\tilde{T}_t)$. For notational convenience, let $h_t = \lambda(T(\tilde{T}_t))$ be the business-time default intensity at the business time

associated with observation date t.[7]

Fixing a given reference entity, let $y_{t,m}$ be the log of the CDS spread observed on date t for maturity $m \in \mathfrak{M} = \{1, 2, 3, 5, 7, 10\}$ years. Let $F_m(h_t; \mu, \kappa^{\mathbb{Q}}, \sigma, \alpha, \rho_t, P_t)$ map h_t to a model-implied log-spread on a CDS of maturity m. For notational compactness, we drop explicit reference to the date-t riskfree discount function and recovery rate, and simply write $F_{t,m}(h_t; \mu, \kappa^{\mathbb{Q}}, \sigma, \alpha)$.

We assume that CDS log-spreads are observed with noise. Without measurement noise, the spread would be a deterministic function of underlying parameters and the state, which would lead to stochastic singularity in filtering and smoothing.[8] In the option pricing literature, this approach has been used by Eraker (2004) and Johannes et al. (2009), among others. Our measurement equation is

$$y_{t,m} = F_{t,m}(h_t; \mu, \kappa^{\mathbb{Q}}, \sigma, \alpha) + \zeta\, \varepsilon_{t,m}^{(y)} \tag{6}$$

Measurement errors $\varepsilon_{t,m}^{(y)}$ are assumed to be i.i.d. standard normal random variables. The new parameter ζ scales the pricing errors. For parsimony, we impose homoscedasticity across both time and contract maturity.

The model has two state evolution equations. The first governs the distribution of increments to the business time clock. Let

$$\chi_{t+1} = T(\tilde{T}_{t+1}) - T(\tilde{T}_t) \overset{\text{iid}}{\sim} IG(\Delta, \alpha\Delta^2) \tag{7}$$

be the inverse Gaussian increment to the business clock between observations t and $t + 1$. The second governs the evolution of the default intensity under the physical measure. We apply a first order Euler discretization scheme to SDE (1) with stochastic time-increments to obtain

$$h_{t+1} = h_t + (\mu - \kappa^{\mathbb{P}} h_t)\chi_{t+1} + \sigma\, \sqrt{h_t \chi_{t+1}}\, \varepsilon_{t+1}^{(\lambda)} \tag{8}$$

The $\{\varepsilon_t^{(\lambda)}\}$ are i.i.d. standard normal random variables, independent of $\{\varepsilon_{t,m}^{(y)}\}$. We assume that both $\kappa^{\mathbb{P}}$ and μ are strictly positive, so that h_t is bounded nonnegative and stationary.[9] We refer to equations (6)–(8) as the state-space representation.

Our framework extends the basic SMP algorithm of Section 3.1 along two dimensions. First, our measurement equation is of dimension $\#\mathfrak{M}$, as we have one observation for each contract in maturity set $\mathfrak{M}$. As in the standard SMP algorithm, we introduce a

---

[7]This should not be confused with the calendar time default intensity $\tilde{\lambda}(\tilde{T}_t)$.

[8]In our setting, measurement error could alternatively be justified by misspecification or estimation error in the fitted riskfree term-structure P_t or by the presence of time-varying liquidity premia in the CDS market.

[9]Discretization introduces the possibility of negative realizations of the default intensity. We avoid this in the MCMC simulation by simple truncation at zero.

linear auxiliary model for the measurement equation with mixing variable $k_t^{(y)}$. Second, and of greater consequence, the state equation (8) for h_t has state-dependent volatility, as $V[h_{t+1}|h_t] = \sigma^2 h_t \chi_{t+1}$. As FFBS requires state-independent volatility, we introduce a linear auxiliary model for the state equation (8) with a new mixing variable $k_t^{(\lambda)}$. For parsimony, the two mixing variables share the same support $\{1, \ldots, \mathcal{K}\}$, as well as mean and volatility constants (ν_k, ξ_k).[10]

We assume that log-spread $y_{t,m}$ is conditionally independent of $k_t^{(\lambda)}$ given $k_t^{(y)}$. The auxiliary model for the measurement equation is conditionally Gaussian with

$$p^a(y_t|h_t, k_t^{(y)}, \Theta) = \prod_{m \in \mathfrak{M}} p^a(y_{t,m}|h_t, k_t^{(y)}, \Theta)$$

Hence, the auxiliary model is given by the following mixture of Gaussian densities:

$$p^a(y_t|h_t, \Theta) = \sum_{k=1}^{\mathcal{K}} p^a(k_t^{(y)} = k|h_t, \Theta) \prod_{m \in \mathfrak{M}} p^a(y_{t,m}|h_t, k_t^{(y)} = k, \Theta)$$

where

$$p^a(k_t^{(y)} = k|h_t, \Theta) = \frac{\phi(h_t; \nu_k, \xi_k)}{w(h_t)} \quad \text{for } w(x) = \sum_{k=1}^{\mathcal{K}} \phi(x; \nu_k, \xi_k)$$

As in the standard SMP algorithm, we assume that the auxiliary model for the measurement equation is linear and Gaussian conditional on the mixture indicator. We linearize the measurement equation (6) locally at node k, and define

$$p^a(y_{t,m}|h_t, k_t^{(y)} = k, \Theta) \sim N(F_{t,m}(\nu_k; \Theta) + F'_{t,m}(\nu_k; \Theta) \cdot (h_t - \nu_k), \zeta^2)$$

Observe that the coefficients in the mean of the auxiliary distribution vary over time and across maturity, as well as across nodes.

We now introduce the auxiliary model for the state evolution equation (8). Conditional on $k_t^{(\lambda)}$ and h_{t-1}, h_t is independent of the measurement equation mixing variable $k_t^{(y)}$. The auxiliary model is conditionally linear and Gaussian. We linearize state equation (8) at node k, and define

$$p^a(h_t|h_{t-1}, k_t^{(\lambda)} = k, \chi_t, \Theta) \sim N(\mu \chi_t + (1 - \kappa^{\mathbb{P}} \chi_t)h_{t-1}, \sigma^2 \chi_t \nu_k).$$

As the drift of the true evolution equation is affine in h_{t-1}, the mean of the auxiliary model can be matched to the true mean without conditioning on the node. Only the variance of

---

[10]In our implementation, we achieve higher acceptance rates if the time-series is divided into contiguous blocks of, say, 40–50 observations each. The node constants are tuned on a block-by-block basis.

the auxiliary model depends on $k^{(\lambda)}$.

The auxiliary model for state evolution equation is decomposed in the usual way:

$$p^a(h_t|h_{t-1}, \chi_t, \Theta) = \sum_{k=1}^{\mathcal{K}} p^a(h_t|h_{t-1}, k_t^{(\lambda)} = k, \chi_t, \Theta) \cdot p^a(k_t^{(\lambda)} = k|h_{t-1}, \Theta) \tag{9}$$

The state $k_t^{(\lambda)}$ is assumed to be conditionally independent of $k_t^{(y)}$ with

$$p^a(k_t^{(\lambda)} = k|h_{t-1}, \Theta) = \frac{\phi(h_{t-1}; \nu_k, \xi_k)}{w(h_{t-1})} \tag{10}$$

We introduce additional notation so that the algorithm can be presented compactly. The data vector is $Y = (y_1, \ldots, y_{\mathcal{T}})$ where $y_t = (y_{t,m})_{m \in \mathfrak{M}}$ is the vector of CDS log-spreads observed at time t for all maturities $m \in \mathfrak{M}$. The parameter vector is $\Theta = (\mu, \kappa^{\mathbb{P}}, \kappa^{\mathbb{Q}}, \sigma, \alpha, \zeta)$. The state vector X contains four components: default intensities $X^{(\lambda)} = (h_1, \ldots, h_{\mathcal{T}})$; time-change increments $X^{(\chi)} = (\chi_1, \ldots, \chi_{\mathcal{T}})$; mixture variables for the observation equation $X^{(y)} = \vec{k}^{(y)} = (k_1^{(y)}, \ldots, k_{\mathcal{T}}^{(y)})$ with $k_t^{(y)} \in \{1, \ldots, \mathcal{K}\}$; and mixture variables for the state evolution equation $X^{(x)} = \vec{k}^{(\lambda)} = (k_1^{(\lambda)}, \ldots, k_{\mathcal{T}}^{(\lambda)})$ with $k_t^{(\lambda)} \in \{1, \ldots, \mathcal{K}\}$. For any subset $V \subset X$, let $X^{(-V)}$ denote $X \backslash V$.

In each iteration of the MCMC procedure, we follow these steps:

Draw the mixture indicator variables. By Bayes' law and the assumption of conditional independence,

$$p^a(k_t^{(y)}|X^{(-\vec{k}^{(y)})}, Y, \Theta) \propto p^a(y_t|h_t, k_t^{(y)}, \Theta) \cdot p^a(k_t^{(y)}|h_t, \Theta)$$
$$p^a(k_t^{(\lambda)}|X^{(-\vec{k}^{(\lambda)})}, Y, \Theta) \propto p^a(h_t|h_{t-1}, k_t^{(\lambda)}, \chi_t, \Theta) \cdot p^a(k_t^{(\lambda)}|h_{t-1}, \Theta)$$

The complete conditional posterior distribution for each mixture indicator is multinomial and is easily sampled.

Generate proposal default intensity states. We apply Bayes' law to the auxiliary model for the state equation:

$$p^a(X^{(\lambda)}|X^{(-\lambda)}, Y, \Theta) \propto p^a(Y|X^{(\lambda)}, \vec{k}^{(y)}, \vec{k}^{(\lambda)}, \Theta) \cdot p^a(X^{(\lambda)}|\vec{k}^{(y)}, \vec{k}^{(\lambda)}, X^{(\chi)}, \Theta)$$
$$\propto p^a(Y|X^{(\lambda)}, \vec{k}^{(y)}, \Theta) \cdot p^a(\vec{k}^{(y)}|X^{(\lambda)}, \Theta) \cdot p^a(X^{(\lambda)}|\vec{k}^{(\lambda)}, X^{(\chi)}, \Theta)$$

where

$$p^a(Y|X^{(\lambda)}, \vec{k}^{(y)}, \Theta) = \prod_{t=1}^{\mathcal{T}} \prod_{m \in \mathfrak{M}} p^a(y_{t,m}|h_t, k_t^{(y)}, \Theta)$$

$$p^a(\vec{k}^{(y)}|X^{(\lambda)}, \Theta) = \prod_{t=1}^{\mathcal{T}} \frac{\phi(h_t; \nu_{k_t}, \xi_{k_t})}{w(h_t)}$$

$$p^a(X^{(\lambda)}|\vec{k}^{(\lambda)}, X^{(\chi)}, \Theta) = p^a(h_0|k_1^{(\lambda)}, \Theta) \prod_{t=2}^{\mathcal{T}} p^a(h_t|h_{t-1}, k_t^{(\lambda)}, \chi_t, \Theta)$$

The proposal distribution is

$$\check{p}(X^{(\lambda)}|X^{(-\lambda)}, Y, \Theta) \propto p^a(X^{(\lambda)}|X^{(-\lambda)}, Y, \Theta) \cdot \prod_{t=1}^{\mathcal{T}} w(h_t)$$

The standard FFBS algorithm can be used to sample jointly default intensity states from the proposal distribution $\check{p}$.

Accept or reject proposal draw. By repeated application of Bayes' law:

$$
\begin{aligned}
\frac{p(X^{(\lambda)}|X^{(-\lambda)}, Y, \Theta)}{\check{p}(X^{(\lambda)}|X^{(-\lambda)}, Y, \Theta)} &\propto \frac{p(Y|X^{(\lambda)}, \Theta) \cdot p(X^{(\lambda)}|X^{(\chi)}, \Theta) \cdot p^a(\vec{k}^{(y)}|X^{(\lambda)}, \Theta) \cdot p^a(\vec{k}^{(\lambda)}|X^{(\lambda)}, \Theta)}{p^a(Y|X^{(\lambda)}, \vec{k}^{(y)}, \Theta) \cdot p^a(X^{(\lambda)}|\vec{k}^{(\lambda)}, \Theta) \cdot p^a(\vec{k}^{(y)}|X^{(\lambda)}, \Theta) \cdot \prod_{t=1}^{\mathcal{T}} w(h_t)} \\
&\propto \frac{p(Y|X^{(\lambda)}, \Theta) \cdot p(X^{(\lambda)}|X^{(\chi)}, \Theta)}{p^a(Y|X^{(\lambda)}, \Theta) \cdot p^a(X^{(\lambda)}|X^{(\chi)}, \Theta) \cdot \prod_{t=1}^{\mathcal{T}} w(h_t)}
\end{aligned}
$$

Hence, the MH acceptance probability is given by:

$$\min\left\{1, \frac{p(Y|\check{X}^{(\lambda)}, \Theta) \cdot p(\check{X}^{(\lambda)}|X^{(\chi)}, \Theta) \cdot p^a(Y|\check{X}^{(\lambda)}, \Theta) \cdot p^a(\check{X}^{(\lambda)}|X^{(\chi)}, \Theta) \cdot \prod_{t=1}^{\mathcal{T}} w(\check{X}_t^{(\lambda)})}{p(Y|X^{(\lambda)}, \Theta) \cdot p(X^{(\lambda)}|X^{(\chi)}, \Theta) \cdot p^a(Y|X^{(\lambda)}, \Theta) \cdot p^a(X^{(\lambda)}|X^{(\chi)}, \Theta) \cdot \prod_{t=1}^{\mathcal{T}} w(h_t)}\right\}$$

$$(5')$$

where $\check{X}^{(\lambda)} = (\check{h}_1, \dots, \check{h}_{\mathcal{T}})$ is the proposal draw and $X^{(\lambda)} = (h_1, \dots, h_{\mathcal{T}})$ is the current value. Relative to the standard SMP acceptance probability in (5), expression (5') introduces the ratio of kernels $p^a(\check{X}^{(\lambda)}|X^{(\chi)}, \Theta)$ and $p(\check{X}^{(\lambda)}|X^{(\chi)}, \Theta)$ to account for the second auxiliary model for the state equation. Calculation of the acceptance probabilities is straightforward, as the kernels for true model p involve only Gaussian densities, and the auxiliary kernels p^a are mixtures of Gaussian densities.

Draw the parameters. The conditional posteriors for the parameters $\kappa^{\mathbb{P}}$ and ζ^2 are conjugate under the choice of normal prior and inverse gamma prior, respectively. Thus, the Gibbs sampler can be used for these parameters. The remaining parameters $(\mu, \kappa^{\mathbb{Q}}, \sigma, \alpha)$ have complete conditional posteriors that depend intractably both on the

14

time-series of default intensity states $(h_t)_{t=1,\ldots,\mathcal{T}}$ and the observed data Y. We sample these parameters via the MH normal random walk algorithm. For all parameters we assume flat uninformative priors. We restrict $\kappa^{\mathbb{P}}$ and α to be strictly positive.

Draw the time-change increments. By Bayes' law and the conditional independence of the increments,

$$p(\chi_t|X^{(-\chi)}, Y, \Theta) \propto p(h_t|h_{t-1}, \chi_t, \Theta) \cdot p(\chi_t|\Theta)$$

It is straightforward to show that these kernels give a generalized inverse Gaussian (GIG) density from which we can sample directly in a Gibbs sampler step.

4 Empirical results

We estimate the CIR and time-changed CIR models on daily data for CDS spreads on ten single-name reference entities for the period from the beginning of 2002 to the end of 2009. The sample period spans the financial crisis of 2007–09. We retain the period from the beginning of 2010 to the end of 2011 for out-of-sample forecasting exercises in Section 5. The sample of reference entities includes U.S. and European nonfinancial corporations of varying credit quality: Alcoa, Anadarko, CenturyLink, Clear Channel, Ford, Lennar, Limited Brands, RadioShack, Sprint Nextel, and Tyson Foods. All have liquid trading in CDS. For each name, spreads on CDS of maturities 1, 2, 3, 5, 7 and 10 years are taken from the Markit database.[11] In the upper panels of Figures 3 and 4 we plot time-series of 1, 5 and 10 year CDS spreads for Alcoa and Ford, which are among the most heavily-traded names in our sample.

Even for these liquid names, it is not unusual to find missing observations in CDS data. In Table 1 we report the number of missing observations in our eight year sample by reference entity and maturity. As a practical matter, it would be infeasible to estimate the model without a convenient and rigorous means to accommodate missing observations. In MCMC estimation, we can simply treat missing observations as latent variables and sample from their complete conditional posteriors. This approach avoids additional assumptions on the missing data points and does not interfere with the law of motion imposed by the observed data points.[12]

Model-implied CDS spreads depend on the riskfree discount function P_t and on recovery rates. We parameterize P_t using the Svensson yield curve, for which we rely on daily

---

[11]The time-series for Sprint Nextel Corp. begins on August 24, 2005. Data for the other nine firms span the entire sample period.

[12]See Tsay (2010, §12.6) on the handling of missing data via data augmentation.

parameter estimates provided by the Federal Reserve Board.[13] Recovery rates are taken from the Markit database.

We estimate model specifications both with and without time-change in default intensity using the algorithm of Section 3.2. Recall that the model without time-change is nested in the time-change model as the limiting case $\alpha = \infty$. After discarding the burn-in period, we simulate chains with 100,000 draws from the invariant distribution for each model. The parameter estimates and the smoothed state estimates are calculated as means of draws from the joint posterior distribution of parameters and states.

The parameter estimates of CDS pricing models are reported in Tables 2.a–2.j. The tables report for each firm the mean, standard deviation, 5% quantile and 95% quantile of the posterior distribution for the state-space model in equations (6)–(8). All parameters are estimated with high precision except for the speed of mean reversion under the physical measure. This exception is not unexpected. In maximum likelihood estimation of a CIR specification for interest rates, Chapman and Pearson (2000) and Phillips and Yu (2009) show that $\kappa^{\mathbb{P}}$ has a large standard error in reasonable sample sizes. The speed of mean reversion under the risk-neutral measure ($\kappa^{\mathbb{Q}}$) is for all firms negative even at the 95% quantiles. This finding is qualitatively consistent with Duffee (1999), Pan and Singleton (2008) and Jacobs and Li (2008), and underscores the necessity of a pricing methodology that remains valid under non-stationarity.

The estimates of the parameter α provide strong evidence for stochastic time-change in default intensities. For two of the firms (Lennar and Sprint), the mean of the posterior distribution of α is near 20. At such values, stochastic time-change has essentially no effect on the term-structure of credit spreads. Nonetheless, as illustrated in Figure 2, the kurtosis of daily changes in CDS spreads is nonetheless an order of magnitude larger than the kurtosis in the CIR model without time-change. For RadioShack, the mean and even the 95% quantile of the posterior for α is below 2.0. At such values, time-change has a noticeable effect on credit spreads. For the remaining seven firms in our sample, the estimated α lie between these extremes.

Alcoa and Ford are typical in this regard, with estimated α between 7 and 8. In the middle panels of Figures 3 and 4, we plot smoothed estimates of the calendar-time default intensity ($\tilde{\lambda}_t$) for these two firms for the models with and without time-change. The smoothed time-series track quite closely across the two models, except when the estimated CIR default intensity dips below 1bp. Intuitively, the close alignment is a consequence of the small effect of time-change on the model-implied term-structure of spreads and the sufficiency of the observed term-structure to pin down the latent default intensity.

In the lower panels, we plot smoothed estimates of the time-change increments (χ_t)

---

[13]Daily Svensson curves are publicly available for download at the Federal Reserve Board website. For details on data sources and methodology, see Gürkaynak et al. (2007).

for the time-changed model. We scale by Δ so that the increments have unit mean. Not surprisingly, we find very large outliers during the financial crisis of 2007–09 with business time ticking 15 times faster than calendar time for the most active days for Alcoa and up to about 25 times faster for Ford. Counter to our model assumption of Lévy time change, the smoothed time-series suggest serial correlation in time-change increments. We leave this as an avenue for future research.

The moments of the estimated residuals in the measurement equations and state equation for the default intensity provide metrics of in-sample fit for the CDS pricing models with and without time-change. Table 3 reports the posterior mean, standard deviation, skewness and kurtosis of the smoothed distribution of the stacked measurement equation residuals, $\varepsilon_{t,m}^{(y)}$. A correct specification implies that these residuals are standard normal with mean zero, standard deviation one, skewness zero and kurtosis three. Comparison of the moments suggests that stochastic time-change offers modest benefit in fitting CDS spreads. The most notable improvement is a reduction in skewness for nine of the ten reference entities.

It is in the time-series dimension where we see a clear distinction between the models with and without time-change. Table 4 reports the moments of the smoothed distribution of the innovations $(\varepsilon_t^{(\lambda)})$ in the default intensity state equation. Across all firms in our sample, introducing stochastic time-change brings the mean, standard deviation, skewness and kurtosis of the state equation innovations closer to the moments of the standard normal distribution. In the case of Alcoa, for example, the mean falls from 0.05 for the CIR model to 0.01 for the time-changed model, variance increases from 0.67 to 0.82, skewness decreases from 0.34 to 0.03, and kurtosis falls from 4.27 to 3.22. Indeed, introducing time-change essentially eliminates skewness and reduces kurtosis to below 3.25 for every firm in the sample.

We use the estimated parameters for Alcoa to illustrate how stochastic time-change alters the implications of the baseline CIR model. In Figure 5, we plot the term structure of CDS spreads on three dates, one during the credit boom (18 March 2005), one in the early phase of the crisis (19 June 2008), and one in the late phase of the crisis (19 March 2009).[14] The five-year CDS spreads on these dates for Alcoa were 19bp, 93bp and 851bp, respectively. For each model and date, model parameters are fixed to the mean of the posterior distribution for the model, as reported in Table 2.a. The business time default intensity is taken from the estimated smoothed distribution of h_t for the date and model. Despite the two-orders-of-magnitude difference across dates in the initial condition (h_t), the two models produce very similar term structures, which suggests that the two models will

---

[14]Dates are chosen to be representative of the range of observed spreads over the sample period. Each of these dates immediately precedes a CDS settlement date, so the nominal maturity of a CDS on these dates equals the exact actual maturity.

yield similar fit to the cross-section of observed spreads on each sample date.

The fitted models differ, however, in the forecast distribution for spreads at a future horizon. We again fix parameters for each model to the mean of the posterior distribution for Alcoa. For simplicity, we fix a riskless rate of 3% and recovery of 40%. For the CIR model, we simulate the quantiles of the forecast distribution of $\text{CDS}(\lambda_\delta)$ for $\delta = 1/250$ (i.e., a horizon of one trading day) given initial condition λ_0. Similarly, for the time-changed model we simulate the quantiles of the forecast distribution of $\text{CDS}(\lambda(T_\delta); \alpha)$. We do this for two values of λ_0, one below the long-run mean (5 bp) and one above the long-run mean (50 bp), and report results in Table 5. Consistent with the findings on kurtosis of spread changes (Figure 2), we find that the time-change model assigns higher probability to very large changes in the spread over the short horizon.

We return briefly to consider whether empirical results are consistent with two orthogonality assumptions imposed in Section 2:

- There is no leverage effect, i.e., the innovation $\varepsilon_t^{(\lambda)}$ in the default intensity state equation is uncorrelated with time-change increments χ_t.

- The default intensity process is independent of the riskfree short rate process, i.e., $\varepsilon_t^{(\lambda)}$ is uncorrelated with the change in the federal funds rate from $t-1$ to t.

To test these assumptions, we report in Table 6 the posterior means and standard deviations of the respective time-series correlations. We find the correlation between $\varepsilon_t^{(\lambda)}$ and changes in the short rate is economically small and statistically insignificant for all ten names in our sample. The correlation between $\varepsilon_t^{(\lambda)}$ and χ_t is positive for all ten names and in some cases statistically significant, but generally of second-order economic magnitude. For all firms except Sprint, the posterior mean correlation is under 3%.

5   Out-of-sample performance

In this section, we assess out-of-sample model performance. For each observation date in the out-of-sample period from January 2010 through December 2011, we calculate the day-ahead forecast density implied by the models with and without time-change in default intensity. For each model and each date, the forecast density is evaluated at the realized next-day observation. The test statistic, due to Diebold and Mariano (1995) and Amisano and Giacomini (2007), allows for a two-sided test of the null hypothesis that the two models perform equally well.

Construction of forecast densities presents new challenges. The MCMC algorithm of Section 3.2 delivers a smoothed estimator of the default intensity, whereas for this application online filtering is required. As in our MCMC estimator, the filtering algorithm must accommodate nonlinearity in the measurement equation and non-Gaussian state evolution.

Particle filtering is well-suited to such challenges, but the most tractable filters, such as the SIR filter of Gordon et al. (1993), are ill-suited to our setting. Since the latent time-change process has serially independent increments, propagation of state variables from t to $t+1$ using only time-t information leads to particle degeneration, because the filter cannot "anticipate" abrupt changes in CDS spreads.[15] Contrasted to such non-adapted filters are fully-adapted filters, in which propagation is conditioned on the observation at $t+1$. Here the challenge is tractability. To navigate between these obstacles, we develop a partially-adapted particle filter using an auxiliary linearized model in the spirit of our approach in Section 3.2. To reduce the dimensionality of the problem, we fix all model parameters to the in-sample posterior means estimated in Section 4.

5.1 Partially-adapted particle filter on a linearized model

The particle filter is a recursive algorithm that constructs a discrete approximation to the filtered distribution of the state variables at each observation date:

$$x_t^{(i)} \overset{\text{iid}}{\sim} p(x_t|Y^t) \text{ for all } t = \mathcal{T} + 1, \ldots, \mathcal{T}^*$$

where $Y^t = (y_1, \ldots, y_t)$ are observations available up to time t, x_t is the state vector, $\mathcal{I}$ is the number of particles $i = 1, \ldots, \mathcal{I}$, and $\mathcal{T}^* - \mathcal{T}$ is the number of trading days in the out-of-sample period. To reduce notation, explicit dependence on model parameters is omitted throughout this section.

In the auxiliary particle filter (APF) of Pitt and Shephard (1999), the sample of particles at observation time t is constructed as the marginal distribution for x_t in the joint distribution for (x_t, x_{t-1}) given Y^t, for which the kernel can be decomposed as

$$p(x_t, x_{t-1}|Y^t) \propto p(x_t|x_{t-1}, y_t) \cdot p(y_t|x_{t-1}) \cdot p(x_{t-1}|Y^{t-1}). \tag{11}$$

Fully-adapted sampling from this distribution is conducted in two steps. First, the particles $\{x_{t-1}^{(i)}\}_{i=1,\ldots,\mathcal{I}}$ that serve as a discrete approximation to $p(x_{t-1}|Y^{t-1})$ are resampled with weights proportional to $p(y_t|x_{t-1}^{(i)})$. This step requires pointwise evaluation of the required density and induces consistency of the "old" particles $x_{t-1}^{(i)}$ with the "new" observation y_t. Second, the resampled time-$t-1$ particles are propagated to time t by sampling from $p(x_t|x_{t-1}, y_t)$. As is often the case in application, in our model setting we cannot easily evaluate $p(y_t|x_{t-1})$ for the resampling step, nor can we sample directly from $p(x_t|x_{t-1}, y_t)$.

---

[15]Particle degeneration arises when only a small number of particles have significant weight in the resampling stage. See Lopes and Tsay (2011) for a review of particle filter methods with discussion of particle degeneration.

We substitute $x_t = (h_t, \chi_t)$ in equation (11) and re-write the first two kernels as

$$p(h_t, \chi_t | h_{t-1}, \chi_{t-1}, y_t) \cdot p(y_t | h_{t-1}, \chi_{t-1})$$
$$= p(h_t | \chi_t, h_{t-1}, \chi_{t-1}, y_t) \cdot p(y_t | \chi_t, h_{t-1}, \chi_{t-1}) \cdot p(\chi_t | h_{t-1}, \chi_{t-1}, Y^{t-1}).$$

In our model, the χ_t are iid, so $p(\chi_t | h_{t-1}, \chi_{t-1}, Y^{t-1}) = p(\chi_t)$. Conditional on h_{t-1}, the CDS spread at t does not depend on χ_{t-1}, so $p(y_t | \chi_t, h_{t-1}, \chi_{t-1}) = p(y_t | \chi_t, h_{t-1})$. Similarly, conditional on h_{t-1}, h_t does not depend on χ_{t-1}, so $p(h_t | \chi_t, h_{t-1}, \chi_{t-1}, y_t) = p(h_t | \chi_t, h_{t-1}, y_t)$. Combining these terms, we re-write equation (11) as

$$p(h_t, \chi_t, h_{t-1}, \chi_{t-1} | Y^t) \propto p(h_t | \chi_t, h_{t-1}, y_t) \cdot p(y_t | \chi_t, h_{t-1}) p(\chi_t) \cdot p(h_{t-1}, \chi_{t-1} | Y^{t-1}) \quad (12)$$

As the propagation kernel $p(h_t | \chi_t, h_{t-1}, y_t)$ is intractable, we introduce an auxiliary linear model.[16] Begin by writing

$$p(h_t | \chi_t, h_{t-1}, y_t) \cdot p(y_t | \chi_t, h_{t-1}) = p(y_t | h_t, \chi_t, h_{t-1}) \cdot p(h_t | \chi_t, h_{t-1}) \quad (13)$$

In our setting, this is simplified by noting that

$$p(y_t | h_t, \chi_t, h_{t-1}) = p(y_t | h_t) = \prod_{m \in \mathfrak{M}} p(y_{t,m} | h_t).$$

In the auxiliary model, the measurement equation is linearized as

$$y_{t,m} = \gamma_{t,m} + \beta_{t,m} h_t + \zeta \, \varepsilon_{t,m}^{(y)} \quad (6')$$

so that $p^a(y_{t,m} | h_t) \sim N(\gamma_{t,m} + \beta_{t,m} h_t, \zeta^2)$. The choice of constants $\gamma_{t,m}$ and $\beta_{t,m}$ determines how well the auxiliary linear model approximates the true nonlinear model. The auxiliary model state equations are assumed to be the same as those for the true model in (7) and (8). In the appendix, we derive a proposal propagation density $\breve{p}(h_t | \chi_t, h_{t-1}, y_t)$ and auxiliary evaluation weights $p^a(y_t | \chi_t, h_{t-1})$ such that equation (13) can be replaced by

$$\breve{p}(h_t | \chi_t, h_{t-1}, y_t) \cdot p^a(y_t | \chi_t, h_{t-1}) = p^a(y_t | h_t) \cdot p(h_t | \chi_t, h_{t-1}) \quad (13')$$

The proposal propagation density is Gaussian (so sampling is trivial), and the auxiliary evaluation weights are similarly convenient to compute.

We arrive at the following algorithm for sampling $\{h_t, \chi_t\}$ conditional on particles $\{h_{t-1}^{(i)}\}_{i=1,\dots,\mathcal{I}}$ and on Y^t:

---

[16] The possibility of constructing a proposal distribution on a linearization of the true model was suggested by Pitt and Shephard (1999, §3.3).

1. Draw particles $\{\chi_t^{(i)}\}_{i=1,\dots,\mathcal{I}}$ blindly from the unconditional IG distribution for χ_t.

2. Draw particles $\{\dot{h}_{t-1}^{(i)}, \dot{\chi}_t^{(i)}\}_{i=1,\dots,\mathcal{I}}$ by resampling $\{h_{t-1}^{(i)}, \chi_t^{(i)}\}_{i=1,\dots,\mathcal{I}}$ with weights proportional to $p^a(y_t|\chi_t^{(i)}, h_{t-1}^{(i)})$.

3. Propagate $\{\breve{h}_t^{(i)}\}_{i=1,\dots,\mathcal{I}}$ by sampling from $\breve{p}(h_t|\dot{\chi}_t^{(i)}, \dot{h}_{t-1}^{(i)}, y_t)$.

4. To restore consistency with the law of the true model, generate $\{h_t^{(i)}, \ddot{h}_{t-1}^{(i)}, \chi_t^{(i)}\}_{i=1,\dots,\mathcal{I}}$ by resampling from $\{\breve{h}_t^{(i)}, \dot{h}_{t-1}^{(i)}, \dot{\chi}_t^{(i)}\}_{i=1,\dots,\mathcal{I}}$ with weights proportional to[17]

$$\omega_t^{(i)} = \frac{p(\breve{h}_t^{(i)}|\dot{\chi}_t^{(i)}, \dot{h}_{t-1}^{(i)}, y_t) \cdot p(y_t|\dot{\chi}_t^{(i)}, \dot{h}_{t-1}^{(i)})}{\breve{p}(\breve{h}_t^{(i)}|\dot{\chi}_t^{(i)}, \dot{h}_{t-1}^{(i)}, y_t) \cdot p^a(y_t|\dot{\chi}_t^{(i)}, \dot{h}_{t-1}^{(i)})} = \frac{p(y_t|\breve{h}_t^{(i)})}{\breve{p}(y_t|\breve{h}_t^{(i)})} \tag{14}$$

5. Discard the sample of χ_t, and regenerate by sampling $\chi_t^{(i)}$ from $p(\chi_t|h_t^{(i)}, \ddot{h}_{t-1}^{(i)})$. As noted in Section 3.2, the conditional distributed is GIG.

This algorithm is fully-adapted for h_t with respect to the auxiliary model. Even though the auxiliary model diverges from the true model, with reasonable choices of $\gamma_{t,m}$ and $\beta_{t,m}$ we can still expect that the realization of y_t will inform the propagation of h_{t-1} to h_t. We impose

$$\beta_{t,m} = F'_{t,m}(\widehat{h}_t)$$
$$\gamma_{t,m} = F_{t,m}(\widehat{h}_t) - \beta_{t,m}\widehat{h}_t$$

where $\widehat{h}_t$ is given by the inverse of model-implied 5-year CDS spread observed at time t: $\widehat{h}_t = F_{t,5}^{-1}(y_{t,5})$.

Sampling of χ_t in Step 1 is non-adapted, but the regenerated sample in Step 5 is fully adapted, i.e., it utilizes all the information content of y_t under the true model. The regeneration step is permissible only because χ_t does not enter the observation equation (i.e., because y_t is independent of χ_t when conditioned on h_t). It is straightforward to show that the regenerated $\{\chi_t\}_{i=1,\dots,\mathcal{I}}$ has the same conditional distribution as the sample after Step 4, but regeneration improves the efficiency of sampling. We include the final step to facilitate examination of the time-series of filtered states in Section 5.2.

Finally, it is trivial to modify the algorithm for the baseline model without time-change, because the pricing model is a special case of the model with time-change. We can simply fix $\chi_t^{(i)} = \Delta$ in Step 1 and drop Step 5.

---

[17]We denote the resampled h_{t-1} as $\{\ddot{h}_{t-1}\}_{i=1,\dots,\mathcal{I}}$ in order to avoid confusion with the original cloud of particles $\{h_{t-1}\}_{i=1,\dots,\mathcal{I}}$, which are conditioned only on Y^{t-1}.

5.2 Density forecast analysis

In this section we present the out-of-sample model selection analysis based on the Diebold and Mariano (1995) and Amisano and Giacomini (2007) test statistic. We begin with a brief description of the test. The average predictive likelihood for the out-of-sample points $t = \mathcal{T} + 1, \ldots, \mathcal{T}^*$ is given by:

$$\mathcal{L}_{\mathcal{M}} = \frac{1}{\mathcal{T}^* - \mathcal{T}} \sum_{t=\mathcal{T}+1}^{\mathcal{T}^*} \log\left(p(y_t|Y^{t-1}, \Theta_{\mathcal{M}})\right)$$

where $\mathcal{M} \in \{TC, CIR\}$ denotes a model specification with "CIR" for the baseline model without time-change and "TC" for the model with time-change, p denotes the predictive likelihood evaluated at the future realized CDS log-spread at time t, given model parameters $\Theta_{\mathcal{M}}$, which are fixed at the in-sample MCMC estimates. The predictive likelihood can be computed from the filtering distribution as follows:

$$p(y_t|Y^{t-1}, \Theta_{\mathcal{M}}) = \int p(y_t|x_t, \Theta_{\mathcal{M}}) \, p(x_t|Y^{t-1}, \Theta_{\mathcal{M}}) \, dx_t$$

$$p(x_t|Y^{t-1}, \Theta_{\mathcal{M}}) = \int p(x_t|x_{t-1}, \Theta_{\mathcal{M}}) \, p(x_{t-1}|Y^{t-1}, \Theta_{\mathcal{M}}) \, dx_{t-1}$$

The particle filter provides a sample $\{x_{t-1}^{(i)}\}_{i=1,\ldots,\mathcal{I}}$ from the filtering distribution $p(x_{t-1}|Y^{t-1}, \Theta_{\mathcal{M}})$. We propagate this sample forward via the state evolution equations by drawing from the transition density $p(x_t|x_{t-1}^{(i)}, \Theta_{\mathcal{M}})$ to obtain a sample $\{\hat{x}_t^{(i)}\}_{i=1,\ldots,\mathcal{I}}$ from the distribution $p(x_t|Y^{t-1}, \Theta_{\mathcal{M}})$. The predictive likelihood for observation t is approximated as the Monte Carlo integral

$$\widehat{\mathcal{L}}_{\mathcal{M},t} = \log\left(\frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} p(y_t|\hat{x}_t^{(i)}, \Theta_{\mathcal{M}})\right).$$

Finally, we average across time to obtain

$$\widehat{\mathcal{L}}_{\mathcal{M}} = \frac{1}{\mathcal{T}^* - \mathcal{T}} \sum_{t=\mathcal{T}+1}^{\mathcal{T}^*} \widehat{\mathcal{L}}_{\mathcal{M},t}.$$

Intuitively, a model that performs better out-of-sample has a higher value of average predictive likelihood. Under the null hypothesis of equal performance the Amisano and Giacomini (2007) statistic

$$U = \sqrt{\mathcal{T}^* - \mathcal{T}} \frac{(\widehat{\mathcal{L}}_{TC} - \widehat{\mathcal{L}}_{CIR})}{\sqrt{\frac{1}{\mathcal{T}^* - \mathcal{T} - 1} \sum_{t=\mathcal{T}+1}^{\mathcal{T}^*} \left(\widehat{\mathcal{L}}_{TC,t} - \widehat{\mathcal{L}}_{CIR,t}\right)^2 - \left(\widehat{\mathcal{L}}_{TC} - \widehat{\mathcal{L}}_{CIR}\right)^2}} \tag{15}$$

has an asymptotic standard normal distribution.

The partially-adapted APF of Section 5.1 is implemented for each model and each firm with $\mathcal{I} = 100{,}000$ particles. We confirm that our particle filter avoids significant particle degeneration. In the case of the time-changed model for Ford, for example, the range of the efficient sample size (as a fraction of the number of particles) for the partially-adapted APF is $(68.3\%, 99.8\%)$ with the 1%-quantile of 80.6%. We have implemented the standard SIR filter as well, and in this case find a range of $(1.7\%, 88.4\%)$ along with the 1%-quantile of 15.6%.

The test statistic U is estimated for each of the ten names in our sample and reported in Table 7. For nine of the names (all firms but Lennar), the coefficient is positive and significant at the 1% level. In Figures 6 and 7, we contrast the typical case of Ford with the exceptional case of Lennar. We plot CDS spreads (top graph), the filtered estimates of default intensity (middle graph), and the filtered estimates of time-change increments (bottom graph) for the out-of-sample period covering 2010 and 2011.[18] The filtered time-series of χ_t for Ford exhibits significant volatility with multiple spikes, whereas for Lennar this time-series is roughly constant with only two minor instances of higher activity. In each case, the spikes in χ_t coincide with relatively large movements in the one-year CDS spread and the filtered default intensity. Thus, adding stochastic time-change to the baseline CIR model improves the forecast performance when the CDS spread exhibits volatility in volatility, but may lead to overfit when volatility is relatively constant.

6 Conclusion

Stochastic volatility has been studied extensively in markets for equities, interest rates and commodities, but there has been relatively little empirical work to date on stochastic volatility in credit markets. We appear to be the first to estimate a model in which stochastic volatility in default intensity is induced by stochastic time change. Nonlinearity in the CDS pricing function and state-dependent variance in the evolution of the default intensity present significant challenges for estimation. We overcome these difficulties using a new Bayesian MCMC estimation method that builds on earlier work by Stroud, Müller, and Polson (2003). For out-of-sample density forecasts, we also develop an auxiliary particle filter. Although computationally intensive, both methodologies performed well on our data, and may have application to related problems in modeling interest rate or volatility swaps.

We estimate our model on firm-level time-series of single-name credit default swap spreads. In-sample, stochastic time-change is found to be a statistically and economically significant feature of the data for all firms in our sample. In out-of-sample density forecasts

---

[18]To be precise, the plotted value at date t for the filtered estimate is the mean of the sample of $\{x_t^{(i)}\}_{i=1,\ldots,\mathcal{I}}$ from the partially-adapted APF.

tests, we find that the time-changed model outperforms the baseline CIR model at high levels of significance for nine of our ten firms.

Smoothed estimates of the increments of stochastic time show large outliers during the financial crisis of 2007–09 with business time ticking substantially faster than calendar time for the most volatile days in our sample. Intuitively, large jumps in the business clock allow the model to accommodate large daily changes in the default intensity (i.e., relative to volatility). While the time-changed model offers only modest improvement in fitting the term-structure of observable spreads on any given date, it greatly improves the fit in the time-series dimension. The findings have important implications for risk-management applications, because the time-changed model allows for much higher kurtosis in the loss distribution over short horizons, and so higher quantiles in the tail (i.e., higher value-at-risk). For similar reasons, the findings also suggest that the widely-used CIR model will tend to underprice deep out-of-the-money options on CDS.

While the model imposes a variety of independence restrictions, only one appears to be materially at odds with the data. Counter to our model assumption of Lévy time change, the smoothed time-series of state variables exhibit serial correlation in time-change increments. The pricing methodology can easily be extended to allow for such volatility clustering, but our MCMC and particle filter methodologies would be quite challenging without the tractability afforded by serial independence. We leave this as an avenue for future research.

Appendix: Proposal and resampling weights in linearized model

Application of the proposed adapted filter requires sampling from $\breve{p}(h_t|\chi_t, h_{t-1}, y_t)$ and evaluation of the weights $p^a(y_t|\chi_t, h_{t-1})$. We begin with the kernels on the right-hand-side of equation $(13')$. Since $p^a(y_{t,m}|h_t) \sim N(\gamma_{t,m} + \beta_{t,m}h_t, \zeta^2)$ and the measurement errors are independent across maturity, we have

$$
p^a(y_t|h_t) = \prod_{m \in \mathfrak{M}} \phi\left(y_{t,m}; \gamma_{t,m} + \beta_{t,m}h_t, \zeta\right)
$$

$$
= \left(\frac{1}{\sqrt{2\pi}\zeta}\right)^{\#\mathfrak{M}} \prod_{m \in \mathfrak{M}} \exp\left(-\frac{1}{2}\left(\frac{y_{t,m} - \gamma_{t,m} - \beta_{t,m}h_t}{\zeta}\right)^2\right)
$$

$$
= \left(\frac{1}{\sqrt{2\pi}\zeta}\right)^{\#\mathfrak{M}} \prod_{m \in \mathfrak{M}} \exp\left(-\frac{1}{2}\left(\frac{h_t - \nu_{t,m}}{\xi_{t,m}}\right)^2\right) = \prod_{m \in \mathfrak{M}} \frac{1}{\beta_{t,m}} \phi\left(h_t; \nu_{t,m}, \xi_{t,m}\right)
$$

where $\nu_{t,m} \equiv (y_{t,m} - \gamma_{t,m})/\beta_{t,m}$ and $\xi_{t,m} \equiv \zeta/\beta_{t,m}$. From the state equation (8), we have $p(h_t|\chi_t, h_{t-1}) = \phi\left(h_t; \nu_{t,*}, \xi_{t,*}\right)$ where $\nu_{t,*} \equiv h_{t-1} + (\mu - \kappa^{\mathbb{P}}h_{t-1})\chi_t$ and $\xi_{t,*}^2 \equiv \sigma^2 h_{t-1}\chi_t$, where the asterisk indicates parameters of the state equation. For notational convenience in combining these kernels, we define the set $\mathfrak{N} = \mathfrak{M} \cup \{*\}$.

The right-hand-side of equation $(13')$ is

$$p^a(y_t|h_t)p(h_t|\chi_t, h_{t-1}) = \left(\prod_{m \in \mathfrak{M}} \frac{1}{\beta_{t,m}} \phi\left(h_t; \nu_{t,m}, \xi_{t,m}\right) \right) \phi\left(h_t; \nu_{t,*}, \xi_{t,*}\right)$$

$$= \left(\frac{1}{\sqrt{2\pi}\zeta} \right)^{\#\mathfrak{M}} \frac{1}{\sqrt{2\pi}\xi_{t,*}} \exp\left(-\frac{1}{2} \sum_{m \in \mathfrak{N}} \left(\frac{h_t - \nu_{t,m}}{\xi_{t,m}} \right)^2 \right) \quad (16)$$

For $m \in \mathfrak{N}$, let $\xi_{t,-m}^2 \equiv \prod_{j \in \mathfrak{N}\setminus\{m\}} \xi_{t,j}^2$, and let $\eta_{t,m}$ be the weight $\eta_{t,m} = \xi_{t,-m}^2 / \sum_{j \in \mathfrak{N}} \xi_{t,-j}^2$. Let $\bar{\xi}_t^2$ be the harmonic sum of the variances, i.e.,

$$\bar{\xi}_t^2 \equiv \frac{\prod_{m \in \mathfrak{N}} \xi_{t,m}^2}{\sum_{m \in \mathfrak{N}} \xi_{t,-m}^2} = \left(\sum_{m \in \mathfrak{N}} \frac{1}{\xi_{t,m}^2} \right)^{-1}$$

It is straightforward to sum the fractions within the exp() term of equation (16) as

$$\sum_{m \in \mathfrak{N}} \left(\frac{h_t - \nu_{t,m}}{\xi_{t,m}} \right)^2 = \left(\frac{h_t - \sum_{m \in \mathfrak{N}} \eta_{t,m}\nu_{t,m}}{\bar{\xi}_t} \right)^2 + \frac{1}{\bar{\xi}_t^2} \left(\sum_{m \in \mathfrak{N}} \eta_{t,m}\nu_{t,m}^2 - \left(\sum_{m \in \mathfrak{N}} \eta_{t,m}\nu_{t,m} \right)^2 \right)$$

Observe that h_t appears only in the first term, which we embed in the proposal density as

$$\breve{p}(h_t|\chi_t, h_{t-1}, y_t) = \phi\left(h_t; \sum_{m \in \mathfrak{N}} \eta_{t,m}\nu_{t,m}, \bar{\xi}_t \right)$$

Intuitively, the proposal is the posterior distribution for an unobserved state given a set of independent Gaussian signals with realization $\{\nu_{t,m}\}_{m \in \mathfrak{N}}$. The mth signal receives weight $\eta_{t,m}$, which is proportional to its precision $1/\xi_{t,m}^2$. The precision of the posterior, $1/\bar{\xi}_t^2$, is the sum of signal precisions.

To enforce identity $(13')$, remaining terms in equation (16) are captured in the resampling weight

$$p^a(y_t|\chi_t, h_{t-1}) = \left(\frac{1}{\sqrt{2\pi}\zeta} \right)^{\#\mathfrak{M}} \frac{\bar{\xi}}{\xi_{t,*}} \exp\left(-\frac{1}{2} \frac{\sum_{m \in \mathfrak{N}} \eta_{t,m}\nu_{t,m}^2 - \left(\sum_{m \in \mathfrak{N}} \eta_{t,m}\nu_{t,m} \right)^2}{\bar{\xi}_t^2} \right)$$

The resampling weight is low when the variation across signals $\nu_{t,m}$ is large. Thus, a particle carrying state (χ_t, h_{t-1}) is penalized in the resampling step when it implies large differences across the $\#\mathfrak{M} + 1$ equations in the likelihood of y_t.

References

Carol Alexander and Andreas Kaeck. Capital structure and asset prices: Some effects of bankruptcy procedures. *Journal of Banking and Finance*, 32(6):1008–1021, June 2008.

Gianni Amisano and Raffaella Giacomini. Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business and Economics Statistics*, 25(2):177–199, 2007.

Torben G. Andersen, Luca Benzoni, and Jesper Lund. An empirical investigation of continuous-time equity return models. *Journal of Finance*, 57(3):1239–1284, 2002.

Ole E. Barndorff-Nielsen and Albert Shiryaev. *Change of Time and Change of Measure*, volume 13 of *Advanced Series on Statistical Science and Applied Probability*. World Scientific, 2010.

Bradley P. Carlin, Nicholas G. Polson, and David S. Stoffer. A Monte Carlo approach to nonnormal and nonlinear state-space modeling. *Journal of the American Statistical Association*, 87(418):493–500, 1992.

C.K. Carter and R. Kohn. On Gibbs sampling for state space models. *Biometrika*, 81(3): 541–553, 1994.

David A. Chapman and Neil D. Pearson. Is the short rate drift actually nonlinear? *Journal of Finance*, 55(1):355–388, February 2000.

Siddhartha Chib and Edward Greenberg. Markov chain Monte Carlo simulation methods in econometrics. *Econometric Theory*, 12(3):409–431, 1996.

Ovidiu Costin, Michael B. Gordy, Min Huang, and Pawel J. Szerszen. Expectations of functions of stochastic time with application to credit risk modeling. *Mathematical Finance*, forthcoming.

Francis X. Diebold and Roberto S. Mariano. Comparing predictive accuracy. *Journal of Business and Economics Statistics*, 13(3):252–263, 1995.

Joost Driessen. Is default event risk priced in corporate bonds? *Review of Financial Studies*, 18:165–195, 2005.

Gregory R. Duffee. Estimating the price of default risk. *Review of Financial Studies*, 12 (1):197–226, Spring 1999.

Darrell Duffie and Kenneth J. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720, 1999.

Darrell Duffie and Kenneth J. Singleton. *Credit Risk: Pricing, Measurement, and Management.* Princeton University Press, Princeton, 2003.

Bjørn Eraker. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance*, 59(3):1367–1403, 2004.

Sylvia Frühwirth-Schnatter. Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2):183–202, 1994.

D. Geman and S. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F on Radar and Signal Process*, 140(2):107–113, April 1993.

Michael B. Gordy and Søren Willemann. Constant proportion debt obligations: A post-mortem analysis of rating models. *Management Science*, 58(3):476–492, March 2012.

Refet S. Gürkaynak, Brian Sack, and Jonathan H. Wright. The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291–2304, 2007.

Kris Jacobs and Xiaofei Li. Modeling the dynamics of credit spreads with stochastic volatility. *Management Science*, 54(6):1176–1188, June 2008.

Eric Jacquier, Nicholas G. Polson, and Peter E. Rossi. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122(1):185–212, September 2004.

Robert Jarrow and Stuart Turnbull. Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*, 50(1):53–86, March 1995.

Robert A. Jarrow, David Lando, and Fan Yu. Default risk and diversification: Theory and empirical implications. *Mathematical Finance*, 15(1):1–26, January 2005.

Michael S. Johannes, Nicholas G. Polson, and Jonathan R. Stroud. Optimal filtering of jump diffusions: Extracting latent states from asset prices. *Review of Financial Studies*, 22(7):2759–2799, 2009.

Christopher S. Jones. The dynamics of stochastic volatility: evidence from underlying and options market. *Journal of Econometrics*, 116(1–2):181–224, 2003.

Matthew Leeming, Søren Willemann, Arup Ghosh, and Rob Hagemans. Standard corporate CDS handbook: Ongoing evolution of the CDS market. Technical report, Barclays Capital, February 2010.

Hedibert F. Lopes and Ruey S. Tsay. Particle filters and Bayesian inference in financial econometrics. *Journal of Forecasting*, 30:168–209, 2011.

Rafael Mendoza-Arriaga and Vadim Linetsky. Time-changed CIR default intensities with two-sided mean-reverting jumps. *Annals of Applied Probability*, 24(2):811–856, 2014.

N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6): 1087–1092, 1953.

Jun Pan and Kenneth J. Singleton. Default and recovery implicit in the term structure of sovereign CDS spreads. *Journal of Finance*, 63(5), October 2008.

Peter C.B. Phillips and Jun Yu. Simulation-based estimation of contingent-claims prices. *Review of Financial Studies*, 22(9):3669–3705, September 2009.

Michael K. Pitt and Neil Shephard. Analytic convergence rates and parameterization issues for the Gibbs sampler applied to state space models. *Journal of Time Series Analysis*, 20(1):63–85, January 1999.

Jonathan R. Stroud, Peter Müller, and Nicholas G. Polson. Nonlinear state-space models with state-dependent variances. *Journal of the American Statistical Association*, 98(462): 377–386, 2003.

Ruey S. Tsay. *Analysis of Financial Time Series*. John Wiley & Sons, Hoboken, NJ, third edition, 2010.

Table 1: **Counts of missing observations in CDS data.**
The table reports the number of missing observations in the estimation sample from the beginning of 2002 to the end of 2009 (with the time series for Sprint Nextel starting on August 24, 2005).

	Maturity					
	1y	2y	3y	5y	7y	10y
Alcoa	0	0	0	0	0	8
Anadarko	6	32	0	0	0	8
CenturyLink	61	115	12	3	11	107
Clear Channel	0	1	0	0	0	0
Ford	0	0	0	0	0	0
Lennar	120	301	71	2	218	233
Limited Brands	18	42	10	3	0	19
RadioShack	16	17	4	0	1	4
Sprint Nextel	0	0	0	0	0	0
Tyson Foods	39	147	22	30	25	208

Table 2.a: **Parameter estimates of CDS pricing models: Alcoa.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior distribution for the CDS pricing model in equations (6)-(8). The results are reported for model specifications without time-change (CIR model) and with time-change in default intensity (CIR + IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
Alcoa: CIR Model				
$\kappa^{\mathbb{P}}$	0.4794	0.3698	0.0366	1.1952
σ	0.1877	0.0030	0.1826	0.1925
$\mu * 100$	0.0829	0.0009	0.0814	0.0844
$\kappa^{\mathbb{Q}}$	-0.2526	0.0068	-0.2634	-0.2413
ζ	0.1627	0.0011	0.1609	0.1645
Alcoa: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	0.6590	0.5188	0.0483	1.6713
σ	0.2238	0.0019	0.2207	0.2269
$\mu * 100$	0.0688	0.0009	0.0673	0.0702
$\kappa^{\mathbb{Q}}$	-0.3787	0.0068	-0.3903	-0.3679
ζ	0.1583	0.0010	0.1566	0.1600
α	7.1439	0.4384	6.4713	7.9060

Table 2.b: **Parameter estimates of CDS pricing models: Anadarko.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior distribution for the CDS pricing model in equations (6)-(8). The results are reported for model specifications without time-change (CIR model) and with time-change in default intensity (CIR + IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
Anadarko: CIR Model				
$\kappa^{\mathbb{P}}$	0.6569	0.3852	0.0915	1.3452
σ	0.0975	0.0042	0.0910	0.1051
$\mu * 100$	0.1643	0.0012	0.1623	0.1662
$\kappa^{\mathbb{Q}}$	-0.0318	0.0035	-0.0382	-0.0265
ζ	0.1627	0.0011	0.1610	0.1645
Anadarko: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	1.7103	1.2431	0.1488	4.0767
σ	0.2596	0.0025	0.2555	0.2636
$\mu * 100$	0.1096	0.0018	0.1067	0.1128
$\kappa^{\mathbb{Q}}$	-0.3398	0.0094	-0.3547	-0.3223
ζ	0.1535	0.0010	0.1519	0.1552
α	5.5933	0.3528	5.0505	6.2100

Table 2.c: **Parameter estimates of CDS pricing models: CenturyLink.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior distribution for the CDS pricing model in equations (6)-(8). The results are reported for model specifications without time-change (CIR model) and with time-change in default intensity (CIR + IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
CenturyLink: CIR Model				
$\kappa^{\mathbb{P}}$	1.1616	0.7956	0.1122	2.6471
σ	0.2531	0.0036	0.2471	0.2589
$\mu * 100$	0.2095	0.0031	0.2044	0.2145
$\kappa^{\mathbb{Q}}$	-0.3253	0.0104	-0.3423	-0.3078
ζ	0.2002	0.0013	0.1980	0.2024
CenturyLink: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	1.8575	1.3345	0.1640	4.3793
σ	0.2922	0.0023	0.2884	0.2960
$\mu * 100$	0.1558	0.0034	0.1505	0.1618
$\kappa^{\mathbb{Q}}$	-0.5531	0.0132	-0.5730	-0.5302
ζ	0.1944	0.0013	0.1923	0.1965
α	5.9933	0.4425	5.3364	6.7968

Table 2.d: **Parameter estimates of CDS pricing models: Clear Channel.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior distribution for the CDS pricing model in equations (6)-(8). The results are reported for model specifications without time-change (CIR model) and with time-change in default intensity (CIR + IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
Clear Channel: CIR Model				
$\kappa^{\mathbb{P}}$	0.2310	0.1806	0.0171	0.5820
σ	0.2332	0.0040	0.2266	0.2398
$\mu * 100$	0.2594	0.0064	0.2493	0.2702
$\kappa^{\mathbb{Q}}$	-0.3094	0.0115	-0.3275	-0.2904
ζ	0.2689	0.0017	0.2661	0.2719
Clear Channel: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	0.1759	0.1373	0.0133	0.4422
σ	0.2662	0.0038	0.2601	0.2725
$\mu * 100$	0.1983	0.0056	0.1889	0.2081
$\kappa^{\mathbb{Q}}$	-0.5013	0.0172	-0.5301	-0.4738
ζ	0.2640	0.0017	0.2613	0.2669
α	3.8124	0.4316	3.1516	4.5369

Table 2.e: **Parameter estimates of CDS pricing models: Ford.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior distribution for the CDS pricing model in equations (6)-(8). The results are reported for model specifications without time-change (CIR model) and with time-change in default intensity (CIR + IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
Ford: CIR Model				
$\kappa^{\mathbb{P}}$	0.3466	0.2579	0.0283	0.8446
σ	0.3207	0.0028	0.3162	0.3254
$\mu * 100$	0.4686	0.0094	0.4529	0.4834
$\kappa^{\mathbb{Q}}$	-0.4997	0.0086	-0.5136	-0.4859
ζ	0.1824	0.0012	0.1805	0.1844
Ford: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	0.3299	0.2508	0.0263	0.8129
σ	0.3390	0.0028	0.3345	0.3436
$\mu * 100$	0.4147	0.0112	0.3958	0.4318
$\kappa^{\mathbb{Q}}$	-0.6419	0.0127	-0.6626	-0.6213
ζ	0.1801	0.0012	0.1782	0.1821
α	7.5902	0.6288	6.6347	8.6549

Table 2.f: **Parameter estimates of CDS pricing models: Lennar.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior distribution for the CDS pricing model in equations (6)-(8). The results are reported for model specifications without time-change (CIR model) and with time-change in default intensity (CIR + IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
Lennar: CIR Model				
$\kappa^{\mathbb{P}}$	0.4378	0.3252	0.0356	1.0646
σ	0.2576	0.0057	0.2479	0.2670
$\mu * 100$	0.2983	0.0037	0.2924	0.3047
$\kappa^{\mathbb{Q}}$	-0.1736	0.0114	-0.1927	-0.1545
ζ	0.1802	0.0013	0.1782	0.1823
Lennar: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	0.6873	0.5193	0.0537	1.6897
σ	0.3197	0.0037	0.3137	0.3258
$\mu * 100$	0.2593	0.0027	0.2549	0.2639
$\kappa^{\mathbb{Q}}$	-0.3330	0.0103	-0.3500	-0.3158
ζ	0.1771	0.0012	0.1751	0.1791
α	18.1031	1.4843	15.8002	20.6781

Table 2.g: **Parameter estimates of CDS pricing models: Limited Brands.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior
distribution for the CDS pricing model in equations (6)-(8). The results are reported for model
specifications without time-change (CIR model) and with time-change in default intensity (CIR +
IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
Limited Brands: CIR Model				
$\kappa^{\mathbb{P}}$	0.3738	0.2655	0.0343	0.8747
σ	0.1334	0.0068	0.1218	0.1442
$\mu * 100$	0.2292	0.0041	0.2227	0.2361
$\kappa^{\mathbb{Q}}$	-0.0922	0.0088	-0.1063	-0.0771
ζ	0.2399	0.0016	0.2374	0.2425
Limited Brands: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	0.9273	0.7077	0.0722	2.2911
σ	0.2531	0.0034	0.2477	0.2587
$\mu * 100$	0.1518	0.0037	0.1453	0.1576
$\kappa^{\mathbb{Q}}$	-0.3907	0.0148	-0.4178	-0.3680
ζ	0.2334	0.0015	0.2309	0.2358
α	3.8126	0.3055	3.3556	4.3503

Table 2.h: **Parameter estimates of CDS pricing models: RadioShack.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior
distribution for the CDS pricing model in equations (6)-(8). The results are reported for model
specifications without time-change (CIR model) and with time-change in default intensity (CIR +
IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
RadioShack: CIR Model				
$\kappa^{\mathbb{P}}$	0.6419	0.4726	0.0538	1.5468
σ	0.1907	0.0025	0.1866	0.1947
$\mu * 100$	0.0681	0.0022	0.0644	0.0717
$\kappa^{\mathbb{Q}}$	-0.3902	0.0082	-0.4036	-0.3764
ζ	0.2660	0.0018	0.2632	0.2689
RadioShack: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	1.3374	1.0143	0.1068	3.2928
σ	0.1968	0.0025	0.1928	0.2008
$\mu * 100$	0.0388	0.0018	0.0358	0.0418
$\kappa^{\mathbb{Q}}$	-0.6591	0.0167	-0.6856	-0.6314
ζ	0.2561	0.0016	0.2534	0.2588
α	1.7946	0.0901	1.6592	1.9488

Table 2.i: **Parameter estimates of CDS pricing models: Sprint.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior
distribution for the CDS pricing model in equations (6)-(8). The results are reported for model
specifications without time-change (CIR model) and with time-change in default intensity (CIR +
IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
Sprint: CIR Model				
$\kappa^{\mathbb{P}}$	0.5126	0.4062	0.0367	1.3034
σ	0.2404	0.0032	0.2351	0.2456
$\mu * 100$	0.1407	0.0018	0.1378	0.1437
$\kappa^{\mathbb{Q}}$	-0.3992	0.0101	-0.4154	-0.3820
ζ	0.1836	0.0017	0.1809	0.1864
Sprint: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	0.5894	0.4706	0.0422	1.5102
σ	0.2543	0.0028	0.2498	0.2591
$\mu * 100$	0.1320	0.0018	0.1291	0.1348
$\kappa^{\mathbb{Q}}$	-0.4586	0.0106	-0.4769	-0.4424
ζ	0.1822	0.0016	0.1796	0.1848
α	21.3563	2.2668	17.8964	25.3523

Table 2.j: **Parameter estimates of CDS pricing models: Tyson Foods.**
The table reports the mean, standard deviation, 5%- and 95%-quantiles of the MCMC posterior
distribution for the CDS pricing model in equations (6)-(8). The results are reported for model
specifications without time-change (CIR model) and with time-change in default intensity (CIR +
IG Time Change Model).

	Mean	Std Dev	5% Quantile	95% Quantile
Tyson Foods: CIR Model				
$\kappa^{\mathbb{P}}$	0.7200	0.5119	0.0630	1.6877
σ	0.2469	0.0034	0.2412	0.2524
$\mu * 100$	0.2753	0.0027	0.2708	0.2798
$\kappa^{\mathbb{Q}}$	-0.2400	0.0075	-0.2524	-0.2279
ζ	0.1683	0.0011	0.1665	0.1702
Tyson Foods: CIR + IG Time Change Model				
$\kappa^{\mathbb{P}}$	1.1482	0.8454	0.0946	2.7624
σ	0.2916	0.0027	0.2870	0.2960
$\mu * 100$	0.2384	0.0026	0.2339	0.2427
$\kappa^{\mathbb{Q}}$	-0.3805	0.0091	-0.3949	-0.3651
ζ	0.1647	0.0011	0.1630	0.1665
α	12.1615	0.8741	10.8246	13.6870

Table 3: **Moments of measurement equation residuals.**
The table reports the posterior mean of mean, std. deviation, skewness and kurtosis of measurement equation residuals, $\varepsilon^{(y)}$, calculated at each iteration of MCMC algorithm for model specifications without time-change (CIR model) and with time-change in default intensity (CIR + IG Time Change Model). A correct specification implies that the residuals are standard normal with mean zero, standard deviation one, skewness zero and kurtosis three.

	Mean	Std Dev	Skewness	Kurtosis
Alcoa				
CIR Model	-0.0157	0.9998	-0.1532	4.3142
CIR-IG TC Model	-0.0051	1.0000	0.0274	4.4102
Anadarko				
CIR Model	-0.0048	1.0000	-0.4178	3.4865
CIR-IG TC Model	-0.0044	0.9999	-0.1429	3.4420
CenturyLink				
CIR Model	-0.0169	0.9998	-0.6530	4.0214
CIR-IG TC Model	-0.0041	0.9999	-0.4849	4.0898
Clear Channel				
CIR Model	-0.0112	0.9999	-0.2315	3.7204
CIR-IG TC Model	-0.0049	1.0000	-0.0709	3.6722
Ford				
CIR Model	-0.0125	0.9999	-0.4358	4.3982
CIR-IG TC Model	-0.0048	0.9999	-0.3012	4.4348
Lennar				
CIR Model	-0.0097	0.9999	-0.3456	3.9502
CIR-IG TC Model	-0.0040	0.9999	-0.1962	3.8164
Limited Brands				
CIR Model	-0.0068	0.9999	-0.6464	3.5192
CIR-IG TC Model	-0.0028	1.0000	-0.4513	3.2429
RadioShack				
CIR Model	-0.0208	0.9998	-0.5183	3.5148
CIR-IG TC Model	-0.0022	1.0000	-0.2493	3.3394
Sprint				
CIR Model	-0.0122	0.9999	-0.0525	3.2900
CIR-IG TC Model	-0.0047	1.0000	-0.0032	3.3974
Tyson Foods				
CIR Model	-0.0134	0.9999	0.0039	3.6866
CIR-IG TC Model	-0.0042	1.0000	0.1442	3.7100

Table 4: **Moments of state equation innovations.**
The table reports the posterior mean of mean, std. deviation, skewness and kurtosis of state equation residuals, $\varepsilon^{(\lambda)}$, calculated at each iteration of MCMC algorithm for model specifications without time-change (CIR model) and with time-change in default intensity (CIR + IG Time Change Model). A correct specification implies that the residuals are standard normal with mean zero, standard deviation one, skewness zero and kurtosis three.

	Mean	Std Dev	Skewness	Kurtosis
Alcoa				
CIR Model	0.0463	0.6674	0.3405	4.2667
CIR-IG TC Model	0.0108	0.8230	0.0316	3.2158
Anadarko				
CIR Model	0.0159	0.9196	0.1544	3.2541
CIR-IG TC Model	−0.0018	0.8354	0.0232	3.1892
CenturyLink				
CIR Model	0.0492	0.7081	0.2956	3.3331
CIR-IG TC Model	−0.0024	0.8671	0.0242	3.1424
Clear Channel				
CIR Model	0.0420	0.8925	0.1341	3.4407
CIR-IG TC Model	0.0015	0.9581	0.0177	3.0508
Ford				
CIR Model	0.0306	0.8265	0.1025	3.4196
CIR-IG TC Model	0.0001	0.9351	0.0156	3.0696
Lennar				
CIR Model	0.0436	0.7893	0.2021	3.3338
CIR-IG TC Model	0.0127	0.8083	0.0392	3.2169
Limited Brands				
CIR Model	0.0308	0.9215	0.1456	3.1110
CIR-IG TC Model	−0.0023	0.9208	0.0173	3.0802
RadioShack				
CIR Model	0.0359	0.7439	0.1330	3.1776
CIR-IG TC Model	−0.0040	0.9290	0.0065	3.0734
Sprint				
CIR Model	0.0728	0.7659	0.4195	4.4997
CIR-IG TC Model	0.0398	0.8156	0.0451	3.2393
Tyson Foods				
CIR Model	0.0406	0.7290	0.2760	3.4170
CIR-IG TC Model	0.0003	0.8201	0.0351	3.2036

Table 5: **Forecast distribution for 1-day horizon.**
For the CIR model, we report the quantiles of the forecast distribution of $\mathrm{CDS}(\lambda_\delta)$ for $\delta = 1/250$ (i.e., a horizon of one trading day) given initial condition λ_0 and the posterior mean parameters estimated for Alcoa. For the time-changed model we report the quantiles of the forecast distribution of $\mathrm{CDS}(\lambda(T_\delta); \alpha)$. Two values of λ_0 are considered, one below the long-run mean (5 bp) and one above the long-run mean (50 bp). Parameters for each model are fixed to the mean of the posterior distribution. A riskless rate is fixed at 3% and recovery rate at 40%.

λ_0 (bp)	Quantiles of 5-year $\mathrm{CDS}(\lambda(T_\delta))$ in basis points								
	0.001	0.01	0.1	0.25	0.5	0.75	0.9	0.99	0.999
CIR Model									
5	17.2	17.6	18.9	20.0	21.5	23.3	25.2	28.9	32.2
50	42.7	47.2	54.1	58.5	63.6	69.1	74.3	83.9	90.9
CIR-IG TC Model									
5	17.5	17.5	21.2	22.5	23.1	23.7	24.7	32.7	61.8
50	17.5	40.7	68.8	72.5	74.3	76.1	79.3	104.5	177.6

Table 6: **Validation of independence assumptions.**
For the time-changed model, we report the posterior mean and standard deviation of correlations between innovations $\varepsilon_t^{(\lambda)}$ in the default intensity process and time-change increments χ_t, and between $\varepsilon_t^{(\lambda)}$ and changes in the federal funds rate ($d\,\mathrm{FFR}_t = \mathrm{FFR}_t - \mathrm{FFR}_{t-1}$).

	Corr $\left[\varepsilon^{(\lambda)}, \chi\right]$		Corr $\left[\varepsilon^{(\lambda)}, d\,\mathrm{FFR}_t\right]$	
	Mean	Std Dev	Mean	Std Dev
Alcoa	0.0248	0.0125	0.0033	0.0204
Anadarko	0.0209	0.0122	0.0023	0.0218
CenturyLink	0.0212	0.0110	0.0048	0.0209
Clear Channel	0.0284	0.0145	0.0000	0.0216
Ford	0.0251	0.0138	-0.0083	0.0213
Lennar	0.0231	0.0128	0.0109	0.0225
Limited Brands	0.0235	0.0125	0.0010	0.0214
RadioShack	0.0197	0.0109	0.0024	0.0221
Sprint	0.0503	0.0184	-0.0014	0.0279
Tyson Foods	0.0276	0.0122	0.0030	0.0204

Table 7: **Out-of-sample density forecast performance results.**
The table reports Amisano and Giacomini (2007) test statistics defined in equation (15). Under the null hypothesis of equal performance, the statistics have an asymptotic standard normal distribution. The positive (negative) values support the time-changed (standard CIR) model.

	U
Alcoa	10.98
Anadarko	6.47
CenturyLink	17.65
Clear Channel	47.62
Ford	19.44
Lennar	-14.73
Limited Brands	63.07
RadioShack	39.79
Sprint	9.45
Tyson Foods	23.60

Figure 1: **Daily changes in log-spreads.**
Daily changes in the log-spread for the on-the-run 5 year CDX.NA.IG index. Index roll dates
(marked on the x-axis) are dropped to avoid contamination from changes in index composition.
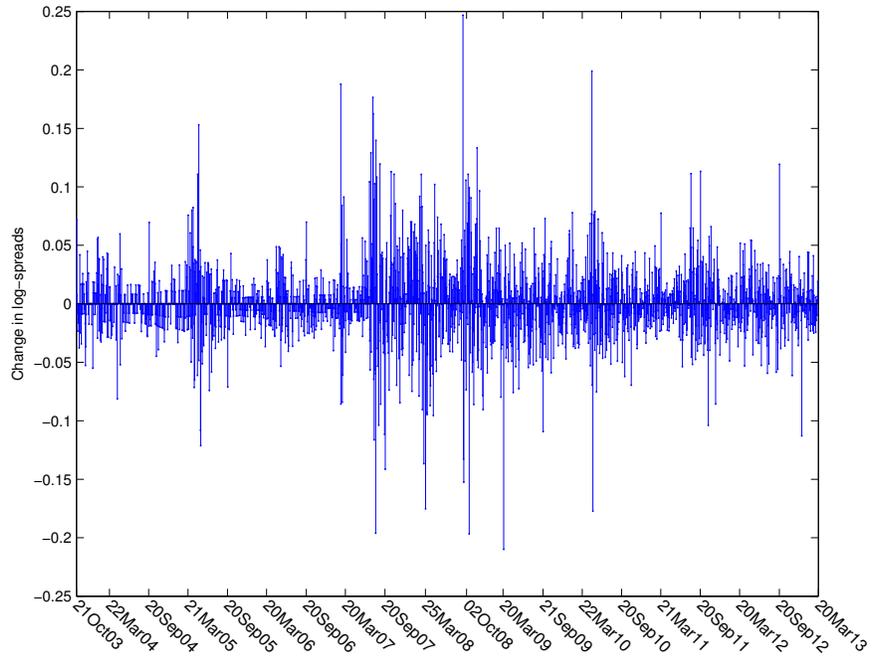
Figure 2: **Kurtosis of change in the CDS spread.**
Stationary CIR model under business time with parameters $\kappa^{\mathbb{P}} = \kappa^{\mathbb{Q}} = 0.2, \mu = 0.004, \sigma = 0.1$, and $R = 0.4$. Black dotted line marks limit as $\delta \to \infty$ of kurtosis of spread change in CIR model. Both axes on log-scale. Moments of the calendar time increments are obtained by simulation with 4 million trials.
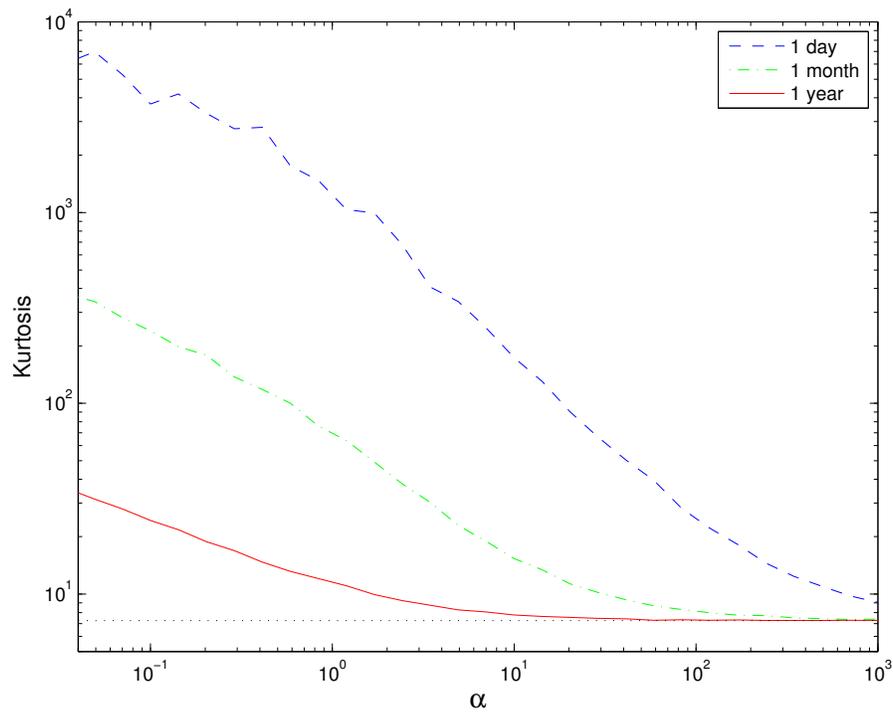
Figure 3: **In-sample CDS spreads and smoothed state estimates: Alcoa.**
We plot CDS spreads (top panel, log-scale), smoothed estimates of the default intensity (middle panel, log-scale) and smoothed estimates of time-change increments (bottom panel) for Alcoa. Smoothed estimates of the default intensity are obtained as the posterior mean of the MCMC chain for the model without time-change (CIR) and with time-change in default intensity (CIR + IG Time Change). The smoothed estimates of time-change increments for the model with time-change are scaled by Δ for ease of comparison against model without time-change.
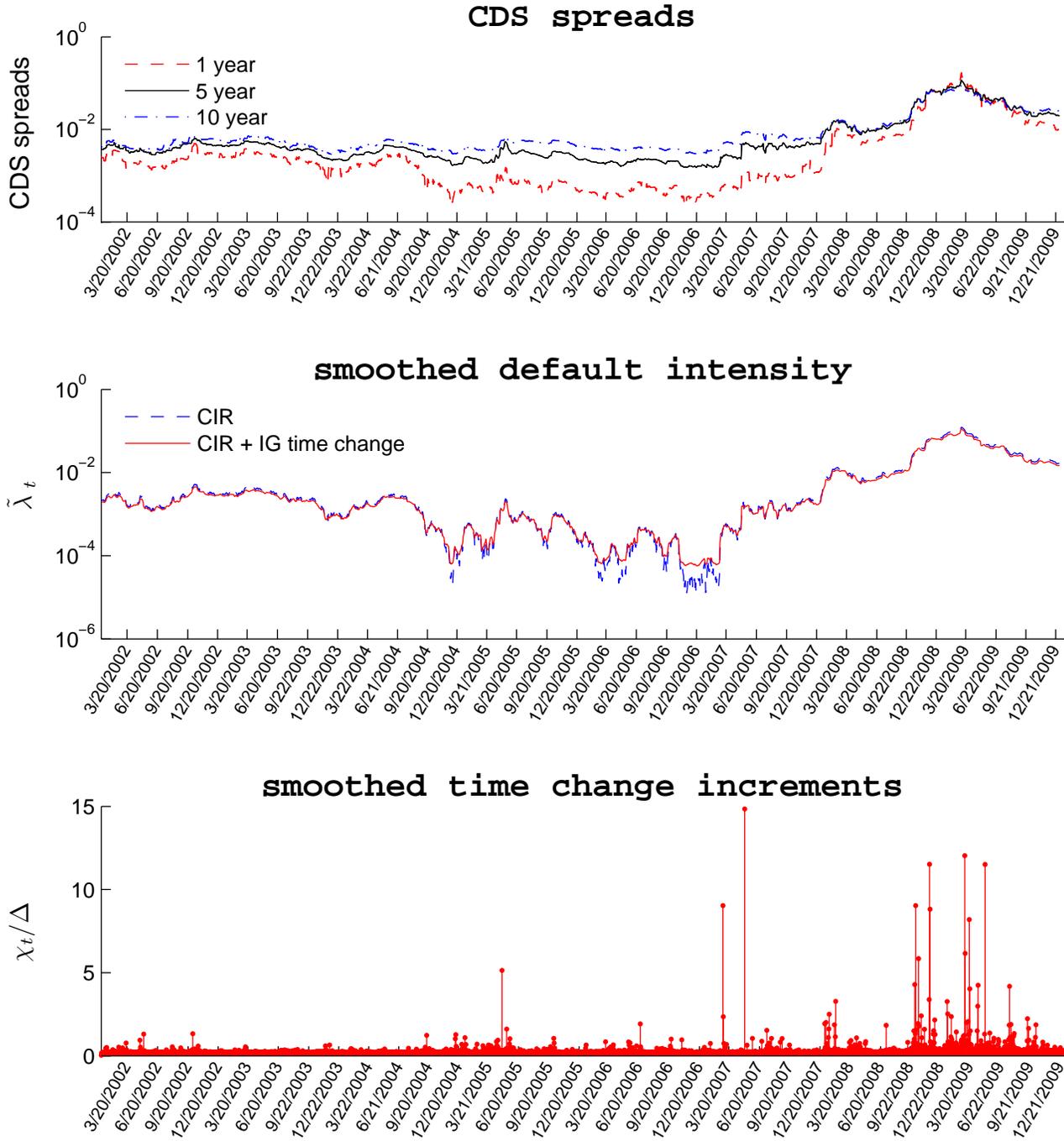
Figure 4: **In-sample CDS spreads and smoothed state estimates: Ford.**
We plot CDS spreads (top panel, log-scale), smoothed estimates of the default intensity (middle panel, log-scale) and smoothed estimates of time-change increments (bottom panel) for Ford. Smoothed estimates of the default intensity are obtained as the posterior mean of the MCMC chain for the model without time-change (CIR) and with time-change in default intensity (CIR + IG Time Change). The smoothed estimates of time-change increments for the model with time-change are scaled by Δ for ease of comparison against model without time-change.
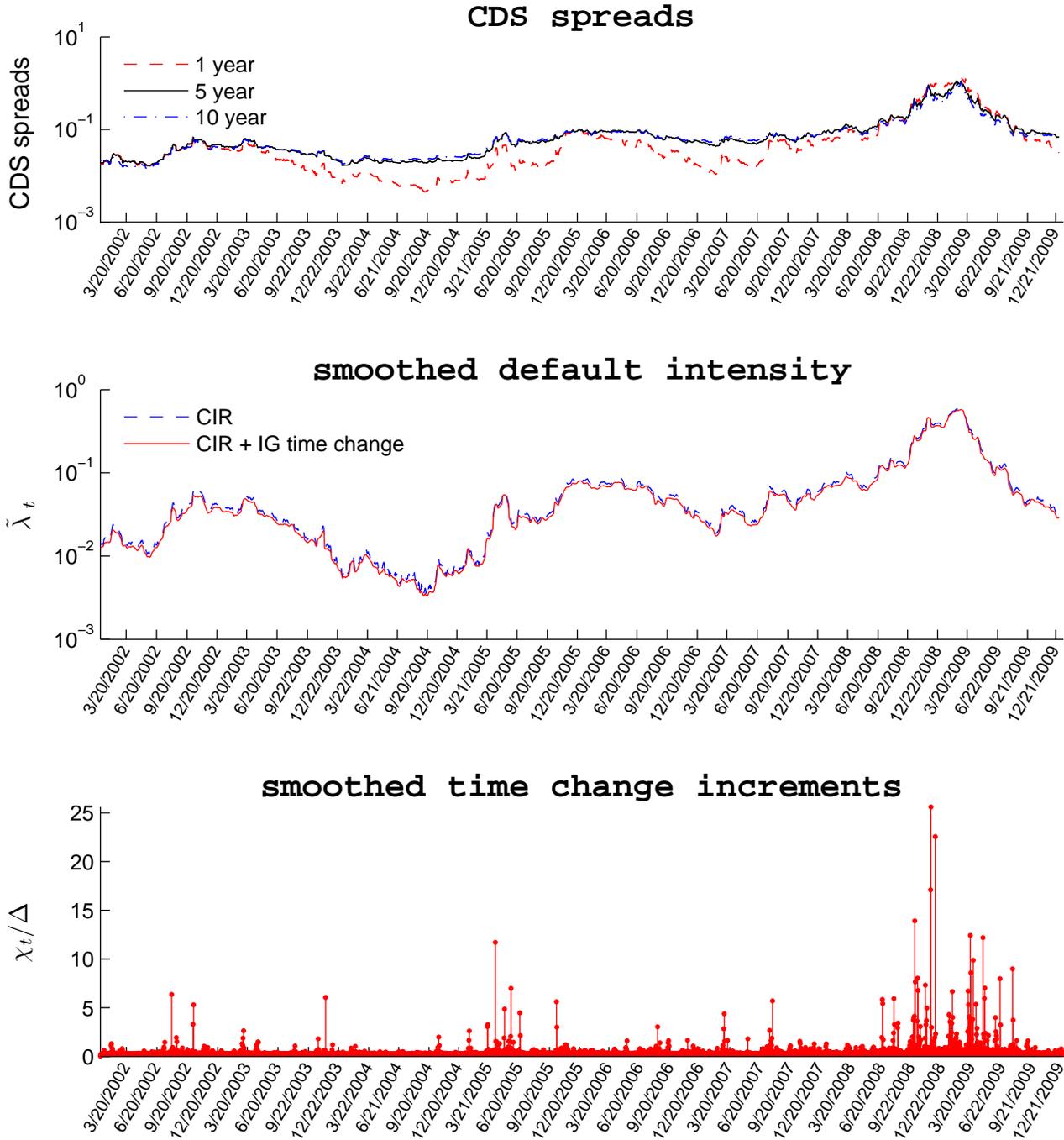
Figure 5: **Model-implied term structure of spreads for Alcoa.**
Term-structure of credit spreads for Alcoa on three dates for fitted model with time-change (solid lines) and without time-change (dashed lines). Spreads plotted in basis points on a log scale.
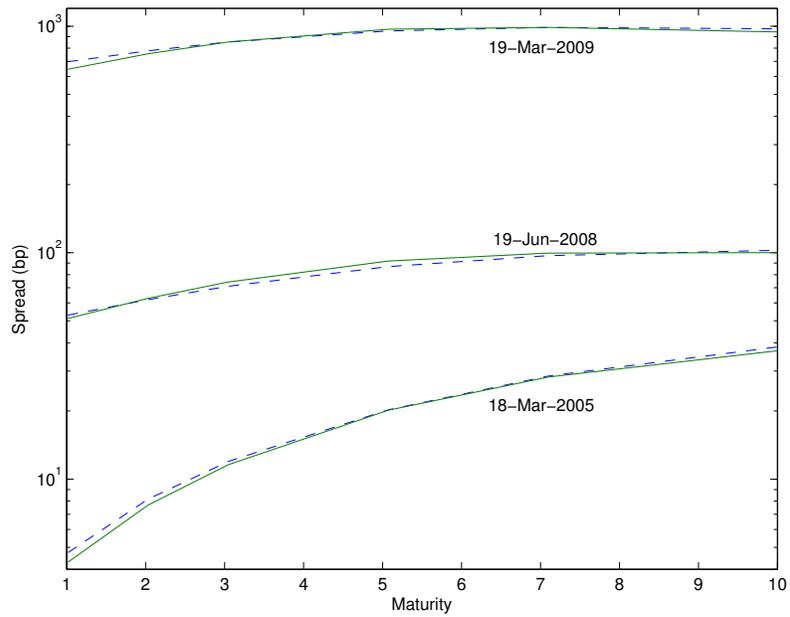
Figure 6: **Out-of-sample CDS spreads and filtered state estimates: Ford.**
We plot CDS spreads (top panel, log-scale), filtered estimates of the default intensity (middle panel, log-scale) and filtered estimates of time-change increments (bottom panel) for Ford. Filtered estimates of the default intensity are obtained as the mean of the filtered distribution produced by the particle filter for the model without time-change (CIR) and with time-change in default intensity (CIR + IG Time Change). The filtered estimates of time-change increments for the model with time-change are scaled by Δ for ease of comparison against model without time-change.
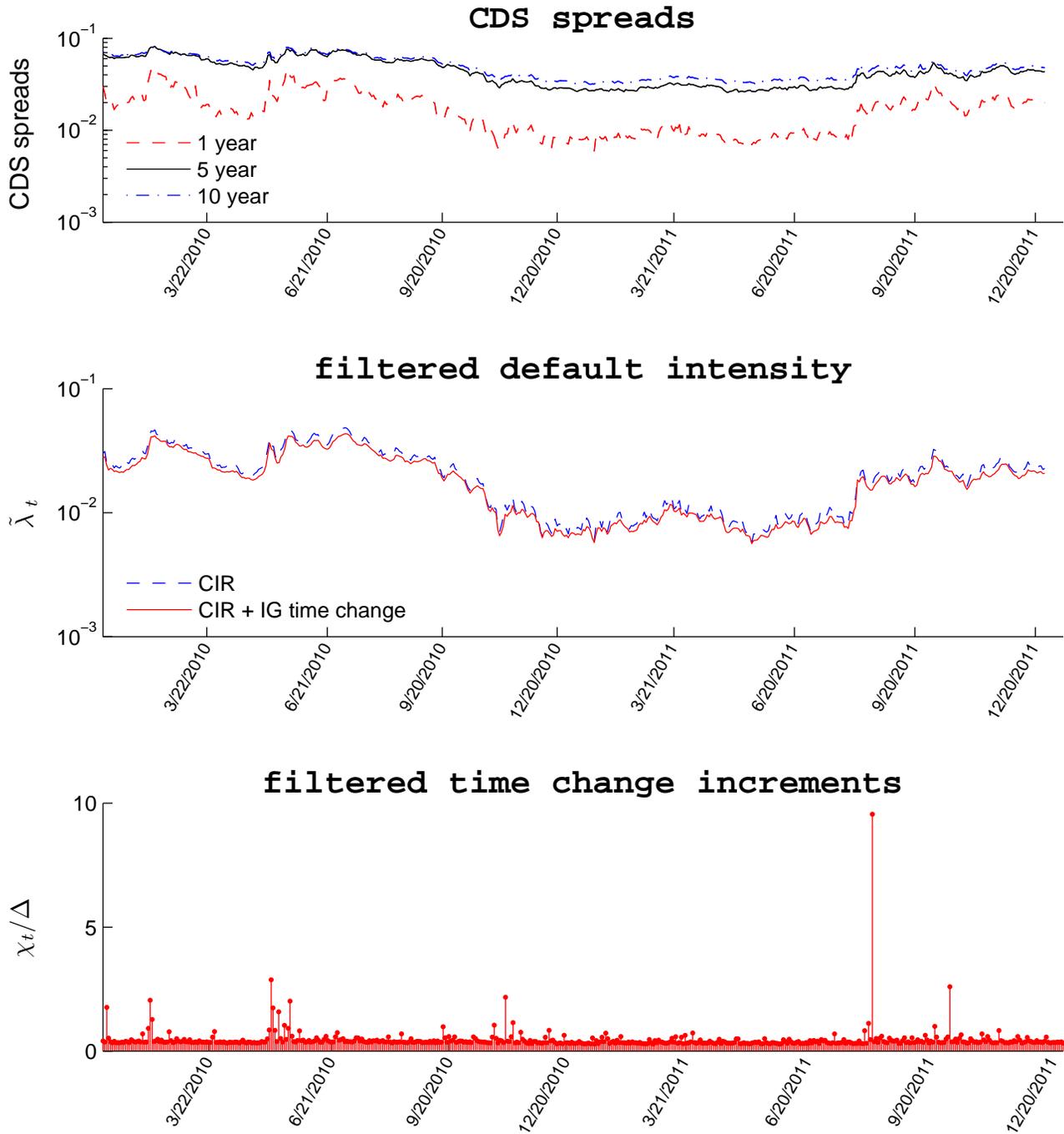
Figure 7: **Out-of-sample CDS spreads and filtered state estimates: Lennar.**
We plot CDS spreads (top panel, log-scale), filtered estimates of the default intensity (middle panel, log-scale) and filtered estimates of time-change increments (bottom panel) for Lennar. Filtered estimates of the default intensity are obtained as the mean of the filtered distribution produced by the particle filter for the model without time-change (CIR) and with time-change in default intensity (CIR + IG Time Change). The filtered estimates of time-change increments for the model with time-change are scaled by Δ for ease of comparison against model without time-change.

45