International Journal of Theoretical and Applied Finance

Credit Risk Modeling with Misreporting and Incomplete Information

Agostino Capponi*

California Institute of Technology
Division of Engineering and Applied Science
1200 E. California Blvd. M.C. 256-80
Pasadena, CA 91125
acapponi@caltech.edu

Jakša Cvitanić†

California Institute of Technology
Department of Humanities and Social Sciences
1200 E. California Blvd. M.C. 228-77
Pasadena, CA 91125
cvitanic@hss.caltech.edu

Received (Day Month Year)
Revised (Day Month Year)

We propose a structural model for the valuation of defaultable securities of a firm which models the effect of deliberate misreporting done by insiders in the firm and unobserved by others. We derive exact formulas for equity and bond prices and approximate expressions for the conditional default probability, recovery rate, and credit spread under the proposed credit risk framework. We propose a novel estimation approach to structural model estimation which accounts for noisy observed asset values. We apply the proposed method to calibrate a simple version of our model to the case of Parmalat and show that the model is able to recover a certain amount of misreporting during the years of accounting irregularities.

Keywords: Credit risk; incomplete information; calibration

1. Introduction

Structural models of credit risk represent an elegant framework for modeling valuation of risky debt. Those models make explicit assumptions about the dynamics of

*Research has been supported by a research fellowship granted by the Social and Information Sciences Laboratory at Caltech.
†Research supported in part by NSF grants DMS 04-03575 and DMS 06-31298 and through the Programme "GUEST" of the National Foundation For Science, Higher Education and Technological Development of the Republic of Croatia. We are solely responsible for any remaining errors, and the opinions, findings and conclusions or suggestions in this article do not necessarily reflect anyone's opinions but the authors'.

a firm's assets and its capital structure, which are then used to determine the occurrence of default. The literature on structural models goes back to Merton (1974), where the firm defaults if, at the time of servicing the debt, its assets are below its outstanding debt. A more general approach was introduced by Black and Cox (1976) who relax the Merton's assumption and model default as the first passage time of the firm's asset value below a certain threshold. Further generalizations treat coupon bonds, the effect of bond indenture provisions, see Geske (1977), stochastic interest rates, see Longstaff and Schwartz (1995) and Collin-Dufresne (2001), and endogenous default barriers optimally triggered by equity owners when the asset fall to a sufficiently low level, see Leland and Toft (1996).

All those models are based on the assumption that the firm's dynamic is regulated by a diffusion process and that the value of the firm can be observed directly. Since diffusion processes have continuous sample paths and default is the first hitting time of a barrier, then default is a predictable stopping time. This leads to an underestimation of the short-term credit spread, which are by far lower than those observed in the market. This major flaw of structural models has given rise to alternative approaches to credit risk modeling. One possibility is to extend structural models, for example by including jumps. Another very popular method is the so-called intensity based approach, also known as reduced form approach. The intensity based approach does not model default in terms of assets and liabilities of the firm, but defines the time of default as the first jump-time of an exogenously given counting process. The advantage is that the default event becomes an inaccessible stopping time, thereby removing the disturbing feature of strong underestimation of short-term credit spreads. Intensity-based models, however, are often criticized because they lose the micro-economic interpretation of the default time.

There are also hybrid models which combine the best features of both approaches. Duffie and Lando (2001) reconcile the structural and intensity-based framework by observing that the key difference between them is the information set assumed to be known. More specifically, they propose a model with endogenous default threshold, but in which the market only observes noisy or delayed accounting reports from which investors have to draw inference of the true asset value of the firm. This model creates a non-zero instantaneous hazard rate of default, thus implying a non-zero short term credit spread. Several works have appeared after this seminal paper, all focusing on how the available information set impacts the term structure of credit spreads. Cetin et al. (2004) and Guo et al. (2006) propose an approach in which the market is assumed to only partially observe, and possibly with a lag, relevant information concerning the state of the firm. Giesecke and Goldberg (2004) add incompleteness to structural models by assuming that the default barrier is a stochastic process, thus investors cannot deduce the distance to default from the firm's fundamentals as in Merton or Black-Cox models. Frey and Runggaldier (2007) consider a model in which the intensity is driven by unobserved state processes, and the calculation of measures of risk such as default probability leads to a nonlinear filtering problem. In all these cases, default becomes an inaccessible stopping time for the market, thus yielding a reduced form credit risk model.

Our paper belongs to this branch of the literature and focuses on the role of accounting information on pricing. We explicitly model the dynamics of the misreporting process as dependent on the actual performance of the firm and impacting both the disclosure of the market value of the assets and the future evolution of the firm's asset value. We calibrate a simple version of our model to the data for the Parmalat company around its bankruptcy. The results indicate that the amount of misreporting was not negligible, and that by ignoring it, the model would have resulted in a large overestimation of the firm's volatility.

The rest of the paper is organized as follows. Section 2 describes the components of the proposed model with incomplete information and misreporting. Section 3 gives explicit formulas for bond and equity prices under a Merton framework. Since such formulas are not directly implementable, we provide in Section 4 implementable expressions for bond and equity prices. Section 5 provides formulas for the default measures resulting from our credit risk framework. Section 6 presents results of the calibration of a simple version of the model to the Parmalat case. Section 7 concludes the paper. More technical results and proofs are provided in the appendix.

2. A structural model with Incomplete and Distorted Information

2.1. *The effect of accounting quality on pricing*

A recent branch of credit risk literature has focused on the quality of accounting information and its effect on the term structure of credit spreads. Yu (2005) proves that accounting noise is actually priced in the market by showing empirically that a risk premium is charged to the credit spreads of firms that adopt less transparency. Cherubini and Manera (2006) model the effect of deliberate misreporting on accounting statements through the introduction of a probability of fraud which the market updates whenever new information about balance sheet is issued. Brigo and Morini (2006) consider the effect of accounting reliability by modeling the ratio between the level of default barrier and the value of company assets as a random variable, where pessimistic scenarios, possibly corresponding to fraud in accounting, are associated with larger values of this ratio. None of the above studies models explicitly the dynamics of misreporting. We introduce a credit risk framework which incorporates the misreporting event as an intrinsic feature, and is estimable using market and accounting data.

Misreporting may arise if the market has incomplete knowledge of the manager's objective function, since then he may be better off with the option to misreport, see Fisher and Verrecchia (2000). We work under this assumption since the exact nature of the manager's compensation, his time horizon, and his litigation risk and reputation costs associated with biased reporting are often unavailable to the market.

Although it is not easy to estimate the managerial risk of misreporting, recent studies have started addressing this issue. For example, Wang (2007) proposes a bivariate probit model to recover the probability of committing fraud from the probability of detected fraud. As an illustration, Figure 1 reports the estimated probability of misreporting obtained using the methodology in Wang (2007) for Tyco, a well known case of misreporting in the United States history. The values of the predictors used in the model have been taken from Edgar database on a three-month basis for the period ranging from January 2001 to December 2003. Such time frame includes the misreporting period which covers the years 2001 and 2002, and this is well captured by the model which shows a much higher probability of misreporting (around 80%).

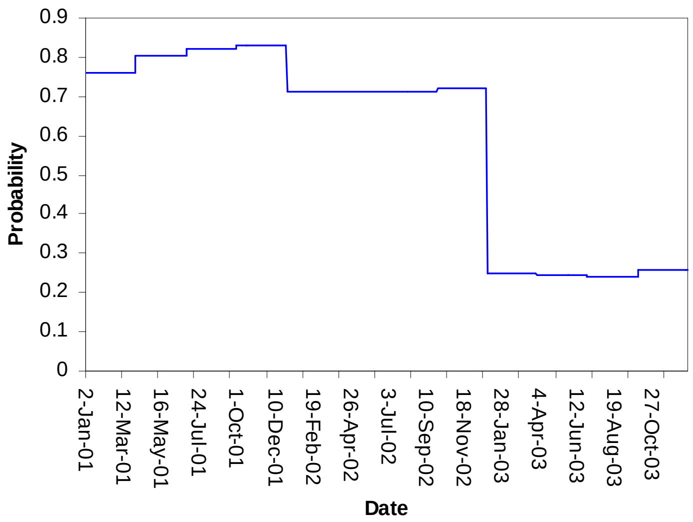

Fig. 1. The probability of fraud as a function of time obtained for Tyco using the model 17 proposed in (Wang (2007))

The discussion above supports our argument that it is worth considering misreporting risk as an additional risk factor in credit models (beside volatility risk and accounting noise). The next subsection defines in detail the proposed framework.

2.2. *Model Definition*

We consider a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ with the following system of stochastic difference equations, a generalized version of a *Hidden Markov Model*:

$$V_k = e^{x_k} \tag{2.1}$$

$$x_k = x_{k-1} + (\mu(\theta_{k-1}) - 0.5\sigma(\theta_{k-1})^2)\Delta_k + \sigma(\theta_{k-1})v_k \tag{2.2}$$

$$\theta_k = \Gamma(\theta_{k-1}, x_k, \varrho, w_k) \tag{2.3}$$

$$z_k = x_k + h(\theta_{k-1}) + \nu(\theta_{k-1})u_k \tag{2.4}$$

Here, V_k describes the evolution of the asset value of the firm and is modeled as a discretized geometric Brownian motion, with x_k being the log-asset value process, and $x_0 = \mathcal{N}(\mu_0, \sigma_0^2)$, and drift μ and volatility σ both depending on the parameter θ. Moreover, z_k describes the released observation to the outsiders, with Δ_k denoting the time between consecutive observations, typically associated with release of balance sheet reports which often occurs on a quarterly basis. We assume that $\{v_k\}$, $\{w_k\}$, $\{u_k\}$ are independent sequences of i.i.d Gaussian random variables with zero mean and unit variance.

Moreover, $\Theta = \{\theta^{(1)} = 1, \theta^{(2)} = 2, \ldots, \theta^{(m)} = m\}$ is a finite set of integer modes of cardinality m, Γ is a measurable mapping $\Theta \times \mathbb{R} \times \mathbb{R} \to \mathbb{M}$, and h is a measurable mapping $\Theta \to \mathbb{R}$ assumed to be time-invariant for notational simplicity only.

The interpretation is as follows. We can think of θ as a random variable which designates the *report model* used by the manager of the firm to release the observations to the investors. Such report model affects both the future evolution of the actual asset value, see eq.(2.1), and the value released to the outsiders by the manager of the firm, see eq.(2.4). More precisely, the released log-asset value z_k depends on the report model θ_{k-1} in place during the time interval $[t_{k-1}, t_k]$ through the function h which models the amount of misreporting associated with a given report model. Depending on whether $h(\theta_{k-1})$ is positive or negative, an overstatement or an understatement of the actual performance of the firm will occur when the report model θ_{k-1} is selected by the manager. The situation of no distortion occurring can be modeled by having $h(\theta_{k-1}) = 0$. The parameter $\nu(\theta_{k-1})$ captures the variance of accounting noise associated with the report model $\nu(\theta_{k-1})$.

The set of report models is assumed to be finite. Equation (2.3) models the choice of the current report model θ used by the firm's manager as dependent on the last report model, the state of the firm at the time immediately preceding the release of the observation and other factors affecting misreporting represented by the vector ϱ. Such factors can be the quality of corporate governance, the litigation and reputation costs associated with getting caught, or the manager bonus triggering threshold. As already mentioned in Section 2.1, outsiders do not always have a perfect knowledge of managerial objectives. We model this lack of information with a gaussian random variable w_k, and assume that the market estimates the model θ used by the manager using the function Γ which depends on the optimal manager's choice of the report model. Such choice is the solution of an optimization problem, where the manager maximizes his expected utility function of misreporting, typically depending on the manager's stock ownership and equity compensation minus his disutility of getting caught. Such optimization problem depends on the true state x_k, the previous report model θ_{k-1} and w_k. The dependence on the previous report model is introduced in our model to denote the fact that the misreporting event also depends on the past managerial behavior.

3. Merton Model with Incomplete Information

3.1. *Notation and terminology*

The following notation will be used henceforth:

- $n(x; \mu, \sigma)$: gaussian density with mean μ and standard deviation σ
- $\lambda^{i,j}(y) := P(\theta_k = j | \theta_{k-1} = i, x_k = y)$: mode switching probabilities
- $\lambda^{j}(y) := P(\theta_0 = j | x_0 = y)$: prior mode probability
- $p(x|y, l) := P(x_k = x | \theta_{k-1} = l, x_{k-1} = y)$: mode dependent transition density
- $p(x) := P(x_0 = x)$: the initial density on x_0
- $p_{k|k}(x) := P(x_k = x | \mathscr{F}^z_k)$: the posterior density
- $p^l_{k|k}(x) := P(x_k = x, \theta_{k-1} = l | \mathscr{F}^z_k)$: the mode conditioned posterior density
- $p^l_{k,k}(x) := P(x_k = x, \theta_{k-1} = l, \mathscr{F}^z_k)$: the unnormalized posterior density
- $L^l_{k|k}(x) := p(z_k | x_k = x, \theta_{k-1} = l)$: the mode conditioned measurement likelihood
- $E_f[g(W)] = \int_{\mathbb{R}} f(w) g(w) dw$: expectation of $g(W)$ with respect to the density f
- $\mathscr{F}^x_k = \sigma(x_1, x_2, \ldots, x_k)$: the insider filtration
- $\mathscr{F}^z_k = \sigma(z_1, z_2, \ldots, z_k)$: the outsider filtration
- $\hat{x}_{l|k} := E[x_l | \mathscr{F}^z_k] \qquad l > k$
- $\sigma^2_{l|k} := E[(x_l - \hat{x}_{l|k})^2 | \mathscr{F}^z_k] \qquad l > k$

3.2. *The Pricing Framework*

Assuming a Merton-type structural model, we propose a valuation framework in which bond and equity prices can be calculated as risk-neutral conditional expectations. We assume a fixed maturity T of the debt, and that the default event can only occur at maturity, which happens if the actual asset value is below the nominal value of the debt, assumed to be constant. We define the random variable

$$\varsigma = \begin{cases} T & \text{if } V_T \leq K \\ \infty & \text{if } V_T > K \end{cases} \tag{3.1}$$

and denote by $\mathscr{F}^{z,\varsigma}_t = \mathscr{F}^z_t \vee \sigma(s \wedge \varsigma, s \leq t)$ the sigma algebra generated by the observations enlarged with the information generated by the default indicator random variable ς. Before proceeding with the analysis, we state a useful result from JeanBlanc and Rutkowski (2000).

Proposition 1 (Projection Formula). *Let A be a bounded, $\mathscr{F}^z_t$-measurable random variable. Then for every $t \leq T$:*

$$E[\mathbf{1}_{\varsigma > T} A | \mathscr{F}^{z,\varsigma}_t] = \mathbf{1}_{\varsigma > t} \frac{E[\mathbf{1}_{\varsigma > T} A | \mathscr{F}^z_t]}{P(\varsigma > t | \mathscr{F}^z_t)} \tag{3.2}$$

We apply the pricing methodology proposed in Coculescu et al. (2006) to our credit risk framework. Denoting by P^* the pricing risk-neutral probability, we define the risk-neutral estimate of the variable V_T as the $\mathscr{F}^z_t$ measurable random variable

$$\hat{V}_T = \frac{1}{Z_T} E^{P^*}[\mathbf{1}_{V(T)>K} V_T | \mathscr{F}^z_T] \tag{3.3}$$

where $Z_T = P^*(V_T > K|\mathscr{F}^z_T)$, and define a defaultable contingent claim as an integrable, $\mathscr{F}^z_T$ measurable random variable of the form:

$$d_T = \mathbf{1}_{V_T>K} f(\hat{V}_T) + \mathbf{1}_{V_T \le K} g(V_T) \tag{3.4}$$

The main difference with the complete information models is that the defaultable claims are assumed to be evaluated using the estimate $\hat{V}_T$ when the firm is not in the default state and the true value is only observed at default. The price of a defaultable claim is then computed in Coculescu et al. (2006) as

$$d_t = e^{-r(T-t)} E^{P^*}[\mathbf{1}_{V_T>K} f(\hat{V}_T) + \mathbf{1}_{V(T)\le K} g(V_T) | \mathscr{F}^{z,\varsigma}_t] \tag{3.5}$$

In our case, if the defaultable claim is a bond, $f(x) = K$, $g(x) = x$, with K being the nominal value of the debt. Therefore, we get that the time t-price of the bond is:

$$\begin{aligned} B(t,T) &= e^{-r(T-t)} E^{P^*}[K\mathbf{1}_{V_T>K} + \mathbf{1}_{V_T\le K} V_T | \mathscr{F}^{z,\varsigma}_t] \\ &= e^{-r(T-t)}(K - E^{P^*}[(K - V_T)^+ | \mathscr{F}^z_t]) \end{aligned} \tag{3.6}$$

If the defaultable claim is equity, $f(x) = (x-K)$ and $g(x) = 0$. Therefore, we obtain the time t price of the equity is:

$$E(t,T) = e^{-r(T-t)} E^{P^*}[(\hat{V}_T - K)\mathbf{1}_{V_T>K} | \mathscr{F}^{z,\varsigma}_t] \tag{3.7}$$

We have:

$$\begin{aligned} \mathbf{1}_{V_T>K}(\hat{V}_T - K) &= \mathbf{1}_{V_T>K}\left(\frac{E^{P^*}[\mathbf{1}_{V_T>K} V_T | \mathscr{F}^z_T]}{P^*(V_T > K)} - K\right) \\ &= E^{P^*}[\mathbf{1}_{V_T>K} V_T | \mathscr{F}^{z,\varsigma}_T] - \mathbf{1}_{V_T>K} K \end{aligned} \tag{3.8}$$

where the first equation is obtained using definition (3.3), while the second equation follows from the projection formula (3.2). Therefore,

$$\begin{aligned} E(t,T) &= e^{-r(T-t)} E^{P^*}[E^{P^*}[\mathbf{1}_{V_T>K} V_T | \mathscr{F}^{y,\varsigma}_T] | \mathscr{F}^z_t] - E^{P^*}[\mathbf{1}_{V_T>K} K | \mathscr{F}^z_t] \\ &= e^{-r(T-t)} E^{P^*}[(V_T - K)^+ | \mathscr{F}^z_t] \end{aligned} \tag{3.9}$$

We next discuss the choice of the pricing measure P^*. Our market is incomplete since the measurement equation of our filtering model exhibits jumps of m possible different sizes, where m is the number of report models used by manager. Assuming that the released log-asset value z_k is the only traded asset, this means that the jump risk cannot be hedged away. In addition, we have a discrete-time model with continuously valued normally distributed noise. The specific measure P^* has to be inferred from market prices, either via modeling of the market price of risk or of the dynamics of the Radon-Nikodim derivative, see Runggaldier (2004) for a survey.

3.3. *Exact Filtering*

The calculation of the equity and bond prices requires the computation of a conditional expectation, which means that the time T filtering density $p_{T|t}(x)$ of our filtering model given by eq.(2.2-2.4) needs to be evaluated. This in turn requires the calculation of the time t filtering density $p_{t|t}(x)$. Such calculation is presented in Appendix A, and it results in

$$p_{t|t}(x) = \frac{\sum_{l=1}^{m} p^l_{t,t}(x)}{\int_{\mathbb{R}} \sum_{r=1}^{m} p^r_{t,t}(y)dy} \tag{3.10}$$

where $p^l_{t,t}(x)$ is recursively defined as

$$p^l_{t,t}(x) = L^l_{t|t}(x) \sum_{r=1}^{m} \int_{\mathbb{R}} \lambda^{r,l}(y) p(x|y,l) p^r_{t-1|t-1}(y)dy \tag{3.11}$$

3.4. *Equity and Bond Prices*

In order to simplify the exposition, the pricing formulas derived in this section assume a filtering model in which $\mu(i) = \mu$, $\sigma(i) = \sigma$, i.e. the expected return rate and the volatility of the asset is not affected by the report model used. We follow an argument similar to the one used in Merton jump diffusion model, i.e. replace μ with r and assume that the biases $h(i)$ and the jump risks $\nu(i)$ are the same under the historical and risk neutral measure. This is because the jumps in the released log-asset value model managerial bias, which depends on the vector ϱ of parameters including factors specific to the executive such the reputation and litigation costs associated with misreporting. Such risk is "nonsystematic" risk and cannot be diversified away, thus we assume the same form under both measures.

Let us denote $\tau = T - t$ and introduce functions $d_1(x)$ and $d_2(x)$ as

$$\begin{aligned} d_1(x) &= \frac{x - \log(K) + (r + 0.5\sigma^2)\tau}{\sigma\sqrt{\tau}} \\ d_2(x) &= d_1(x) - \sigma\sqrt{\tau} \end{aligned} \tag{3.12}$$

The price of the bond is given by the following Proposition, proven in Appendix C.

Proposition 2. *The time t price of a bond maturing at T is given by:*

$$B(t,T) = Ke^{-r\tau} - E_{p_{t|t}}[Ke^{-r\tau}N(-d_2(Y)) - e^y N(-d_1(Y))] \tag{3.13}$$

where $p_{t|t}$ is the density given in eq.(3.10), Y is a random variable with density $p_{t|t}$, and the functions d_1 and d_2 are defined in eq.(3.12).

Similarly, the equity price $E(t,T)$ is given by

Proposition 3. *The price of the equity at time t is given by:*

$$E(t,T) = E_{p_{t|t}}[e^Y N(d_1(Y)) - Ke^{-r\tau}N(d_2(Y))] \tag{3.14}$$

4. Bond and Equity Price Computation

Although the filtering density (3.11) can be explicitly obtained, it is not amenable to an efficient implementation. First of all, the recursive expression (3.11) shows that an exponentially increasing number of terms have to interact to obtain the unnormalized density at time t_k. Additionally, it involves the evaluation of non gaussian integrals due to the appearance of the terms $\lambda^{j,l}$ and such integrals may in general be computationally expensive to evaluate. Furthermore, the normalization step required to obtain the posterior density in eq.(3.10) involves an evaluation of m spatial integrations, thus increasing the computational burden even further. This makes therefore practically impossible to compute exactly bond and equity prices in our credit risk framework, since they both require to take expectation with respect to the actual filtering density. In the next subsection we describe a filtering scheme which may be used to approximate the density (3.10) and then compute equity and bond prices under the proposed filtering approximation scheme.

4.1. *The filtering approximation scheme*

We employ the approximation scheme proposed in Capponi (2008) to compute an approximate filtering density. Such methodology provides an approximation for the unnormalized density $p^l_{k,k}(x)$ using a weighted sum of gaussian densities selected from a predefined finite base set as the solution of a convex programming problem. Let us denote the base set B of gaussian densities as

$$B = \{n_{i,j}(x)\}_{i \in I, j \in J} \tag{4.1}$$

where I and J are two finite set of indices and $n_{i,j}(x)$ stands for the gaussian density with mean μ_i and standard deviation σ_j. Moreover, we require

$$\begin{aligned} \mu_{i_1} &\neq \mu_{i_2}, \qquad \forall i_1 \neq i_2 \\ \sigma_{j_1} &\neq \sigma_{j_2} \qquad \forall j_1 \neq j_2 \end{aligned} \tag{4.2}$$

meaning that the means and covariances of the gaussian densities in B are all different.

We next show how an approximation to the actual unnormalized density $p^l_{k,k}(x)$ at time t_k can be obtained from an existing set of approximations to the unnormalized density $\{p^r_{k-1,k-1}(y)\}_{r=1,\ldots,m}$ available from time t_{k-1}.

For any $r = 1, \ldots, m$, let us denote by $\hat{p}^r_{k-1,k-1}(y)$ such an approximation which would have been computed in the previous step of the approximation procedure and is given by a weighted sum of gaussian densities as

$$\hat{p}^r_{k-1,k-1}(y) \approx \sum_{(i,j) \in I \times J} w^r_{(i,j)} n_{i,j}(y) \tag{4.3}$$

Our goal is to approximate the actual filter density $p^l_{k,k}(x)$ at time t_k given in eq.(3.11). This is done in two separate steps. The output of the first step is an approximate density $\check{p}^l_{k,k}(x)$ computed as follows:

$$\begin{aligned}\check{p}^l_{k,k}(x) &:= L^l_{k|k}(x)\sum_{r=1}^{m}\int_{\mathbb{R}}\lambda^{r,l}(y)p(x|y,l)\hat{p}^r_{k-1,k-1}(y)dy \\ &\approx L^l_{k|k}(x)\sum_{r=1}^{m}\int_{\mathbb{R}}\lambda^{r,l}(y)p(x|y,l)p^r_{k-1,k-1}(y)dy \\ &= p^l_{k,k}(x)\end{aligned} \tag{4.4}$$

where the approximation step consists in replacing each term $p^r_{k-1,k-1}(y)$ with its previously computed approximation $\hat{p}^r_{k-1,k-1}(y)$. We can write out explicitly each term in the approximation, thus obtaining

$$\check{p}^l_{k,k}(x) = L^l_{k|k}(x)\sum_{r=1}^{m}\sum_{(i,j)} w^r_{(i,j)} \cdot \int_{\mathbb{R}}\lambda^{r,l}(y)p(x|y,l)n_{i,j}(y)dy \tag{4.5}$$

Eq. (4.5) shows that we have obtained an approximation density $\check{p}^l_{k,k}(x)$ for $p^l_{k,k}(x)$ consisting of a larger number of components, which is no longer a mixture of gaussians due to the appearance of the state-dependent mode probability $\lambda^{r,l}(y)$. To this purpose we approximate $\check{p}^l_{k,k}(x)$ further before propagating it to the next step $k+1$. This is the second step of the methodology, which solves the programming problem (P2) in Appendix B with density $\boldsymbol{p}$ replaced by

$$\check{\boldsymbol{p}}^l_{k,k} = (\check{p}^l_{k,k}(x_1), \ldots, \check{p}^l_{k,k}(x_q)) \tag{4.6}$$

where $x_1 < x_2 \ldots < x_q$ is a set of points chosen in such a way that $p^l_{k,k}(x)$ concentrates most of its probability mass inside the interval $[x_1, x_q]$ as detailed in Capponi (2008). If $\mathbf{w}^*$ denotes the optimal solution to (P2), the approximation density at step k, which is propagated to step $k+1$ is

$$\hat{p}^l_{k,k}(x) = \sum_{(i,j)} w^l_{(i,j)} n_{i,j}(x) \tag{4.7}$$

The associated mode conditioned normalized density is given by

$$\hat{p}^l_{k|k}(x) = \frac{\sum_{(i,j)} w^l_{(i,j)} n_{i,j}(x)}{\sum_{r=1}^{m}\sum_{(i,j)} w^r_{(i,j)}} \tag{4.8}$$

and the associated normalized density is given by

$$\hat{p}_{k|k}(x) = \frac{\sum_{l=1}^{m}\sum_{(i,j)} w^l_{(i,j)} n_{i,j}(x)}{\sum_{r=1}^{m}\sum_{(i,j)} w^r_{(i,j)}} \tag{4.9}$$

Expression (4.9) shows that the density approximation $\hat{p}_{k|k}(x)$ returned from our filter approximation scheme is a gaussian mixture. We summarize the procedure with a block diagram of the estimator in Figure 4.1. The above described approximation scheme has been shown to be extremely accurate and computationally efficient in Capponi (2008). Moreover, the introduced error can be analytically controlled

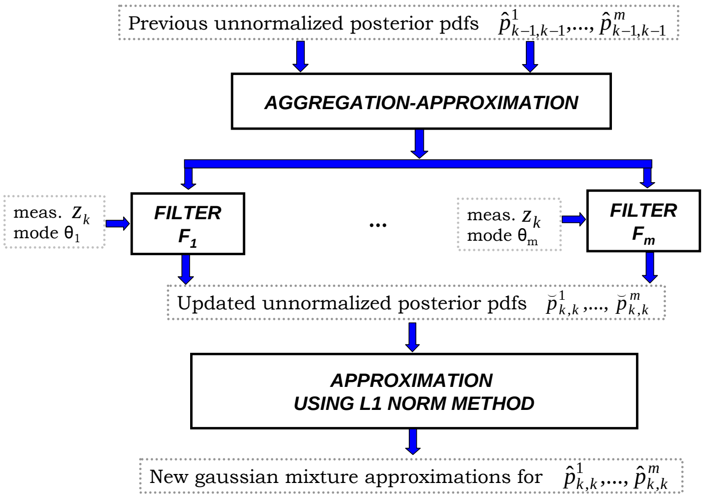

Fig. 2. One cycle of the estimator.

through an upper bound on the total variation distance between the actual unnormalized filter density $p^l_{k,k}(x)$ and the approximate density $\hat{p}^l_{k,k}(x)$ provided in Capponi (2008), where the total variation distance is defined by

$$\int_{\mathbb{R}} |p^l_{k,k}(x) - \hat{p}^l_{k,k}(x)|dx \tag{4.10}$$

4.2. *Approximate Bond and Equity Prices*

We provide a directly implementable expression for the price of equity and debt in our proposed model. The initial value of the firm's asset at time t is estimated using the filtering procedure described earlier, thus the density of the initial log-asset value is a gaussian mixture. The formulas will be given as equalities, although they are to be considered as approximations due to the filtering density of the initial log-asset value being approximated by a gaussian mixture in our non-linear filtering model. We now state an obvious result as a lemma which will be extensively used hereafter:

Lemma 1. *Let Z be a random variable having a density of the form*

$$f_Z(z) = \sum_{i=1}^{n} w_i f_{Z_i}(z), \quad \sum_{i=1}^{n} w_i = 1 \tag{4.11}$$

where $f_{Z_i}(z)$ is the pdf of a measurable random variable Z_i. Then the expectation of a function f of Z may be expressed in terms of the expectations of the function f of the component random variables Z_i, each of them taken with respect to the *probability density* f_{Z_i}

$$E[f(Z)] = \sum_{i=1}^{n} w_i E[f(Z_i)] \tag{4.12}$$

Let

$$\{\hat{x}^j_{t|t}, \hat{\sigma}^j_{t|t}, \mu^j_t\} \tag{4.13}$$

be the means, variances and weights of the components of the Gaussian mixture density for the log-asset value at the initial time t, computed using the filtering methodology described in Section 4.1 along with the received observations $\mathscr{F}^z_t$. For reading purposes, we denote the variance $(\hat{\sigma}^j_{t|t})^2$ of the j-th mixture component as $\hat{\sigma}^{2,j}_{t|t}$. It follows, from eq.(2.2) and from the assumptions made in Section 3.4 that both drift and volatility are mode independent, that the conditional density of the log-asset value x_T at maturity has a gaussian mixture density specified by

$$\hat{p}_{T|t}(x) = \sum_j \mu^j_t n(x; \hat{x}^j_{T|t}, \hat{\sigma}^j_{T|t}) \tag{4.14}$$

where $\hat{x}^j_{T|t} = \hat{x}^j_{t|t} + (r - 0.5\sigma^2)\tau$ and $\hat{\sigma}^j_{T|t} = \sqrt{\hat{\sigma}^{2,j}_{t|t} + \sigma^2\tau}$. Using Lemma 1 we have:

$$E[e^{-r\tau}(K - V_T)^+ | \mathscr{F}^z_t] = e^{-r\tau} \sum_j \mu^j_t E[(K - V^j_{T|t})^+] \tag{4.15}$$

where $V^j_{T|t}$ is the exponential of a gaussian random variable with mean $\hat{x}^j_{T|t}$ and standard deviation $\hat{\sigma}^j_{T|t}$ as defined above. Moreover, let

$$d^j_1 = \frac{\hat{x}^j_{t|t} - \log(K) + (r - 0.5\sigma^2)\tau}{\sqrt{\hat{\sigma}^{2,j}_{t|t} + \sigma^2\tau}}, \quad d^j_2 = d^j_1 - \left(\sqrt{\hat{\sigma}^{2,j}_{t|t} + \sigma^2\tau}\right) \tag{4.16}$$

Then the price of the bond is given by the following Proposition, proven in Appendix D

Proposition 4. *The price of the bond at time t is given by:*

$$B(t,T) = Ke^{-r\tau} - \sum_j \mu^j_t (Ke^{-r\tau} N(-d^j_2) - e^{\hat{x}^j_{t|t} + 0.5\hat{\sigma}^{2,j}_{t|t}} N(-d^j_1)) \tag{4.17}$$

A similar calculation can be done for the equity price $E(t,T)$ leading to

Proposition 5. *The price of the equity at time t is given by:*

$$E(t,T) = \sum_j \mu^j_t (e^{\hat{x}^j_{t|t} + 0.5\hat{\sigma}^{2,j}_{t|t}} N(d^j_1) - Ke^{-r\tau} N(d^j_2))$$

5. Computation of Default Measures

Credit risk models have in common the goal of explaining two quantities, default probability and loss given default, whose product can then be used to compute the credit spread. In the next subsections we give closed form expressions for conditional default probability, recovery rate, and credit spreads under our framework, and assuming that the proposed filtering approximation scheme is used to approximate the density of the initial asset value.

5.1. *Default Probability*

The default may occur only at time T, and it occurs if $V_T \leq K$. We define the default probability as

$$PD(t,T) = \mathcal{P}(V_T \leq K | V_t > K, \mathscr{F}_t^z) \tag{5.1}$$

A calculation presented in Appendix F shows that the default probability in our model is given by a weighted sum of conditional default probabilities, where each weight w is proportional to the weight of the gaussian density in the mixture and the distance to default of the asset value estimated using the mean of such density.

Proposition 6. *The probability of default at time T as seen at time t is given by:*

$$PD(t,T) = \sum_j w_t^j \mathcal{P}(Y_j \leq -d_2(\hat{x}_{t|t}^j) | X_j \leq dd_t^j) \tag{5.2}$$

where $w_t^j = \frac{\mu_t^j N(dd_t^j)}{\sum_{i=1}^m \mu_t^i N(dd_t^i)}$, $dd_t^j = \frac{\hat{x}_{t|t}^j - \log(K)}{\hat{\sigma}_{t|t}^j}$, d_2 *is defined in eq.(3.12) and* (X_j, Y_j) *has a bivariate normal distribution function with positive correlation* $\rho_{X_j Y_j}$*:*

$$\mu_{X_j} = \mu_{Y_j} = 0, \sigma_{X_j}^2 = 1, \ \sigma_{Y_j}^2 = 1 + \frac{\hat{\sigma}_{t|t}^{2,j}}{\sigma^2 \tau}, \ \rho_{X_j Y_j} = \frac{\hat{\sigma}_{t|t}^j}{\sqrt{\sigma^2 \tau + \hat{\sigma}_{t|t}^{2,j}}} \tag{5.3}$$

We can think of X_j and Y_j as two positively correlated random variables, where X_j is measurable using the information available by time t and reveals information about the distance to default of the current asset value estimated using the mean of the j-th gaussian density in the approximation, while Y_j is measurable using the information available at maturity T and denotes a default event if it is smaller than the default threshold.

The default probability computed using our model reduces to the Merton default probability as the uncertainty around the true value x_t gets to zero, and the conditional pdf $p(x_t | \mathscr{F}_t^z)$ approaches a sum of delta functions $\sum_j \delta(x_t - \hat{x}_{t|t}^j)$. Then we obtain

$$\begin{aligned} PD(t,T) &= \sum_j w_t^j \lim_{\sigma^j_{t|t}\to 0} \mathcal{P}(Y_j \le -d_2(\hat{x}^j_{t|t})|X_j \le dd^j_t) \\ &= \sum_j w_t^j \lim_{\hat{\sigma}^j_{t|t}\to 0} \mathcal{P}(Y_j \le -d_2(\hat{x}^j_{t|t})) \\ &= \sum_j w_t^j \mathcal{P}(Y \le -d_2(x_t)) \\ &= N(-d_2(x_t)) \end{aligned} \quad (5.4)$$

where the second line follows because Y_j becomes uncorrelated from X_j as $\hat{\sigma}^j_{t|t} \to 0$ since $\rho_{X_jY_j} \to 0$. The last line of eq.(5.4) corresponds to the probability of default in the standard Merton model.

5.2. *Expected Recovery Rate*

When default occurs, the recovery rate $RR(t,T)$ is given by the ratio of the asset value to the nominal value of the debt, V_T/K. However, this is true only if $V_T < K$, otherwise no default happens and no recovery can be observed. More formally, the expected recovery rate, $RR(t,T)$, in Merton model is defined as:

$$RR(t,T) = E\left[\frac{V_T}{K}\middle| V_T < K\right] \quad (5.5)$$

Altman et al. (2000) give an explicit expression of the recovery rate in terms of the ratio of two standard Gaussian probability functions, i.e.

$$RR(t,T) = \frac{V_t}{K} e^{r\tau} \frac{N(-d_1(\log(V_t))}{N(-d_2(\log(V_t)))} \quad (5.6)$$

In our model, we have to condition on the set of received observations, thus leading to the following definition of recovery rate at default:

$$RR(t,T) = E\left[\frac{V_T}{K}\middle| V_t > K, V_T < K, \mathscr{F}^z_t\right] \quad (5.7)$$

The detailed calculation is presented in Appendix G. Here, only the final result is stated:

Proposition 7. *The recovery rate at time T as seen at time t is given by:*

$$RR(t,T) = e^{r\tau} \sum_j w_t^j \frac{e^{\hat{x}^j_{t|t} + \frac{\hat{\sigma}^{2,j}_{t|t}}{2}}}{K} \frac{\mathcal{P}(\xi_j \le dd^j_t, \psi_j \le -d_1(\hat{x}^j_{t|t})}{\mathcal{P}(X_j \le dd^j_t, Y_j \le -d_2(\hat{x}^j_{t|t}))} \quad (5.8)$$

where

$$w_t^j = \mu_t^j \frac{\mathcal{P}(X_j \le dd^j_t, Y_j \le -d_2(\hat{x}^j_{t|t}))}{\sum_{i=1}^m \mu_t^i \mathcal{P}(X_i \le dd^i_t, Y_i \le -d_2(\hat{x}^i_{t|t}))} \quad (5.9)$$

and (ξ_j, ψ_j) *is a bivariate gaussian with negative correlation coefficient* $\rho_{\xi_j\psi_j}$*:*

$$\mu_{\xi_j} = -\hat{\sigma}^j_{t|t},\ \mu_{\psi_j} = \frac{\hat{\sigma}^{2,j}_{t|t}}{\sigma\sqrt{\tau}}, \sigma^2_{\xi_j} = 1,\ \sigma^2_{\psi_j} = 1 + \frac{\hat{\sigma}^{2,j}_{t|t}}{\sigma^2\tau},\ \rho_{\xi_j\psi_j} = -\frac{\hat{\sigma}^j_{t|t}}{\sqrt{\sigma^2\tau + \hat{\sigma}^{2,j}_{t|t}}} \tag{5.10}$$

while (X_j, Y_j) *is the bivariate defined in eq.(5.3)*

As already pointed out in Subsection 5.1, when the model consists only of eq.(2.2), $\hat{\sigma}_{t|t} = 0$ and the conditional pdf $p(x_t|\mathscr{F}^z_t)$ becomes the delta function $\delta(x_t - \hat{x}_{t|t})$. Moreover, both (X, Y) and (ξ, ψ) become uncorrelated pairs since their correlation coefficient is zero. Therefore, the cumulative distribution function of X cancels the cumulative distribution function of ξ, and both Y and ψ converge in distribution to a standard Gaussian. Hence, we recover the expected default rate in the standard Merton model given by eq.(5.6).

5.3. *The term structure of credit spreads*

The bond price $B(t, T)$ can be expressed as

$$B(t, T) = e^{-r\tau} E[K\mathbf{1}_{\{V_T > K\}} + V_T \mathbf{1}_{\{V_T \le K\}} | \mathscr{F}^z_t] \tag{5.11}$$

From the definition of credit spread

$$CS(t, T) = -\frac{1}{\tau} \log\left(\frac{B(t, T)}{Ke^{-r\tau}}\right) \tag{5.12}$$

it follows immediately using equation (5.11) that

$$CS(t, T) = -\frac{1}{\tau} \log(1 - PD(t, T)LGD(t, T)) \tag{5.13}$$

where $LGD(t, T) := 1 - RR(t, T)$ is the *loss given default*. Using the previously derived results, we obtain:

Proposition 8. *The credit spread at time* T *as seen at time* t *is given by:*

$$CS(t, T) = -\frac{1}{\tau} \log\left[1 - \frac{\sum_j \mu^j_t \mathcal{P}(X_j \le dd^j_t, Y_j \le -d_2(\hat{x}^j_{t|t}))}{\sum_{i=1}^m \mu^i_t N(dd^i_t)} + \frac{\sum_j \mu^j_t \frac{1}{K} e^{r\tau + \hat{x}^j_{t|t} + \frac{\hat{\sigma}^{2,j}_{t|t}}{2}} \mathcal{P}(\xi_j \le dd^j_t, \psi_j \le -d_1(\hat{x}^j_{t|t}))}{\sum_{i=1}^m \mu^i_t N(dd^i_t)}\right] \tag{5.14}$$

where (X_j, Y_j) *is the bivariate gaussian defined in eq.(5.3),* (ξ_j, ψ_j) *is the bivariate gaussian defined in eq.(5.10).*

Differently from Merton model, the credit spread for short maturities does not approach zero in our credit risk framework, but is given by

$$\lim_{\tau \to 0} CS(t, \tau) = \frac{\sigma^2}{2} \frac{\sum_j \mu^j_t n^j_t}{\sum_i \mu^i_t N(dd^i_t)} \tag{5.15}$$

where

$$n_t^j = \frac{1}{\sqrt{2\pi\hat{\sigma}^{2,j}_{t|t}}} e^{-\frac{(\log(K)-\hat{x}^j_{t|t})^2}{2\hat{\sigma}^{2,j}_{t|t}}} \tag{5.16}$$

Eq.(5.15) can be easily derived by taking the limits of eq.(5.14) as $\tau \to 0$. Eq.(5.15) may be rewritten as:

$$\lim_{\tau\to 0} CS(t,\tau) = \frac{\sigma^2}{2}\sum_j w_t^j \frac{n_t^j}{N(dd_t^j)} \tag{5.17}$$

where

$$w_t^j = \frac{\mu_t^j N(dd_t^j)}{\sum_i \mu_t^i N(dd_t^i)} \tag{5.18}$$

We notice that eq.(5.17) has the functional form of a weighted sum of instantaneous hazard rates, since each term $\frac{n_t^j}{N(dd_t^j)}$ is the ratio of the survival density conditioned on the j-th gaussian density being the actual density of the log-asset value and the corresponding survival probability. In this context, each term $\frac{n_t^j}{N(dd_t^j)}$ can be thought as a quantitative measure of the likelihood that the asset value falls below K in the short term if the estimate $\hat{x}^j_{t|t}$ is believed to be the correct log-asset value estimate.

6. Calibration Results

6.1. *The estimation procedure*

We consider a simple case of our model in eq.(2.1-2.4) where at any time, the manager can either report the log-asset value correctly or bias it by a fixed amount h. We further assume constant drift μ and volatility σ in eq.(2.2) and constant variance ν in the measurement equation (2.4). Since we are assuming the same distortion throughout the period, we do not have the model selection equation (2.3) under this framework. The filter density is then gaussian with mean $\hat{x}_{k|k}$ and variance $\sigma^2_{k|k}$ given by the Kalman filter as

$$(\hat{x}_{k|k}, \sigma^2_{k|k}) = \mathbf{KF}(\hat{x}_{k|k-1}, \hat{\sigma}^2_{k|k-1}, z_k, h, \nu) \tag{6.1}$$

where

$$\hat{x}_{k|k-1} = \hat{x}_{k-1|k-1} + (\mu - 0.5\sigma^2)\Delta_k \tag{6.2}$$

$$\hat{\sigma}^2_{k|k-1} = \hat{\sigma}^2_{k-1|k-1} + \sigma^2\Delta_k \tag{6.3}$$

$$G = \frac{\hat{\sigma}^2_{k|k-1} h}{\hat{\sigma}^2_{k|k-1} h + \nu^2} \tag{6.4}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + G\cdot(z_k - \hat{x}_{k|k-1}) \tag{6.5}$$

$$\hat{\sigma}^2_{k|k} = (1-G)\hat{\sigma}^2_{k|k-1} \tag{6.6}$$

Moreover, it is easily seen that under these assumptions the price of equity $E(t,T)$ in eq.(3.14) reduces to

$$E(t,T) = e^{\hat{x}_{t|t}+0.5\hat{\sigma}^2_{t|t}} N(d_1(\hat{x}_{t|t})) - Ke^{-r\tau} N(d_2(\hat{x}_{t|t})) \tag{6.7}$$

We next describe a maximum likelihood estimator procedure which recovers the system parameters from a time series of market prices of equity. Differently from Duan (1994, 2000), who implies the exact asset values from equity prices, we imply the the reported values from equity prices. Let us first assume that a sample of released observations $z_1, \ldots, z_n$ was observed on the market at discrete dates $\{t_1, t_2, \ldots, t_n\}$ and denote by $\Delta_i = t_i - t_{i-1}$ the time between consecutive observations. Eq.(2.4) implies that the log-likelihood function is given by:

$$LL(z_1, z_2, \ldots, z_n) = -\frac{n}{2}\log(2\pi) - n\log(\sigma) - \sum_{i=1}^{n} \frac{(z_i - \hat{z}_{i|i-1})^2}{2\hat{\nu}^2_{i|i-1}} \tag{6.8}$$

where $\hat{z}_{i|i-1}$ and $\hat{\nu}^2_{i|i-1}$ are given by

$$\hat{z}_{i|i-1} = \hat{x}_{i|i-1} + h \qquad \hat{\nu}^2_{i|i-1} = \hat{\sigma}^2_{i|i-1} + \nu^2 \tag{6.9}$$

However, when we do not directly observe z_i, we need to imply it from equity data and this is done using the approach described next. Eq.(6.7) shows that the equity price at time t_i for a maturity T_i can be expressed as

$$\begin{aligned} E(t_i, T_i) &= f(\hat{x}_{i|i}, \hat{\sigma}^2_{i|i}, \mu, \sigma, h, \nu, r) \\ &= g(z_i, \hat{x}_{i|i-1}, \hat{\sigma}^2_{i|i-1}, \mu, \sigma, h, \nu, r) \end{aligned} \tag{6.10}$$

where f and g are two deterministic functions. The latter equation follows from the updating step of the Kalman filter given in equation (6.5) and (6.6).

Expressing the equity at time t_i in terms of z_i is particularly convenient because it would allow to imply z_i through inversion of g, in case g is invertible. This turns out to be the case since the first order derivative of g with respect to z_i

$$\frac{\partial g(y_i, \hat{x}_{i|i-1}, \hat{\sigma}^2_{i|i-1}, \mu, \sigma, h, \nu)}{\partial z_i} = e^{\hat{x}_{i|i}+0.5\hat{\sigma}^2_{i|i}} N(d_3^i) G \tag{6.11}$$

with d_3^i defined by

$$d_3^i = \frac{\log(\hat{x}_{i|i}/K_i) + (r + 0.5\sigma^2)\tau_i - \hat{\sigma}^2_{i|i}}{\sqrt{\hat{\sigma}^2_{i|i} + \sigma^2\tau_i}} \tag{6.12}$$

is positive. Here τ_i is the difference $T_i - t_i$ and K_i represents the amount of outstanding debt at time t_i. Therefore, we can recursively imply the observations z_i from the market prices of equities $e_1, \ldots, e_i$, where e_j denotes the equity price at time t_j, using the three-step procedure described next:

(1) Predict the mean and variance of the true log-asset value using the time propagation formulas of the Kalman filter given in eq.(6.2) and (6.3).

(2) Imply z_i from g, i.e. numerically compute

$$z_i = g^{-1}(e_i; \hat{x}_{i|i-1}, \hat{\sigma}^2_{i|i-1}, \mu, \sigma, h, \nu, r) \tag{6.13}$$

where the notation $g^{-1}(e_i; \hat{x}_{i|i-1}, \hat{\sigma}^2_{i|i-1}, \mu, \sigma, h, \nu, r)$ indicates that in the inversion we are considering g as a function only of the variable z_i and keeping the remaining parameters fixed.

(3) Update the mean and variance of the true log-asset value using the correction formulas of the Kalman filter given in eq.(6.5) and (6.6).

Let $p_i = (\hat{x}_{i|i-1}, \hat{\sigma}^2_{i|i-1}, \mu, \sigma, h, \nu)$. Then we can rewrite the log-likelihood function as

$$LL(e_1, e_2, \ldots, e_n) = -\frac{n}{2}\log(2\pi) - n\log(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\frac{(g^{-1}(e_i; p_i) - \hat{y}_{i|i-1})^2}{2\hat{s}^2_{i|i-1}} - \sum_{i=1}^{n}\log(G \cdot N(d_3^i)) - \sum_{i=1}^{n}(\hat{x}_{i|i} + 0.5\hat{\sigma}^2_{i|i}) \tag{6.14}$$

where $\hat{z}_{i|i-1}$ is computed using eq.(6.9) in the step (1) of the above procedure. The value $z_i = g^{-1}(e_i; p_i)$ is implied from step (2), while $\hat{x}_{i|i}$ and $\hat{\sigma}^2_{i|i}$ are computed in step (3) along with d_3^i.

6.2. *Application to the Parmalat case*

We apply the above methodology to estimate h, μ, σ, ν^2 for the case of Parmalat, an Italian food firm which experienced a crisis during the years 2002-2003 and resulted in the largest bankruptcy in European history. We chose Parmalat since this has been already investigated in the credit risk literature, see Brigo and Morini (2006) and Cherubini and Manera (2006), and therefore it allows us to make a comparison of our results with theirs.

During those two years Parmalat was repeatedly announcing issuance of bonds despite its balance sheet statements reporting huge amount of available cash liquidity that was not used. On December 8, 2003, it was suddenly discovered that its claimed liquidity of four billion euros did not exist, and that eight million Euros in bonds of investors' money had evaporated as well. An illustration of Parmalat financial distress is illustrated in Figure 3 using CDS, equity and debt data.

In Figure 3 the default risk is estimated as the difference between the quasi-debt, i.e. the value of debt discounted with the risk free rate, and the fair value of debt, computed discounting debt with the risk free rate plus the 5-year CDS spread. The bottom plot of Figure 3 shows that the crisis was announced by a decline in the stock price combined with a simultaneous increase of the default risk. Moreover, the top plot shows a steep increase of the credit spread curves at around the time of default.

We collected the daily stock data from January 1, 2002 to December 1, 2003 and computed the daily value of equity multiplying the number of outstanding shares

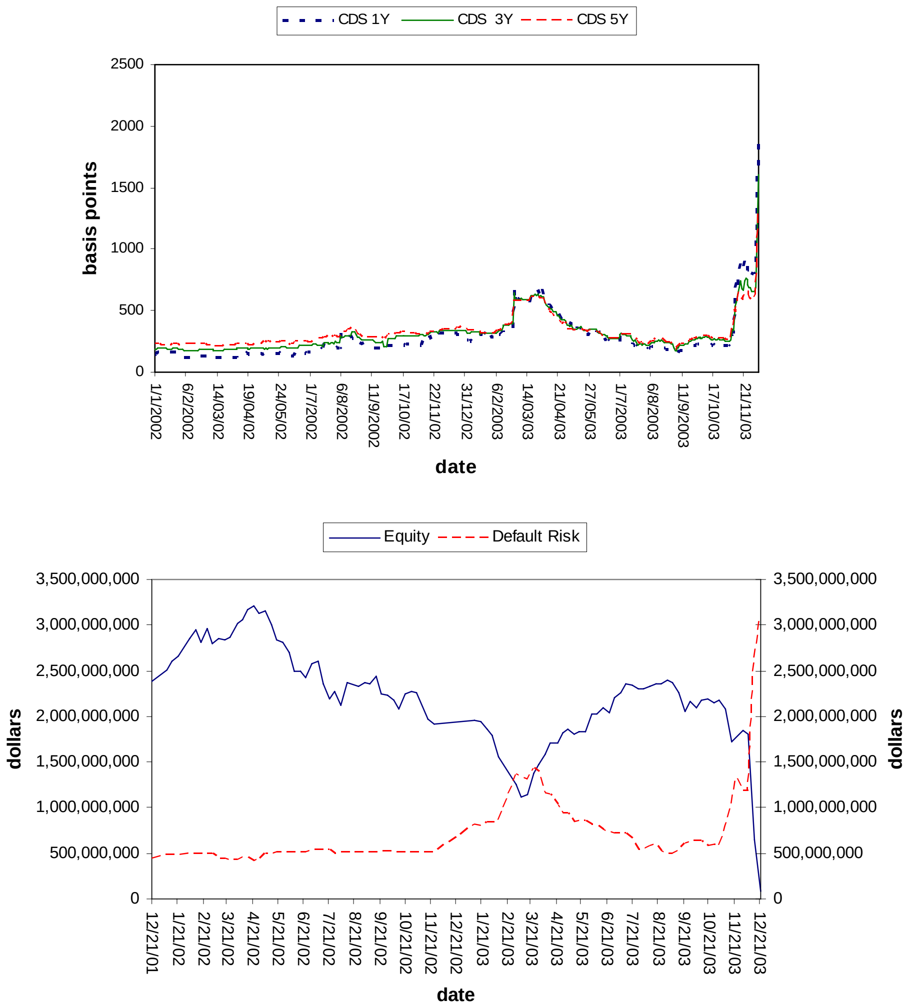

Fig. 3. Parmalat crisis

by the stock price. At each time t_i, we assume that the outstanding debt K_i has 5-year maturity and proxy it with the long term debt recovered from the balance sheet statements. We use the 5-year treasury yield as the discount factor. We run the calibration procedure using our method and compare the parameter estimates with the ones obtained by running the maximum likelihood estimation procedure by Duan on the standard Merton model. The estimates reported in Table 1 show that the hidden parameter h plays an important role in the specification of the model. If it were omitted, as it happens for the Merton model, then the reduced value of the firm due to accounting misreporting would simply be explained by an increase of the asset volatility which rises to 17 %. Our estimate of asset volatility of 11 % matches closely with the estimates of asset volatility found by different calibration procedures, see Cherubini and Manera (2006) and Brigo and Morini (2006). Our estimate 0.2052 of the distortion factor recalls the findings of Cherubini and Manera (2006), who perform a historical analysis of risk neutral probabilities of fraud via a modification of Merton model and extract an implied probability of fraud of about 0.2. Although there are differences in the assumptions and analysis between our work and theirs, making it hard to compare, using the simplest version of our model we find qualitatively similar results to the ones they obtain with the full power of their model.

7. Conclusions

In this paper we have proposed a framework for modeling deliberate distortion of the reported firm's asset values. The distortion is applied by insiders of the firm such as managers and is not directly observable by the outside market which can only infer it from available data. The amount of distortion depends on the current performance of the firm and it also has implications on the future management of the firm. The estimate of the log-asset value of the firm under the market information becomes a non-linear filtering problem and the conditional log-asset value density cannot be directly computed due to number of terms increasing exponentially in time, with each of them being computationally intensive to evaluate. In order to deal with the exploding computational complexity of the problem, we have employed a practical and implementable filtering approximation scheme which approximates the actual log-asset value density with a gaussian mixture. We have derived explicit and computable expressions for bond and equity prices, conditional default probability, recovery rate and credit spread under the proposed framework using the above mentioned approximation scheme.

We have presented a novel estimation procedure for a simplified version of our model and applied it to the Parmalat case. Such estimation procedure implies the market observations of the reported log-asset value, instead of the actual log-asset value which is only known to insiders, from equity prices. The obtained results

	MLE Estimates	
Parameters	Proposed Model	Merton Model
μ	-1.3% (0.0055)	-0.8% (0.054)
σ	11.2% (0.0022)%	17% (0.007)
h	0.2052 (0.0024)	
ν	0.002 (0.0002)%	

Table 1. Parameter estimates

are in line with the existing findings in the literature, and confirm the importance of modeling accounting distortion in credit risk models and evidence that their omission may result in exaggeratedly large estimates of asset volatility.

Appendix A. Filter Equations

Since $\{\theta_k\}$ take values in a finite set and Γ is a mapping into Θ, equation (2.3) induces state-dependent mode transition probabilities as follows:

$$\begin{aligned}\lambda^{i,j}(x) &:= P(\theta_k = j|\theta_{k-1} = i, x_k = x)\\ &= \int_{\mathbb{R}} P(w_k, \theta_k = j|\theta_{k-1} = i, x_k = x)dw_k\\ &= \int_{\mathbb{R}} P(\theta_k = j|\theta_{k-1} = i, w_k, x_k = x)f(w_k)dw_k\\ &= \int_{\mathbb{R}} \mathbf{1}_{\{j=\Gamma(i,x,w_k)\}} f(w_k)dw_k \end{aligned}\tag{A.1}$$

We next develop a recursive expression for the unnormalized posterior density, instead of working directly with its normalized counterpart. Such density is obtained through an interaction of m Bayesian filters, with each filter being an unnormalized posterior density

$$p^l_{k,k}(x) := P(x_k = x, \theta_{k-1} = l, \mathscr{F}^z_k), \qquad l = 1, \ldots, m \tag{A.2}$$

The unnormalized prediction density is defined as

$$p^l_{k,k-1}(x) := P(x_k = x, \theta_{k-1} = l, \mathscr{F}^z_{k-1}) \tag{A.3}$$

and can be developed as

$$\begin{aligned}&\int_{\mathbb{R}} \sum_{r=1}^{m} P(x_k = x, x_{k-1} = y, \theta_{k-1} = l, \theta_{k-2} = r, \mathscr{F}^z_{k-1})dy\\ &= \sum_{r=1}^{m} \int_{\mathbb{R}} \lambda^{r,l}(y)p(x|y,l)p^r_{k-1,k-1}(y)dy\end{aligned} \tag{A.4}$$

where it is easily seen from eq.(2.2) that $p(x|y,l)$ is a gaussian density with mean $y + (\mu(l) - 0.5\sigma^2(l))\Delta_k$ and variance $\sigma(l)^2\Delta_k$ The above decomposition steps follow from straightforward application of Bayes rule. The unnormalized posterior density may then be obtained as

$$\begin{aligned}p^l_{k,k}(x) &= L^l_{k|k}(x)p^l_{k,k-1}(x), \ k \geq 2,\\ p^l_{1,1}(x) &= L^l_{1|1}(x) \int_{\mathbb{R}} p(x|u,l)\pi(u,l)p(u)du\end{aligned} \tag{A.5}$$

where the correction term $L^l_{k|k}(x)$ is a gaussian density with mean $x_k + h(\theta_{k-1})$ and variance $\nu^2(\theta_{k-1})$, which is easily seen from eq.(2.4). The mode conditioned posterior density is then obtained as

$$p^l_{k|k}(x) = \frac{p^l_{k,k}(x)}{\int_{\mathbb{R}} \sum_{r=1}^{m} p^r_{k,k}(y)dy} \tag{A.6}$$

The posterior density may then be obtained from the interacting bayesian filter through normalization as

$$p_{k|k}(x) = \frac{\sum_{l=1}^{m} p_{k,k}^{l}(x)}{\int_{\mathbb{R}} \sum_{r=1}^{m} p_{k,k}^{r}(y)dy} \tag{A.7}$$

Appendix B. The Density Approximation Method

Recall from Section 4.1 that the base set B of gaussian densities is defined as

$$B = \{n_{i,j}(x)\}_{i \in I, j \in J} \tag{B.1}$$

where I and J are two finite set of indices and $n_{i,j}(x)$ stands for the gaussian density with mean μ_i and standard deviation σ_j

Let $p(x)$ be the density which we wish to approximate. We choose a *training* set

$$\boldsymbol{X} = (x_1, x_2, \dots, x_q) \tag{B.2}$$

where $x_i \in \mathbb{R}$. Let us define the matrix

$$\phi(\boldsymbol{X}) = \begin{pmatrix} \mathbf{n}_1(x_1) & \mathbf{n}_2(x_1) & \dots & \mathbf{n}_{|I|}(x_1) \\ \mathbf{n}_1(x_2) & \mathbf{n}_2(x_2) & \dots & \mathbf{n}_{|I|}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{n}_1(x_q) & \mathbf{n}_2(x_q) & \dots & \mathbf{n}_{|I|}(x_q) \end{pmatrix} \tag{B.3}$$

where $\mathbf{n}_i(x_l) = (n_{i,1}(x_l), n_{i,2}(x_l), \dots, n_{i,|J|}(x_l))$, i.e. a row vector whose j-th entry is the gaussian density in B with mean μ_i and standard deviation σ_j evaluated at x_l.

Moreover, we assume that $q < |I| \times |J|$, i.e. the size of the training set is strictly smaller than the cardinality of the base set B. Let $\mathbf{p} = (p(\mathbf{x}_1), \dots, p(\mathbf{x}_q))'$. The linear system

$$\mathbf{p} = \phi \mathbf{z} \tag{B.4}$$

is solvable and overdetermined as $q < |I| \times |J|$. Although we could solve the system and then approximate the density $p(x)$ with $\sum_{(i,j)} z_{(i,j)} n_{i,j}(x)$, we notice that such approach would require to propagate a number of gaussian densities equal to the size q of the training set, and therefore it would scale linearly with the size of the training set, making a real time implementation computationally intensive. Our goal is to approximate $\mathbf{p}$ using a short linear combination of gaussian densities and at the same time not commit too great an error. Therefore, we look for the sparsest representation of $p(x)$ in the following sense:

$$\min ||\boldsymbol{v}||_0 \quad \text{subject to } ||\mathbf{p} - \phi \boldsymbol{v}||_2 \leq \iota \tag{P1}$$

(P1) is a mathematical programming problem with decision variables $\boldsymbol{v}$, and $||\boldsymbol{v}||_0$ denotes the number of non-zero entries of the vector $\boldsymbol{v}$, i.e.

$$||\boldsymbol{v}||_0 = |\{(i,j) : v_{(i,j)} \neq 0\}| \tag{B.5}$$

where $\upsilon_{i,j}$ is the entry of the $|I| \times |J|$ dimensional vector $\boldsymbol{\upsilon}$ multiplying the gaussian density $n_{i,j}$. If $\boldsymbol{\upsilon}^*$ is the solution of (P1), we would approximate $p(x)$ with

$$p(x) \approx \sum_{(i,j)} \upsilon_{(i,j)} n_{i,j}(x) \tag{B.6}$$

However, this is of little practical use, since the optimization problem (P1) is non-convex and generally impossible to solve as its solution usually requires an intractable combinatorial search. To this purpose, we look for the convex penalty function which is as closest as possible to $||\boldsymbol{\upsilon}||_0$. The closest convex function which does not charge for zero coefficients, and charges proportionally more for small coefficients than for large coefficients turns out to be the l_1 norm. Therefore, we propose to solve the following optimization problem

$$\min ||\mathbf{w}||_1 \quad \text{subject to } ||\mathbf{p} - \phi\mathbf{w}||_2 \leq \epsilon \tag{P2}$$

with decision variable $\mathbf{w}$. The problems (P1) and (P2) differ only in the choice of the objective function, with the latter using an l_1 norm as a proxy for the sparsity count. However, unlike (P1), (P2) is a convex *second order cone* programming problem, and can be solved efficiently in polynomial time using standard optimization algorithms. Therefore, we will approximate the actual density p as

$$p(x) \approx \sum_{(i,j)} w_{(i,j)} n_{i,j}(x) \tag{B.7}$$

A set of experiments presented in (Capponi (2008)) show that the number of non zero components of $\mathbf{w}$ is very small if the density p is reasonably sparse.

Appendix C. Bond and Equity Prices

We recall from Section 3 that the time t price of a bond with maturity T is given by

$$B(t,T) = e^{-r\tau}(K - E^{P^*}[(K - V_T)^+ | \mathscr{F}_t^z]) \tag{C.1}$$

Following the assumptions made in Section 3.4 the probability measure P^* is obtained replacing μ with r and assuming that the biases $h(i)$ and the jump risks $\nu(i)$ are the same under the historical and risk neutral measure. We then have

$$\begin{aligned} E[e^{-r\tau}(K - V_T)^+ | \mathscr{F}_t^z] &= e^{-r\tau} \int_{\mathbb{R}} \int_{\mathbb{R}} P(x_T = x | x_t = y) p_{t|t}(y) (K - e^x)^+ dy dx \\ &= e^{-r\tau} \int_{\mathbb{R}} p_{t|t}(y) \int_{\mathbb{R}} P(x_T = x | x_t = y)(K - e^x)^+ dx dy \end{aligned} \tag{C.2}$$

where we have used the assumption made in Section 3.4 that the volatility risk $\sigma(l)$ is mode-independent. The inner integral can be explicitly computed since $P(x_T = x | x_t = y)$ is a gaussian density with mean $\mu_y = y + (r - 0.5\sigma^2)\tau$ and variance $\sigma^2\tau$, thus

$$\int_{\mathbb{R}} P(x_T = x | x_t = y)(K - e^x)^+ dx = E_{n_{\mu_y, \sigma^2\tau}}[(K - e^X)^+] \tag{C.3}$$

where $n_{\mu_y,\sigma^2\tau}$ indicates the gaussian density with mean μ_y and variance $\sigma^2\tau$ and X is a random variable with density $n_{\mu_y,\sigma^2\tau}$.

We have

$$\begin{aligned} E_{n_{\mu_y,\sigma^2\tau}}[(K-e^X)^+] &= E_{n_{\mu_y,\sigma^2\tau}}[K\mathbf{1}_{\{K\geq e^X\}}] - E_{n_{\mu_y,\sigma^2\tau}}[e^X\mathbf{1}_{\{K\geq e^X\}}] \\ &:= I_1 - I_2 \end{aligned} \tag{C.4}$$

It is easy to see that I_1 is given by:

$$I_1 = KN(-d_2(y))$$

where $d_2(y) = \frac{\mu_y-\log(K)}{\sigma\sqrt{\tau}}$. We compute I_2 as:

$$\begin{aligned} I_2 &= \int_{-\infty}^{-d_2(y)} e^{\mu_y+\sigma\sqrt{\tau}z} n(z)dz \\ &= e^{y+r\tau}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-d_2(y)} e^{-\frac{1}{2}(z-\sigma\sqrt{\tau})^2}dz \\ &= e^{y+r\tau}N(-d_1(y)) \end{aligned} \tag{C.5}$$

where $d_1(y) = d_2(y) + \sigma\sqrt{\tau}$. Therefore,

$$E_{n(x;\mu_y,\sigma^2\tau)}[(K-e^x)^+] = Ke^{-r\tau}N(-d_2(y)) - e^yN(-d_1(y)) \tag{C.6}$$

Thus, we have that

$$E[e^{-r\tau}(K-V_T)^+|\mathscr{F}_t^z] = E_{p_{t|t}}[Ke^{-r\tau}N(-d_2(Y)) - e^YN(-d_1(Y))] \tag{C.7}$$

where Y is a random variable with density $p_{t|t}$ and the price of the bond is then given by

$$B(t,T) = Ke^{-r\tau} - E_{p_{t|t}}[Ke^{-r\tau}N(-d_2(Y)) - e^yN(-d_1(Y))] \tag{C.8}$$

Appendix D. Approximate Bond and Equity Prices

We recall from Section 4.2 that the price of the bond at time t is given by

$$B(t,T) = Ke^{-r\tau} - \sum_j \mu_t^j E[(K-V_{T|t}^j)^+] \tag{D.1}$$

where $V_{T|t}^j$ is the exponential of a gaussian random variable with mean $\hat{x}_{T|t}^j = \hat{x}_{t|t}^j + (r-0.5\sigma^2)\tau$ and standard deviation $\hat{\sigma}_{T|t}^j = \sqrt{\hat{\sigma}_{t|t}^{2,j} + \sigma^2\tau}$. We have

$$\begin{aligned} E[(K-V_{T|t}^j)^+] &= E[K\mathbf{1}_{\{K\geq V_{T|t}^j\}}] - E[V_{T|t}^j\mathbf{1}_{\{K\geq V_{T|t}^j\}}] \\ &:= I_1^j - I_2^j \end{aligned} \tag{D.2}$$

It is easy to see that I_1^j is given by:

$$I_1^j = KP\left(\mathcal{N}(0,1) \leq \frac{\log(K) - \hat{x}^j_{t|t} - (r - 0.5\sigma^2)\tau}{\sqrt{\hat{\sigma}^{2,j}_{t|t} + \sigma^2\tau}}\right)$$
$$= KN(-d_2^j) \tag{D.3}$$

where $d_2^j = \frac{\hat{x}^j_{t|t} - \log(K) + (r-0.5\sigma^2)\tau}{\sqrt{\hat{\sigma}^{2,j}_{t|t} + \sigma^2\tau}}$. We need a little bit of extra work for I_2^j:

$$I_2^j = \int_{-\infty}^{-d_2^j} \exp\left\{\hat{x}^j_{t|t} + (r - 0.5\sigma^2)\tau + \sqrt{\hat{\sigma}^{2,j}_{t|t} + \sigma^2\tau z}\right\} n(z)dz$$
$$= e^{\hat{x}^j_{t|t} + 0.5\hat{\sigma}^{2,j}_{t|t} + r\tau} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_2^j} e^{-\frac{1}{2}(z - \sqrt{\hat{\sigma}^{2,j}_{T|t} + \sigma^2\tau})^2} dz$$
$$= e^{\hat{x}^j_{t|t} + 0.5\hat{\sigma}^{2,j}_{t|t} + r\tau} N(-d_1^j) \tag{D.4}$$

where $d_1^j = d_2^j + \sqrt{\hat{\sigma}^{2,j}_{t|t} + \sigma^2\tau}$. Thus

$$E[e^{-r\tau}(K - V^j_{T|t})^+] = Ke^{-r\tau}N(-d_2^j) - e^{\hat{x}^j_{t|t} + 0.5\hat{\sigma}^{2,j}_{t|t}} N(-d_1^j) \tag{D.5}$$

Therefore, plugging eq.(D.5) into eq.(D.1), we have that the bond price is given by

$$B(t,T) = Ke^{-r\tau} - \sum_j \mu_t^j (Ke^{-r\tau}N(-d_2^j) - e^{\hat{x}^j_{t|t} + 0.5\hat{\sigma}^{2,j}_{t|t}} N(-d_1^j)) \tag{D.6}$$

Appendix E. Bivariate Integrals

It is possible to establish a relation between integrals of the form

$$\int_{-\infty}^{x_L} e^{ax} N(bx + c) n(x) dx \tag{E.1}$$

for some constants a, b, c and the bivariate normal distribution function. This is done as follows:

$$\int_{-\infty}^{x_L} e^{ax} N(bx + c) n(x) dx$$
$$= \int_{-\infty}^{x_L} \int_{-\infty}^{c} e^{ax} n(bx + y) n(x) dy dx$$
$$= \frac{e^{\frac{a^2}{2}}}{2\pi} \int_{-\infty}^{x_L} \int_{-\infty}^{c} e^{-\frac{(1+b^2)(x-a)^2}{2} - \frac{(y+ab)^2}{2} - b(y+ab)(x-a)} dy dx \tag{E.2}$$

This can be related to the bivariate normal distribution as follows:

$$B(a,b,c) := \int_{-\infty}^{x_L} e^{ax} N(bx + c) n(x) dx = e^{\frac{a^2}{2}} \mathcal{P}(X \leq x_L, Y \leq c) \tag{E.3}$$

where (X, Y) has a bivariate normal distribution with

$$\mu_X = a,\ \mu_Y = -ab,\ \sigma_x^2 = 1,\ \sigma_j^2 = 1 + b^2,\ \rho = -\frac{b}{\sqrt{1+b^2}} \tag{E.4}$$

Appendix F. Probability of Default

Set $l = r - 0.5\sigma^2$. Then, we have

$$\begin{aligned} PD(t,T) &= \mathcal{P}(V_T \le K | V_t \ge K, \mathscr{F}_t^z) \\ &= \frac{\mathcal{P}(V_T \le K, V_t \ge K | \mathscr{F}_t^z)}{\mathcal{P}(V_t \ge K | \mathscr{F}_t^z)} \\ &= \frac{\mathcal{P}(X_T \le \log(K), X_t \ge \log(K) | \mathscr{F}_t^z)}{\mathcal{P}(X_t \ge \log(K) | \mathscr{F}_t^z)} \\ &= \frac{\mathcal{P}(l\tau + \sigma\mathcal{N}(0,\tau) \le -X_t + \log(K), X_t > \log(K) | \mathscr{F}_t^z)}{\mathcal{P}(X_t \ge \log(K) | \mathscr{F}_t^z)} \end{aligned} \tag{F.1}$$

where the last equality follows because we can decompose $X_T = X_t + \mathcal{N}(l\tau, \sigma^2\tau)$ as it can be easily checked. Denote $dd_t^j = \frac{\hat{x}_{t|t}^j - \log(K)}{\hat{\sigma}_{t|t}^j}$. Developing the probability (F.1) using the associated density functions, we obtain that the denominator is given by:

$$\mathcal{P}(X_t \ge \log(K) | \mathscr{F}_t^z) = 1 - \sum_{j=1}^m \mu_t^j N(-dd_t^j) = \sum_{j=1}^m \mu_t^j N(dd_t^j) \tag{F.2}$$

while the numerator $\mathcal{P}(l\tau + \sigma\mathcal{N}(0,\tau) \le -X_t + \log(K), X_t > \log(K) | \mathscr{F}_t^z)$ can be computed as follows:

$$\begin{aligned} &= \int_{\log(K)}^{\infty} \left(\sum_{j=1}^m \mu_t^j n(x; \hat{x}_{t|t}^j, \hat{\sigma}_{t|t}^{2,j}) \right) \left(\int_{-\infty}^{-x+\log(K)} n(y; l\tau, \sigma^2\tau) dy \right) dx \\ &= \sum_{j=1}^m \mu_t^j \int_{\log(K)}^{\infty} n(x; \hat{x}_{t|t}^j, \hat{\sigma}_{t|t}^{2,j}) N\left(\frac{-x + \log(K) - l\tau}{\sigma\sqrt{\tau}} \right) dx \\ &= \sum_{j=1}^m \mu_t^j \int_{\log(K)}^{\infty} \hat{\sigma}_{t|t}^j n\left(\frac{x - \hat{x}_{t|t}^j}{\hat{\sigma}_{t|t}^j} \right) N\left(\frac{-x + \log(K) - l\tau}{\sigma\sqrt{\tau}} \right) dx \\ &= \sum_{j=1}^m \mu_t^j \int_{-\infty}^{dd_t^j} n(y) N\left(\frac{y\hat{\sigma}_{t|t}^j - \hat{x}_{t|t}^j + \log(K) - l\tau}{\sigma\sqrt{\tau}} \right) dy \end{aligned} \tag{F.3}$$

where the last line follows from a change of variable $y = \frac{\hat{x}_{t|t}^j - x}{\hat{\sigma}_{t|t}^j}$. Using the result in Appendix E, we can conclude that the numerator is a convex combination of bivariate cumulative distribution functions, i.e.

$$\mathcal{P}(V_T \le K, V_t \ge K | \mathscr{F}_t^z) = \sum_{j=1}^m \mu_t^j \mathcal{P}(X_j \le dd_t^j, Y_j \le -d_2(\hat{x}_{t|t}^j)) \tag{F.4}$$

where $dd_t^j = \frac{\hat{x}^j_{t|t} - \log(K)}{\hat{\sigma}^j_{t|t}}$, d_2 is the function defined in eq.(3.12), and (X_j, Y_j) has a bivariate normal distribution function with positive correlation $\rho_{X_j Y_j}$:

$$\mu_{X_j} = \mu_{Y_j} = 0, \sigma^2_{X_j} = 1,\ \sigma^2_{Y_j} = 1 + \frac{\hat{\sigma}^{2,j}_{t|t}}{\sigma^2 \tau},\ \rho_{X_j Y_j} = \frac{\hat{\sigma}^j_{t|t}}{\sqrt{\sigma^2 \tau + \hat{\sigma}^{2,j}_{t|t}}} \tag{F.5}$$

Defining $w_t^j = \frac{\mu_t^j N(dd_t^j)}{\sum_{i=1}^m \mu_t^i N(dd_t^i)}$, we obtain

$$PD(t,T) = \sum_j w_t^j \mathcal{P}(Y_j \le -d_2(\hat{x}^j_{t|t}) | X_j \le dd_t^j) \tag{F.6}$$

Appendix G. Recovery Rate

Set $l = (r - 0.5\sigma^2)$, $Z_\tau = l\tau + \sigma \mathcal{N}(0, \tau)$. We have

$$\begin{aligned} RR(t,T) &:= \frac{1}{K} E[V_T | V_T < K, V_t > K, \mathscr{F}_t^z] \\ &= \frac{1}{K} E[e^{X_t + Z_\tau} | X_t > \log(K), Z_\tau < \log(K) - X_t, \mathscr{F}_t^z] \\ &= \frac{\int_{\log(K)}^\infty e^x \int_{-\infty}^{\log(K)-x} e^{l\tau + \sigma\sqrt{\tau} y} n(y) dy \sum_{j=1}^m \mu_t^j n(x; \hat{x}^j_{t|t} \hat{\sigma}^{2,j}_{t|t}) dx}{\mathcal{P}(X_t + Z_\tau < \log(K), X_t > \log(K) | \mathscr{F}_t^z)} \end{aligned} \tag{G.1}$$

The denominator can be decomposed as follows:

$$\mathcal{P}(X_t + Z_\tau < \log(K), X_t > \log(K) | \mathscr{F}_t^z) = PD(t,T) P(X_t > \log(K) | \mathscr{F}_t^z) \tag{G.2}$$

where both of the quantities above have been computed in Appendix D. The numerator in the last line of equation (G.1) can be written as:

$$\begin{aligned} &= e^{r\tau} \sum_{j=1}^m \mu_t^j \int_{\log(K)}^\infty e^x n(x; \hat{x}^j_{t|t}, \hat{\sigma}^{2,j}_{t|t}) N\left(\frac{\log(K) - x - l\tau - \sigma^2 \tau}{\sigma\sqrt{\tau}} \right) dx \\ &= e^{r\tau} \sum_{j=1}^m \mu_t^j e^{\hat{x}^j_{t|t}} \int_{-\infty}^{dd_t^j} e^{-y \hat{\sigma}^j_{t|t}} n(y) N\left(\frac{-\hat{x}^j_{t|t} + y\hat{\sigma}^j_{t|t} + \log(K) - l\tau - \sigma^2\tau}{\sigma\sqrt{\tau}} \right) dy \end{aligned} \tag{G.3}$$

where $dd_t^j = \frac{\hat{x}^j_{t|t} - \log K}{\hat{\sigma}^j_{t|t}}$. The last line of the derivation above follows from the change of variable $y = \frac{\hat{x}^j_{t|t} - x}{\hat{\sigma}^j_{t|t}}$. Using the result in Appendix E, we obtain that the numerator is again a convex combination of bivariate gaussian cumulative distribution functions given by

$$e^{r\tau} \sum_{j=1}^m \mu_t^j e^{\hat{x}^j_{t|t} + \frac{\hat{\sigma}^{2,j}_{t|t}}{2}} \mathcal{P}(\xi_j \le dd_t^j, \psi_j \le -d_1(\hat{x}^j_{t|t})) \tag{G.4}$$

where d_1 is the function defined in eq.(3.12), and (ξ_j, ψ_j) is a bivariate gaussian with negative correlation coefficient $\rho_{\xi_j \psi_j}$:

$$\mu_{\xi_j} = -\hat{\sigma}^j_{t|t},\ \mu_{\psi_j} = \frac{\hat{\sigma}^{2,j}_{t|t}}{\sigma\sqrt{\tau}}, \sigma^2_{\xi_j} = 1,\ \sigma^2_{\psi_j} = 1 + \frac{\hat{\sigma}^{2,j}_{t|t}}{\sigma^2\tau},\ \rho_{\xi_j\psi_j} = -\frac{\hat{\sigma}^j_{t|t}}{\sqrt{\sigma^2\tau + \hat{\sigma}^{2,j}_{t|t}}} \tag{G.5}$$

Defining $w^j_t = \frac{\mu^j_t \mathcal{P}(X_j \le dd^j_t, Y_j \le -d_2(\hat{x}^j_{t|t}))}{\sum_{i=1}^m \mu^i_t \mathcal{P}(X_i \le dd^i_t, Y_i \le -d_2(\hat{x}^i_{t|t}))}$, we obtain

$$RR(t,T) = e^{r\tau} \sum_{j=1}^m w^j_t \frac{e^{\hat{x}^j_{t|t} + \frac{\hat{\sigma}^{2,j}_{t|t}}{2}}}{K} \frac{\mathcal{P}(\xi_j \le dd^j_t, \psi_j \le -d_1(\hat{x}^j_{t|t}))}{\mathcal{P}(X_j \le dd^j_t, Y_j \le -d_2(\hat{x}^j_{t|t}))} \tag{G.6}$$

References

[1] E. Altman, A. Resti and A. Sironi, Analyzing and explaining default recovery rates, *Technical report: The International Swaps and Derivatives Association* (2000).

[2] F. Black and J. Cox, Valuing corporate securities: some effects of bond indenture provisions, *Journal of Finance* **31** (1976) 351–367.

[3] D. Brigo and M. Morini, CDS Calibration with tractable structural models under uncertain credit quality, *Risk Magazine* **19** (2006).

[4] A. Capponi, A convex optimization approach to Filtering in Jump Non Linear Systems, *Technical Report, California Institute of Technology*, available at `http://www.cs.caltech.edu/~acapponi/papers/HybridEstimation` (2008).

[5] U. Cetin, R. Jarrow, P. Protter and Y. Yildirim, Modeling credit risk with partial information, *Annals of Applied Probability* **14** (2004) 1167–1178.

[6] U. Cherubini and M. Manera, Hunting the Living Dead: A peso problem in Corporate Liabilities Data, *ICFAI Journal of Financial Risk Management* **3** (2006) 44–56.

[7] D. Coculescu, H. Geman and M. Jeanblanc, Valuation of Default Sensitive Claims Under Imperfect Information, *Finance and Stochastics* **12** (2008), 195-218.

[8] R. Collin-Dufresne, Do credit spreads reflect stationary leverage ratios?, *Journal of Finance* **56** (2001) 1929–1957.

[9] J. Duan, Maximum Likelihood Estimation Using Price Data of the Derivative Contract, *Mathematical Finance* **4** (1994) 155-167.

[10] J. Duan, Correction: Maximum Likelihood Estimation Using Price Data of the Derivative Contract, *Mathematical Finance* **10** (2000) 461- 462.

[11] D. Duffie and D. Lando, Term structure of credit spreads with incomplete accounting information, *Econometrica* **63** (2001) 633–664.

[12] R. Frey and W. Runggaldier, Credit Risk and Incomplete Information: a Nonlinear-Filtering Approach, Preprint (2007).

[13] R. Geske, The valuation of corporate liabilities as compound options, *Journal of Financial and Quantitative Analysis* **12** (1977) 541–552.

[14] K. Giesecke and L. R. Goldberg, Forecasting default in the face of uncertainty, *Journal of Derivatives* **12** (2004) 11–25.

[15] X. Guo, R. A. Jarrow and Y. Zeng, Credit risk with incomplete information, *Working paper*, Cornell University (2006).

[16] M. JeanBlanc and M. Rutkowski, Modelling Default risk: An Overview, *Mathematical Finance: Theory and Practise, Higher Education Press, Beijing* **1** (2000) 71–269.

[17] H. E. Leland and K. Toft, Optimal capital structure, endogeneous bankruptcy, and the term structure of credit spreads, *Journal of Finance* **51** (1996) 987–1019.

[18] F. Longstaff and E. Schwartz, A simple approach for valuing risky fixed and floating rate debt, *Journal of Finance* **50** (1995) 789–819.
[19] R. C. Merton, On the pricing of corporate debt: the risk structure of interest rates, *Journal of Finance* **29** (1974) 449–470.
[20] W. J. Runggaldier, Estimation via stochastic filtering in financial market models, In *Mathematics of Finance (G.Yin and Q.Zhang eds.), Contemporary Mathematics, America Mathematical Society* **351** (2004) 309–318.
[21] P. Fisher and R. Verrecchia, R., Reporting bias, *The Accounting Review* **75** (2000) 229–245.
[22] F. Yu, Accounting transparency and the term structure of credit spreads, *Journal of Financial Economics* **75** (2005) 53–84.
[23] T. Wang, Real Investment and Corporate Securities Fraud, *Working paper*, University of Minnesota (2007).