Università
Ca'Foscari
Venezia

Ca' Foscari
Dorsoduro 3246
30123 Venezia

Master's Degree programme – Second Cycle (*D.M. 270/2004*) in Economia-Economics

Erasmus Mundus QEM: Models and Methods of Quantitative Economics

Final Thesis

The Role of Volatility Persistence in Default Probability Prediction: A Bayesian Model

Supervisor
Ch. Prof. Roberto Casarin

Graduand
Mariam Petrosyan
Matriculation Number 861174

Academic Year
2016 / 2017

Abstract

This thesis studies a new specification of stochastic volatility (SV) process in the prediction of probability of default in one year, and in estimation of credit default swap spreads. The novelty of the model consists in allowing for both stochastic interest rate and volatility and in the extension of the SV dynamics process to an AR (2) process. The estimation is carried out by exploiting Bayesian methods via implementation of Gibbs sampling for the state space model of stochastic volatility and returns. The model is compared with the structural credit risk model with stochastic interest rate and AR (1) stochastic volatilities presented in Rodriguez et al. (2014) and with a model that gives white noise specification of the volatilities on the basis of Bayes factors, using financial data series of three firms.

Keywords: *Stochastic volatility, Structural credit risk models, Risk analysis, MCMC algorithms, Bayes factors, State space models, Kalman filter*

JEL classification: *C11, C13, G01, G32*

List of Tables

List of Figures

Contents

1 Introduction

In this thesis a new stochastic volatility model is developed to capture the persistence in volatility when predicting the default and estimating CDS spreads. The goal of this study is to propose a model that would provide a better forecast for default events.

Given that the event of default means the inability of corporations of meeting their obligations towards the creditors, the risk of default is tantamount to the credit risk. Hence the assessment of creditworthiness of corporations is carried out as an essential part of risk management since most of the risks in both financial and industrial sectors are related to the risk of corporations going default.

Taking into consideration that banks and other financial intermediaries are primarily engaged in lending activities, the development of models that would adequately capture asset volatilities and provide reliable prediction for default probabilities is of particular importance for financial sector. This significance is enhanced by the fact that the activities of financial intermediaries are mainly financed with deposits from customers and borrowings from investors, and by the interconnected and complex nature of financial markets which create contagion effects by this leading to consequences that are not only sector vise but may afflict the entire economic system. This has been witnessed during the latest financial crisis of 2007-2008, when the sub prime mortgage loan market collapsed as a consequence of high default rates caused by unprecedentedly high ratio of the debt to assets that had been maintained by banks. The repercussions were severe for the international financial system with as a major player as the investment bank Lehman Brothers going bankrupt. Risk hedging tools like credit default swaps that were designed to spread the risk appeared to concentrate it instead. And as banks were mostly independent and were left to implement their internal risk assessment models, they were assigning lower risk to their assets as result having low capital ratios and thus becoming more vulnerable to adverse economic events (”The origins of the financial crisis: Crash course”,2013).

In the aftermath new stricter regulations were proposed inducing higher minimum

capital requirements and restricting the range of transactions that banks could engage in. It became clear as well that the models that had been used in credit risk management were not performing sufficiently good, hence needed to be modified an developed to be able to capture the risks prevalent in interconnected financial markets. This in turn would result in new pricing models for credit derivatives that would more accurately reflect the credit worthiness of corporations.

Credit risk measurement models can be differentiated based on whether default is viewed as an endogenous event driven by changes in the internal structure of firms' assets and liabilities, or it is assumed that defaults are exogenous and default arrival times are being modeled based on observed data of historical default events. Models that are categorized under the first approach are called structural credit risk (SCR) models whereas the ones developed within latter approach are known as intensity based or reduced form models. A third category of models emerges from the combination of the previously mentioned classes of models. It is known as incomplete information and in this case it is assumed that investors can not observe the firms' asset values directly because of noisy or delayed accounting reports, so what is actually observed is the asset values plus the noise term. In this way the "incomplete information" is introduced into structural models.

This thesis studies a SCR model, the fundamental reference model of which was proposed by Merton (1974). Here assumptions are made about the dynamics of the firm's assets, its capital structure and debt, and default occurs when asset values fall below the debt value. Merton (1974) views firms' capital as an implicit call option on the assets with a strike price equal to debt value. Thus, equity can be assessed using Black and Scholes (1973) option pricing formula. Merton proposed that volatility of asset values is constant over time and the interest rates are deterministic.

Since the first introduction of Merton's model researchers have proposed numerous extensions of it each of which altered one or more of basic model's underlying assumptions regarding asset price dynamics , volatility or interest rate processes. The constant volatilities assumption, for instance, has been generalized by incorporation of time varying

conditional volatility, the main two model groups of which are autoregressive conditional heteroskedastic (ARCH) -type models and stochastic volatility (SV) models. In the former case volatility is modeled as a deterministic function of past returns, whereas in the case of SV models conditional volatility is subject to random shocks which renders it unobservable and latent (Rachev et al., 2008).

A further extension to the baseline model is the incorporation of uncertainty in the interest rate process (Rodriguez et al., 2014). The role of interest rate uncertainty in valuing risky debts has been discussed by Shimko (1993) and Longstaff et al. (1995) where they find empirical evidence in favor of incorporation of stochastic interest rates.

This thesis is developed in the framework of both stochastic volatilities and stochastic interest rates and estimates default probabilities and CDS spreads in SCR model. In model specification and inference it mainly follows the procedure provided in Rodriguez et al. (2014) by extending the stochastic volatility process to autoregressive of order 2 (AR(2)) process with the purpose of detecting the role of volatility persistence in default prediction. The estimation is carried out by exploiting Bayesian methods via implementation of Gibbs sampling for the state space model of stochastic volatility and returns and Metropolis Hastings acceptance rejection sampling algorithms. The model is compared on the basis of Bayes factors with the structural credit risk model with stochastic interest rate and AR (1) stochastic volatilities presented in Rodriguez et al. (2014) and with a model that gives white noise specification of the volatilities, considering financial data series of three firms: Lehman Brothers, Citigroup and Ford Motor Company. An accurate inference on the degree of volatility persistence is an important economic issue since it will allow to estimate the impact of a transitory shock on the price of a contingent claim (Lamoureux and Lastrapes, 1990)

The rest of the thesis is organized as follows: section 2 presents a literature review on three main approaches to credit risk modeling and on incorporating stochastic interest rates and stochastic volatilities into structural credit risk model; section 3 defines model dynamics, section 4 describes the inference methodology, and finally sections 5 and 6

present some empirical results and conclude the thesis respectively.

2 Literature review

In this section a more in debt literature review on the credit risk modeling approaches discussed in the introduction and their modifications is represented.

2.1 Main approaches to credit risk modeling

2.1.1 Structural approach

As the focus of this study is development of model for default prediction within structural approach, this section begins with an extensive literature review on SCR models. This class of models owes its name to the fact that within this approach default is considered as an endogenous event and is linked to the capital structure of the firm. The seminal paper on SCR model was presented by R. Merton (1974) , where he developed a model for pricing of corporate liabilities based on the Black and Scholes option pricing formula (Black and Scholes, 1973). While to develop the model Merton makes assumptions regarding "efficient" markets, the most consequential assumptions pertain to the dynamics of asset prices, default arrival times and interest rate structure. Namely, Merton assumes that the process governing assets' values dynamics, V_t, is a geometric Brownian motion and is given by the following partial differential equation[1]:

$$dV_t = \mu V_t dt + \sigma V_t dW_t, \qquad t \in \mathbb{R}^+ \tag{2.1}$$

Where μ is the instantaneous rate of return on the asset per unit time, σ is the instantaneous volatility of the return on the assets and $\{W_t\}_{t\geq 0}$ is a Wiener process with $W_0 = 0$ (Merton, 1974). As can be inferred from 2.1, in this model it is assumed that the volatility is constant over time. Further, by definition the relationship between firm's equity, assets and debt is given by $E_t = V_t - D_t$. Thus, having equity as a function of asset values, Merton (1974) derives the following parabolic partial differential equation for equity

[1]In the original text Merton (1974) assumes a payout on assets, so the first term on the right hand side of the equation is $(\mu V_t - c)dt$

pricing:

$$\frac{1}{2}\sigma^2 V_t^2 \frac{\partial^2 E_t}{\partial V_t^2} + rV_t \frac{\partial E_t}{\partial V_t} - rE_t - \frac{\partial E_t}{\partial \tau} = 0 \tag{2.2}$$

Where r is the interest rate which is assumed to be constant over time and known with certainty and τ is the time to maturity of an assumed debt. Merton (1974) obtains the boundary conditions for equation (2.2) by first considering the basic balance sheet identity, that is, $V_t = K_t(V_t) + E_t(V_t)$, where equity and debt are expressed as functions of current asset values, and then pointing out that neither debt nor equity can take negative values and hence the following initial condition holds:

$$E_t(0) = K_t(0) = 0 \tag{2.3 a}$$

Further, asset values are always greater than or equal to equity (and debt), from which the following regularity condition is derived (Merton, 1974)

$$E_t/V_t \leq 1 \tag{2.3 b}$$

which provides the second boundary condition with $0 \leq V_t \leq \infty$ (Merton, 1974).
Finally, the third constraint is obtained by observing that at maturity debt holders will either get paid K, if firm's assets exceed obligations or if the firm fails to make promised payments, the payout will be equal to the value of total assets: $D_T = \min(V_T, K)$. The investors, on the other hand, will receive the equity with the value:

$$E_T = \max(0, V_T - K) \tag{2.3 c}$$

Equation (2.2) with boundary condition (2.3 c) are the same as the pricing equations for European call option on a non divined paying stock ,presented in Black and Scholes (1973) with spot price corresponding to the asset values and the exercise price to K. In this manner Merton (1974) shows that Black and Scholes (1973) equation can be exploited

to have an analytical expression for pricing equity:

$$E_t = V_t \Phi(d_t) - K e^{-r(T-t)} \Phi(d_t - \sigma \sqrt{T-t}) \tag{2.4}$$

where

$$d_t = \frac{\ln(\frac{V_t}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma \sqrt{T-t})}$$

and Φ is the standard Gaussian distribution function.

Equipped with this analogy between European call option and equity it is possible not only to have the pricing formula for the latter, but also have analytical expressions for predicted default probabilities and CDS spreads. Defining default as the event when $V_t < K$, Merton's model provides the following expression for the probability of default (Duan and Fulop, 2009):

$$P(V_t < K) = \Phi\left(\frac{\ln(K/V_t) - (\mu - \frac{1}{2}\sigma^2)(T-t)}{\sigma \sqrt{T-t}}\right) \tag{2.5}$$

and the pricing formula for CDS spreads is given as:

$$S_t = -\frac{1}{T-t} \ln\left(\frac{v_t}{K}\Phi(-d_t) + e^{r(T-t)}\Phi(d_t - \sigma \sqrt{T-t})\right) - r \tag{2.6}$$

While being a milestone in credit risk modeling, Merton's model (1974) is too simplistic in its initial form and has been altered since its first introduction and a number of models have been proposed by researchers who reconsider different assumptions underlying the basic model. The constant interest rate assumption has been challenged for instance by Shimko et al. (1993), Longstaff and Shwartz (1995), Bryis and Van de Varenne (1997), Kim et al. (1993) by introduction of stochastic interest rates. Accordingly, Shimko et al. (1993) and Longstaff and Shwartz (1995) incorporate stochastic interest rates into Merton's (1974) model by implementation of Vasicek's (1977) model. In both studies default is assumed to take place prior to maturity and before the firm exhausts all of its assets, in contrast to the assumptions made in Merton (1974), where on of the crucial assumptions is that default can occur only at the maturity of the debt under consideration. Among

other contributions, most importantly, these papers discuss and empirically document that defaults are triggered not only by the capital structure of firms, but as well by interest rate structure. Additionally making an assumption regarding the correlation between risk free interest rate and asset returns they study the effect of that correlation on CDS spread valuation. Kim et al. (1993), on the other hand, assume a Cox, Ingersoll and Ross (1985) type stochastic process for risk free interest rate. They too assume correlation between interest rate and asset returns.

Leland and Toft (1996) contribute by incorporating taxes and bankruptcy costs. In their model it's assumed, that firm owners decide whether to recognize bankruptcy or not based on the amount and maturity of debt as well on tax benefits and bankruptcy costs. Geske (1977) contributes by considering sequential coupon payments. Zhang et al. (2005) incorporate both stochastic volatility and jumps in asset values and concentrate their study on CDS spread pricing. They illustrate empirically that in this manner the model is significantly more able to capture the observed movements of CDS spreads and hence provide accurate forecasts, as opposed to previous works where even in case of stochastic interest rates, endogenous default barriers (Leland and Toft, 1996) and mean reverting leverage ratios the models have little power in estimation of default risks which underlie CDS spread valuation. The sensitivity analysis carried out by Zhang et al. (2005) indicates positive correlation between spreads and volatility, long run mean of volatility process, whereas higher mean reversion coefficient of volatilities reduces spreads, and the effect of jumps in asset is asymmetric with negative jumps tending to have a larger effect than positive ones. Delianedis and Geske (1999) consider credit spreads as indicators of credit risk and study the factors that underlie the pricing of spreads such as taxes, jumps in the asset values, liquidity factors and market risk factors. They price default spreads as components of credit spreads and in this manner they make two important observations. First, corporate riskiness is not fully explained by default risk which is evinced by the difference between corporate spreads and model induced default spreads. The second observation made by Delianedis and Geske (2003) is that failing to consider jumps in the asset value dynamics significantly reduces predicted defaults resulting in underestimation of default risk (Delianedis and Geske,2003). A comprehensive empirical analysis of seven structural credit risk models can be found in Huang and Huang, 2012. They report similar results as

those of Delianedis and Geske (2003) and in addition they calibrate the models to match historical default events. A good reference for empirical examination of a number of structural credit risk models is provided by Eom et al. (2004). Here the authors test Merton's basic model against its alternations developed by Geske (1977), Longstaff and Schwartz (1995), Leland and Toft (1996) and Collin-Dufresne and Goldstein (2001).

2.1.2 Reduced form approach

Intensity based or reduced form approach dates back to Pye (1974) and Litterman and Iben (1991). As opposed to the structural models described above, which link default explicitly to the first time that assets fall below a certain level, reduced-form or intensity based models are built based on the assumption that default is an exogenous event and its arrival time is given by an intensity parameter which typically is estimated by calibrating the model to the market data or based on historical observations, and firm's value may, though not necessarily, be included as an auxiliary covariate (Bielcki and Rutkowsky, 2002). Most commonly the event of the default is modeled as the first arrival time of a Poisson process which in the basic version has a constant mean arrival time λ, which is called the intensity (Duffie and Singleton, 2003). Within this set up, the probability that the default will not occur during succeeding t years is given by the exponential function $p(t) = exp(-t\lambda)$ and the probability of (instantaneous) defaulting at time $t + dt$ conditional on no default at time t is given by λdt. In addition, the expected time to default is given by $1/\lambda$ (Duffie and Singleton, 2003).

A simple extension to the above described model is assuming for a deterministic, but continuously varying intensity, in which case the survival probability for the next t years will be given by $e^{-\int_0^t \lambda(t)dt}$. Apart of being used for default prediction per se, λ is an indicator of the credit quality of a firm, and its variations over time express changes in its creditworthiness. It is more reasonable to assume that the intensity will vary differently based on the arrived information. This suggests that more realistically λ should be given as a stochastic process. According to this setting, conditional on the path of the intensity $\lambda(t), t \geq 0$ defaults arrive as Poisson process the time varying intensity of which is given

by $\lambda(t)$ (Duffie and Singleton, 2003):

$$p(t) = E[P(\tau > t|\lambda(s) : 0 \leq s \geq t)] = E[e^{-\int_0^s \lambda(s)ds}] \tag{2.7}$$

The intensity λ can be modeled in various ways. Two common instances are mean reverting intensities with jumps and Cox Ingersoll Ross (CIR) intensities. In the former case λ is assumed to be reverting around its long run mean in between jumps , whereas in the latter case, the dynamics of the intensity is given in a similar way as in CIR interest rate model:

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

where W_t is a Wiener process, γ is the long run mean, κ is the adjustment rate and σ is the volatility coefficient. Other examples of intensity models include , for instance, implying for a (possibly) multi-dimensional state process underlying the intensity and giving dynamics for that latent variable.

A comprehensive step by step development of an intensity model for simultaneously predicting defaults and estimating interest rates is given in Lando (2004), together with a general guideline on the model calibration. Bielcki and Rutkowsky (2002) provide an extensive coverage on both mathematical apparatus underlying the estimation of intensity models and empirical illustration of the theory together with practical methods of risk management within the scope of reduced form approach. The focus of model implementation in Bielcki and Rutkowsky (2002) is pricing of basket derivatives and prediction of default in case of correlation between default arrival times for multiple firms. Other noteworthy references for reduced form approach include: Jarrow and Turnbull (1995), where they build the model in discrete time approach, Jarrow et al.(1997), an extension to the former model with a Markov process for default (bankruptcy), Collin-Dufrense et al. (2003 a), Madan and Unal (1998),Hull and White (1995).

2.1.3 Incomplete information approach

Incomplete information approach was first introduced by Duffie and Lando (2001) as a combination of the two competing approaches of structural and reduced form models. It has been discussed in Collin-Dufresne et al (2003), Cetin et al. (2004), and Jarrow and Protter (2004). The motive for developing this approach was to explain the discrepancy between the term structure of credit spreads that is implied by structural models and the one actually observed in market. More precisely, according to structural models as the maturity of debts goes to zero, so should do the credit spreads (the difference between risky debt yield and risk free bond yield), which is in contrast what has actually been observed in the market. Though this "puzzle" has been attempted to be solved by adding jump factors in asset dynamics to account for unobservable components (see Zhou, 1997), Duffie and Lando (2001) argue that jumps only are not able to capture the difference between the observable (noisy) and real asset values. Thus, they transform a structural model into reduced form model by introducing the "incomplete" information.

In particular, the dynamics of log asset values observed by debt holders is given as $Y_t = Z_t + U_t$ where Z_t is the true process governing asset values and is given as $Z_t = Z_0 + mt + \sigma W_t$, with W_t distributed as a Wiener process, and U_t is normally distributed, independently of W_t, and can be thought of as the accounting noise. The crucial difference between this model and structural models with perfect information when firm's value process is adapted with regard to the information available at time t , and imperfect information is that in the former case the probability of defaulting within a h short period of time converges to zero, as $h \to 0$, whereas under incomplete information setting there exists a non zero limit for that probability and this limit is defined as the default intensity (Duffie and Lando, 2001). Further, Duffie and Lando (2001) provide an explicit formula for intensity λ_t of default (poisson) time τ as follows:

$$\lambda_t = \frac{1}{2}\sigma^2 f_x(t, \underline{\nu}), \qquad 0 \leq t \leq \tau$$

where f_x is the derivative of conditional (on observable factors, Y_t) density function of Z_t, and $\underline{\nu}$ is the boundary asset value at which $f = 0$. In this manner the predictable default

(stopping) times are transformed into unpredictable stopping process with the above given intensity (Jarrow et al., 2004).

A good comparative analysis of structural and reduced form approaches and a discussion of Duffie and Lando (2001) model is presented in Jarrow and Protter (2004). The key point highlighted there is that the fundamental difference between the three presented approaches is in the level of information that is assumed to be available to the researcher. From this variation in the information set steam the assumptions and implications regarding predictability of default events. They conclude thus that the question regarding which model to choose pins down to the question which information set should be used. They argue that since credit risk assessment is conducted by market participants for whom the detailed information regarding firms' balance sheets is obscure, preference should be given to reduced form models. On the other hand, if credit risk modeling is performed by the managers with the purpose of optimal capital structuring, then structural approach should be applied.

2.2 Incorporation of stochastic volatility and stochastic interest rates into structural credit risk models

This section gives a review on the literature on the extensions of Merton's structural credit risk model with the focus on incorporation of non constant and non deterministic interest rates and volatilities.

Since the introduction of Merton's model it has been acknowledged by researchers that the assumptions it makes are too simplistic and are not in accord with empirical evidence. In particular, much has been discussed the assumption regarding deterministic character of risk free interest rates. The modification of the basic model with introduction of stochastic interest rate has been highlighted both from the perspective of first modeling interest rate in a more realistic way, since interest rates are in fact stochastic, and second it has been documented that there is a correlation between risk free interest rates and credit spreads (Lando,2004). One of the first studies on this topic has been presented by Shimko et al. (1993). Here the dynamics of interest rates is given according to Vasicek's

model (1977):

$$dr_t = \rho(\gamma - r)dt + \sigma_r dW_{rt} \tag{2.8}$$

where ρ and γ are respectively the mean reverting coefficient and long run mean of interest rate and W_{rt} is Wiener process, which is correlated with the Wiener process controlling asset values process, W_{vt}. They demonstrate that in most cases high volatility of interest rates as well stronger correlation between interest rates and asset values induce larger credit spreads, which indicates that failing to take into account stochastic nature of interest rates may lead to underpricing of credit spreads and underestimation of credit risk. A similar observation is made by Longstaff and Swartz (1995) and Lando (2004). Saa-Requejo and Santa Clara (1997) in addition assume that default boundaries are stochastic. Briys and Varenne (1997) model interest rate with time dependent deterministic coefficients.

Another oversimplification in Merton's (1974) assumptions that has been widely discussed is the constant volatility hypothesis. The two main approaches to account for changing nature of volatility process have been modeling it through observation induced and parameter driven processes (Shephard, 1996). In the former case volatilities can be directly estimated as a function of values of observable variables, for instance asset values or rates of return. This class of models includes the Autoregressive Conditional Heteroskedasticity models, where variances are modeled as functions of squares of past values of the observed variable. An early introduction of this class of models has been suggested by Engle (1982) where system dynamics is defined through the following equations:

$$y_t = \epsilon_t h_t^{1/2}, \qquad h_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \cdots + \alpha_p y_{t-p}^2$$

Of the advantages of ARCH type models Engle (1982) highlights their ability to capture the changing nature of forecast variances and their usefulness as a good approximation to models with a more complex, non standard disturbances. This class of models has been deemed compelling also from the perspective of estimation since it is straightforward by the very construction of models also existing estimation techniques of autoregressive and moving average models can be used for ARCH models (Shephard, 1996) An extensive

coverage of ARCH type models and their application in finance is provided in Shephard (1996) and Bollersev et al. (1992). Duan (1995) exploits GARCH (Generalized ARCH) setting in option pricing context.

The second approach for introducing heteroskedasticity and the main focus of this thesis, parameter driven volatility models , is based on the assumption that volatility is a stochastic process that is governed by some unobservable latent factors.

First introduction of SV models dates back to Taylor (1982). Where the stochastic process governing returns is designed as a product two independent stochastic processes:

$$y_t = \alpha + \sigma_t \epsilon_t$$

where α and σ_t are the long run mean and the conditional standard deviation of the returns respectively, with the standard deviation fluctuating over time, and ϵ_t is a process with zero mean and unit variance. Further, $\ln \sigma_t$ is modeled as AR(1) process.

The seminal work in incorporating stochastic volatilities in Black and Scholes (1973) option pricing is Hull and White (1987) where the system dynamics of asset values and volatilities is given by the following continuous time diffusion processes:

$$dV_t = \mu V_t dt + \sigma_t V_t dW_t \qquad d\sigma_t^2 = \alpha \sigma_t^2 dt + \tau \sigma_t^2 dZ_t \tag{2.9}$$

Another milestone in options pricing (and consequently structural credit risk modeling) with stochastic volatility is the paper by Heston (1993) where the system dynamics differs from that of Hull and White (1987) by the assumption that volatility follows an Ornstein-Uhlenbeck (O-U) process as follows:

$$d\sigma_t = -\alpha \sigma_t dt + \tau dZ_t \tag{2.10}$$

Further, a number of authors (see, for example, Chesney and Scott (1989), Scott (1987,1991) and Wiggins (1987)) have modeled the dynamics of log volatility with a mean reverting

O-U process:

$$d\ln\sigma_t = \phi(\alpha - \ln\sigma_t)dt + \tau dZ_t \tag{2.11}$$

Other specifications of the dynamics of volatility with their corresponding closed form solutions of option pricing can be found in Johnson and Shanno (1987), Stein and Stein (1991) Bailey and Stulz (1989).

While continuous time diffusion models have been widely applied in option pricing for the purposes of data analysis and inference on the behavior of volatilities either ARCH type or discrete time specifications of SV models have been in use. One specific discrete time SV model considered by Chesney and Scott (1989) and Duffie and Singelton is[2] :

$$\ln(V_t) = \ln(V_{t-1}) + \mu + \sigma_{t-1}\epsilon_{1t} \qquad \ln(\sigma_t) = \alpha + \phi(\ln(\sigma_{t-1}) - \alpha) + \tau\epsilon_{2t} \tag{2.12}$$

ϵ_{1t} and ϵ_{2t} follow iid standard Gaussian distribution and correlations between them in general case is non zero. The fact that in the logarithm asset value equation instead of the contemporaneous one there is the lagged value of volatility is due to Euler approximation of continuous time dynamics of assets in (2.9) (Taylor, 1994). Taylor (1994) suggests considering a simpler approximation for asset prices in (2.9): $\ln(V_t) = \ln(V_{t-1}) + \sigma_t\epsilon_{1t}$. Rodriguez et al. (2014) use this specification and Euler discretization of Vasicek (1973) model for logarithm interest rates and define model dynamics as follows:

$$\begin{aligned} \ln V_t &= \ln V_{t-1} + \mu + \sigma_t\epsilon_{1t} \\ \ln\sigma_t &= \alpha + \phi(\ln\sigma_{t-1} - \alpha) + \tau\epsilon_{2t} \\ \ln r_t &= \gamma + \rho(\ln r_{t-1} - \gamma) + \kappa\epsilon_{3t} \end{aligned} \tag{2.13}$$

where the error terms $\epsilon_{it}, \quad i = 1, 2, 3$ are assume to follow iid standard Gaussian distribution with null correlation among them.

Regarding the relation between ARCH type and SV models it can be said that while in contrast with ARCH models estimation of stochastic volatility models is substantially

[2]See appendix 6 for derivation of discrete time specification from the continuous one.

harder to conduct, they are believed to better capture the complexity of information flow (Shephard,1996). Empirical investigation on comparative performances of ARCH type and SV models can be found in Kim et al. (1998), where application of both models on stock indexes returns reveals that SV models perform better, Danielsson (1998), where comparison is conducted between multivariate models and he too finds that SV models generally produce higher marginal likelihood values. As noted by Fleming and Kirby (2003), at least in the context of portfolio optimization and risk management , empirical studies speak in favor of SV models.

3 Model specification

As has been stated, this thesis studies a model developed within the context of Merton's (1974) structural model. Accordingly, at each point t in times the balance sheet is assumed to composed of debt K_t, which matures m periods later and equity E_t both entities are traded in the market, moreover, equity is homogeneous in that different classes of priorities of shares are not considered. On the other hand, for asset values we naturally must have for all times t that $V_t = K_t + E_t$

3.1 Defining the system dynamics

In modeling system dynamics, this thesis closely follows Rodriguez et al. (2014) with an extension to the dynamics of log volatility processes , namely modeling them as AR(2) process. In this manner to design the system dynamics the following modification of (2.13) is considered:

$$\begin{aligned} \ln V_t &= \ln V_{t-1} + \mu + \sigma_t \epsilon_{1t} \\ \ln \sigma_t &= \alpha + \phi_1(\ln \sigma_{t-1} - \alpha) + \phi_2(\sigma_{t-2} - \alpha) + \tau \epsilon_{2t} \\ \ln r_t &= \gamma + \rho(\ln r_{t-1} - \gamma) + \kappa \epsilon_{3t} \end{aligned} \tag{3.1}$$

with $t = 2, 3, \ldots$ for the log asset values and log interest rates equations, and $t = 3, 4, \ldots$ for log volatility process. The disturbances ϵ_{it}, $i = 1, 2, 3$ are uncorrelated iid standard Gaussian random variables. Further, mean-reverting AR(1) and AR(2) processes of log

interest rates and log volatilities are assumed stationary, that is, the autoregression coefficients are restricted to corresponding stationary regions: $\{\rho \in \mathbb{R} : -1 < \rho < 1\}$ and $\{\phi_1, \phi_2 \in \mathbb{R} : |\phi_2| < 1, \phi_2 - \phi_1 < 1,\ \phi_1 - \phi_2 > -1\}$.

Finally, for the initial conditions the following assumptions hold:

$$\ln V_1 \sim \mathsf{N}(\xi_v, \varpi_v^2), \qquad (\ln \sigma_1, \ln \sigma_2) \sim \mathsf{N}_2(\boldsymbol{\xi_\sigma}, \Sigma_\sigma),$$

$$lnr_1 \sim \mathsf{N}(\xi_r, \varpi_r^2). \tag{3.2}$$

where ξ_v and ϖ_v^2 are set by researcher, the parameters of distribution of $\ln r_1$ are equal to the steady state mean and variance of log interest rate process, that is γ and $\kappa^2/(1-\rho^2)$, and finally the first two elements of log volatilities' process are assumed to follow a bivariate normal distribution with the following mean vector and covariance matrix:

$$\xi_\sigma = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}, \qquad \Sigma_\sigma = \tau^2 \begin{pmatrix} \frac{1-\phi_2}{(1+\phi_2)((1-\ \phi_2)^2-\phi_1^2)} & \frac{\phi_1}{(1+\phi_2)((1-\phi_2)^2-\phi_1^2)} \\ \frac{\phi_1}{(1+\phi_2)((1-\phi_2)^2-\phi_1^2)} & \frac{1-\phi_2}{(1+\phi_2)((1-\phi_2)^2-\phi_1^2)} \end{pmatrix}. \tag{3.3}$$

3.2 Pricing and default probabilities

Recalling that in the context of structural models equity is viewed as a contingent claim on the firm's assets with an exercise price equal to its debt value, the current value of equity becomes the expected discounted value of the payoff of implied option, which is $V_t - K_t(m)$ if $V_t > K_t(m)$ and zero otherwise. Thus, we have:

$$E_t = E_{\mathcal{Q}_t}\left(\exp\left\{-\sum_{s=t+1}^{t+m} r_s\right\} \max\left\{V_t - K_t(m), 0\right\}\right), \tag{3.4}$$

where the expectation is taken with respect to the risk neutral measure $\mathcal{Q}_t$, which is Gaussian distribution with $\tilde{r}_t - \tilde{\sigma}_t^2$ mean and $\tilde{\sigma}_t^2$ variance (Rodsriguez et al., 2014). Rodriguez et al. (2014) apply Taylor approximation to (3.4) and provide the following modification of Black Scholes and Merton (1974) formula for equity pricing:

$$E_t = g_{r_t,K_t(m),\theta}(V_t) = V_t\Phi\left(\frac{\ln V_t - \ln K_t(m) + \tilde{r}_t(m) + \tilde{\sigma}_t^2(m)/2}{\tilde{\sigma}_t(m)}\right)$$
$$- K_t(m)\exp\left\{-\tilde{r}_t(m)\right\}\Phi\left(\frac{\ln V_t - \ln K_t(m) + \tilde{r}_t(m) - \tilde{\sigma}_t^2(m)/2}{\tilde{\sigma}_t(m)}\right) \quad (3.5)$$

where θ denotes the parameter vector, and Φ is the cdf of standard Gaussian distribution. $\tilde{r}_t(m)$ and $\tilde{\sigma}_t(m)$ are the expected interest rate and accumulated volatility over the next m periods and the expectation is taken conditional on the information available at time t:

$$\tilde{r}_t(m) = E\left\{\sum_{s=t+1}^{t+m} r_s \,\middle|\, \mathcal{I}_t\right\} = \sum_{u=1}^{m}\exp\left\{\gamma(1-\rho^u) + \frac{\kappa^2}{2}\frac{1-\rho^{2u}}{1-\rho^2} + \rho^u \ln r_t\right\} \quad (3.6.1)$$

and

$$\tilde{\sigma}_t^2(m) = E\left\{\sum_{s=t+1}^{t+m} \sigma_s^2 \,\middle|\, \mathcal{I}_t\right\} = \sum_{u=1}^{m}\exp\left\{2E(\ln\sigma_{t+u}|\mathcal{I}_t) + 2Var(\sigma_{t+u}|\mathcal{I}_t)\right\}, \quad (3.6.2)$$

where the conditional variances and expectations of $\ln\sigma_{t+u}$ are computed recursively from the initial conditions (3.3) and taking into account the dynamics of log volatilities. Bauwens and Lubrano (1997) provide a modification of BSM option pricing formula with Student t disturbance of asset values and volatilities modeled as asymmetric GARCH(1,1) process, though the closed form formula is not provided, and computation is performed by Monte Carlo approximation. Heston and Nandi (2000) provide a closed form formula in the setting of GARCH specification and and assuming correlation between errors of asset returns and volatilities. As in the structural models by assumption default event is associated with asset values falling below some threshold, at each $t+m$ point in time having as that threshold the debt value $K_t(m)$ maturing at $t+m$, the probability of default will be given by the following formula (Rodriguez et al. (2014) and Duan and Fulop, (2009)

)[3]:

$$\mathsf{Pr}\left(V_{t+m} < K_t(m)\right) = \Phi\left(\frac{\ln K_t(m) - \ln V_t - \mu m}{\tilde{\sigma}_t(m)}\right) \tag{3.7}$$

The formula (2.6) for pricing credit default swaps becomes:

$$s_t = -10000 \ln\left(1 - \frac{E_t + K_t(m)\exp\{-\tilde{r}_t(m)\} - V_t}{K_t(m)\exp\{-\tilde{r}_t(m)\}}\right) \tag{3.8}$$

4 Calibration

In the model calibration method the thesis closely follows Rodriguez et al. (2014) and their main reference Kim et al. (1998). For Monte Carlo Markov Chain (MCMC) sampling algorithms main references are Robert and Casella (2004) and Haario et al. (2005)

4.1 Likelihood functions

Considering the null correlation assumption between the error terms of asset values, interest rates and volatilities processes, the joint likelihood function of the model is given by the product of likelihood functions of the aforementioned processes:

$$\begin{aligned} p\left(\sigma_1, \ldots, \sigma_n, ; v_1, \ldots, v_n; r_1, \ldots, r_n | K_1(m), \ldots, K_n(m), \mu, \alpha, \phi_1, \phi_2\tau, \gamma, \rho, \kappa\right) \propto \\ p(v_1, \ldots, v_n \mid \mu, \sigma_1, \ldots, \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \alpha, \beta, \tau, \gamma, \rho, \kappa) \\ p(r_1, \ldots, r_n \mid r_1, \gamma, \rho, \kappa) p(\sigma_1, \ldots, \sigma_n \mid \alpha, \beta, \tau) \end{aligned}$$

Since asset values are not observed, the crucial step in conducting the model estimation is obtaining likelihood function for the observed data, namely, equity prices. As the dynamics is specified only for the asset values, the likelihood of equity prices is derived by exploiting the well known arguments on the likelihood of the transformed data. Namely, having the transformation $Y = T(X|\theta)$, where Y and X are the n -dimensional data

[3]Appendix 6 contains the derivation the first order derivative of default probability function with regard to volatility in order to see the dependence of the former on the latter

vectors, the likelihood of the transformed data Y is expressed by (Duan, 1994):

$$F(Y;\theta) = F_X(T^{-1}(Y;\theta))|J(T^{-1}(Y;\theta))|$$

where J stands for the Jacobian of the transformation function. Further, according to equation (2.4) in Duan (1994) in case of one-to-one and element-by-element transformation from variable X to Y, that is when the Jacobian of the transformation function is a diagonal matrix, the following formula for the likelihood of transformed (observable) data is derived that is for $y_t = g(x_t;\theta)$:

$$F(Y;\theta) = F_X(g^{-1}(y_1;\theta),\ldots,g^{-1}(y_t;\theta);\theta)\left[\prod_{t=1}^{n}\left|\frac{dg(g^{-1}(y_t;\theta);\theta)}{dx_t}\right|\right]^{-1} \tag{4.1}$$

On the other hand, the likelihood function for asset values, given parameters , interest rates and volatilities, is derived straightforwardly from (3.1) and (3.1) (Rodriguez et al., 2014):

$$\begin{aligned} p(v_1,\ldots,v_n \mid \mu,\sigma_1,\ldots,\sigma_n) = {} & \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\varpi_v}\left(\prod_{t=2}^{n}\frac{1}{\sigma_t}\right)\left(\prod_{t=1}^{n}\frac{1}{v_t}\right) \\ & \exp\left\{-\frac{[\ln v_1 - \xi_v]^2}{2\varpi_v^2}\right\}\exp\left\{-\sum_{t=2}^{n}\frac{[\ln v_t - \ln v_{t-1} - \mu]^2}{2\sigma_t^2}\right\}, \end{aligned} \tag{4.2}$$

Now, substituting in (4.1) $g(x_t;\theta)$ for the BS pricing function (3.5), x_t for the asset values v_t and F_X for the likelihood (4.2), for observed equity prices we will have the following likelihood function:

$$\begin{aligned} & p(e_1,\ldots,e_n \mid \mu,\sigma_1,\ldots,\sigma_n,r_1,\ldots,r_n,K_1(m),\ldots,K_n(m),\alpha,\phi,\tau,\gamma,\rho,\omega) = \\ & \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\varpi_v}\frac{1}{\prod_{t=2}^{n}\sigma_t}\exp\left\{-\sum_{t=2}^{n}\frac{\left[\ln g^{-1}_{r_t,K_t(m),\theta_t}(e_t) - \ln g^{-1}_{r_{t-1},K_{t-1}(m),\theta_{t-1}}(e_{t-1}) - \mu\right]^2}{2\sigma_t^2}\right\} \\ & \exp\left\{-\frac{\left[\ln g^{-1}_{r_1,K_1(m),\theta_1}(e_1) - \xi_v\right]^2}{2\varpi_v^2}\right\}\left[\prod_{t=1}^{n}\frac{1}{g^{-1}_{r_t,K_t(m),\theta_t}(e_t)}\right]\left[\prod_{t=1}^{n}\left.\frac{\partial g_{r_t,K_t(m),\theta_t}}{\partial v_t}\right|_{v_t=g_\theta^{-1}(e_t)}\right]^{-1}, \end{aligned} \tag{4.3}$$

where the inverse function g^{-1} can be computed numerically with Newton-Raphson method and it's derivative with respect to asset prices is (Rodriguez et al., 2014):

$$\left.\frac{\partial g_{r_t,K_t(m),\theta_t}}{\partial v_t}\right|_{v_t=g_\theta^{-1}(e_t)} = \Phi\left(\frac{\ln g^{-1}_{r_t,K_t(m),\theta_t}(e_t) - \ln K_t(m) + \tilde{r}_t(m) + \tilde{\sigma}_t^2(m)/2}{\tilde{\sigma}_t(m)}\right)$$

Further, exploiting Bayes' rule, for interest rates we have:

$$\begin{aligned} p(r_1,\ldots,r_n \mid \gamma,\rho,\kappa) = & \\ & p(r_1 \mid \gamma,\rho,\kappa)p(r_2 \mid r_1,\gamma,\rho,\kappa)\cdot\ldots\cdot p(r_n \mid r_1,\ldots r_{n-1},\gamma,\rho,\kappa) \\ & = \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\kappa^{n-1}\varpi_r}\frac{1}{\prod_{t=1}^n r_t}\exp\left\{-\frac{[\ln r_1-\xi_r]^2}{2\varpi_r^2}\right\} \\ & \qquad \exp\left\{-\sum_{t=2}^n \frac{[\ln r_t-\gamma-\rho(\ln r_{t-1}-\gamma)]^2}{2\kappa^2}\right\}. \end{aligned} \tag{4.4}$$

In the same fashion the likelihood function of log volatilities conditional on the parameters of the model is given as:

$$\begin{aligned} p(\ln\sigma_1,\ldots,\ln\sigma_n \mid \alpha,\phi_1,\phi_2,\tau) = & \left(\frac{1}{\sqrt{2\pi}}\right)^n \mid \Sigma_\sigma \mid^{-1/2} \exp\left\{-\frac{1}{2}(z_2-\xi_\sigma)'\Sigma_\sigma^{-1}(z_2-\xi_\sigma)\right\} \\ & \frac{1}{\tau^{n-2}}\exp\left\{-\sum_{t=3}^n \frac{[\ln\sigma_t-\alpha-\phi_1(\ln\sigma_{t-1}-\alpha)-\phi_2(\ln\sigma_{t-2}-\alpha)]^2}{2\tau^2}\right\}, \end{aligned} \tag{4.5}$$

where $z_2 = (\ln\sigma_1, \ln\sigma_2)'$ is the column vector of the first two elements of log volatilities vector and is distributed according to the initial conditions (3.3)

4.2 On Metropolis Hastings and Gibbs sampling

To conduct the inference on the model two random walk Monte Carlo Markov Chain algorithms,the Gibbs and Metropolis-Hastings sampling methods, have been combined in this thesis, following Rodriguez et al. (2014). Gibbs sampling is used when there is a need to sample a k , $k \geq 2$, dimensional X vector from some form of multivariate distribution.

Now, since in most cases this function is not a commonly known function, direct generation is either impossible or difficult. The idea behind Gibbs sampling is, having some initial values, to generate one component at time, conditional on the remaining component and at each step use the updated values for conditioning. More formally, the algorithm is given in the following way (Robert and Casella, 2004): given that the univariate conditional distributions $f_i(x_i|x_{\setminus i})$ exist and given $x^{(t)} = (x_1^{(t)}, \dots x_k^{(t)})$ generate

1. $X_1^{(t+1)} \sim f_1(x_1|x_2^{(t)} \dots x_k^{(t)})$
2. $X_2^{(t+1)} \sim f_2(x_2|x_1^{(t+1)} x_3^{(t)} \dots x_k^{(t)})$
   $\vdots$
3. $X_k^{(t+1)} \sim f_k(x_k|x_1^{(t+1)} \dots x_{k-1}^{(t+1)})$

Here the vector x represents the parameter vector and functions $f_1 \dots f_k$ the conditional posteriors.

On the other hand, Metropolis-Hastings algorithm, which is in fact a generalization of Gibbs algorithm, is implemented when the conditional posteriors are not of known form and it is impossible to make random draws from these distributions. The main principle here is to generate the random variables from some well known or symmetric proposal distribution $q(.|x)$ which is easy to simulate from . The main requirement here is that the ratio of the density of interest (target density) and the proposal density is known up to some constant, independently of x (Robert and Casella, 2004). Then the algorithm goes as follows: Given $x^{(t)}$

1. Generate x^* from $q(x^*|x^{(t)})$
2. Take

$$X^{(t+1)} = \begin{cases} x^* & \text{with probability} \quad \rho(x^{(t)}, x^*), \\ x^{(t)} & \text{with probability} \quad 1 - \rho(x^{(t)}, x^*) \end{cases}$$

where

$$\rho(x^{(t)}, x^*) = \min \left\{ \frac{f(x^*)}{f(x)} \frac{q(x|x^*)}{q(x^*|x)}, 1 \right\}$$

In the case of model considered in the thesis Metropolis Hastings algorithm has been implemented even when the conditional posterior distributions of the parameters are well known and easy to generate from. The necessity of this approach comes from the fact that the asset values are not directly observable, but instead are obtained from an approximated pricing formula, so in order to make more robust inference, at each sampling nod asset values are updated and parameters are accepted based on the likelihood function of the observable data, that is the equity values.

4.3 Hyperpriors

Following Rodriguez et al. (2014), conditional on the specified prior distributions the initial values of log volatilities, expected interest rates and asset returns, parameters of each of the processes are first sampled from their corresponding full conditional posterior distributions, then after a certain number of Gibbs sampling iterations, having obtained the an initial sample, the Metropolis Hastings algorithm is implemented with proposal densities given by the full conditionals and the acceptance probability equal to the ratio of the likelihood function (4.3) with new and old values of the parameters.

More specifically, beginning with the parameters of the interest rate, the priors of long run mean γ and conditional variance κ^2 are given as independent normal and inverse Gamma : $\gamma, \kappa^2 \propto \mathsf{N}(\xi_\gamma, \varpi_\gamma^2) IG(\frac{a_\kappa}{2}, \frac{b_\kappa}{2})$ with ξ_γ selected to represent an annualized 5% risk free rate and ϖ_γ^2 moderately large and $a_\kappa = 2$ and $b_\kappa = 0.2$ (Rodriguez et al., 2014). The same priors are assumed for the mean and variance of the log volatilities process: $\alpha, \tau^2 \propto \mathsf{N}(\xi_\alpha, \varpi_\alpha^2) IG(\frac{a_\tau}{2}, \frac{b_\tau}{2})$ with with $\xi_\alpha = \ln(0.2/\sqrt{52})$ and $(a_\tau, b_\tau) = (2, 0.4)$. In all of the above mentions cases the posterior distributions have the same functional forms as the priors and the posterior parameters are calculated as given in appendix 6

On the other hand prior distributions for the autoregression coefficients are given as

follows:

$$\pi(\rho) \propto \mathsf{N}_{(-1,1)}(0.95, 1)$$
$$\pi(\phi) \propto \mathsf{N}_S(\xi_\phi, \varpi_\phi^2)$$

where $\phi \equiv (\phi_1, \phi_2)'$, S is the stationary region for AR(2) process and ξ_ϕ and ϖ_ϕ^2 are adjusted for each data set estimation in order to maintain an optimal rate of acceptance. In the case of autoregression coefficients the training sample as well is obtained through Metropolis Hastings algorithm, since in this case the posterior kernels are not a Gaussian (bivariate) distribution, but are given by the following functions (following Chib and Greenberg (1995)):

$$p(\phi|\ln\sigma, \tau^2) \propto \Psi(\phi, \tau^2)\{f_{N_2}(\phi|\hat{m}_\phi, \hat{\Sigma}_\phi)I(\phi \in S)\} \tag{4.6}$$

$$p(\rho|\ln r, \kappa^2) \propto f_{\ln r_1}(\rho, \kappa^2)\{f_N(\phi|\hat{m}_\rho, \hat{\sigma}_\rho)I(\rho \in (-1, 1))\} \tag{4.7}$$

where $\Psi(\phi, \tau^2)$ and $f_{\ln r_1}$ are the distribution of $(\ln\sigma_1, \ln\sigma_2)'$ and $\ln r_1$ respectively and are given in section 3.1 by (3.1) and(3.3), and posterior means and variances of ρ and ϕ are given in appendix 6. To sample from (4.6) and(4.7), one needs to make draws from respective univariate and bivariate normal distributions f_N and f_{N_2} and if the generated values lie within corresponding stationary regions, accept the new values with probabilities

$$\min\left[\frac{\Psi(\phi^*\tau^2)}{\Psi(\phi, \tau^2)}, 1\right] \quad \textit{and}$$

$$\min\left[\frac{f_{\ln r_1}(\rho^*\kappa^2)}{f_{\ln r_1}(\rho, \kappa^2)}, 1\right]$$

respectively.

Finally, for the mean of the asset returns, μ, a normal prior $\mathsf{N}(\xi_\mu, \varpi_\mu^2)$ is assumed with mean $\xi_\mu = 0.02/52$ (roughly, the long-term average of the S&P500 index) and variance

$\varpi_\mu^2 = 0.2^2/52$ (Rodriguez et al., 2014).

4.4 Normal mixture representation of the returns process

This section discusses an algorithm suggested by Kim et al. (1998) for an efficient Monte Carlo Markov Chain sampling for state space model of asset returns and log volatilities. First, consider the following modification of the dynamics of asset returns and volatilities from system (3.1):

$$y_t^* = \ln \sigma_t + \ln \epsilon_{1t}, \quad \ln \sigma_t = \alpha + \phi_1(\ln \sigma_{t-1} - \alpha) + \phi_2(\sigma_{t-2} - \alpha) + \tau\epsilon_{2t}. \tag{4.8}$$

where $y_t^* = \ln(\ln(V_t/V_{t-1}) - \mu)$ is the the logarithm of mean corrected return at time t. Although this linearization of the measurement equation in (4.8) makes it possible to implement Kalman filtering for the sampling of log volatilities, error terms in the measurement equation are not distributed according to Gaussian distribution , which means that Kalman filter and the corresponding smoothing algorithms will not be optimal so one should work with an approximate filtering algorithm which may produce inconsistent parameter estimates (Shephard, 1994).

This problem of Gibbs sampling from non Gaussian state space models and specifically for stochastic volatility estimation has been discussed, among other authors, by (Shephard, 1994), where the authors discuss approaches tackling other types of non standard distributions of errors,Carter and Kohn (1994), where errors are designed as normal mixtures, Kim et al. (1998), where a detailed algorithm for stochastic volatility is presented. The basic idea behind the algorithms presented in the papers mentioned is to decompose the errors in such a way that it is straightforward to implement Gibbs sampling on the parameters of the decomposition and conditional on those parameters have Gaussian distribution for the errors.

In this manner, following Kim et al., (1998) the log error terms in measurement equation in (4.6) are approximated by mixture of seven Gaussian variables, with the

weights and parameters of the components given as follows[4]:

Table 1: Parameters of normal mixture approximation

k	Pr(k=i)	m_i	$\sqrt{s_i}$
1	0.00730	-5.7002	1.4490
2	0.00002	-4.9186	1.2949
3	0.10556	-2.6216	0.6534
4	0.25750	-1.1793	0.3157
5	0.34001	-0.3255	0.1600
6	0.24566	0.2624	0.0851
7	0.04395	0.7537	0.0418

4.5 Sampling log volatilities

Having normal mixture representation for the errors of measurement equation in (4.8), conditional on the parameters of the model, it is possible to implement linear Gaussian Kalman filter to sample log volatilities. For this purpose as a preliminary step, below (4.6) is presented in vector form:

$$y_t^* = \boldsymbol{x}\boldsymbol{z_t} + s_{1t} \quad \boldsymbol{z_t} = \boldsymbol{\alpha} + \boldsymbol{\Phi}(\boldsymbol{z_{t-1}} - \boldsymbol{\alpha}) + \boldsymbol{v_t} \tag{4.9}$$

where $\boldsymbol{z_t} \equiv (\ln\sigma_t, \ln\sigma_{t-1})', t = 2, \ldots,$ $\boldsymbol{x} = (1,0)$,$\boldsymbol{\alpha} = (\alpha,\alpha)'$, $\boldsymbol{v_t} = (\tau\epsilon_{2t}, 0)'$, $\boldsymbol{\Phi} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix}$, and s_{1t} is the mixture variable. Within this setting, conditional on the mixture parameters a particular Kalman filtering and smoothing algorithm can be implemented as detailed in Kim and Nelson (1999), chapter 8. Having the initial conditions as $z_{0|0} = \boldsymbol{\alpha}$ and $P_{0|0} = \Sigma_\sigma$, to guarantee stationarity, Kalman filter first runs recursively forward in time over the following equations (Mumtaz and Blake, 2012)

[4]The figures are obtained from Kim et al. (1998), adjusting for the fact that they model not log volatilities but log variances

$$
\begin{aligned}
z_{t|t-1} &= \boldsymbol{\alpha} + \Phi(z_{t-1|t-1} - \boldsymbol{\alpha}) \\
P_{t|t-1} &= \Phi P_{t-1|t-1}\Phi' + Q \\
\eta_{t|t-1} &= y_t^* - \boldsymbol{x} z_{t|t-1} - m_i \\
f_{t|t-1} &= \boldsymbol{x} P_{t|t-1} \boldsymbol{x}' + s_i \\
z_{t|t} &= z_{t|t-1} + K\eta_{t|t-1} \\
P_{t|t} &= P_{t|t-1} - K\boldsymbol{x}P_{t|t-1}
\end{aligned} \tag{4.10}
$$

where $Q = \begin{pmatrix} \tau^2 & 0 \\ 0 & 0 \end{pmatrix}$ is the covariance matrix of $\boldsymbol{v_t}$ and $K = P_{t|t-1}\boldsymbol{x}' f_{t|t-1}^{-1}$

After the last recursion $z_T = (\ln \sigma_{n+1}, \ln \sigma_n)'$ is generated from bivariate normal distribution with mean $z_{T|T}$ and variance $P_{T|T}$.

Next, the smoothing is performed through the following backward recursive equations, to collect the remaining elements of the vector $\boldsymbol{z}$ from bivariate normal distribution with the following means and covariance matrices:

$$
\begin{aligned}
z_{t|t,z_{t+1}^*} &= z_{t|t} + K^*(z_{t+1}^* - \boldsymbol{\alpha} - \Phi^* z_{t|t}) \\
P_{t|tz_{t+1}^*} &= P_{t|t} - K^*\Phi^* P_{t|t}
\end{aligned} \tag{4.11}
$$

where Φ^* is the first row of the matrix Φ and $K^* = P_{t|t}\Phi^{*'}(\Phi^* P_{t|t}\Phi^{*'} + \tau^2)^{-1}$. This slight modification of the usual Kalman smoothing equations is due to the fact that in (4.9) the second row of state equation is identity and the covariance matrix of the error terms $\boldsymbol{v_t}$ is singular, hence if we condition on the entire vector of the state variable generated at the previous step, the smoothing updating covariances would be singular , rendering generation of z_t using (4.11) impossible (Kim and Nelson, 1999). Consequently, only the first element of z_{t+1} should be used as a conditioning factor. At the final step only the first element of the generated $z_t, t = T, T-1, \ldots, 1$, is kept for inference.

After obtaining a sufficiently large sample of log volatilities through Kalman filter, for the remaining iterations of Gibbs a block wise random walk Metropolis Hastings algorithm is applied for sampling log volatilities within which at each step a block of the vector of log volatilities is updated through a normal distribution with mean equal to previously generated values of the elements of the corresponding block and proposal variance for all the blocks is selected so as to maintain an optimal acceptance rate (above 10 % and below 40 %). The newly generated values are accepted with the probability $\min\{\Delta_z, 1\}$ where

$$\Delta_z = \frac{p(\ln\sigma_1, \ldots, \ln\sigma_{t-1} \ln\sigma_t^*, \ldots, \ln\sigma_{t+i}^* \ldots \ln\sigma_n \mid \alpha, \phi_1, \phi_2, \tau)}{p(\ln\sigma_1, \ldots, \ln\sigma_{t-1} \ln\sigma_t, \ldots, \ln\sigma_{t+i} \ldots \ln\sigma_n \mid \alpha, \phi_1, \phi_2, \tau)} \frac{p(e_1, \ldots, e_n \mid, \sigma_1, \ldots, \sigma_{t-1}\sigma_t^*, \ldots, \sigma_{t+i}^* \ldots \sigma_n, \boldsymbol{r}, \boldsymbol{K}, \Theta)}{p(e_1, \ldots, e_n \mid, \sigma_1, \ldots, \sigma_{t-1}\sigma_t, \ldots, \sigma_{t+i} \ldots \sigma_n, \boldsymbol{r}, \boldsymbol{K}, \Theta)}$$

where $\boldsymbol{r}$, $\boldsymbol{K}$, and Θ are the vectors of interest rates, debt values and parameters of the model respectively. The full details of MCMC algorithm is given in the appendix **??**

5 Empirical application

The application of the model has been carried out on three companies of which one from industrial sector (Ford Motor Company) and two from financial sector (Citigroup and Lehman Brothers). As in the case of Rodriguez et al. (2014), the data frequency is one week and the prediction is carried out for default probabilities and CDS spreads associated with one year debt. The data regarding debt value has been obtained from the balance sheets of the companies and was calculated as the sum of short term and half of the long term liabilities. The one day London Interbank Offering Rates (LIBOR) have been used as risk free interest rates. Finally, 20 000 samples have been generated through MCMC algorithm with the number of burn in samples equal to 6 000.

5.1 The role of deposits in default prediction

In addition to expanding the model of Rodriguez et al. (2014) to AR (2) process, this thesis considers a different debt value by including customer deposits in case of the two banks considered (Citigroup and Lehman Brothers). To provide a preliminary evidence on the

importance of including the deposits, a replication of the base model has been performed for the two banks with and without including the deposits. In the relevant literature it isn't explicitly stated whether customer deposits should be included, but rather short term debt, that is not only short term borrowings, is considered [5]. While the rationale behind considering only short term borrowings is that in that category are included the bonds with maturity less than one year and the payable interest on the bonds of all maturities and the model considers the default of the company from the perspective of meeting its obligations to bondholders, including deposits is justifiable given that we are interested in estimating the probability of default from the perspective of the company meeting its short term obligations in general not only with regard to the bonds. This is particularly important, since the results regarding prediction of default probabilities change when different values for the barrier are considered. In case of Citigroup including deposits in the default barrier augments them in average by 3.45 times, that coefficient being around two at the initial points of the data set and growing as the deposits grow.

On the other hand, in case of Lehman Brothers inclusion of the deposits don't seem to have a considerable impact on the estimated quantities. This can bes explained by the fact that deposits constitute a small proportion of the company's liabilities also by already high leverage ratio (debt to equity ratio in case of Lehman Brothers in average is 13 versus 3 in case of Citigroup). Figure 1 and 2 illustrate estimation output for Citigroup and Lehman Brothers on weakly data for the time period of May 2003 - August 2008.

Furthermore, the estimated one-year CDS spreads are better calibrated to the observed market prices of the spreads with the same maturity when deposits are included. An implication of this is that, from the investment making perspective, failing to consider the deposits in the analysis would provide misleading information on whether the company's riskiness is under or over estimated by the market and to which extent. It can be concluded thus that to make a proper analysis on the riskiness of the banks it is necessary

[5] In the paper under study the authors give reference to Gray et al. (2008), where the authors refer to Crouhy et al. (2000) and KMV (1999).In Crouhy et al. (2000) the "default point" is set at par value of current liabilities plus half of the long term debt, referring to KMV model (1999), where again short term liabilities and not only borrowings are considered.

to include deposits in companies' liabilities

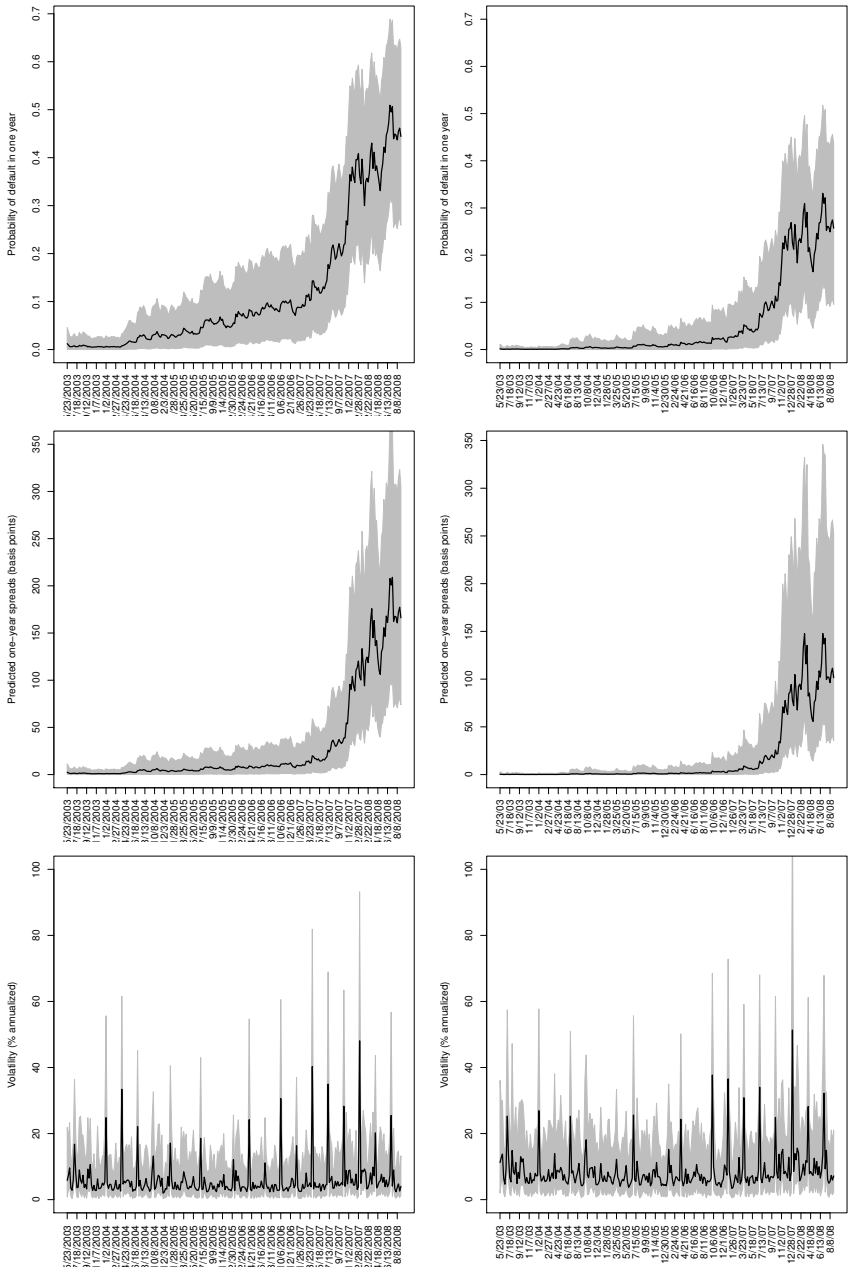

Figure 1: Estimated probability of default (first row), one year CDS spreads predicted (second row) and estimated annual volatility (third row) of Citigroup with (left) and without (right) deposits included in the short term debt

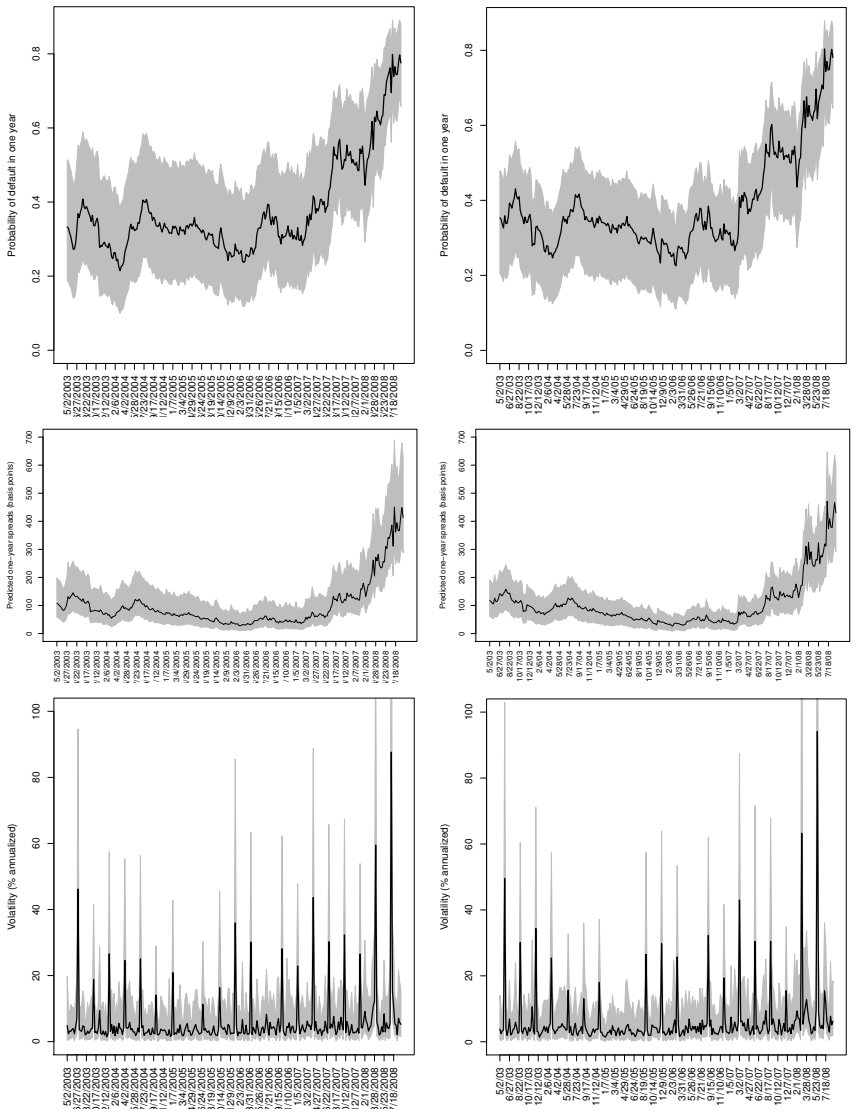

Figure 2: Estimated probability of default (first row), one year CDS spreads predicted (second row) and estimated annual volatility (last row) of Lehman Brothers with (left) and without (right) deposits included in the short term debt

5.2 Model application on real financial data

This section discusses the results of model application on financial data obtained for two banks, Citigroup and Lehman Brothers and one firm, Ford Motor Company, and compares results for three models: the SV AR (2) model, the SV AR(1) model and White Noise SV

model. The data covers the period from May 2003 to August 2008 and thus includes the financial crisis of 2007-2008.

The observed data on equity, debt value and one year credit default spread prices (in basis points) are presented in figures 3 and 4. At first glance a noteworthy difference is revealed in the dynamics of the data sets of Ford Motor Company and Citigroup. First, there is an apparent distinction in the stability of equity prices of these two companies with those of Ford being more volatile having two downward jumps in 2005, the same period as when the company's bonds' ratings cut to junk status(Schneider,2005), and in late 2007, early 2008, during the financial crisis, whereas in the case of Citigroup equity prices fall sharply only in the latter period. There are noticeable differences also in the trends of the default barriers, with Ford's debt values in fact having overall no upward or downward trend and on the contrary, Citigroup's liabilities growing until 2007, which again coincides with the financial crisis.

Figure 3 also displays the one year CDS spreads which, being an indicator of companies' overall riskiness, in case of both firms jump up at avout the period of the crisis. Figure 4 displays the financial data on Lehman Brothers. Although the shapes of the liabilities' dynamics resembles that of the Citigroup, there is a significant difference in the levels. Lehman Brothers has been considered as a badly lead company for the period prior to the financial crisis, this is reflected in the values of CDS spreads, which reach the level of 700 basis points in May 2008, four months before its bankruptcy was announced. CDS spreads have a sudden upward shift during the crisis also in case of Citigroup. It should be remembered however that, as has been mentioned earlier, CDS spreads reflect and are affected by not only individual credit risk of companies, but also depend on market circumstances, such as contagion effects, overall instability of the market (particularly, the financial sector during the crisis).

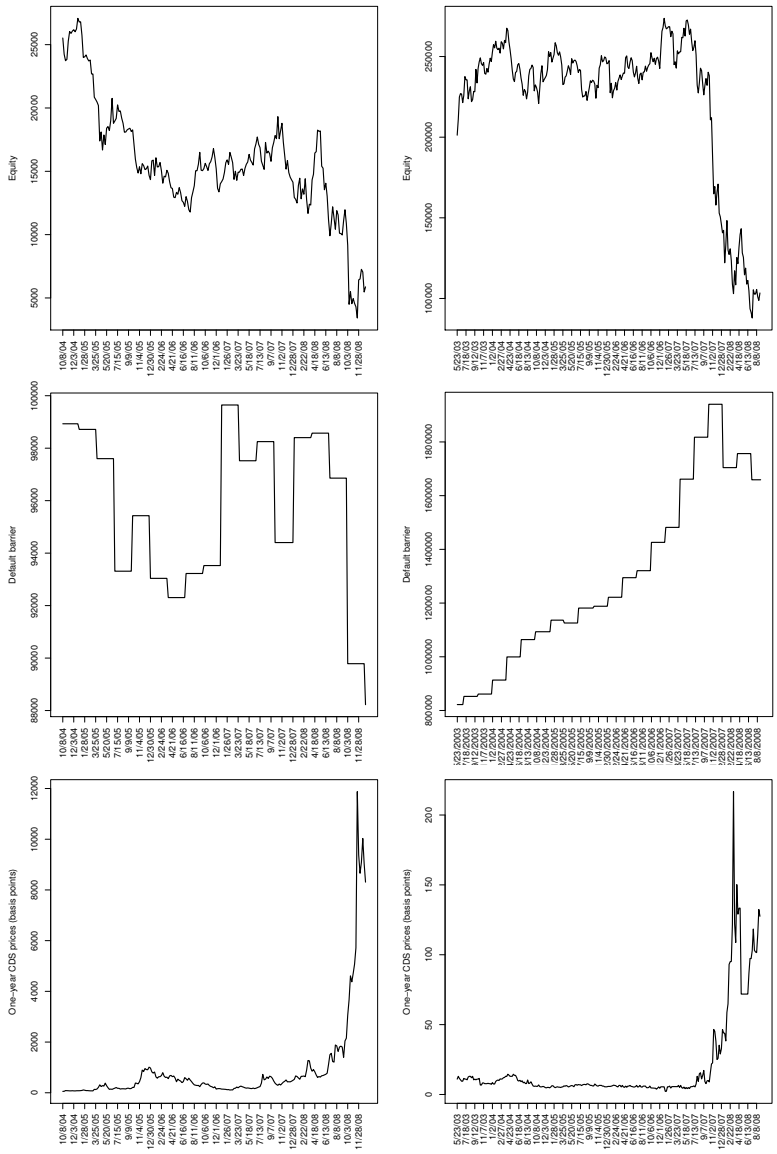

Figure 3: Equity prices (first row), promised payments (second row) and and one-year market CDS spreads (last row) for Ford (left column) and Citigroup (right column).

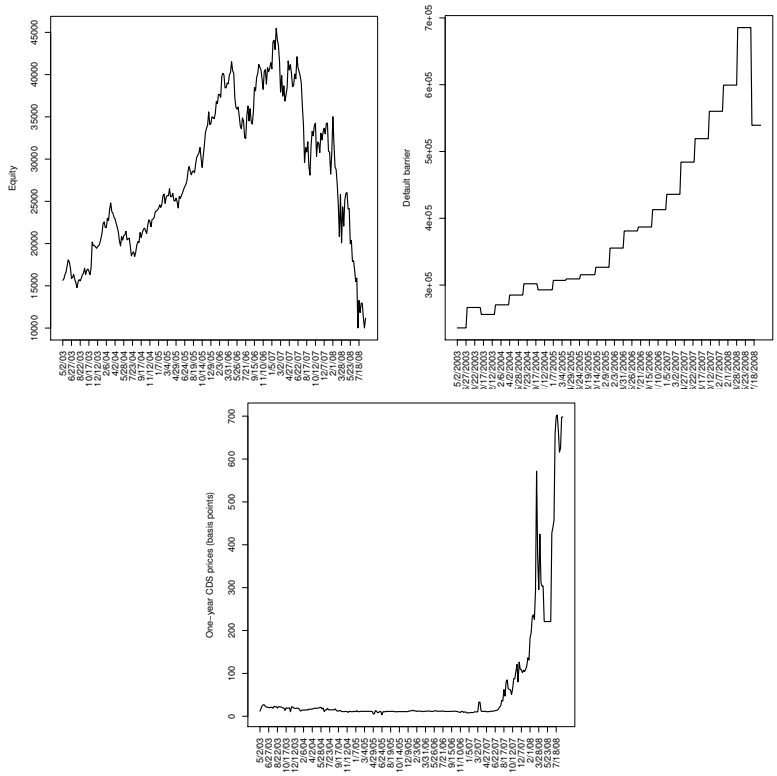

Figure 4: Equity prices (first row), promised payments (second row) and and one-year market CDS prices (third row) for Lehman Brothers .

Further, figures 5 and 6 display the outputs of estimation of three models (SV-AR(2), SV-AR(1), SV-WN) performed on the financial data of Citigroup and Ford Motor Company and Lehman Brothers. First, considering the estimated asset values trajectories it can be observed that all three models produce almost identical output for both Lehman Brothers and Citigroup, with a more considerable difference between model implied values at the end of 2007 an in 2008 in case of Citigroup. On the other hand, in case of Ford, the three models deliver somewhat less similar values. Considering the fact asset values are not directly observable but calibrated from observed values of equity and debt value,one possible explanation for this phenomena is that in case of the two banks the

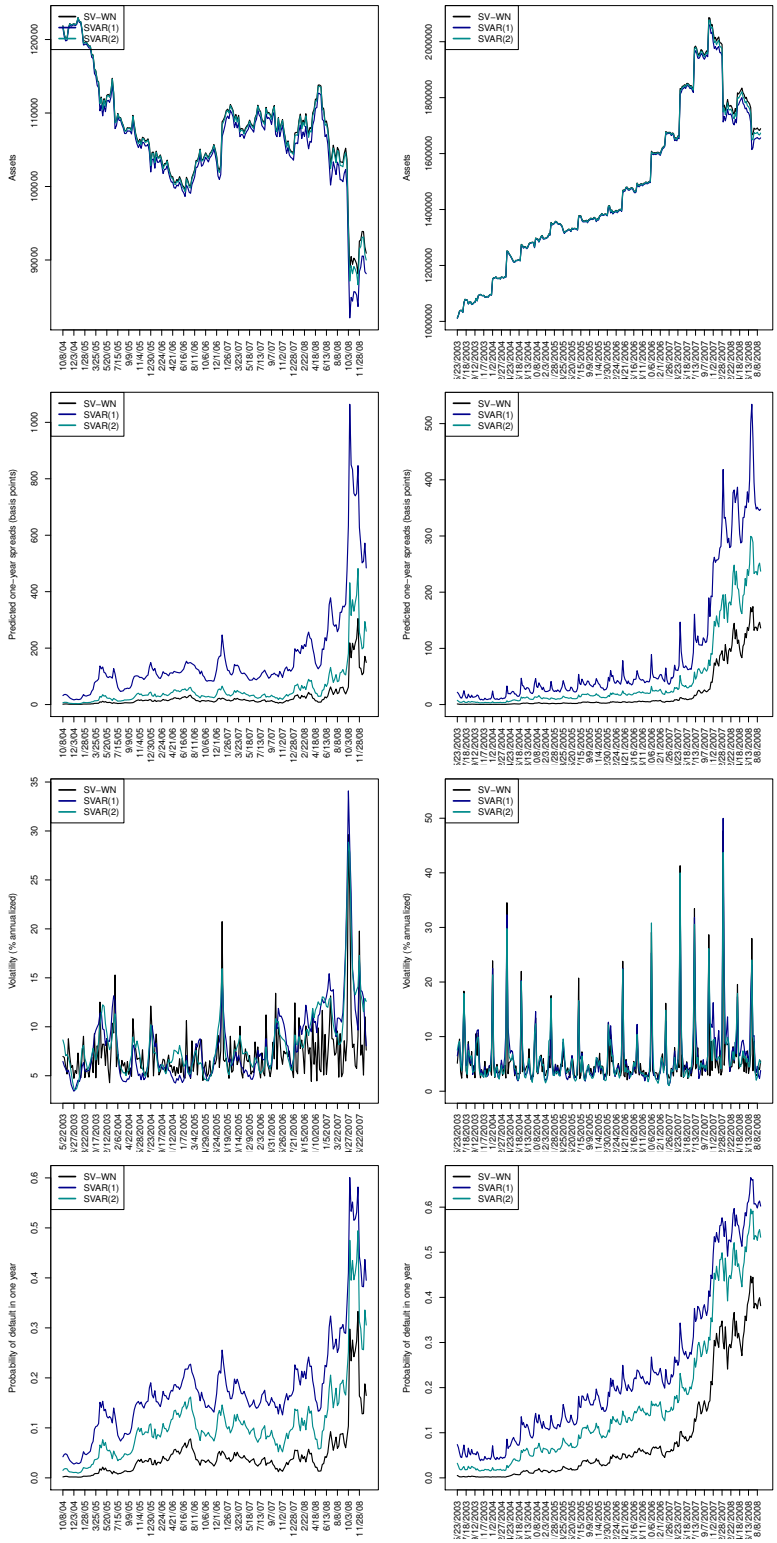

Figure 5: Estimated asset values, 1-year CDS spreads, volatilities and default probabilities for Ford (left) and Citigroup (right) .

retrieved balance sheet values of debt are more accurate[6], that is less contaminated by noise or delayed accounting reports. An observation that speaks in favor of this assump-

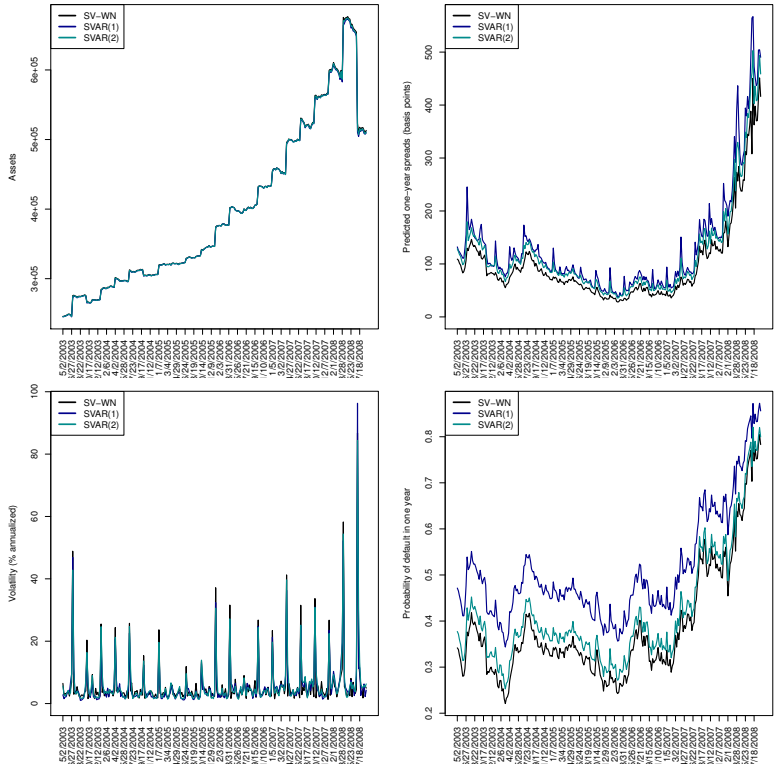

Figure 6: Estimated asset values, 1-year CDS spreads, volatilities and default probabilities for Lehman Brothers.

tion is that at the time of prior to its default, Lehman Brother's balance sheet entries were scrutinized in detail and thus even in the period of high overall instability of the market observed accounting records were less noisy, so model implied values of assets are much closer to each other than in the case of Citigroup.

Furthermore, in the case of CDS spreads all three models have similar shapes but they differ dramatically in levels in case of Ford and Citigroup. In the former case SV-AR (1) predicts spreads that are closer to those observed in the market, while in the latter,

[6]In particular, Previts (2002) mentions the lack of disclosure regarding M&A activities of Ford.

SV-AR(2) seems to have a better fit. Next, predicted default probabilities follow the same path as the CDS spreads and interestingly enough, all three model predict close to 80 % default probability for Lehman Brothers at the period close to its bankruptcy. As before, discrepancy between model provided predictions is larger in case of Ford. And lastly as can

Table 2: Posterior means and 90 % credible intervals in square brackets

Lehman Brothers			
	SV-AR(1)	SV-AR(2)	SV-WN
ϕ_1	0.4597	0.568	-
	[0.4076 ,0.5214]	[0.414, 0.720]	
ϕ_2	-	-0.204	-
		[-0.3510, -0.0470]	
α	-5.5896	-5.444	-5.4869
	[-5.7776, -5.3132]	[-5.700, -5.207]	[-5.626, -5.348]
τ^2	0.8016	0.69	0.9302
	[0.6261, 0.9363]	[0.538, 0.867]	[0.7953, 1.0719]
γ	-8.04322	-8.062	-8.083
	[-8.6294 ,-7.5840]	[-8.705, -7.546]	[-8.7758, -7.5352]
ρ	0.9932	0.994	0.9944
	[0.9877, 0.9979]	[0.989, 0.998]	[0.9889, 0.9987]
κ^2	0.00181	0.0018	0.0018
	[0.0016 0.0021]	[0.0016,0.0021]	[0.00157, 0.0021]
μ	0.00065	0.000287	0.000202
	[0.0002,0.0011]	[-0.0001 0.0007]	[-0.0001, 0.0005]

be expected, considering a higher complexity of asset structure and more interconnected nature of banking system, the volatilities of the two banks have more frequent and higher jumps than that of Ford.And, as in the case of asset values, the three models produce closer estimates in case of the banks. This goes in hand in hand with the posterior estimates of the volatility of volatility which is significantly higher for Citigroup and Lehman Brothers compared to Ford, for all three models (see tables 2 and 3)

Tables 2,3 and 4 contain the posterior means and corresponding 90 % credible intervals for the parameters of the three models considered. The first lag autoregression

Table 3: Posterior means and 90 % credible intervals in square brackets

	Citigroup		
	SV-AR(1)	SV-AR(2)	SV-WN
ϕ_1	0.55785	0.64422	-
	[0.4752, 0.6318]	[0.4589, 0.8239]	
ϕ_2	-	-0.2415	-
		[-0.4581, -0.0249]	
α	-5.1558	-5.1901	-5.26054
	[-5.3817, -4.9413]	[-5.3710,-5.0126]	[-5.3853, -5.1472]
τ^2	0.63002	0.60277	0.7528
	[0.4548,0.8001]	[0.3820, 0.8611]	[0.6171, 0.9033]
γ	-7.85171	-7.83263	-7.82241
	[-8.7741, -7.0858]	[-8.7359, -7.0542]	[-8.7928, -6.9807]
ρ	0.99581	0.99573	0.99601
	[0.9905, 0.9994]	[0.9904, 0.9995]	[0.9906, 0.9995]
κ^2	0.001802	0.001802	0.0018
	[0.0016, 0.0021]	[0.0016, 0.0021]	[0.0016, 0.0021]
μ	0.000572	0.000348	-0.00012
	[0.0001,0.0011]	[-0.0002,0.0008]	[-0.0007, 0.0004]

coefficient is much higher in the case of Ford, which coincides with the trajectory of log volatilities, which happens to have less jumps than the other two firms. The parameter of interest, second lag coefficient of log volatility process is negative in case of all three firms and is insignificant in case of Ford. Corresponding to the downward trend of asset values, long run mean return rate is negative in case of Ford. Overall, the three models have closer values of posterior means for the parameters of interest rate process, that being explained by the fact that interest rates are observable and hence model outputs should be more consistent.

Table 4: Posterior means and 90 % credible intervals in square brackets

	Ford		
	SV-AR(1)	SV-AR(2)	SV-WN
ϕ_1	0.79749	0.85446	-
	[0.7222, 0.8534]	[0.7219, 0.9842]	
ϕ_2	-	-0.074596	-
		[-0.2669, 0.0867]	
α	-4.6638	-4.5279	-4.78542
	[-4.8674,-4.3902]	[-4.7898,-4.2842]	[-4.9215, -4.6382]
τ^2	0.09714	0.0922	0.34519
	[0.0620, 0.1546]	[0.0559, 0.1732]	[0.2409, 0.4999]
γ	-7.6383	-7.60164	-7.58676
	[-8.4276, -7.1054]	[-8.3980, -6.9933]	[-8.3486, -6.9975]
ρ	0.99264	0.9934	0.99328
	[0.9835, 0.9990]	[0.9849, 0.9993]	[0.9845, 0.9992]
κ^2	0.00174	0.001741	0.001742
	[0.0015, 0.0020]	[0.0015, 0.0020]	[0.0015, 0.0020]
μ	-0.00071	-0.000661	-0.0011
	[-0.0017, 0.0002]	[-0.0017, 0.0004]	[-0.0021, -0.0002]

Finally, figures 7 and 8 display the prior and posterior distributions of long run means of asset returns μ, and logarithm volatilities, α , autoregression coefficients ϕ_1, ϕ_2 and variance of the volatility τ^2. It can be inferred form the posterior plots that estimation appears to be quite robust, in that the posterior means are shifted away from the prior ones, moreover, the posterior variances are significantly smaller than those specified in prior distributions. The prior and posterior distributions for the SV-AR(1) and SV-WN models are reported in appendix 6

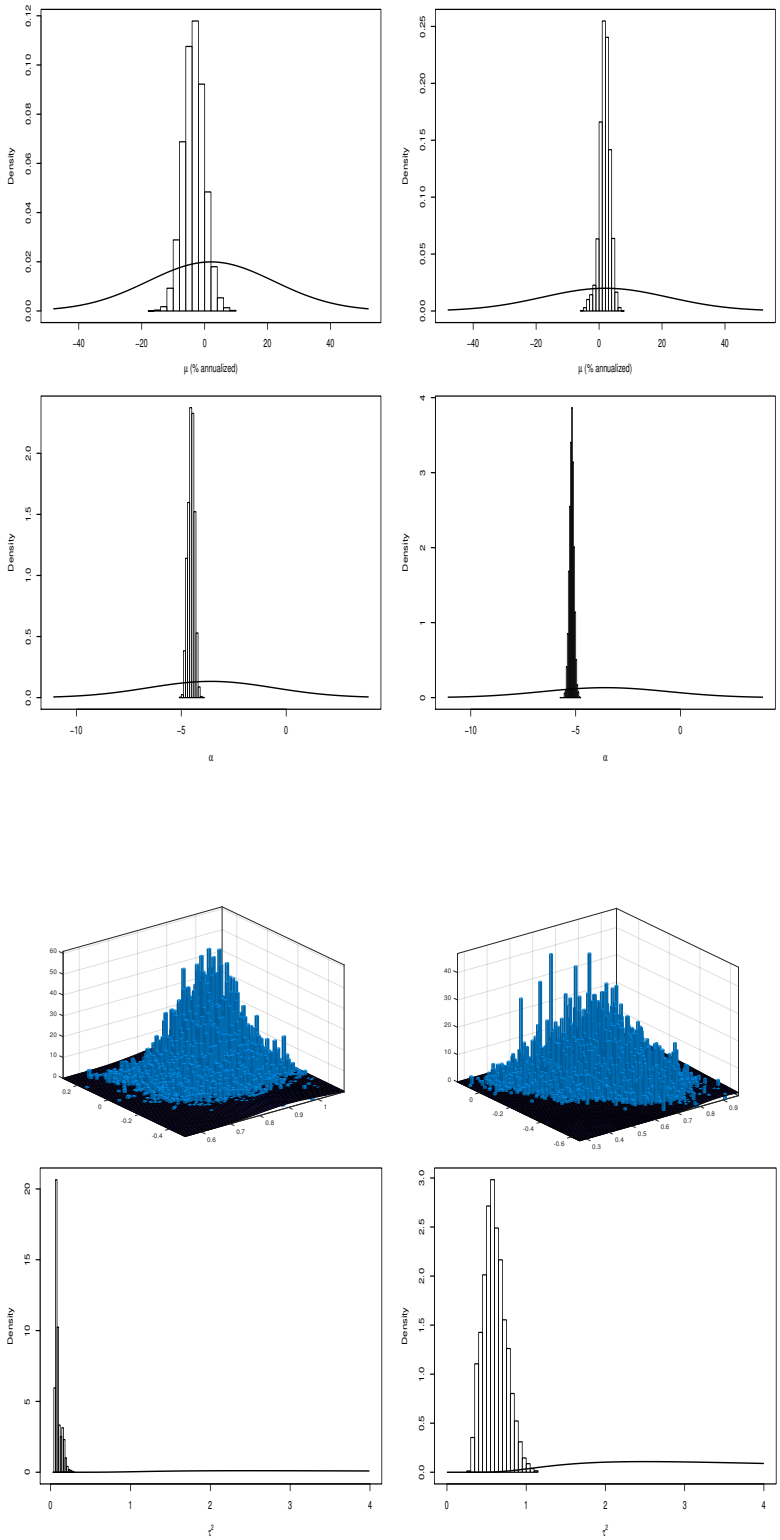

Figure 7: Prior (continuous lines) and posterior distributions (histograms) for the mean expected returns μ (first row) and the parameters for the volatility process α (second row), β (third row) and τ^2 (last row) for Ford (left column) and Citigroup (right column).

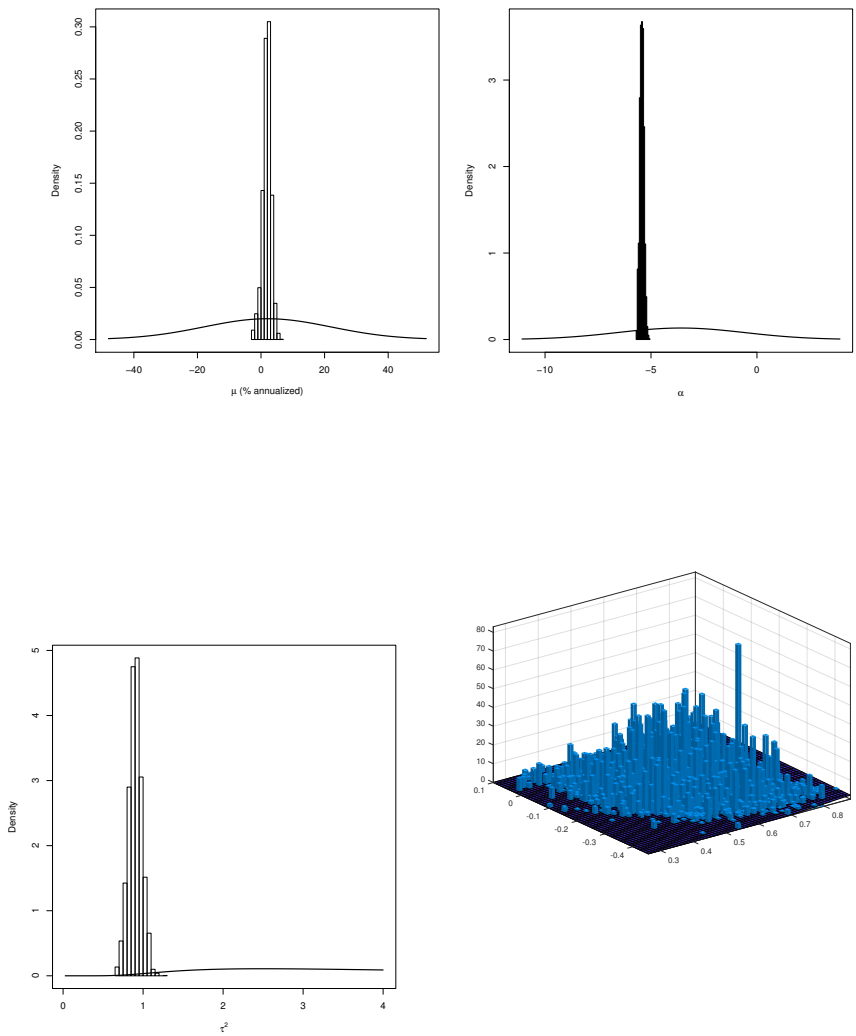

Figure 8: Prior (continuous lines) and posterior distributions (histograms) for the mean expected returns μ (first row) and the parameters for the volatility process α (second row), β (third row) and τ^2 (last row) for Lehman Brothers.

In conclusion it can be stated that the estimation output and model performance depends on the data that the study is carried on. First, as could be expected, the more the available information is closer to the true indicators of firms, the more homogeneous are the results among the three main models considered. Second, which model calibrates to the market data in terms of CDS spreads seems to depend on the sector a firm belongs to. More explicitly, in case of the two banks SV-AR (2) gives a better fit to the CDS spreads than SV-AR (1) which is the contrary in case of the one firm from industrial sector. This could be explained by more interconnected character of transactions prevalent in financial

sector that induce more persistent shocks.

5.3 Bayes factors

This section conducts model comparison based on an MCMC approximation to marginal (integrated) log likelihoods. This approach of using integrated rather maximized likelihoods, also referred to as Bayesian model comparison, was introduced by Jeffrey (1961). Model implied density of data, $P(D|M)$ is calculated integrating over the parameter space (Kass, 1995):

$$P(D|M) = \int P(D|M,\theta)\pi(\theta|H)d\theta$$

where θ ia the parameter set.

Comparing to competing models, say M_1 and M_2 the evidence in favor of the first model against the second is summarized in Bayes factor, the ratio of the posterior odds of the data:

$$BF_{12} = \frac{P(D|M_1)}{P(D|M_2)} \tag{5.1}$$

Jacquier et al. (2003) show that Bayes factor in (5.1) can be written as the expectation of the ratio of full posteriors:

$$BF_{12} = \frac{P(D|M_1)}{P(D|M_2)} = E\Big[\frac{\pi(\theta|M_1)P(D|M_1,\theta)}{\pi(\theta|\theta^*,M_2)P(D|M_2,\theta,\theta^*)}\Big] \tag{5.2}$$

Where the parameter space of M_2 is given by $\theta, \theta*$ and the expectation is taken over the joint posterior of $\{\theta,\theta^*\}$. Equation (5.2), following Jacquier et al. (2003) , is derived first applying Bayes rule to get that $P(D|M_2) = P(\theta,\theta^*|M2)\pi(\theta,\theta^*|M_2/P(\theta,\theta^*|D,M_2)$. By exploiting thus identity it the follows:

$$\begin{aligned} BF_{1|2} = \frac{\int \pi(\theta|M_1)P(D|M_1,\theta)d\theta}{D|M_2} &= \int_{\theta^*}\int_{\theta} \frac{\pi(\theta|M_1)P(D|\theta,M_1)}{P(D|M_2)} d\theta\pi(\theta^*|M_2)d\theta^* \\ &= \int_{\theta^*}\int_{\theta} \frac{\pi(\theta|M_1)P(D|\theta,M_1)P(\theta,\theta^*|D,M_2)\pi(\theta^*)|M_2}{P(D|\theta,\theta^*)\pi(\theta,\theta^*|M_2)} d\theta d\theta^* \\ &= \int_{\theta^*}\int_{\theta} \frac{\pi(\theta|M_1)P(D|\theta,M_1)}{P(D|,\theta,\theta^*,M_2)\pi(\theta|\theta^*,M2)} P(\theta,\theta^*|D)d\theta d\theta^*. \end{aligned} \tag{5.3}$$

Now, considering the independent priors assumption and the fact that for all the com-

mon parameters the models assume the same prior specifications, $\pi(\theta|\theta^*, M_2) = \pi(\theta|M_2) = \pi(\theta)$.The expectation of the remaining ratio, $E[P(D|\theta, M_1)/P(D|\theta, \theta^*, M_2)]$ can be approximated over MCMC draws (Jacquier et al., 2003). Considering computational conveniences, in this thesis it has been chosen to work with the harmonic mean of the ratio of Gibbs outputs of likelihood functions for each model[7]:

$$\hat{E}\left[\frac{P(D|M_1, \theta)}{P(D|M_2, \theta, \theta^*)}\right] = \left[\frac{1}{G}\sum_{1}^{G}\left\{\frac{P(D|\theta^{(i)})}{P(D|\theta^{(i)}, \theta^{*(i)})}\right\}^{-1}\right]^{-1}$$

where G is the number of Gibbs samples and $\{\theta^{(i)}, \theta^{*(i)}\}$ are draws from the conditional posteriors at sampling step i. Table 5 summarizes the comparison results for the three firms:

Table 5: Bayes Factors

	Lehman Brothers	Citigroup	Ford
$BF_{SV-AR(1)\|SV-AR(2)}$	$1.34 \cdot 10^{-19}$	$3.75 \cdot 10^{-19}$	$2.82 \cdot 10^{-12}$
$BF_{SV-WN\|SV-AR(2)}$	$2.60 \cdot 10^{-15}$	$5.72 \cdot 10^{-13}$	$5.82 \cdot 10^{-11}$
$BF_{SV-WN\|SV-AR(1)}$	$5.22 \cdot 10^{-18}$	$2.97 \cdot 10^{-13}$	$1.60 \cdot 10^{-11}$

The figures in the table suggest an overwhelming evidence in favor of SV-AR(2) versus SV-WN and SV-AR(1) , as well as SV-AR(1) vs SV-WN, taking into account the following scale suggested by Kass and Raftery (1995), based on an earlier scale given in Jeffreys (1961,app. B):

[7]For the purpose of model comparison, Rodriguez et al. (2014) use the harmonic mean of the log likelihoods from Gibbs output. While they use as an approximation to the integrated log likelihood, the exact formula requires incorporation of prior and posterior densities of the parameters. In particular, Chib (1995) gives the following formula for estimating marginal likelihood on logarithm scale:

$$\ln \hat{P}(D) = \ln P(D|\theta_*) + \ln \pi(\theta_*) - \ln \hat{P}(\theta_*|D)$$

Where the first term on the right hand side is just the log likelihood of the data computed for some point θ_*, which Chib (1995) suggests to take the posterior means as, the second term is the logarithm of the prior densities and the last term is the logarithm of the posterior densities which can be approximated by the conditional posteriors at each sampling nods. This approach has not been implemented here, since the exact formula for integrated log likelihoods would require computation of posterior densities, which produces infinite values in case of some parameters.

Table 6

Kass-Raftery-Jeffreys scale for model comparison

$2 \ln B_{1\|2}$	$B_{1\|2}$	Conclusion
0 to 2	1 to 3	Barely mentionable
2 to 6	3 to 20	Positive
6 to 10	20 to 150	Strong
> 10	> 150	Very strong

In order to check whether model ordering changes depending on the period of high and low volatility, Bayes factors have been computed for three equal time intervals, thus the last interval starting from the period just before the crash of 2007. The results are summarized in table 7. An interesting result is the one regarding the Bayes factor of comparing SV-AR(1) model against SV-AR(2) in case of Ford Motor Company. Even though on the entire sample the second model is dominant, according to figures in table 5, on the third interval the dynamic Bayes factor speaks in favor of SV-AR(1) model, even though th dominance is not so dramatic. Now, looking at the graph of volatilities of Ford one can see that there one significantly big upward jump at a single point in time with volatilities prior to that period being relatively noisy. This means that volatility in this period is less persistent, which coincides with the dominance of SV-AR(1) over SV-AR(2) at that period.

Table 7: Dynamic Bayes factors

	Lehman Brothers	Ford	Citigroup
$B_{1\|2} - I$	$2.40 \cdot 10^{-9}$	0.02	$1.39 \cdot 10^{-8}$
$B_{1\|2} - II$	$1.88 \cdot 10^{-109}$	2.74	$8.41 \cdot 10^{-9}$
$B_{1\|2} - III$	0	9.15	0
$B_{0\|2} - I$	$2.86 \cdot 10^{-7}$	$5.59 \cdot 10^{-10}$	$6.01 \cdot 10^{-6}$
$B_{0\|2} - II$	$1.97 \cdot 10^{-100}$	$2.55 \cdot 10^{-7}$	$4.42 \cdot 10^{-7}$
$B_{0\|2} - III$	0	$7.09 \cdot 10^{-7}$	0
$B_{0\|1} - I$	$1.23 \cdot 10^{-8}$	$4.8 \cdot 10^{-15}$	$2.07 \cdot 10^{-9}$
$B_{0\|1} - II$	$2.43 \cdot 10^{-67}$	$7.23 \cdot 10^{-13}$	$1.54 \cdot 10^{-8}$
$B_{0\|1} - III$	0	$1.73 \cdot 10^{-12}$	$3.84 \cdot 10^{-113}$

On the other hand , in case of the two banks there are multiple frequent jumps possibly meaning that a high volatility in one period caused higher values of volatilities for the next and more periods. Again, an implication which coincides with the consistent results of tables 5 and 7 for those two banks.

These results are in accord with the observations made in the previous section, namely, that the models' performances depend on the underling data set both in terms of the period the data belongs to (high/low volatility) and the sector (financial/industrial).

6 Conclusion

This thesis has been developed in the context of SCR modeling with stochastic volatility and interest rates. It was built on the model proposed by Rodriguez et al. (2014) and extended it to account for the second lag of volatility process with the purpose of estimating the role of volatility persistence in default predicting.

A novel Bayesian inference method has been developed for AR(2) specification of stochastic volatilities. The model has been applied on real financial data and the second-lag coefficient was significant in case of two firms. For the sensitivity analysis the visual inspection of prior and posterior distributions was taken into account which indicated of the dominance of the information provided by the data over the prior specifications. On the other hand the MCMC trace plots of the sampled parameters illustrate that there is sufficient randomness in the generated values.

The model was as well tested against three alternative specifications, the stochastic volatility autoregressive of order one and white noise stochastic volatility . The considered criteria within Bayesian model comparison strongly spoke in favor of the AR(2) specification of stochastic volatility. Nevertheless, the ordering of the models in terms of the support given by the observed data is sensible to the time period under consideration, in particular, during the periods of sudden jumps of volatilities which are not preceded by large fluctuations the SV-AR (1) model finds better support. An implication of this is that applying the SV-AR (2) model in the periods of rare jumps in volatility values could lead to underestimation of default probability. On the other hand, in periods of large scale turbulences in markets, SV-AR(2) will lead to more accurate prediction of default probabilities.

However, overall, non of the models produced CDS prices that were sufficiently close to the market values. One reason for this might be failing to incorporating jumps in asset values process, the importance of which has been discussed by a number of authors (see, for example, Zhang et al.(2005), Delianedis and Geske (2003) and Huang and Huang

(2012)). On the other hand this would be associated with computational complexities and in addition was not the primary purpose of this study. Other possible improvements and topics for future studies are incorporation of leverage effect, that is assuming asymmetric correlation between asset return and volatility processes (Yu, 2005).

Finally, a word of caution should be said about the similarity in the trajectories of asset values and liabilities. On one hand this could have resulted given the relationship between equity, liabilities and asset prices ($E = V - K$), it could be expected that one could mimic movements of the other. Moreover, for instance considering the data and estimation output of Citigroup, the equity prices of which seem to be more stationary without any conspicuous trend, one can see that implied asset values mimic the trajectory of debt values, that is they have an upward trend. Whereas in the case of Ford it seems to be the contrary, that is the assets follow a similar trajectory with equity prices, which in this case have a downward trend, unlike the debt values the trajectory of which seems to lack any kind of trend.

On the other hand this could mean a misspecification of the, in particular, of the approximation method for asset values. An implication of this is that a more accurate approximation, that incorporates both the equity prices (which provide information about the implied call option on the firm's assets) and the CDS prices (which provide information about the implied put option), along with the call-put parity equation, to derive the implied asset prices (Rodriguez et al., 2014). Another approach would be using the exact pricing formula, rather the approximation. This approach or a polynomial approximation (Durham, 2013) would have decreased the bias that results from inserting the expected interest rates and variances without considering the nonlinearity of the pricing forumla and thus the Jensen effect. To accomplish this would require computation of risk neutral probabilities and would be a more complex task to deal with.

Bibliography

1. Bailey, W., Stulz, R., 1989. The Pricing of Stock Index Options in a General Equilibrium Model. *Journal of Financial and Quantitative Analysis* 24: 1-12.

2. Bauwens, L., and M. Lubrano. 1997. Bayesian Option Pricing Using Asymmetric Garch. In *CORE DP 9759*, Louvain-la-Neuve, pp. 321–342.

3. Bielecki T., and Rutkowski M., 2002. *Credit Risk* : Modeling, Valuation and Hedgin. Berlin, Germany: Springer-Verlag Berlin Heidelberg

4. Black, F., and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81: 637–654.

5. Blake A., and Mmumtaz , H., 2012. *CCBS Technical Handbook* – No. 4 Applied Bayesian econometrics for central bankers.. London, UK: Bank of England

6. Bollerslev , T., Chou, R. Y., and Kroner, K.F., 1992. ARCH Modeling in Finance. A Review of the Theory and Empirical Evidence. *Journal of Econometrics* 52: 5-59

7. ”Briys, E., de Varenne, F.1997. Valuing fixed rate debt: An extention. *Journal of Financial and Quantitative Analysis* 32 :239–248

8. Carter , C. K. and R. Kohn, 1994. On Gibbs Sampling for State Space Models. *Biometrika*, 81 (3): 541-553

9. Cetin, U., Jarrow, R., Protter, P., Yildirim, Y. 2004. Modeling default risk with partial information. *Annals of Applied Probability* 14, 1167–1178

10. Chesney, M. and Scott, L., 1989. Pricing European Currency Options: A Comparison of the Modified Black- Scholes Model and a Random Variance Model. *The Journal of Financial and Quantitative Analysis* 24 (3): 267-284

11. Chib S., and Greenberg , W., 1995. Understanding the Metropolis-Hastings Algorithm. *American Statistical Association. The American Statistician* 49 (4) : 327-335

12. Chib, S., 1995. Marginal Likelihood from the Gibbs Output. *Journal of the American Statistical Association*, 90 (432): 1313-1321

13. Collin-Dufresne, P., and R. Goldstein. 2001. Do Credit Spreads Reflect Stationary Leverage Ratios? *Journal of Finance* 56: 1929–1957.

14. Collin-Dufresne, P., R. Goldstein, and J. Helwege, 2003, Are jumps in corporate bond yields priced? Modeling contagion via the updating of beliefs, Working Paper.

15. Cox J. C., Ingersoll, J. E., Jr., and Ross S., A. 1985. A theory of the Term Structure of Interest Rate. *Econometria* 53(2): 385-405

16. Danielsson, J., 1998. Multivariate stochastic volatility models: estimation and comparison with VGARCH models. *Journal of Empirical Finance* 5:155–173

17. Delianedis, G and Geske,R. L., 2003. Credit Risk and Risk Neutral Default Probabilities: Information About Rating Migrations and Defaults . *EFA 2003 Annual Conference Paper No. 962. Available at SSRN: https://ssrn.com/abstract=424301*

18. Duan, J.-C. 1994. Maximum Likelihood Estimation Using Price Data of the Derivative Contract. *Mathematical Finance* 4: 155–167.

19. Duan, J.-C., 1995. The GARCH Option Pricing Model. *Mathematical Finance. 5* (1): 13-32

20. Duan, J.-C., and A. Fulop. 2009. Estimating the Structural Credit Risk Model When Equity Prices are Contaminated by Trading Noises. *Journal of Econometrics* 150: 288–296.

21. Duffie, D., and Lando, D., 2001. Term Structures Of Credit Spreads With Incomplete Accounting Information. *Econometria* 69 (3): 633-664

22. Duffie, D., and Singelton, K., 2003. *Credit Risk. Pricing, Measurement , and Management*. Scottsdale, Arizona: Princeton University Press

23. Durham, G. 2013. Risk-Neutral Modeling with Affine and Nonaffine Models. *Journal of Financial Econometrics* 11: 650–681.

24. Engle, E. F., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometria* 50 (4): 987-1008

25. EOM, Y. H., J. HELWEGE, and J. Z. HUANG 2004. Structural Models of Corporate Bond Pricing: An Empirical Analysis. *Review of Financial Studies* 17: 499–544

26. Fleming J, and Kirby, C., 2003. A Closer Look at the Relation between GARCH and Stochastic Autoregressive Volatility. *Journal of Financial Econometrics* 1 (3): 365-419

27. G. Pye.,1974. Gauging The Default Premium. *Financial Analyst's Journal* (January-February): 49-50

28. Geske, Robert, 1977, "The Valuation of Corporate Liabilities as Compound Options," *Journal of Financial and Quantitative Analysis*: 541-552

29. Haario H., Saksman, E., and Tamminen , J., 2005. Componentwise Adaptation fo High Dimensional MCMC. Computational Finance 20: 265-273

30. Heston, S. L., and S. Nandi. 2000. A Closed formGARCH Option Valuation Model. *The Review of Financial Studies* 13: 585–625.

31. Heston, S.L., 1993. A Closed -Form Solution for Options with Stochastic Volatility with Application to Bond and Currency Options. *The Review of Financial Studies* 6(2) : 327-343

32. Huang, J.-Z., and M. Huang. 2012. How Much of the Corporate-Treasury Yield Spread is due to Credit Risk? *Review of Asset Pricing Studies* 2: 153–202.

33. Hull, J. C. and A. White, 1995. The Impact of Default Risk on Options and Other Derivative Securities. *Journal of Banking and Finance* 19 (2): 299–322.

34. Jacquier, E., Polson, N. G., and Rossi, P. E., 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics* 122: 185-212

35. Jarrow R. A. and Protter P. ,2004. Structural Versus Reduced Form Models: A New Information Based Perspective. *Journal of Inverstment Management*, 2 (2): 1-10

36. Jarrow, R. A., Lando, D., and Turnbull S., 1997. A Markov Model for the Term Structure of Credit Risk Spreads. *Review of Financial Studies*, 10 (2): 481-523.

37. Jeffreys, H. 1961. *Theory of Probability*, 3rd ed. Oxford Classic Texts in the Physical Sciences. Oxford Univ. Press, Oxford.

38. Johnson, H. and Shanno D., 1987 . Option Pricing when the Variance is Changing. *Journal of Financial and Quantiative Analysis* 22 (2): 369-381

39. Kass R. E., and Raftery A. E., 1995. Bayes Factors. *Journal of American Statistical Association* 90 (430) : 773-795

40. Kim C-J., and Nelson, C. 1999. *State Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications.* USA. The MIT Press.

41. Kim, I. J., K. Ramaswamy, and S. Sundaresan. 1993. Does Default Risk in Coupons Affect the Valuation of Corporate Bonds?: A Contingent Claims Model. *Financial Management, special issue on Financial Distress*

42. Kim, S., Shephard, N. and Chib, S., 1998. Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *The Review of Economic Studies* 65 (3): 361-393.

43. Kloeden, P.E., and E. Platen, 1992. *Numerical Solution of Stochastic Differential Equations.* New York, NY: Springer.

44. Lamoureux, C. G. and Lastrapes , W.D. 1990. Persistence in Variance, Structural Change, and the GARCH Model. *Journal of Business & Economic Statistics* 8 (2): 225-234

45. Lando, D. 2004. *Credit Risk Modling. Theory and Application.* New Jersey, USA: Princeton University Press

46. Leland, H., and K. Toft. 1996. Optimal Capital Structure, Endogenous Bankruptcy and the Term Structure of Credit Spreads. *Journal of Finance* 51: 987–1019.

47. Litterman, R. and Iben, T., 1991, Corporate Bond Valuation and the Term Structure of Credit Spreads. *Financial Analyst's Journal*, Spring: 52-64.

48. Longstaff, F. A., and E. S. Schwartz. 1995. A Simple Approach to Valuing Risky Fixed and Floating Rate Debt. *Journal of Finance* 50: 789–819.

49. Madan, D.B. and H. Unal. 1998. Pricing the Risks of Default. *Review of Derivatives Research* 2: 121-160

50. Merton R., C. 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* 29: 449-470.

51. Previts, G. (Ed), 2002. *Research in Accounting Regulation*, Volume 15 1st Edition. JAI Press

52. Rachev T, Hsu J., Bagasheva B., and Fabozzi F. 2008. *Bayesian Methods in Finance*. New Jersey, USA: John Wiley & Sons

53. Robert, C. P., and G. Casella. 2005. *Monte Carlo Statistical Methods*, 2nd edn. San Francisco, USA: Springer.

54. Rodriguez A., ter Horst E. and Malone S. 2015. Bayesian Inference for a Structural Credit Risk Model with Stochastic Volatility and Stochastic Interest Rates. *Journal of Financial Econometrics* 13(4):839-867

55. Saa-Requejo, J. and Santa-Clara, P., 1999 Bond Pricing with Default Risk. Working Paper. UCLA

56. Schneider, Greg. "GM, Ford Bond Ratings Cut to Junk Status." *The Washington Post*. WP Company, 06 May 2005. Web. 11 June 2017.

57. Scott, L.O. 1987. Option Pricing when the Variance Changes Randomly: Cheory, Estimation and an Application. *Journal of Financial and Quantitative Analysis*, 22, 419-438

58. Scott, L.O. 1991. Random-variance option pricing. *Advances in Future and Options Research* 5: 113–35

59. Shephard, N., 1994. Partial Non-Gaussian State Space. *Biometrika* 81 (1): 115-131

60. Shephard, N., 1996. "Statistical aspects of ARCH and stochastic volatility." *Time Series Models in Econometrics, Finance and Other Fields*, (edited by D.R. Cox, David V. Hinkley and Ole E. Barndorff-Neilsen), 1-67. London: Chapman & Hall, 1-67.

61. Shimko, D. C., Naohiko T. and van Deventer D. 1993. The Pricing of Risky Debt When Interest Rates are Stochastic. *Journal of Fixed Income* 3 (2): 58-65

62. Stein, E. M., and Stein J.C., 1991. Stock Price Distributions with Stochastic Volatility: An Analytical Approach. *The Review of Financial Studies* 4 (4): 727-752

63. Taylor, S. J.,1982. Financial returns modeled by the product of two stochastic processes, a study of daily sugar prices 1961-79, in O. D. Anderson (ed.), *Time Series Analysis: Theory and Practice 1*, North-Holland, sterdam, pp. 203–226.

64. Taylor, S. J.,1994. Modeling Stochastic Volatility: A Review and Comparative Study. *Mathematical Finance* 4(2) : 183-204

65. The origins of the financial crisis: Crash course (2013,September 7). *Retrieved from http://www.economist.com/news/schoolsbrief/21584534-effects-financial-crisis-are-still-being-felt-five-years-article.*

66. Vasicek, O. 1977. An Equilibrium Characterisation of the Term Structure. *Journal of Financial Economics* 5: 177–188.

67. Wiggins, J.B. 1987. Option values under stochastic volatility. Theory and empirical estimates. *Journal of Financial Economics*, 19: 351-372.

68. Yu, J., 2005. On Leverage in a Stochastic Volatility Model. *Journal of Econometrics* 127: 165-178

69. Zhang, B. Y., Z. Hao, and H. Zhu, 2009. Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms. *Review of Financial Studies* 22: 5099–5131.

70. Zhou C., 1997. A Jump-Diffusion Approach to Modeling Credit Risk and Valuing Defaultable Securities. Available at *SSRN: https://ssrn.com/abstract=39800 or http://dx.doi.org/10.2139/ssrn.39800*

A MCMC Sampling Outputs

Citigroup. SV-AR(2), MCMC trace plots

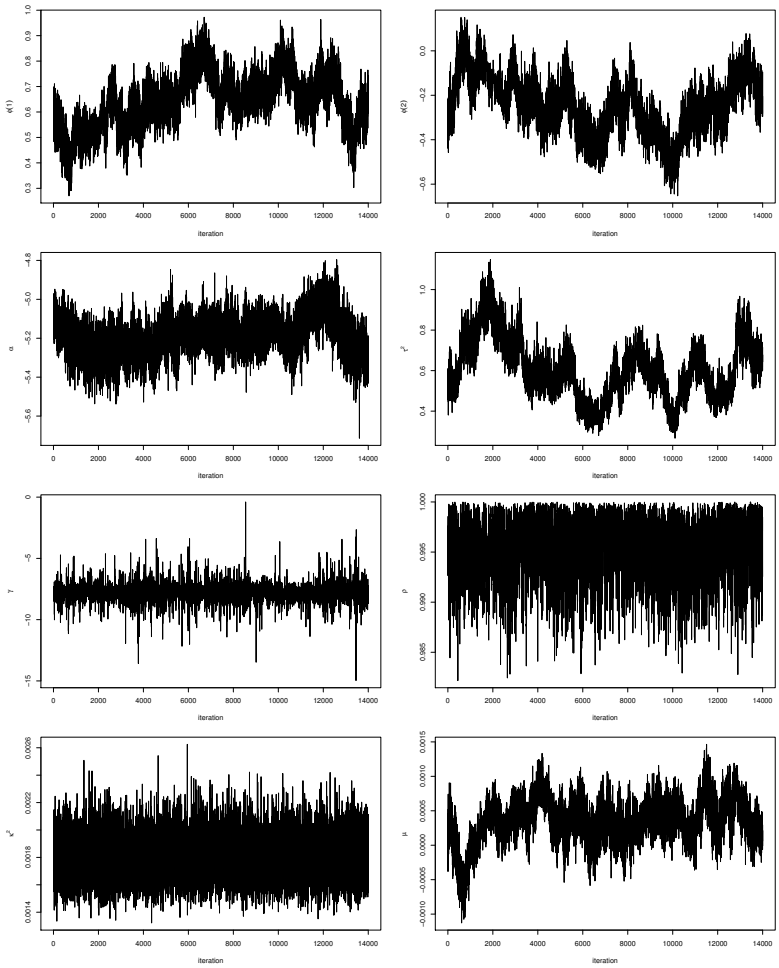

Figure 9: MCMC trace plots, Citigroup SV-AR(2) parameters

Citigroup. SV-AR(1) MCMC trace plots and posterior distributions

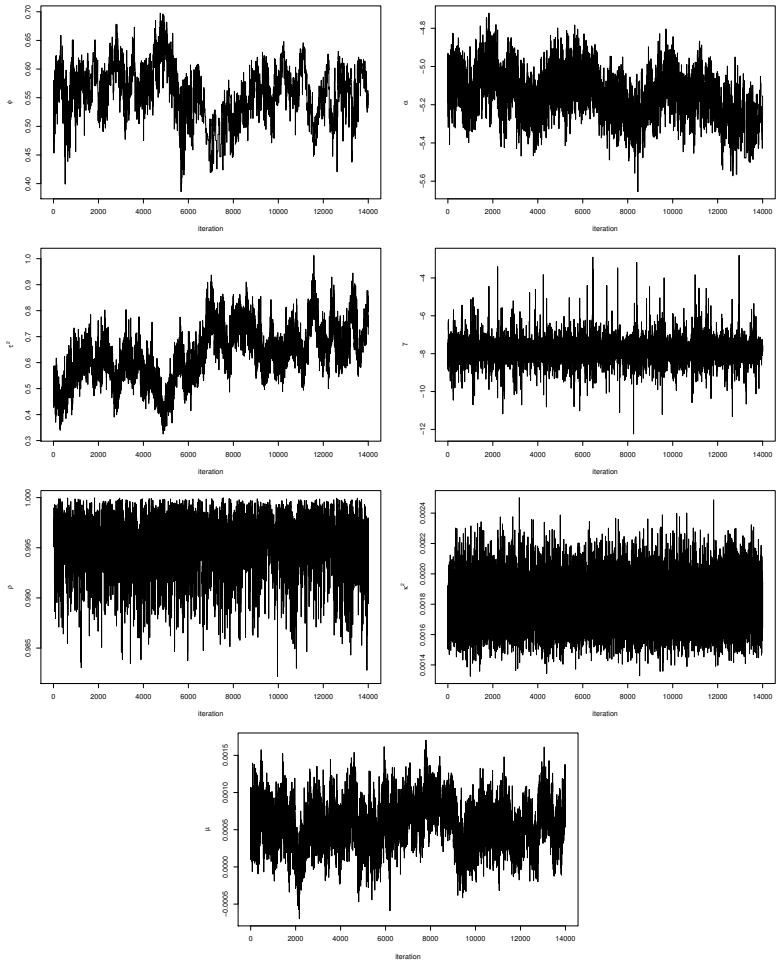

Figure 10: MCMC trace plots, Citigroup. SV-AR(1) parameters

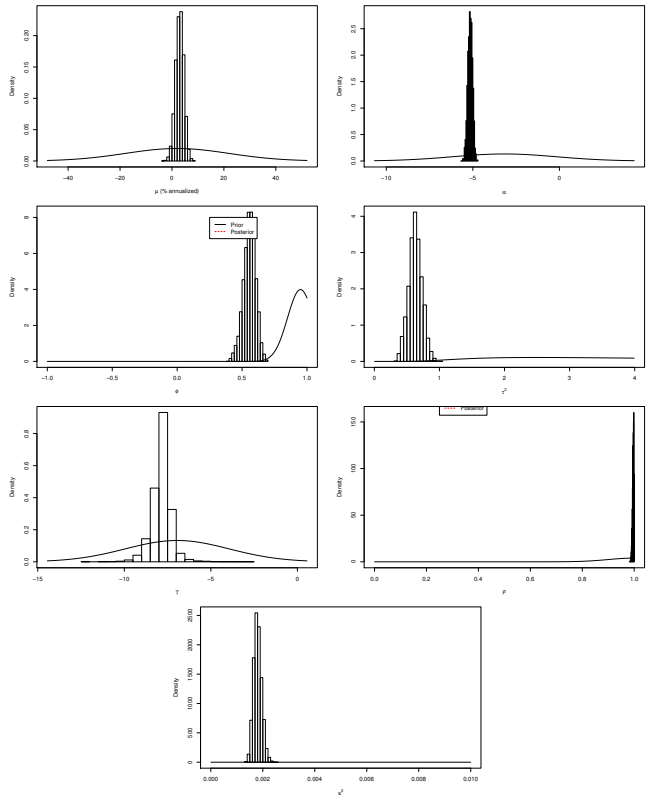

Figure 11: Posterior histograms and prior distributions (solid lines). Citigroup SV-AR(1) parameters

Citigroup. SV-WN MCMC trace plots

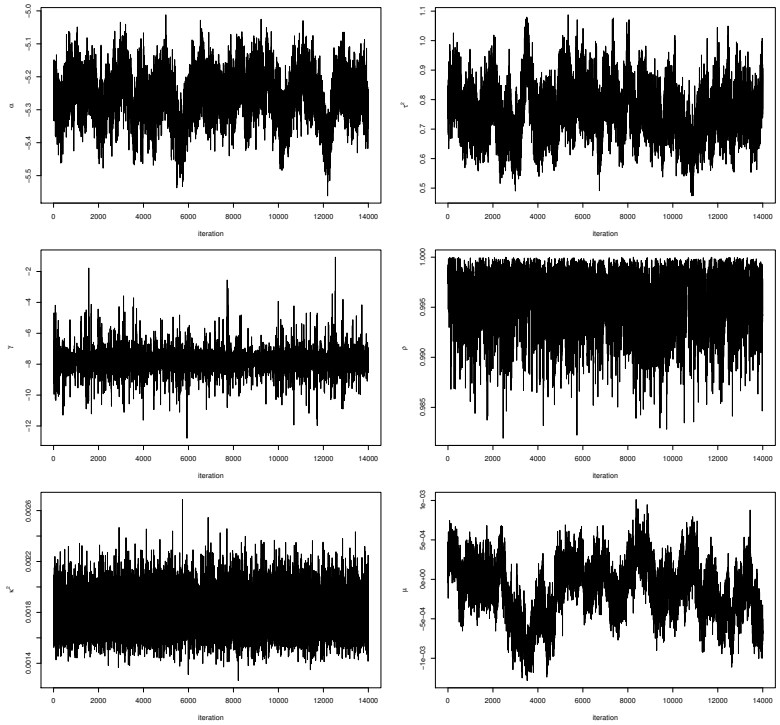

Figure 12: MCMC trace plots, Citigroup. SV -WN parameters

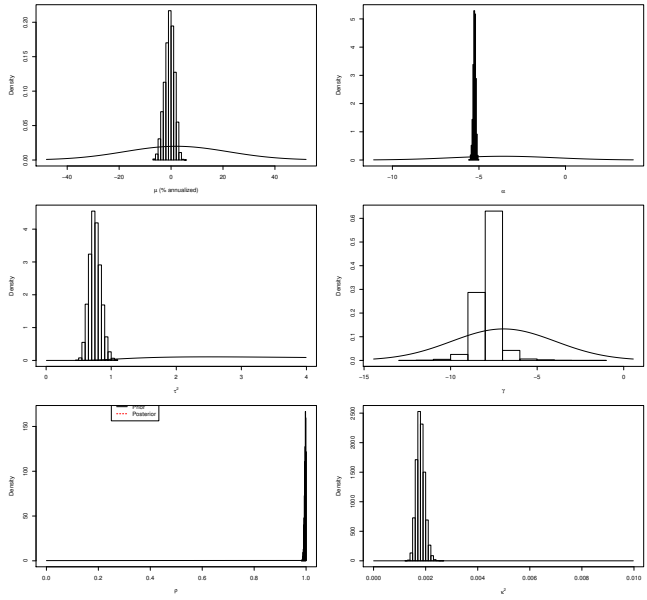

Figure 13: Posterior histograms and prior distributions (solid lines). Citigroup SV-WN parameters

Lehman Brothers SV-AR(2) MCMC traceplots

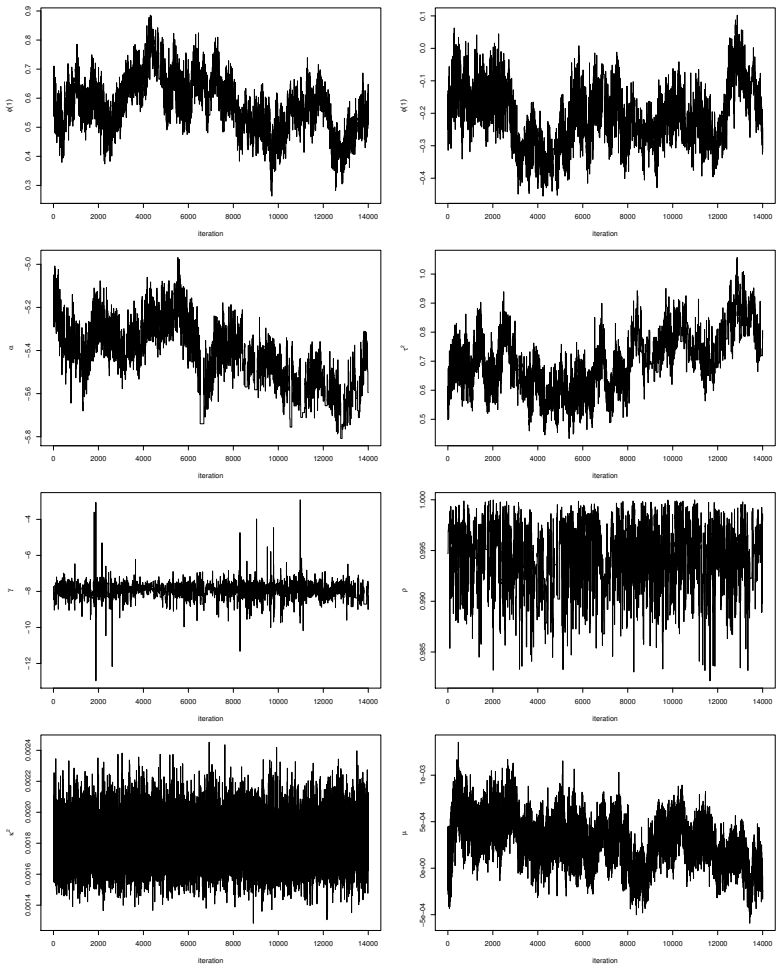

Figure 14: MCMC trace plots, Lehman Brothers SV-AR(2) parameters

Lehman Brothers SV-AR(1) MCMC traceplots and posterior distributions

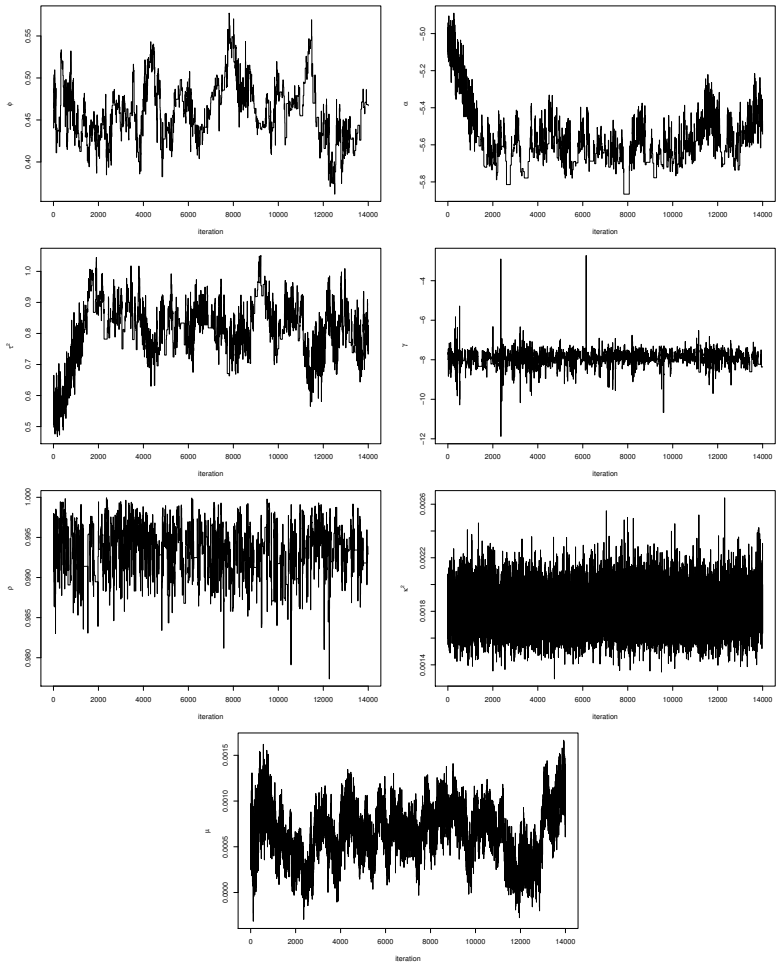

Figure 15: MCMC trace plots, Lehman Brothers. SV-AR(1) parameters

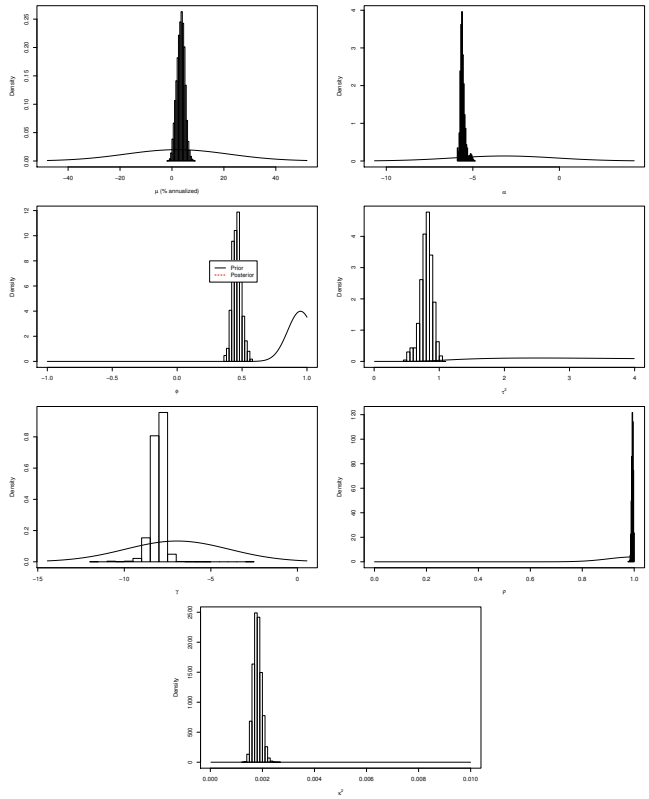

Figure 16: Posterior histograms and prior distributions (solid lines). Lehman Brothers SV-AR(1) parameters

Lehman Brothers. SV-WN MCMC traceplots and posterior distributions

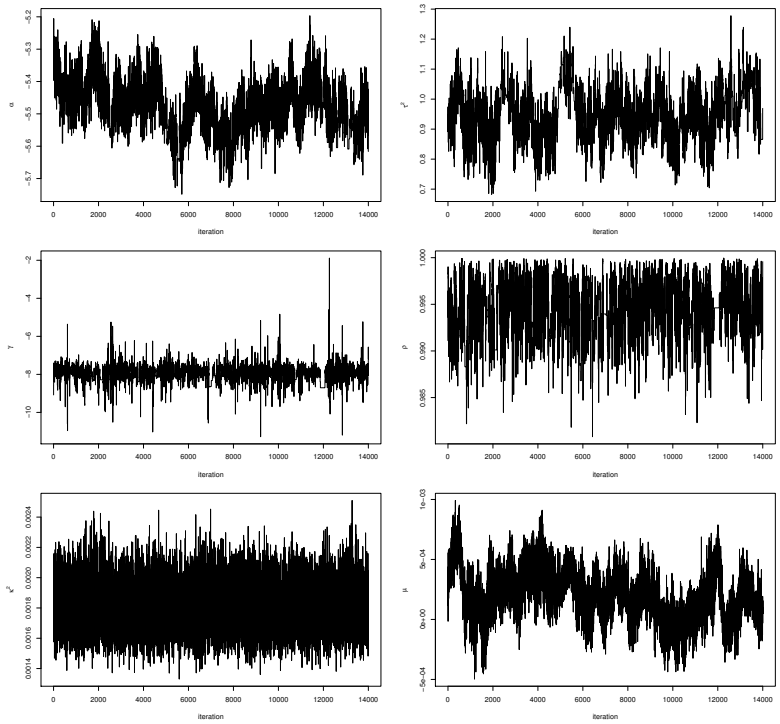

Figure 17: MCMC trace plots, Lehman Brothers. SV -WN parameters

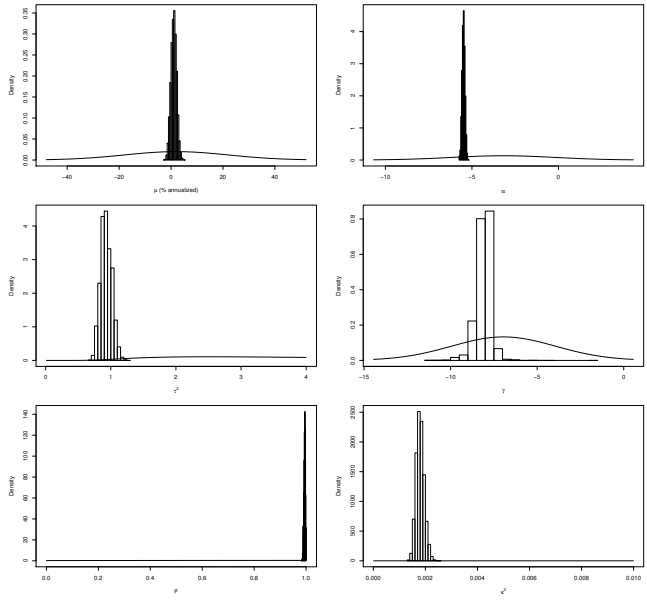

Figure 18: Posterior histograms and prior distributions (solid lines).Lehman Brothers SV-WN parameters

Ford SV-AR(2) MCMC trace plots

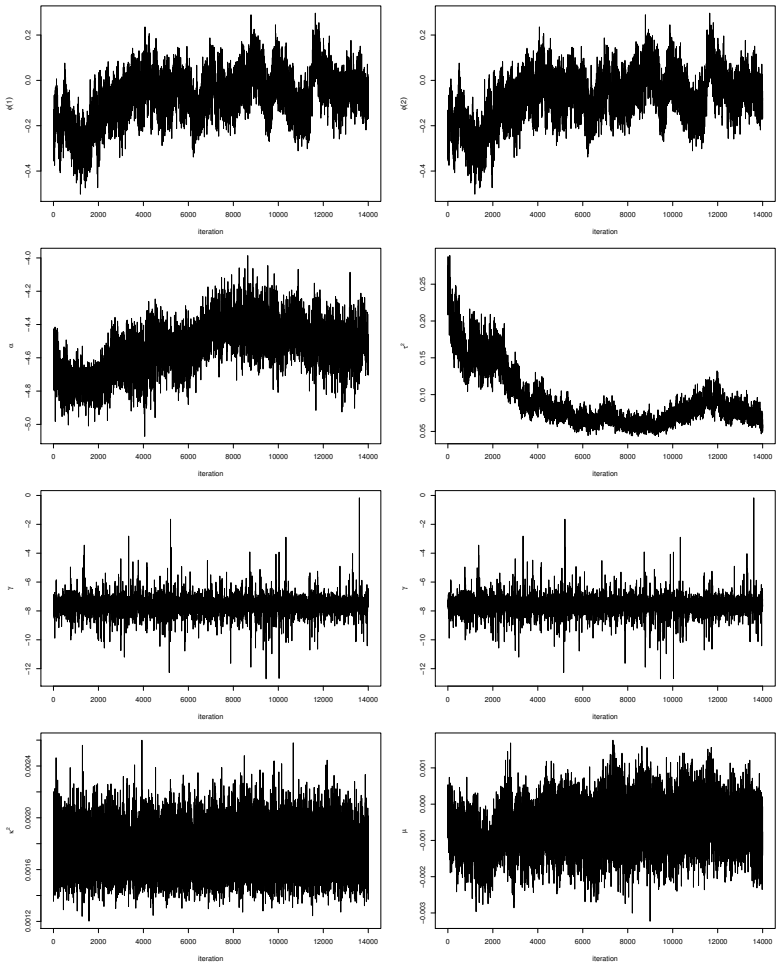

Figure 19: MCMC trace plots, Ford SV-AR(2) parameters

Ford SV-AR(1). MCMC trace plots and posterior distributions

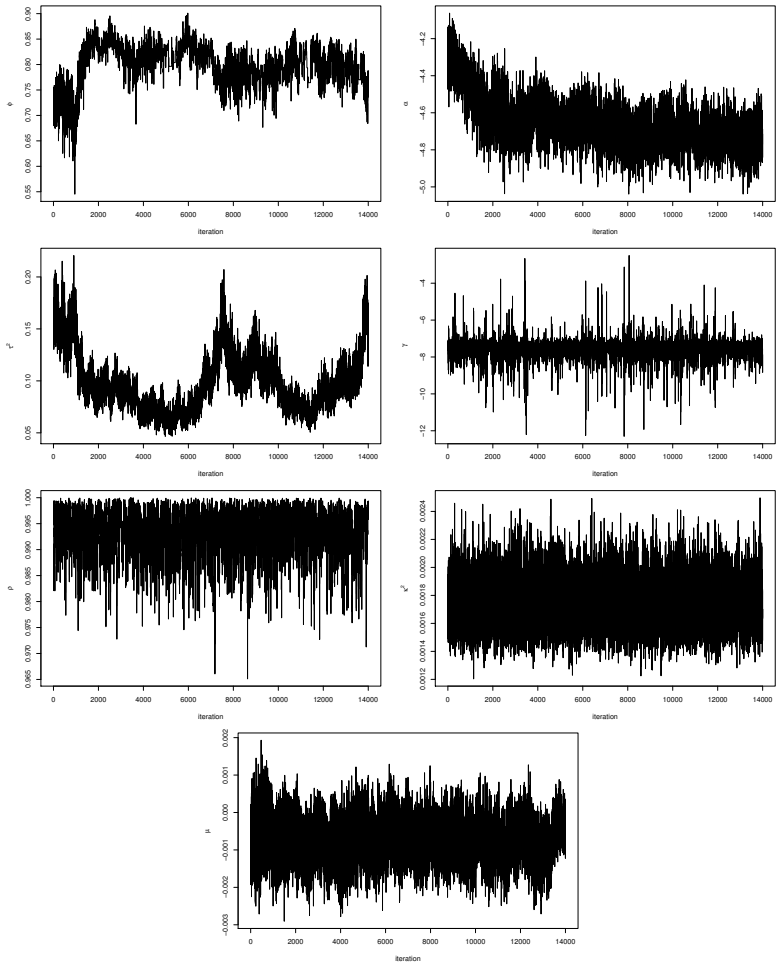

Figure 20: MCMC trace plots, Ford. SV-AR(1) parameters

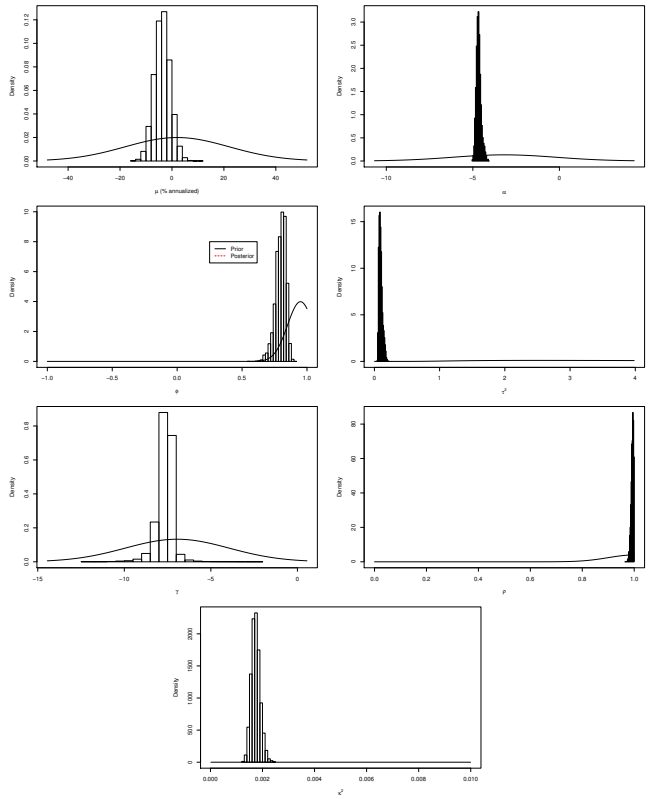

Figure 21: Posterior histograms and prior distributions (solid lines).Ford SV-AR(1) parameters

Ford. SV-WN MCMC traceplots and posterior distributions

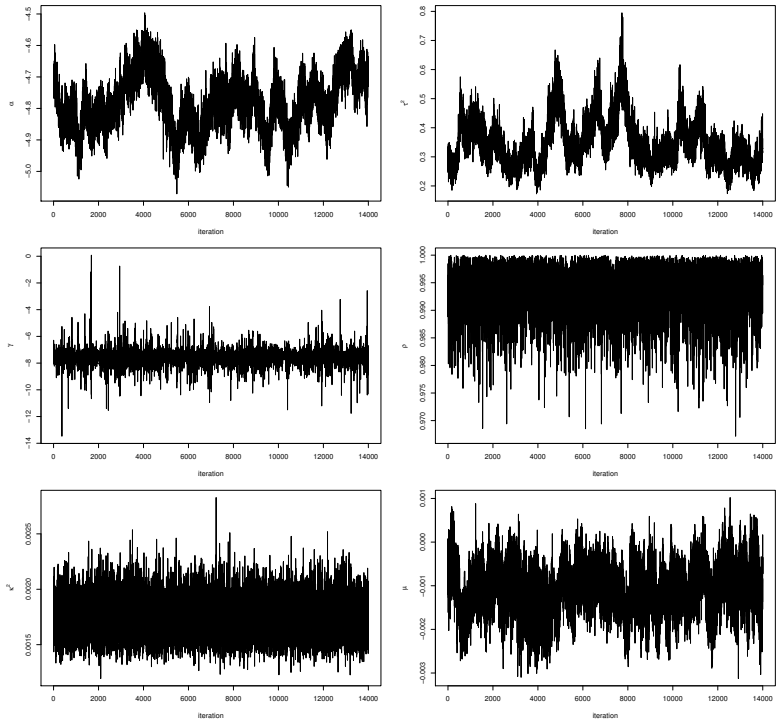

Figure 22: MCMC trace plots, Ford. SV -WN parameters

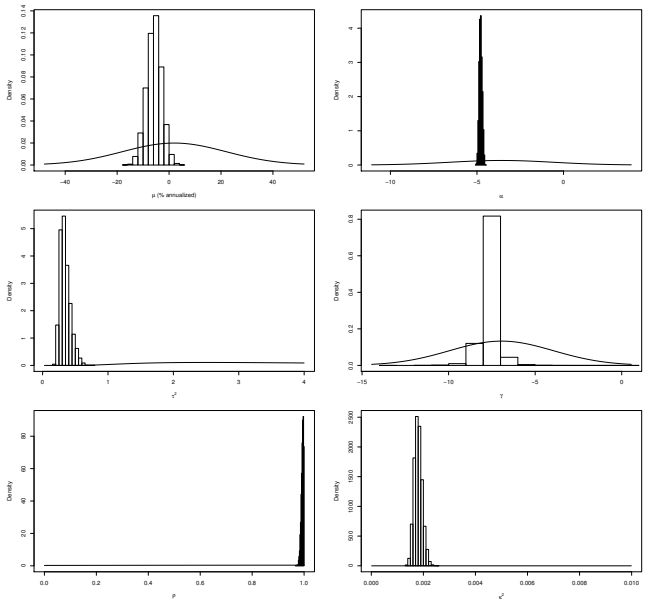

Figure 23: Posterior histograms and prior distributions (solid lines).Ford SV-WN parameters

B Dependence of Default Probability on Volatility

Following Duan (1994) where the default probability is given by the equation (**??**), the derivative of default function with respect to volatility σ is equal to:

$$\frac{\partial P}{\partial \sigma} = \frac{\frac{1}{2}\sigma^2(T-t) - (\ln(F_t/A_t) - \mu(T-t))}{\sigma^2\sqrt{T-t}}\phi\Big(\frac{\ln(F_t/A_t) - (\mu - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\Big)$$

the sign of which depends on the parameters of the model.
On the other hand, under stochastic volatility assumption the partial derivative of default probability wrt volatility at time t, σ_t, applying chain rule, is given by

$$\frac{\partial P}{\partial \sigma_t} = -\frac{\partial \tilde{\sigma}_t(m)}{\partial \sigma_t}\frac{1}{\tilde{\sigma}_t^2(m)}\phi\left(\frac{\ln K_t(m) - \ln V_t - \mu m}{\tilde{\sigma}_t(m)}\right)$$

$$= -\frac{\sum_{u=1}^{m}\beta^u \exp\left\{2\alpha(1-\beta^u) + 2\tau^2\frac{1-\beta^{2u}}{1-\beta^2} + \beta^u \ln\sigma_t^2\right\}}{\sigma_t\sqrt{\sum_{u=1}^{m}\exp\left\{2\alpha(1-\beta^u) + 2\tau^2\frac{1-\beta^{2u}}{1-\beta^2} + \beta^u\ln\sigma_t^2\right\}}}$$
$$\frac{1}{\tilde{\sigma}_t^2(m)}\phi\left(\frac{\ln K_t(m) - \ln V_t - \mu m}{\tilde{\sigma}_t(m)}\right);$$

the sign of which depends on σ_t. Thus default probability can be increasing or decreasing in volatility, depending whether the volatility is positive or negative.

C Deriving Discrete Model from the Continuous

The model given by equation (2.13) is derived from continuous time SCR model with stochastic volatility by applying first order Euler's approximation. Below is demonstrated the discretization of asset's dynamics' and volatility processes.
First consider the following continuous time dynamics for asset prices and log volatility (Chesney and Scott,1989)

$$dS_t = \mu S_t dt + \sigma_t S_t dW_{1t}$$
$$d\ln\sigma_t = \theta(\alpha - \ln\sigma_t)dt + \gamma dW_{2t}$$

Applying Ito's lemma to the stochastic process of asset prices, for log asset prices we have:

$$\begin{aligned} d\ln S_t &= \Big(\frac{\mu S_t}{S_t} - \frac{\sigma_t^2 S_t^2}{2S_t^2}\Big)dt + \frac{S_t\sigma_t}{S}dW_{1t} \\ &= \Big(\mu - \frac{\sigma_t^2}{2}\Big)dt + \sigma_t dW_{1t} \end{aligned} \tag{A.1}$$

Further taking the integral from t to $t + dt$ on both sides of A.1 gives:

$$\begin{aligned} \ln S_{t+dt} &= \ln S_t + \int_t^{t+dt}(\mu - \sigma_u^2/2)du + \int_t^{t+dt}\sigma_u dW_{1u} \\ &\approx \ln S_t + (\mu - \sigma_t^2/2)dt + \sigma(W_{1t+dt} - W_{1t}) \\ &= \ln S_t + (\mu - \sigma_t^2/2)dt + \sigma\sqrt{dt}Z \end{aligned} \tag{A.2}$$

where Z is a standard normal distribution and since W_1 is a Wiener process, $\sqrt{dt}Z$ and $W_{1t+dt} - W_{1t}$ are identically distributed (Kloeden and Platen, 1992).
In a similar fashion we have for $\ln\sigma$:

$$\begin{aligned} \ln\sigma_{t+dt} &\approx \ln\sigma_t + \theta(\alpha - \ln\sigma_t) + \gamma\sqrt{dt}Z \\ &= (1-\theta)\ln\sigma_t + \theta\alpha + \gamma\sqrt{dt}Z \\ &= \alpha(1-\beta) + \beta\ln\sigma_t + \gamma\sqrt{dt}Z \\ &= \alpha + \beta(\ln\sigma_t - \alpha) + \gamma\sqrt{dt}Z \end{aligned} \tag{A.3}$$

Thus by plugging in $dt = 1$ in A.2 and A.3 we get the model specified by 3.1

D Details of the MCMC Algorithm

The details of the Metropolis-Hastings step of the MCMC algorithm for the SV-AR(2) model are provided in this section. The algorithm and posterior distributions are derived based on Kim et al. (1998) , Rodriguez et al. (2014) and Chib and Greenberg (1995). The Gibbs sampling for the pre sample of the parameters is conducted using the same conditional posterior densities, omitting the accept-reject step.

1. Update μ from a Gaussian distribution,

$$\mu \mid \cdots \sim \mathsf{N}\left(\left[\sum_{t=2}^{n}\frac{1}{\sigma_t^2} + \frac{1}{\varpi_\mu^2}\right]^{-1}\left[\sum_{t=2}^{n}\frac{x_t}{\sigma_t^2} + \frac{\xi_\mu}{\varpi_\mu^2}\right], \left[\sum_{t=2}^{n}\frac{1}{\sigma_t^2} + \frac{1}{\varpi_\mu^2}\right]^{-1}\right)$$

where $x_t = \ln V_t/V_{t-1}$ is the return at time t.

2. Update α by first making a proposal α^* from a Gaussian distribution $\alpha^* \mid \cdots \sim \mathsf{N}\left(\tilde{\xi}_\alpha, \tilde{\varpi}_\alpha^2\right)$, where

$$\tilde{\xi}_\alpha = \frac{\frac{(z_1+z_2)(1+\phi_2)(1-\phi_1-\phi_2)}{\tau^2} + \sum_{t=3}^{n} \frac{(1-\phi_1-\phi_2)(z_t-\phi_1 z_{t-1}-\phi_2 z_{t-2})}{\tau^2} + \frac{\xi_\alpha}{\varpi_\alpha^2}}{\frac{2(1-\phi_1-\phi_2)(1+\phi_2)}{\tau^2} + \frac{(n-2)(1-\phi_1-\phi_2)^2}{\tau^2} + \frac{1}{\varpi_\alpha^2}},$$

$$\tilde{\varpi}_\alpha^2 = \frac{1}{\frac{2(1-\phi_1-\phi_2)(1+\phi_2)}{\tau^2} + \frac{(n-2)(1-\phi_1-\phi_2)^2}{\tau^2} + \frac{1}{\varpi_\alpha^2}},$$

and $z_t = \ln \sigma_t$. This proposal is accepted with probability $\min\{1, \Delta_\alpha\}$ where

$$\Delta_\alpha = \frac{p(e_1, \ldots, e_n \mid \mu, \sigma_1, \ldots, \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \alpha^*, \phi_1, \phi_2, \tau, \gamma, \rho, \omega)}{p(e_1, \ldots, e_n \mid \mu, \sigma_1, \ldots, \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \alpha, \phi_1, \phi_2, \tau, \gamma, \rho, \omega)},$$

and the likelihood

$$p(e_1, \ldots, e_n \mid \mu, \sigma_1, \ldots, \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \alpha, \phi_1, \phi_2, \tau, \gamma, \rho, \omega)$$

is given in Equation (4.3).

3. Update $\phi = (\phi_1, \phi_2)'$ by first making a proposal β^* from a Gaussian distribution $\phi^* \mid \cdots \sim \mathsf{N}_2\left(\tilde{\xi}_\beta, \tilde{\varpi}_\beta^2\right)$, where

$$\tilde{\xi}_\phi = \tilde{\varpi}_\phi^2 \Big[\sum_{t=3}^{n} x_t(z_t - \alpha)/\tau^2 + \xi_\phi[\varpi_\beta^2]^{-1}\Big], \qquad \tilde{\varpi}_\phi^2 = \Big[\sum_{t=3}^{n} x_t x_t'/\tau^2 + [\varpi_\phi^2]^{-1}\Big]^{-1}.$$

where $x_t = (z_{t-1} - \alpha, z_{t-2} - \alpha)'$. If stationarity is satisfied, this proposal is accepted with probability $\min\{1, \Delta_\phi\}$ where

$$\Delta_\phi = \frac{\Psi(\phi^*\tau^2)}{\Psi(\phi, \tau^2)} \frac{p(e_1, \ldots, e_n \mid \mu, \sigma_1, \ldots, \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \alpha, \phi^*, \tau, \gamma, \rho, \omega)}{p(e_1, \ldots, e_n \mid \mu, \sigma_1, \ldots, \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \alpha, \phi, \tau, \gamma, \rho, \omega)}.$$

4. Update τ^2 by first making a proposal τ^{*2} from an inverse Gamma distribution,

$\tau^{*2} \mid \cdots \sim \mathsf{IG}\left(\tilde{a}_\tau, \tilde{b}_\tau\right)$ where

$$\tilde{a}_\tau = a_\tau + \frac{n}{2}, \quad \tilde{b}_\tau = b_\tau + \frac{1}{2}\left(z_1(z_1+z_2)(1+\phi_2)(1-\phi_1-\phi_2)) + \sum_{t=3}^{n}\left[z_t - \alpha - x_t'\phi\right]^2\right).$$

This proposal is accepted with probability $\min\{1, \Delta_\tau\}$ where

$$\Delta_\tau = \frac{p(e_1,\ldots,e_n \mid \mu, \sigma_1,\ldots,\sigma_n, r_1,\ldots,r_n, K_1(m),\ldots,K_n(m), \alpha, \phi, \tau^*, \gamma, \rho, \omega)}{p(e_1,\ldots,e_n \mid \mu, \sigma_1,\ldots,\sigma_n, r_1,\ldots,r_n, K_1(m),\ldots,K_n(m), \alpha, \phi, \tau, \gamma, \rho, \omega)},$$

5. Update γ by first making a proposal γ^* from a Gaussian distribution $\gamma^* \mid \cdots \sim \mathsf{N}\left(\tilde{\xi}_\gamma, \tilde{\varpi}_\gamma^2\right)$, where

$$\tilde{\xi}_\gamma = \frac{\frac{(1-\rho)^2 x_1}{\kappa^2} + \sum_{t=2}^{n}\frac{(1-\rho)^2(x_t - \rho x_{t-1})}{\kappa^2} + \frac{\xi_\gamma}{\varpi_\gamma^2}}{\frac{(1-\rho)^2}{\kappa^2} + \frac{(n-1)(1-\rho)^2}{\kappa^2} + \frac{1}{\varpi_\gamma^2}}, \qquad \tilde{\varpi}_\gamma^2 = \frac{1}{\frac{(1-\rho)^2}{\kappa^2} + \frac{(n-1)(1-\rho)^2}{\kappa^2} + \frac{1}{\varpi_\gamma^2}},$$

and $x_t = \ln r_t$. This proposal is accepted with probability $\min\{1, \Delta_\gamma\}$ where

$$\Delta_\gamma = \frac{p(e_1,\ldots,e_n \mid \mu, \sigma_1,\ldots,\sigma_n, r_1,\ldots,r_n, K_1(m),\ldots,K_n(m), \gamma^*, \rho, \kappa, \gamma, \phi, \omega)}{p(e_1,\ldots,e_n \mid \mu, \sigma_1,\ldots,\sigma_n, r_1,\ldots,r_n, K_1(m),\ldots,K_n(m), \gamma, \rho, \kappa, \gamma, \phi, \omega)}.$$

6. Update ρ by first making a proposal ρ^* from a Gaussian distribution $\rho^* \mid \cdots \sim \mathsf{N}_{[0,1]}\left(\tilde{\xi}_\rho, \tilde{\varpi}_\rho^2\right)$, where

$$\tilde{\xi}_\rho = \frac{\sum_{t=2}^{n}\frac{(x_t-\gamma)(x_{t-1}-\gamma)}{\kappa^2} + \frac{\xi_\rho}{\varpi_\rho^2}}{\sum_{t=1}^{n-1}\frac{(x_t-\gamma)^2}{\kappa^2} + \frac{1}{\varpi_\rho^2}}, \qquad \tilde{\varpi}_\rho^2 = \frac{1}{\sum_{t=1}^{n-1}\frac{(x_t-\gamma)^2}{\kappa^2} + \frac{1}{\varpi_\rho^2}}.$$

This proposal is accepted with probability $\min\{1, \Delta_\rho\}$ where

$$\Delta_\rho = \left(\frac{1-\rho^{*2}}{1-\rho^2}\right)^{1/2} \exp\left(\frac{(x_1-\gamma)^2(\rho^{*2}-\rho^2)}{2\kappa^2}\right)$$
$$\frac{p(e_1,\ldots,e_n \mid \mu, \sigma_1,\ldots,\sigma_n, r_1,\ldots,r_n, K_1(m),\ldots,K_n(m), \gamma, \rho^*, \kappa, \gamma, \rho, \omega)}{p(e_1,\ldots,e_n \mid \mu, \sigma_1,\ldots,\sigma_n, r_1,\ldots,r_n, K_1(m),\ldots,K_n(m), \gamma, \rho, \kappa, \gamma, \rho, \omega)}.$$

7. Update κ^2 by first making a proposal κ^{*2} from an inverse Gamma distribution, $\kappa^{*2} \mid \cdots \sim \mathsf{IG}\left(\tilde{a}_\kappa, \tilde{b}_\kappa\right)$ where

$$\tilde{a}_\kappa = a_\kappa + \frac{n}{2}, \qquad \tilde{b}_\kappa = b_\kappa + \frac{1}{2}\left((1-\rho^2)x_1^2 + \sum_{t=2}^{n}\left[x_t - \rho x_{t-1} - \gamma(1-\rho)\right]^2\right).$$

This proposal is accepted with probability $\min\{1, \Delta_\kappa\}$ where

$$\Delta_\kappa = \frac{p(e_1, \ldots, e_n \mid \mu, \sigma_1, \ldots, \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \gamma, \rho, \kappa^*, \gamma, \rho, \omega)}{p(e_1, \ldots, e_n \mid \mu, \sigma_1, \ldots, \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \gamma, \rho, \kappa, \gamma, \rho, \omega)},$$

8. Update a block $\{\sigma_t, \ldots \sigma_{t+k}\}$ at a time by proposing a new value of $z^*_{t+i} = \ln \sigma^*_{t+i}, i = 0 \ldots k$ using a Gaussian random walk where

$$z^*_{t+i} \sim \mathsf{N}(z_{t+i}, \zeta_z^2)$$

This proposal is accepted with probability $\min\{1, \Delta_z\}$ where

$$\Delta_z = \frac{p(\sigma_1, \ldots, \sigma^*_t \ldots \sigma^*_{t+k}, \ldots, \sigma_n \mid \alpha, \beta, \tau)}{p(\sigma_1, \ldots, \sigma_t, \ldots \sigma_{t+k} \ldots, \sigma_n \mid \alpha, \beta, \tau)}$$
$$\frac{p(e_1, \ldots, e_n \mid \mu, \sigma_1, \ldots, \sigma^*_t \ldots \sigma^*_{t+k} \ldots \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \gamma, \rho, \kappa^*, \gamma, \phi)}{p(e_1, \ldots, e_n \mid \mu, \sigma_1, \ldots, \sigma_t, \ldots \sigma_{t+k}, \ldots, \sigma_n, r_1, \ldots, r_n, K_1(m), \ldots, K_n(m), \gamma, \rho, \kappa, \gamma, \phi)}$$