Estimating Implied Recovery Rates from the Term Structure of CDS Spreads

Marcin Jaskowski[1]
Michael McAleer[1,2]

[1] *Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam;*
[2] *Complutense University of Madrid, Kyoto University, Japan, and Tinbergen Institute;*

Estimating Implied Recovery Rates from the Term Structure of CDS Spreads

Marcin Jaskowski[1] and Michael McAleer[*,1,2,3,4]

[1]*Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam*
[2]*Tinbergen Institute, The Netherlands*
[3]*Institute of Economic Research, Kyoto University, Japan*
[4]*Department of Quantitative Economics, Complutense University of Madrid, Spain*

December 2012

Abstract

Credit risk models should reflect the observation that the relevant value of collateral is generally not the average value of the asset over all possible states of nature. In most cases, the relevant value of collateral for the lender is its secondary market value in bad states of nature, where marginal utilities are high. Although the negative correlation between recovery rates and default probabilities is well documented, the majority of pricing models does not allow for correlation between the two. In this paper, we propose a relatively parsimonious reduced-form continuous time model that estimates expected recovery rates and default probabilities from the term structure of CDS spreads. The parameters of the model and latent factors driving recovery risk and default risk are estimated using a Bayesian MCMC algorithm. We find that the Bayesian deviance information criterion (DIC) favors the model with stochastic recovery over constant recovery. We also observe that for companies with a good rating, implied constant recovery rates do not differ much from stochastic recovery. However, if a company is very risky, then forward stochastic recovery rates are significantly lower at longer maturities.

Keywords: Constant recovery, stochastic recovery, implied recovery rate, term structure, CDS spreads.
JEL: G13, G17, G33, E43

---

1 Introduction

In recent years, a number of papers have concentrated on the role of recovery risk in credit risk models. Part of this literature attempts to extract implied recovery rates from observed prices of bonds or CDS spreads. Most of these studies are based on the assumption that recovery rates are constant and are independent of the default probabilities. The assumption of independence is made for tractability. However, this turns out to be a very strong assumption because we have sufficient empirical evidence to believe that recovery rates are stochastic and negatively correlated with default rates (Altman, Brady, Resti, and Sironi (2005), Altman (2006), Acharya, Bharath, and Srinivasan (2007), Bruche and González-Aguado (2010)). In this paper, we address the question of how far this "independence assumption" may be justified.

The basic intuition behind this paper is based on the corporate finance theory of asset fire sales, as described in the seminal paper by Shleifer and Vishny (2012). Their model was the first to make the resale price of collateral endogenous. First, they make the very plausible assumption that assets are specialized, and their first-best use is only within a particular industry. Obviously, when an asset has many alternative uses, it should always have high liquidation value. For instance, commercial real estate can be used in many different ways. However, hardly any asset is so easily redeployable. Most of the assets have no sensible use for firms outside the industry.

The second important assumption is that shocks affecting firms within the same industry are correlated. That means that when one firm has problems meeting its debt obligations, then most probably other firms in the industry will face similar problems. Consequently, firms in times of distress may be forced to sell their assets at significantly depressed prices to outsiders from other industries. Therefore, defaulting firms can impose a negative externality on other firms in the industry, or even on the economy at large.

The insights of Shleifer and Vishny (2012) have been confirmed in many empirical studies. Altman et al. (2005), Altman (2006) shows that the aggregated recovery rates on corporate bonds can be volatile and negatively correlated with aggregate default rates. That is, recovery rates tend to fall at a time when the number of defaults rises. For instance, using yearly data between 1982 and 2009, Altman (2006) shows that a simple linear regression of realized weighted average recovery rates on weighted average default rates gives an $R^2 = 0.5361$ with a significantly negative coefficient.

In a related study, Acharya et al. (2007) show that recovery rates depend more on industry specific conditions than on macroeconomic indicators. They claim that results from Altman et al. (2005), appear to be just a manifestation of omitted variables. Other studies like Campbell, Giglio, and Pathak (2009) provide evidence of spillover effects, in which home foreclosures reduce the prices of other houses in the neighborhood. Benmelech and Bergman (2011) identify another channel for spillover effects of firm bankruptcies. They demonstrate that bankrupt firms in a particular industry impose a negative externality on their non-bankrupt competitors. That happens because liquidations depress the value of collateral, and thereby increase the cost of external debt financing for the rest of the industry. This, in turn, increases the default probabilities of all firms using the same type of asset as collateral. Altogether, it creates an amplifying mechanism that can propagate industry downturns.

Moreover, recent events in the USA, particularly in the housing market, suggest that the collateral shocks have crucial importance to the economy at large. More broadly, all of these studies demonstrate that recovery risk definitely exists and has an economically significant magnitude. Despite all the evidence, the common practice, both in academic circles and in industry, in analysing credit risk models seems to use a constant recovery rate that is usually fixed at a level between 40% and 50%. In other words, the assumption is that the collateral will be equally

easy to sell, no matter what the state of the world.

In this paper, we are concerned with the problem of extracting implied recovery rates from observable prices. The literature on the estimation of implied recovery rates is relatively small and new, especially when compared with the large literature on default probabilities. One of the first attempts to extract implied recovery rates was in Madan, Bakshi, and Zhang (2006). Their model assumed that both default intensity and the recovery rate are driven by the same factor that governs the risk-free rate term structure. In this paper, we also assume that default intensity and recovery rates are driven by a single factor. However, in contrast to Madan et al. (2006), we assume that CDS spreads are driven by a latent factor that is specific to the credit risk of the company rather than to a risk-free rate.

Pan and Singleton (2008) demonstrate that one can exploit the term structure of CDS spreads to identify the parameters of the default intensity and the recovery rate. They find that long-maturity premia are essential for identification as the impact of changes in the recovery rate on short-maturity premia is relatively low. In their model, expected recovery rates are constant. They argue that recovery rates on sovereign bonds are not correlated with the business cycle in the same way as recovery rates for corporate CDS. Therefore, a constant recovery assumption is not so unreasonable for a sovereign debt market.

Schneider, Sögner, and Veža (2011), in a framework similar to Pan and Singleton (2008), simultaneously estimate jumps in default intensities and firm-specific constant recovery rates from a large cross section of US corporate CDS spreads. We augment their approach and extract the parameters of the default intensity process and recovery rate from the term structure of CDS spreads, but also allow the recovery rate to depend on the default intensity.

Das and Hanouna (2009) use data on both CDS spreads and stock price data to infer recovery rates by boot-strapping over the term structure of CDS using a parametric function to link default probabilities with stochastic recoveries. Similarly, we assume a parametric relation between the stochastic recovery rate and default probability. However, we do not calibrate our model to a single time point but perform time series estimation.

Two other papers related to Das and Hanouna (2009) are Le (2007) and Song (2007). Le (2007) infers implied recovery rates in a two-step procedure. Using option prices, he finds the risk-neutral default intensity, and then deduces what should be the corresponding recovery rate from CDS spreads. Song (2007) developed a framework based on cross-sectional no-arbitrage restrictions between different credit derivatives in order to identify constant recovery rates.

Finally, two papers that are most closely related to this paper are Christensen (2007) and Doshi (2011). Both models estimate a stochastic recovery model using CDS data. Christensen (2007) estimates the model for the Ford Motor Corporation using senior CDS contracts, while Doshi (2011) estimates a discrete time model using CDS spreads of different seniorities and for 46 firms.

This paper complements these papers in two ways. First, we estimate a more parsimonious model than either Christensen (2007) or Doshi (2011), which allows us to model default intensity and recovery rate as being driven by the same credit risk specific latent factor. That is, in our case the correlation between recovery and the default intensity does not come from the common dependence on interest rate factors. In terms of the asset fire sales framework, it would be rather difficult to find an economically reasonable mechanism that would link risk-free interest rates with lower or higher recovery rates on bonds. Therefore, we think that it is economically more plausible to model the recovery rate as a function of the same latent process that drives default intensity rather than as a function of the latent process driving the term structure of risk-free interest rates.

Second, we compare the relative performance of the model with stochastic recovery and constant recovery. Specifically, using MCMC methods and a Bayesian model selection criterion, we compare two parsimonious reduced form models in continuous time. The first model is estimated under the assumption that the recovery rates are constant and are independent of default intensities. The second model is estimated with a stochastic recovery rate that is negatively correlated with default intensity of the firm. We assume that both the default probability and the expected recovery rate are driven by the same latent process. In other words, they are in a sense entangled, at least from the perspective of investors choosing defaultable credit instruments. The estimated model uses just one (Cox, Ingersoll, and Ross, 1985, hereafter: CIR) process as a latent factor. Additionally there are two more parameters that drive the recovery rate process. In general, this is a very parsimonious model. We find that the parameter estimates are significant and imply sensible recovery rates.

This paper is organized as follows. Section 2, with a simple example, provides the intuition for the relation of recovery rates under the risk neutral and real world measure. Section 3 presents how the stochastic recovery rate can be accommodated by an affine term structure model. In Section 4 we describe the estimation methodology. Section 5 presents the estimates and Section 6 concludes the paper.

2   Relation Between Stochastic Recovery and Risk Premia

The following simple example should provide some intuition for the relation between the expected recovery rates under the empirical $\mathbb{P}$ and under the pricing $\mathbb{Q}$ measure. From the empirical papers mentioned above, we know that under the $\mathbb{P}$ measure, the higher the number of firms that are liquidated, the lower is the recovery from defaulted bonds. Affine models always assume some sort of market price of risk. In this example, we show that when investors are risk averse and recoveries decrease under the $\mathbb{P}$ measure, then recoveries will decrease under $\mathbb{Q}$.

Assume that there is an investor, two periods and two identical firms. The firms finance themselves by issuing bonds. At time 2, the bond will pay 1 with probability p and will default with probability $1-p$. If the firm defaults, the investor recovers an amount equal to ϕ_1. However, if both firms are liquidated, then the supply of the collateral from defaulted firms on the market is so high that it depresses the price. Then the investor can recover only ϕ_2, which is less than ϕ_1. Assume for simplicity that the investor buys only one bond and $u'(c_0) = 1$, so that the state prices have the following form:

$$\psi_\omega = \frac{p_\omega u'(c_{1,\omega})}{u'(c_0)} = p_\omega u'(c_{1,\omega}).$$

	probabilities under $\mathbb{P}$	payoffs	state prices	probabilities under $\mathbb{Q}$
state 1	(p, p)	1	$\psi_1 = p^2 u'(c_{1;1})$	$\pi_1^{\mathbb{Q}} = \frac{\psi_1}{\sum_i \psi_i}$
state 2	$(1-p, p)$	ϕ_1	$\psi_2 = (1-p)p u'(c_{1;2})$	$\pi_2^{\mathbb{Q}} = \frac{\psi_2}{\sum_i \psi_i}$
state 3	$(p, 1-p)$	1	$\psi_3 = p(1-p) u'(c_{1;3})$	$\pi_3^{\mathbb{Q}} = \frac{\psi_3}{\sum_i \psi_i}$
state 4	$(1-p, 1-p)$	ϕ_2	$\psi_4 = (1-p)^2 u'(c_{1;4})$	$\pi_4^{\mathbb{Q}} = \frac{\psi_4}{\sum_i \psi_i}$

In this example, we replicate the empirical observation that under the physical measure $\mathbb{P}$, the probabilities of

default are negatively correlated with realized recovery rates. The following proposition shows that, in such a case, investors will demand an additional risk premium in order to be compensated for the recovery risk.

Proposition 2.1. *If* $1 > \phi_1 > \phi_2$ *and the investor is risk averse, then*

$$E^{\mathbb{Q}}\left(\phi | default\right) < E^{\mathbb{P}}\left(\phi | default\right).$$

Proof. See the Appendix. □

3 Recovery Rate in an Affine Term Structure Model

The class of affine processes is one of the most popular and widely studied time series models in the empirical finance literature. Its popularity can be largely attributed to analytical tractability, which accommodates stochastic volatility, jumps and correlations among risk factors. As we will show, affine processes can also be used to model recovery rates that are negatively correlated with the probability of default. In this paper, we will make use of the methods described in Duffie, Pan, and Singleton (2003), and summarized and extended in Filipovic (2009).

Specifically, in order to price CDS spreads at different maturities, we will use the transform method derived in Duffie et al. (2003), which gives a closed-form solution for the following expectation:

$$E_t\left[e^{-\int_t^T r(X_s)ds}e^{uY_T}\left(v_0 + v_1 X_T\right)1_{\{\beta X_T < y\}}\right]$$

where X_t is a state variable that follows an affine process, $R(X)$ is a discount rate and an affine function of X, and $e^{uX_T}\left(v_0 + v_1 X_T\right)1_{\{\beta X_T < y\}}$ is the terminal payoff at time T.

3.1 CDS Pricing

3.2 Constant recovery

The CDS price (we adapt here formulae from Duffie (2005), sections 5, 6 and 8) is a ratio of the so-called default leg, $L_t^{default}(T)$, to the fixed leg, $L_t^{fixed}(T)$:

$$cds_t(T) = \frac{L_t^{default}(T)}{L_t^{fixed}(T)}, \tag{1}$$

where $L_t^{default}(T)$ is equal to the expected payment from the CDS issuer to the protection buyer in case of default:

$$L_t^{default}(T) = \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v r_s ds}1_{\{\tau < T\}}\left(1 - \phi_v\right)\right]dv, \tag{2}$$

τ denotes the doubly-stochastic default time, and ϕ is a recovery rate. $L_t^{fixed}(T)$ represents the sum of the discounted CDS premium payments accounting for the probability that the firm survives until the payments are due, plus an accrued premium payment made at default time τ. Let $T_{I(\tau)}$ denote the last time when the premium was paid before the default happened. If default occurs at time τ between premium payments at times T_j and T_{j+1},

and $j = I(\tau)$, then the protection buyer has to pay an accrued premium over the period $\tau - T_{I(\tau)}$, which yields:

$$L_t^{fixed}(T) = \frac{1}{4}\sum_{j=1}^{4T} E_t^{\mathbb{Q}}\left[e^{-\int_t^{\frac{1}{4}j} r_s ds} 1_{\{\tau > T\}}\right] dv + \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v r_s ds} 1_{\{\tau < T\}}\right] \left(v - T_{I(v)}\right) dv. \tag{3}$$

Following Duffie (2005) (sections 5 and 6), we can express the above conditional expectations using a default intensity process, λ_t, to obtain:

$$E_t^{\mathbb{Q}}\left[e^{-\int_t^v r_s ds} 1_{\{\tau < T\}}\left(1 - \phi_v\right)\right] dt = E_t^{\mathbb{Q}}\left[e^{-\int_t^v (r_u + \lambda_u) du} \lambda_v \left(1 - \phi_v\right)\right] dv \tag{4}$$

and

$$E_t^{\mathbb{Q}}\left[e^{-\int_t^\tau r_s ds} 1_{\{\tau > T\}}\right] = E_t^{\mathbb{Q}}\left[e^{-\int_t^T (r_u + \lambda_u) du}\right]. \tag{5}$$

Equations (4) and (5) allow us to express $L_t^{default}(T)$ and $L_t^{fixed}(T)$ as

$$L_t^{default}(T) = \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v (r_s + \lambda_s) ds} \lambda_v \left(1 - \phi_v\right)\right] dv, \tag{6}$$

$$L_t^{fixed}(T) = \frac{1}{4}\sum_{j=1}^{4T} E_t^{\mathbb{Q}}\left[e^{-\int_t^{\frac{1}{4}j} (r_s + \lambda_s) ds}\right] + \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v (r_s + \lambda_s) ds} \lambda_v\right] \left(v - T_{I(t)}\right) dv. \tag{7}$$

In order to obtain these expressions in closed form, we assume that λ_t follows a CIR process under the pricing measure $\mathbb{Q}$:

$$d\lambda_t = \kappa^{\mathbb{Q}}\left(\theta^{\mathbb{Q}} - \lambda_t\right) + \sigma\sqrt{\lambda_t} dW_t^{\mathbb{Q}}. \tag{8}$$

3.3   Stochastic recovery

Assume that the recovery rate is stochastic and depends on the default intensity λ_t. Additionally, assume that there exists a function $\phi(\lambda_t)$ that links the default intensity with the expected recovery rate. The function $\phi(\lambda_t)$ should fulfill three necessary properties. First, it should have the domain on $\mathbb{R}_+$, because CIR is defined on $\mathbb{R}_+$. Second, in order to ensure a negative correlation of default probabilities with recovery rates, we need to use a function that has a negative or nonpositive first derivative. This is a crucial condition. We know that realized recovery rates are negatively related with aggregate default rates. Therefore, we impose the same condition on implied recovery rates. Finally, the values of the function ϕ should be between 0 and 1.

A convenient possibility is an exponential function:

$$\phi(\lambda_t) = \beta_2 + \beta_0 e^{\beta_1 \lambda_t}, \tag{9}$$

The parameters of the recovery function are constrained as follows:

$$\beta_0 \in (0, 1), \ \beta_1 \leq 0, \ \beta_2 \in (0, 1).$$

It is obviously a very special case, but it has one very useful feature. When $\beta_1 \leq 0$, the numerator of the CDS function (1) can be obtained in closed form by means of a Laplace transfrom (see Duffie and Garleanu (2001)).

Any other functional relation of default intensity with expected recovery rate can be implemented using Fourier transform methods (see Filipovic (2009)). However, it would be less precise, and given that we have a panel of 1146×5 data points, it would also be much more computationally time consuming.

3.3.1 Constant risk-free interest rate

We follow Pan and Singleton (2008) and Schneider et al. (2011) and assume that the risk free rate has no impact on credit risk. Therefore, we assume that

$$r = const.$$

The recovery rate is assumed to be a function of the stochastic process $\phi(\lambda_t)$. Here, $L_t^{default}(T)$ is equal to

$$L_t^{default}(T) = \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v\left(1 - \phi(\lambda_v)\right)\right]dv \tag{10}$$

and $L_t^{fixed}(T)$ is

$$L_t^{fixed}(T) = \frac{1}{4}\sum_{j=1}^{4T} E_t^{\mathbb{Q}}\left[e^{-\int_t^{\frac{1}{4}j}\lambda_s ds}\right] + \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v\right]\left(v - T_{I(v)}\right)dv. \tag{11}$$

For $\phi(\lambda_t) = \beta_2 + \beta_0 e^{\beta_1 \lambda_t}$, the numerator of the CDS spread can expressed as follows:

$$L_t^{default}(T) = (1 - \beta_2)\int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v\right]dv - \beta_0 \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v e^{\beta_1 \lambda_t}\right]dv$$

and the two integrands can be expressed in terms of the moment generating functions (as shown in Duffie and Garleanu (2001)):

$$E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v\right] = e^{\Phi(v-t,0)+\Psi(v-t,0)\lambda(t)}\left[\Phi_u(v-t,0) + \Psi_u(v-t,0)\lambda(t)\right], \tag{12}$$

$$E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v e^{\beta_1 \lambda_t}\right] = e^{\Phi(v-t,\beta_1)+\Psi(v-t,\beta_1)\lambda(t)}\left[\Phi_u(v-t,\beta_1) + \Psi_u(v-t,\beta_1)\lambda(t)\right]. \tag{13}$$

The functions $\Phi(v-t,u)$, $\Psi(v-t,u)$, $\Phi_u(v-t,u)$ and $\Psi_u(v-t,u)$ are solutions of certain Riccati equations. These functions for a CIR process are defined in the Appendix. Finally, the CDS price with stochastic recovery rate is equal to

$$cds_t(T) = \frac{\int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v\left(1 - \beta_2 + \beta_0 e^{\beta_1 \lambda(v)}\right)\right]dv}{\frac{1}{4}\sum_{j=1}^{4T} E_t^{\mathbb{Q}}\left[e^{-\int_t^{\frac{1}{4}j}\lambda_s ds}\right] + \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v\right]\left(v - T_{I(v)}\right)dv}. \tag{14}$$

4 Estimation Methodology

The model described in this paper assumes that CDS spreads are driven by unobserved state variables. Here we will use Bayesian simulation methods to estimate the parameters and the latent process λ_t. This method is computationally very intensive, but it has several advantages over other estimation methods for models with latent

processes. MCMC allows us to estimate simultaneously both the parameters of the model and the latent variables, given the observed data.

Let the vector Θ denote the parameters under $\mathbb{Q}$ and $\mathbb{P}$, as well as the parameters in the covariance matrix of the errors, Σ_e, defined by:

$$\Theta = \left(\kappa^{\mathbb{P}}, \kappa^{\mathbb{Q}}, \theta^{\mathbb{P}}, \theta^{\mathbb{Q}}, \sigma, \beta_0, \beta_1, \beta_2, a_0, a_1, a_2\right)'. \tag{15}$$

The generated CDS premia from the model are $y = \{y_t\}_{t=1}^{1146}$, where

$$y_t = \left(y_t\left(T=1\right), y_t\left(T=3\right), y_t\left(T=5\right), y_t\left(T=7\right), y_t\left(T=10\right)\right)$$

are expressed in terms of the latent process, λ_t, and the vector of parameters, Θ. We assume that the panel of CDS premia, computed from equation (14), is observed with an additive iid observation error, $e_t \sim N\left(0, \Sigma_e\left(t\right)\right)$, such that at each point in time t, we have:

$$y_t = cds\left(\lambda_t, \Theta\right) + e_t. \tag{16}$$

The covariance matrix of the error term is a 5×5 diagonal matrix, with entries given by

$$\left(\Sigma_e\right)_{ii} = e^{a_0 + a_1 T_i + a_2 T_i^2}, \tag{17}$$

where $i = 1, ..., 5$ and T_i is the i-th component of the vector of CDS maturities, $T = (1, 3, 5, 7, 10)$. Finally, let $\lambda = \{\lambda_t\}_{t=1}^{1146}$ denote the whole latent process. For time series inference, we discretize equation (6) under the objective measure $\mathbb{P}$:

$$\lambda_t = \lambda_{t-1} + \kappa^{\mathbb{P}}\left(\theta^{\mathbb{P}} - \lambda_{t-1}\right)\Delta + \sigma\sqrt{\lambda_{t-1}\Delta}\varepsilon_\lambda\left(t\right) \tag{18}$$

where Δ is a time step between observations equal to $1/252$. The innovations $\varepsilon_\lambda\left(t\right)$ are iid distributed random variables. The joint posterior density of the parameters and latent state variables are given by the following equation:

$$p\left(\lambda, \lambda_0, \Theta | y\right) \propto p\left(y | \lambda, \lambda_0, \Theta\right) p\left(\lambda | \Theta\right) p\left(\Theta\right) p\left(\lambda_0\right). \tag{19}$$

The density $p\left(y | \lambda, \lambda_0, \Theta\right)$ represents a multivariate normal distribution arising from equation (16) and the transition density $p\left(\lambda | \Theta\right)$ is determined by the latent process from equation (18). Here we use uninformative priors $p\left(\Theta\right)$ and $p\left(\lambda_0\right)$, with normal priors for parameters with support on the real line and gamma priors for parameters with support on the positive real line, both with high variances. Specifically, for $\left(\kappa^{\mathbb{P}}, \kappa^{\mathbb{Q}}, \theta^{\mathbb{P}}, \theta^{\mathbb{Q}}, \sigma\right)$, we use gamma priors, for β_1 we use gamma prior but multiplied by (-1), for (a_0, a_1, a_2) we use normal priors, and for (β_0, β_2) we also use normal priors but truncated to the interval $(0, 1)$. Further details about sampling the parameters and latent process can be found in the Appendix.

8

5 Empirical Results

5.1 Data

The models are estimated for three different firms: ConocoPhillips, General Motors and Campbell Soup. The data on CDS come from the Markit Group, and these are daily data that span approximately 4.5 years from 1 January 2004 until 30 May 2008. ConocoPhillips is the fifth largest private sector energy corporation in the world (A1 rating from Moody's). General Motors is an American multinational automotive corporation. General Motors defaulted in 2009, which is already outside of the time interval covered by the dataset. In an auction on 12 June 2009, the recovery rate for the CDS was settled at a level of 12.5 percent. Campbell Soup is a well-known American producer of canned soups and related products (A2 rating from Moody's).

5.2 ConocoPhillips Case Study

5.2.1 Posterior Parameter Estimates

Table 1 presents the posterior estimates of the parameters both for the model with stochastic and constant recovery rate. What stands out is a very high standard deviation of the parameters measured under the $\mathbb{P}$ measure, that is $\kappa^{\mathbb{P}}$ and $\theta^{\mathbb{P}}$. The parameters that are estimated under the $\mathbb{Q}$ measure are apparently much more precise.

	stochastic recovery model		constant recovery model	
	mean	std.	mean	std.
$\kappa^{\mathbb{P}}$	0.0769	0.1545	0.0709	0.1453
$\kappa^{\mathbb{Q}}$	0.0109	0.0005	0.0106	0.0009
$\theta^{\mathbb{P}}$	0.1988	0.4124	0.4369	0.7989
$\theta^{\mathbb{Q}}$	0.0943	0.0051	0.0752	0.0057
σ	0.0577	0.0091	0.0600	0.0101
β_0	0.3937	0.0410	0.4211	0.0020
β_1	-2.7842	0.3765	-	-
β_2	0.0455	0.0378	-	-
a_0	-14.7590	0.0090	-13.9329	0.0215
a_1	-0.0362	0.0006	0.0032	0.0001
a_2	-0.0095	0.0005	-0.0082	0.0002

Table 1

Results of estimation for ConocoPhillips

We are particularly interested in the estimates of the parameters that govern the implied recovery rate. For the model with stochastic recovery, these are β_0, β_1 and β_2, while for the model with constant recovery, it is just β_0. Figure 1 shows the trace plots and histograms for the beta parameters estimated for ConocoPhillips for the model with stochastic recovery, while Figure 2 presents the estimates for the model with constant recovery.
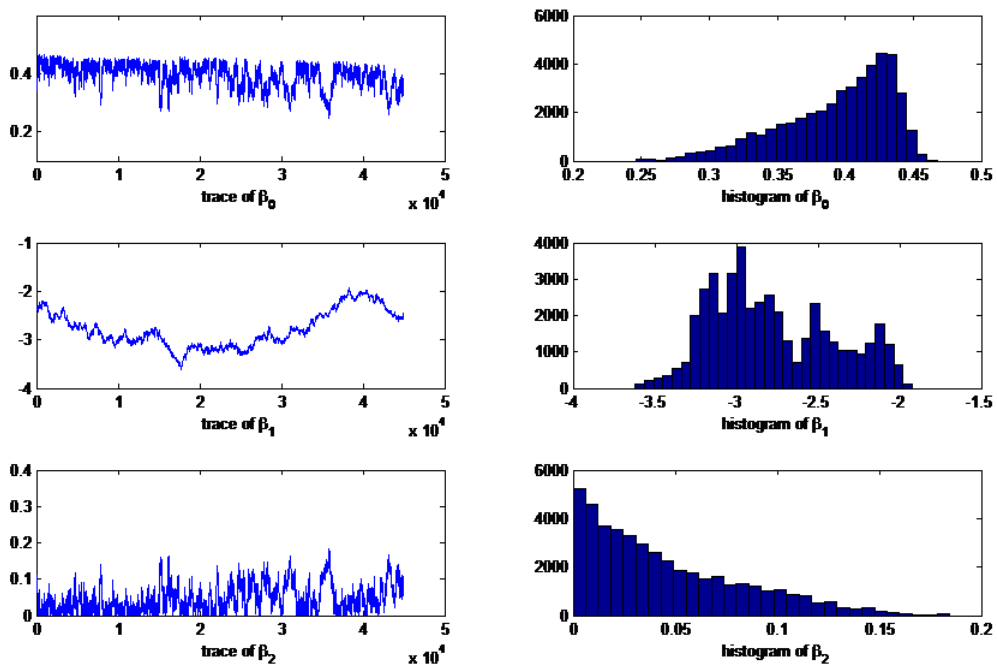
Figure 1: Estimated β parameters for the stochastic model

Using the MCMC output, we find that the β parameters for both models with a 95% confidence will fall into the intervals presented in Table 2.

stochastic recovery model	constant recovery model
$\beta_0 \in (0.2974, 0.4469)$	$\beta_0 \in (0.4170, 0.4251)$
$\beta_1 \in (-3.3665, -2.0520)$	-
$\beta_2 \in (0.0013, 0.1348)$	-

Table 2
95% confidence interval for β derived from the MCMC output

For the model where the recovery is assumed to be constant, we have just one parameter β_0 to estimate. It turns out that, in this case, β_0 can be estimated with very high precision. The posterior constant recovery estimate is

$$\phi_{\text{constant}} = \beta_0 = 0.4210$$

and with a 95% confidence it falls into a very narrow interval, as can be seen in Table 2 and Figure 2. On the other hand, the β parameters in the stochastic model are much less precise. However, one important observation is that the β_1 estimate is significantly less than zero. In fact, the posterior distribution of β_1 is everywhere below zero. This implies that the stochastic recovery model is strongly supported by the data, even if it turns out that it is difficult to pinpoint β_1 exactly. Moreover, the estimate of β_0 for the stochastic model is significantly different from zero.

5.2.2 Forward Recovery Rates and Default Probabilities

Table 3 presents the mean recovery rates and mean default probabilities at different maturities for both models. Forward default probabilities are obtained from the following formula:

$$
\begin{aligned}
PD_t^T &= E_t^{\mathbb{Q}} \left[e^{-\int_t^T \lambda_s ds} \right] & (20) \\
&= e^{\Phi(T-t,0) + \Psi(T-t,0)\lambda_t} & (21)
\end{aligned}
$$

and forward recovery rates in the stochastic model are equal to:

$$
\begin{aligned}
\phi_t^T &= \beta_2 + \beta_0 E_t^{\mathbb{Q}} \left[e^{\beta_1 \lambda_T} \right] & (22) \\
&= \beta_2 + \beta_0 e^{\widetilde{\Phi}(T-t,\beta_1) + \widetilde{\Psi}(T-t,\beta_1)\lambda_t}, & (23)
\end{aligned}
$$

where $\widetilde{\Phi}(T-t,\beta_1)$ and $\widetilde{\Psi}(T-t,\beta_1)$ are solutions to the Riccati equation defined in the Appendix.
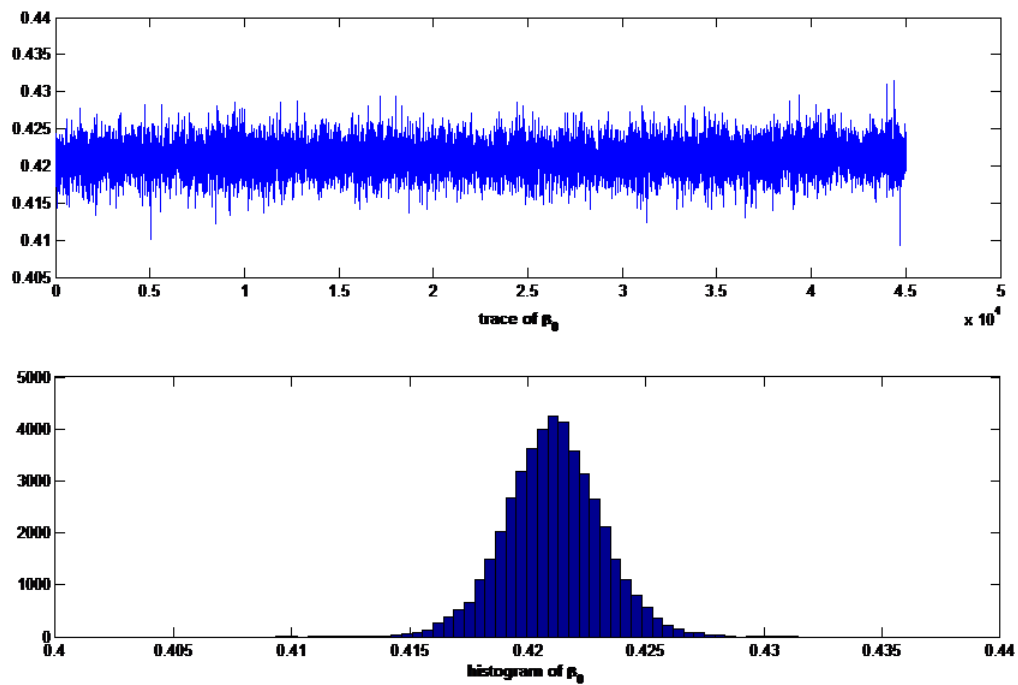
Figure 2: Trace and histogram of β_0 for the model with constant recovery
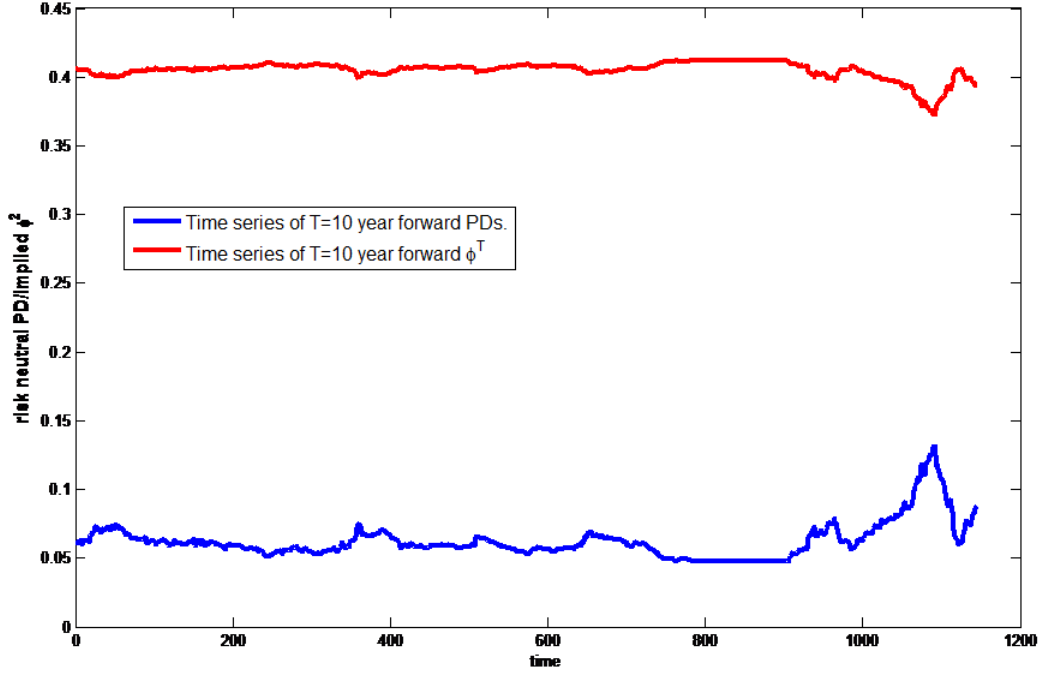
Figure 3: Risk neutral default probabilities and implied forward recovery rates at maturity $T = 10$

<div align="center">ConocoPhillips</div>

	stochastic recovery		constant recovery	
T	forward PD^T	forward ϕ^T	forward PD^T	ϕ
1	0.0022	0.4361	0.0025	0.4207
3	0.0099	0.4312	0.0101	0.4207
5	0.0211	0.4250	0.0221	0.4207
7	0.0355	0.4177	0.0348	0.4207
10	0.0624	0.4051	0.0603	0.4207

<div align="center">Table 3</div>
<div align="center">Mean of PD^T and ϕ^T across all time points</div>

In Table 3 we can see that forward stochastic and constant recovery rates for ConocoPhillips are, in fact, very similar for longer maturities. We have the same impression from Figure 3, where we can see the time series of forward recovery rates at a fixed maturity of $T = 10$. It seems that for ConocoPhillips the implied stochastic recovery rates do not change much over time, so that qualitatively the two models are indistinguishable.

As was pointed out in Johannes and Polson (2003), the MCMC algorithm allows us to quantify estimation risk.

Figure 4: Term structure of forward recovery rates and their 95% confidence interval for the stochastic recovery rate model - ConocoPhillips
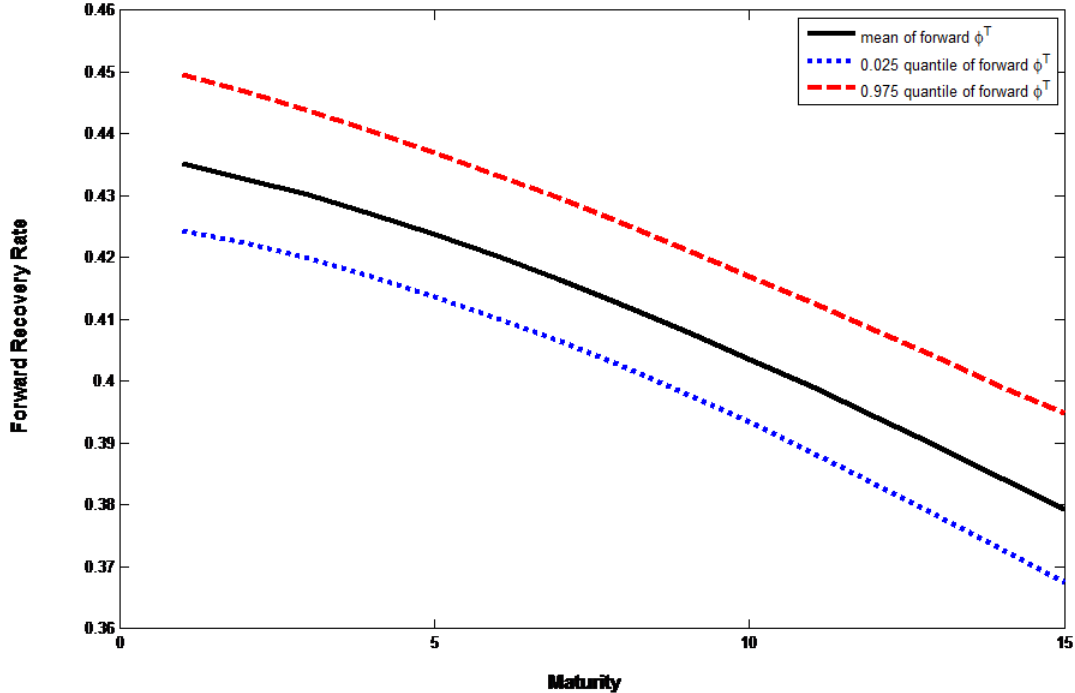
In our case, we would be particularly interested in estimation risk of the recovery rate. Therefore, we use the MCMC output to calculate the 95% confidence intervals for the recovery rate. For the model with constant recovery, it is straightforward because it depends on only one parameter, β_0. As can be seen in Table 2, $\beta_0 \in (0.4170, 0.4251)$ with 95% confidence.

For the model with stochastic recovery, ϕ^T depends on all the parameters and on the realizations of the latent process λ. Altogether, our output of the MCMC algorithm consists of 50000 iterations, but the first 5000 are discarded, so we use information contained in the other 45000 iterations of MCMC to evaluate estimation risk. Specifically, for each iteration step g in the output of MCMC for given $\Theta^{(g)}$ and $\lambda^{(g)}$, we compute the corresponding ϕ^T from equation (21) at different maturities T. That gives us 1146 different forward recovery rates for each each T and each g. We summarize these data by taking an average over all time points from 1 to 1146, and then finding the mean and the 0.025^{th} and 0.975^{th} quantiles of ϕ^T. In this way, we obtain an averaged term structure of recovery rates, together with its 95% confidence interval. Figure 4 plots the results of this procedure.

As can be seen from Figure 4, the 95% confidence interval is relatively narrow. The difference between the 0.975^{th} and 0.025^{th} quantiles of ϕ^T is not wider than 0.03. So, despite the fact that parameters β are difficult to pinpoint exactly, we can see that, on average, forward stochastic rates can be estimated with some precision.

14

5.2.3 Pricing errors

The fact that we use only one factor to explain the term structure of CDS spreads should not be controversial, given that the first principal component accounts for 94% of the variation in all spreads. However, the model does not price all maturities equally well. In Table 4, we can see the pricing errors for different maturities.

	1 year	3 year	5 year	7 year	10 year
stochastic recovery	3.87	2.38	1.45	1.73	2.49
constant recovery	5.74	3.26	0.97	1.34	2.61
mean value of spread	9.17	17.04	25.11	30.64	37.62

Table 4

RMSE at different maturities

The highest root mean square error is for the shortest 1-year maturity and the lowest is for the 5-year maturity, which is also the most liquid one. We observe in Figure 5 that the pricing errors on the 1-year contract are negatively correlated with the 5- and 10-year maturity contract spreads. This may suggest that there is some difficulty in fitting the 1-year spread. The same problem with matching the 1-year spread has also been reported in Pan and Singleton (2008) and Schneider et al. (2011). Schneider et al. (2011) solve this problem by adding a second latent factor and allowing for jumps. We choose the same approach as in Pan and Singleton (2008), and assume only one latent factor. This assumption is made for two reasons. First, in order to identify additional parameters of the recovery rate function, it is necessary to keep the model parsimonious in other dimensions. Second, following the asset fire sales theory, we believe that recovery rates and default probabilities are manifestations of the same economic mechanism. That is, when industry conditions deteriorate and the default probability is increasing, the risk of asset fire sales also increases significantly. The reduced form method to capture the simultaneous worsening of default probability and expected recovery rate is the assumption of one latent factor behind both phenomena.

With alternative models proposed, it is important to compare their relative performance. One can use information criteria. As we have two nested models, it might be possible to use standard information criteria like AIC or BIC. However, these two criteria do not use additional information that the MCMC algorithm provides.

Spiegelhalter, Best, Carlin, and Van Der Linde (2002) have developed a Bayesian alternative to both AIC and BIC. In the Bayesian context, this criterion is more satisfactory than the two former alternatives because it takes into account the prior information, and also provides a natural penalty factor to the log-likelihood function. In a model with latent factors, the parameter space is somewhat arbitrary. Let (Θ, λ) denote the vector of augmented parameters and p be a likelihood function, a multivariate normal distribution arising from the observation equation (16). Then the deviance information criterion (DIC) consists of two parts:
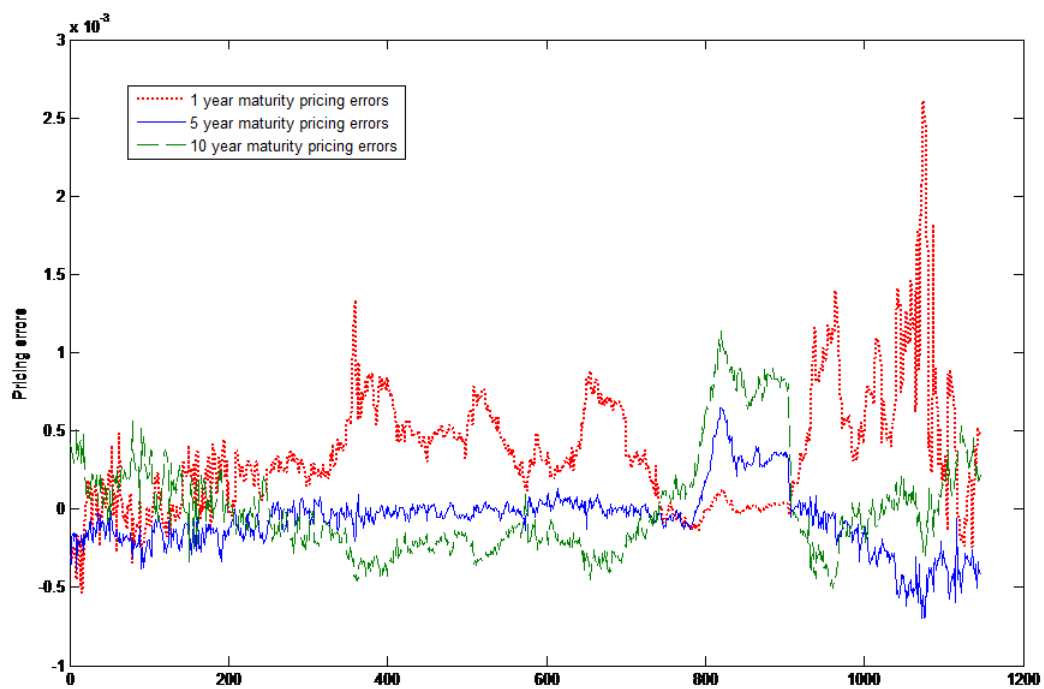
$$DIC = \overline{D} + p_D$$

15

Figure 5: Pricing errors for the model with stochastic recovery rate.

where

$$\overline{D} = E_{(\Theta,\lambda)|y} \left[-2 \ln p\left(y|\lambda, \Theta\right) \right]$$

and

$$
\begin{aligned}
p_D &= \overline{D} - D\left(\overline{\Theta}\right) \\
&= E_{(\Theta,\lambda)|y} \left[-2 \ln p\left(y|\lambda, \Theta\right) \right] + 2 \ln p\left(y|\overline{\lambda}, \overline{\Theta}\right).
\end{aligned}
$$

Here $\overline{D}$ is a Bayesian equivalent of a model fit, and is defined as the posterior expectation of the deviance. More precisely, it is -2 times the log-likelihood value, and it attains smaller values for models with a better fit. In the above, p_D is a measure of complexity, also called the *effective number of parameters*.

For ConocoPhillips, we find that $DIC_{\text{stochastic}} = -84713$ and $DIC_{\text{constant}} = -84249$, so that

$$DIC_{\text{stochastic}} < DIC_{\text{constant}}.$$

Hence, the stochastic recovery model, despite its higher number of parameters and larger uncertainty of the beta parameters, fares better. Therefore, we conclude that a stochastic recovery model is supported by the data for the Bayesian estimates.

5.3   Campbell Soup and General Motors

Table 5 presents the posterior estimates of the parameters for Campbell Soup and General Motors. For these two firms we see that the parameters $\kappa^{\mathbb{P}}$ and $\theta^{\mathbb{P}}$ are much more difficult to estimate precisely than the other parameters under the $\mathbb{Q}$ measure.

| | Campbell Soup | | | | General Motors | | | |
| | stochastic | | constant | | stochastic | | constant | |
	mean	std.	mean	std.	mean	std.	mean	std.
$\kappa^{\mathbb{P}}$	0.1408	0.2289	0.1544	0.2263	0.1580	0.1157	0.1515	0.1129
$\kappa^{\mathbb{Q}}$	0.0117	0.0009	0.0098	0.0020	0.0176	0.0002	0.0135	0.0003
$\theta^{\mathbb{P}}$	0.0699	0.1410	0.0731	0.1181	0.5960	0.3958	0.5277	0.3979
$\theta^{\mathbb{Q}}$	0.0901	0.0065	0.0962	0.0173	0.4858	0.0144	0.6338	0.0092
σ	0.0504	0.0029	0.0652	0.0058	0.1231	0.0027	0.1209	0.0029
β_0	0.4094	0.0132	0.4408	0.0040	0.4191	0.0470	0.5151	0.0005
β_1	-4.7822	0.3586	-	-	-1.0881	0.1468	-	-
β_2	0.0143	0.0122	-	-	0.1896	0.0464	-	-
a_0	-14.9779	0.0048	-14.5412	0.0741	-8.3262	0.0003	-8.1397	0.0011
a_1	-0.0070	0.0002	-0.0077	0.0008	-0.0496	0.0001	-0.3865	0.0001
a_2	-0.0091	0.0009	-0.0098	0.0005	-0.0503	0.0001	-0.0009	0.0001
DIC	-81608		-8.0806		-55713		-53824	

Table 5

Estimates for Campbell Soup and General Motors

Once again, we see that it is more difficult to estimate precisely the β parameters for the model with stochastic recovery. Table 6 presents the 95% confidence intervals for the estimated β parameters for both models. The Deviance Information Criterion in each case favors the model with the stochastic recovery rate.

| Campbell Soup | | General Motors | |
stochastic ϕ	constant ϕ	stochastic ϕ	constant ϕ
$\beta_0 \in (0.37, 0.42)$	$\beta_0 \in (0.43, 0.44)$	$\beta_0 \in (0.34, 0.52)$	$\beta_0 \in (0.51, 0.52)$
$\beta_1 \in (-5.60, -4.24)$	-	$\beta_1 \in (-1.37, -0.85)$	-
$\beta_2 \in (0.00, 0.04)$	-	$\beta_2 \in (0.09, 0.26)$	-

Table 6

95% confidence interval for β

We also observe that the β_1 parameter estimate is again significantly negative for both companies, which may be interpreted as supporting the negative relation between the implied recovery rates and risk-neutral default intensities.

	Campbell Soup				General Motors			
	stochastic		constant		stochastic		constant	
	PD^T	ϕ^T	PD^T	ϕ	PD^T	ϕ^T	PD^T	ϕ
1	0.0020	0.4201	0.0025	0.4408	0.1145	0.5158	0.1040	0.5151
3	0.0086	0.4136	0.0100	0.4408	0.3046	0.4466	0.2822	0.5151
5	0.0191	0.4060	0.0209	0.4408	0.4497	0.3940	0.4235	0.5151
7	0.0331	0.3973	0.0348	0.4408	0.5598	0.3539	0.5346	0.5151
10	0.0596	0.3827	0.0604	0.4408	0.6790	0.3101	0.6587	0.5151

Table 7

Mean of PD^T and ϕ^T across all time points

In Table 7, we can see for Campbell Soup that forward recovery rates for the stochastic model are very close to the estimates of constant recovery. The same is not the case for General Motors. In the right panel of Table 7, we can see that the implied recovery rates for General Motors are very close to each other only for a 1-year maturity. However, the forward recovery rates decline much faster for General Motors than for either Campbell Soup or ConocoPhillips.

We observe in Figure 6 that the 95% confidence interval for General Motors is much broader at longer maturities. The difference between the 0.975^{th} and 0.025^{th} quantiles of the forward recovery rate ϕ^T is only 0.0203 at maturity $T = 1$, but increases to 0.1173 for $T = 10$, while the lower 0.025^{th} quantile is equal to $\phi^{10}_{0.025 \text{ quantile}} = 0.2465$ and the higher is $\phi^{10}_{0.975 \text{ quantile}} = 0.3638$.

Figure 7 presents the time series of forward recovery rates and forward default probabilities at maturity $T = 10$. We see that risk-neutral default probabilities have a larger variance than the implied recovery rates. Contrasting the time series of implied recovery rates $\phi^{T=10}$ for General Motors with the same plot of $\phi^{T=10}$ for ConocoPhillips in Figure 3, we see the qualitative difference between the two. While the time series of ϕ^T for ConocoPhillips seems to be almost flat and is always very close to 0.4, for General Motors it varies between 0.2 and 0.45. This observation strengthens our assertion that a stochastic recovery model is more important for companies that are very risky.

5.3.1   Model implied ϕ^T and realized recovery rate for General Motors

Our dataset covers daily CDS spreads between 1 January 2004 and 30 May 2008, and the estimates of implied recoveries in Table 7 and in the right panel of Figure 5 present averages over this period. We know that after General Motors filed for Chapter 11 reorganization, in an auction on 12 June 2009 the recovery rate for the CDSs was settled at 12.5 percent. The forward recovery rates for General Motors implied by the models are above 12.5 percent, but at longer maturities the lower bound of the implied stochastic recovery is very close to the realized recovery rate for General Motors.

Summarizing, the model that allows for stochastic recovery turns out to be more realistic. In the model with constant recovery, we found that $\phi = 0.5151$, and additionally this parameter has been estimated with very high
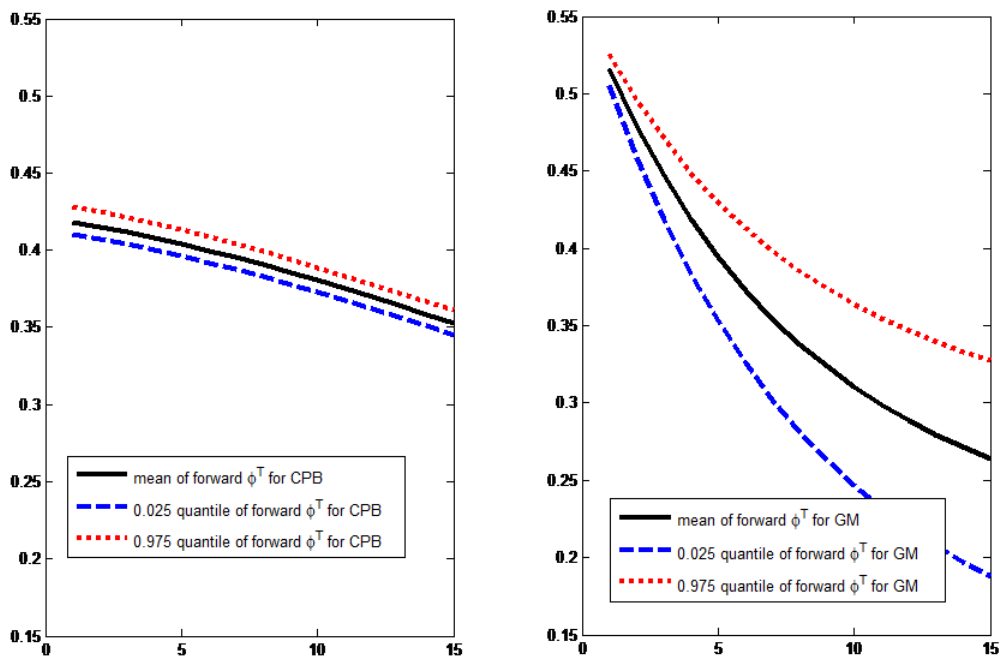
Figure 6: Term structure of forward recovery rates and their 95% confidence interval for the stochastic recovery rate model. Left panel - Campbell Soup (CPB), and right panel - General Motors (GM).
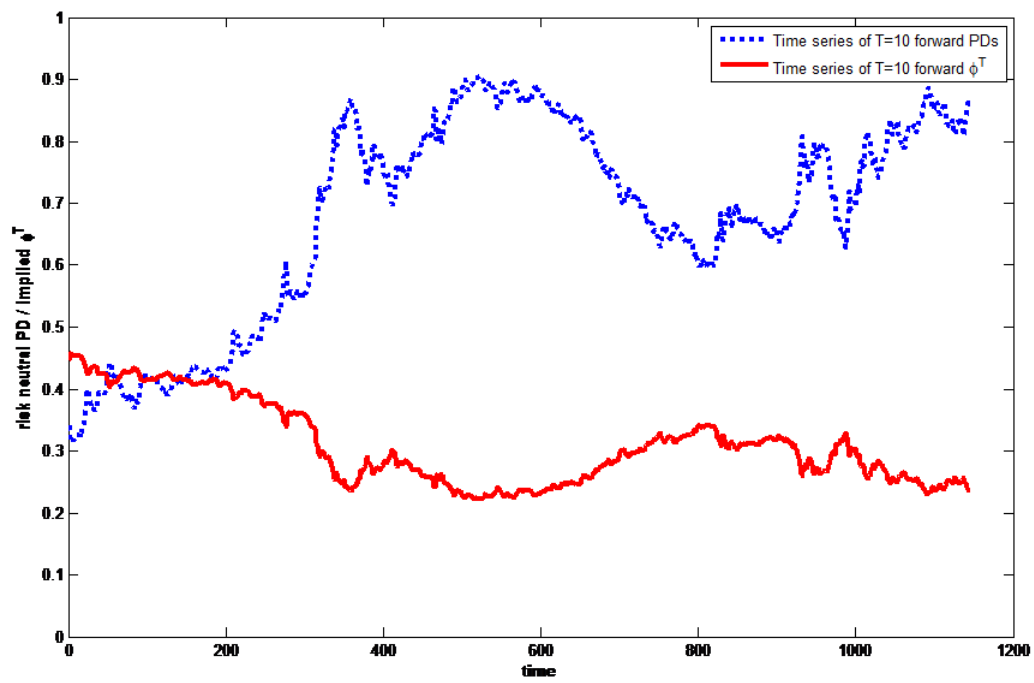
Figure 7: Risk neutral default probabilities and implied forward recovery rates at maturity $T = 10$

precision. On the contrary, the model with stochastic recovery indicates lower forward recovery rates, especially for longer horizons, as can be seen in Figure 6.

6   Conclusion

We augmented the work of Pan and Singleton (2008) and showed how one could extract implied recovery rates under the assumption that they are negatively correlated with risk-neutral default probabilities. We used the framework developed in ? to derive closed-form solutions for CDS prices with stochastic recovery. The parameters of the model and latent factors driving recovery risk and default risk were estimated using a Bayesian MCMC algorithm.

In summary, a model with stochastic recovery received stronger empirical support. First, the parameter driving stochastic recovery, which describes the strength of the negative relation between default intensities and expected recovery rates, is strongly negative. Second, a Bayesian model comparison criterion, namely DIC, which penalizes a larger number of parameters, also supports the model with stochastic recovery rate. Moreover, there is a qualitative reason that favors the stochastic recovery model as it predicted much more realistic recovery rates for General Motors than did the constant recovery model.

In future research we will estimate both models for a larger number of firms and compare these estimates against other determinants of recovery rates in cross-sectional regressions. More specifically, it will be necessary to see whether the variables that the literature has identified as explaining realized recovery rates, can also explain implied recovery rates extracted from CDS spreads. Most importantly, we need to check the firms for which the assumption of stochastic recovery is more important. On the one hand, we expect that the implied stochastic recovery will not differ too much from the implied constant recovery for firms with very good credit ratings. On the other hand, we expect that for risky firms, the implied stochastic recovery will diverge from the implied constant recovery rate.

If this is true, then we may claim that the constant recovery assumption for companies with good credit rating is fairly innocuous. However, for risky companies, the constant recovery assumption may result in serious mispricing.

References

Acharya, V.V., S.T. Bharath, and A. Srinivasan, 2007, Does industry-wide distress affect defaulted firms? evidence from creditor recoveries, *Journal of Financial Economics* 85, 787–821.

Altman, E., 2006, Default recovery rates and lgd in credit risk modeling and practice: an updated review of the literature and empirical evidence, *New York University, Stern School of Business* .

Altman, E.I., B. Brady, A. Resti, and A. Sironi, 2005, The link between default and recovery rates: Theory, empirical evidence, and implications, *The Journal of Business* 78, 2203–2228.

Benmelech, E., and N.K. Bergman, 2011, Bankruptcy and the collateral channel, *The Journal of Finance* 66, 337–378.

Bruche, M., and C. González-Aguado, 2010, Recovery rates, default probabilities, and the credit cycle, *Journal of Banking & Finance* 34, 754–764.

Campbell, J.Y., S. Giglio, and P. Pathak, 2009, Forced sales and house prices, *NBER Working Paper* .

Chen, R.R., and L. Scott, 1995, Interest rate options in multifactor cox-ingersoll-ross models of the term structure, *The Journal of Derivatives* 3, 53–72.

Christensen, J.H.E., 2007, Joint default and recovery risk estimation: an application to cds data, *Working Paper. Available at: http://www.defaultrisk.com/ppcrdrv37.htm* .

Cox, John C, Jr Ingersoll, J. E, and S. A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.

Das, S.R., and P. Hanouna, 2009, Implied recovery, *Journal of Economic Dynamics and Control* 33, 1837–1857.

Doshi, H., 2011, The term structure of recovery rates, *Available at SSRN 1904021* .

Duffie, D., 2005, Credit risk modeling with affine processes, *Journal of Banking & Finance* 29, 2751–2802.

Duffie, D., and N. Garleanu, 2001, Risk and valuation of collateralized debt obligations, *Financial Analysts Journal* 41–59.

Duffie, D., J. Pan, and K. Singleton, 2003, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

Filipovic, D., 2009, Term-structure models: A graduate course, *(Springer)* .

Geyer, A.L.J., 1996, Bayesian estimation of econometric multi factor cox ingersoll ross models of the term structure of interest rates via mcmc methods, *Unpublished Working Paper* .

Johannes, M., and N. Polson, 2003, Mcmc methods for continuous-time financial econometrics, *Available at SSRN 480461* .

Le, A., 2007, Separating the components of default risk: A derivative-based approach, *NYU Working Paper No. FIN-06-008* .

Madan, D., G. Bakshi, and F.X. Zhang, 2006, Understanding the role of recovery in default risk models: Empirical comparisons and implied recovery rates .

Pan, J., and K.J. Singleton, 2008, Default and recovery implicit in the term structure of sovereign cds spreads, *The Journal of Finance* 63, 2345–2384.

Schneider, P., L. Sögner, and T. Veža, 2011, The economic role of jumps and recovery rates in the market for corporate default risk, *Journal of Financial and Quantitative Analysis* 45, 1517–1547.

Shleifer, A., and R.W. Vishny, 2012, Liquidation values and debt capacity: A market equilibrium approach, *The Journal of Finance* 47, 1343–1366.

Song, J., 2007, Loss given default implied by cross-sectional no arbitrage, *Unpublished Working Paper* .

Spiegelhalter, D.J., N.G. Best, B.P. Carlin, and A. Van Der Linde, 2002, Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64, 583–639.

A Appendix

A.1 Proof of Proposition 1

Expected payoffs conditional on default are equal to:

$$
\begin{aligned}
E^{\mathbb{P}}\left(\phi|\text{default}\right) &= p\phi_1 + (1-p)\,\phi_2 \\
E_t^{\mathbb{Q}}\left(\phi|\text{default}\right) &= \frac{\pi_2^{\mathbb{Q}}\phi_1 + \pi_4^{\mathbb{Q}}\phi_2}{\pi_2^{\mathbb{Q}} + \pi_4^{\mathbb{Q}}}
\end{aligned}
$$

so what we want to show is that:

$$
E^{\mathbb{Q}}\left(\phi|\text{default}\right) - E^{\mathbb{P}}\left(\phi|\text{default}\right) < 0
$$

and

$$
\begin{aligned}
E^{\mathbb{Q}}\left(\phi|\text{default}\right) &= \frac{pu'\left(\phi_1\right)\phi_1 + (1-p)\,u'\left(\phi_2\right)\phi_2}{pu'\left(\phi_1\right) + (1-p)\,u'\left(\phi_2\right)} \\
&= \frac{pu'\left(\phi_1\right)\phi_1 + (1-p)\,u'\left(\phi_2\right)\phi_2 + (1-p)\,u'\left(\phi_1\right)\phi_2 - (1-p)\,u'\left(\phi_1\right)\phi_2}{pu'\left(\phi_1\right) + (1-p)\,u'\left(\phi_2\right)} \\
&= \frac{pu'\left(\phi_1\right)\phi_1 - pu'\left(\phi_1\right)\phi_2}{pu'\left(\phi_1\right) + (1-p)\,u'\left(\phi_2\right)} + \frac{u'\left(\phi_1\right)\phi_2 + (1-p)\,u'\left(\phi_2\right)\phi_2 - (1-p)\,u'\left(\phi_1\right)\phi_2}{pu'\left(\phi_1\right) + (1-p)\,u'\left(\phi_2\right)} \\
&= \frac{pu'\left(\phi_1\right)\phi_1 - pu'\left(\phi_1\right)\phi_2}{pu'\left(\phi_1\right) + (1-p)\,u'\left(\phi_2\right)} + \frac{(1-p)\,u'\left(\phi_2\right)\phi_2 + pu'\left(\phi_1\right)\phi_2}{pu'\left(\phi_1\right) + (1-p)\,u'\left(\phi_2\right)} \\
&= \frac{p\phi_1 - p\phi_2}{p + (1-p)\,\frac{u'(\phi_2)}{u'(\phi_1)}} + \phi_2.
\end{aligned}
$$

It follows that:

$$
E^{\mathbb{Q}}\left(\phi|\text{default}\right) - E^{\mathbb{P}}\left(\phi|\text{default}\right) = \underbrace{\left(\frac{1}{p + (1-p)\,\frac{u'(\phi_2)}{u'(\phi_1)}} - 1\right)}_{<0}\left(p\phi_1 - p\phi_2\right)
$$

and for the risk averse investor, we have $u'\left(\phi_2\right) > u'\left(\phi_1\right)$ and

$$
\frac{1}{p + (1-p)\,\frac{u'(\phi_2)}{u'(\phi_1)}} < 1.
$$

It follows that:

$$
E^{\mathbb{Q}}\left(\phi|\text{default}\right) < E^{\mathbb{P}}\left(\phi|\text{default}\right).
$$

Therefore, the inequality is true if the investor is risk averse and when $\phi_1 > \phi_2$ under the $\mathbb{P}$ measure.

A.2 Solution to the Riccati Equation for CIR

A.2.1 Functions $\Phi\left(t,u\right)$, $\Psi\left(t,u\right)$, $\Phi_u\left(t,u\right)$ and $\Psi_u\left(t,u\right)$

The following expectation can be solved by means of the so-called extended affine transform, which follows from Duffie and Garleanu (2001):

$$E\left[e^{\int_0^t q\lambda(s)ds}\lambda_t e^{u\lambda_t}\right] = e^{\Phi(t,u)+\Psi(t,u)\lambda_t}\left[\Phi_u\left(t,u\right)+\Psi_u\left(t,u\right)\lambda_t\right]$$

and the solutions for the functions $\Phi\left(t,u\right)$, $\Psi\left(t,u\right)$, $\Phi_u\left(t,u\right)$ and $\Psi_u\left(t,u\right)$ are:

$$\Phi\left(t,u\right) = \frac{m\left(a_1 c_1 - d_1\right)}{b_1 c_1 d_1}\log\left(\frac{c_1 + d_1 e^{b_1 t}}{c_1 + d_1}\right) + \frac{m}{c_1}t \tag{24}$$

$$\Psi\left(t,u\right) = \frac{1 + a_1 e^{b_1 t}}{c_1 + d_1 e^{b_1 t}} \tag{25}$$

$$\Phi_u\left(t,u\right) = \frac{\partial}{\partial u}\Phi\left(t,u\right) \tag{26}$$

$$\Psi_u\left(t,u\right) = \frac{\partial}{\partial u}\Psi\left(t,u\right) \tag{27}$$

where

$$c_1 = \frac{-n + \sqrt{n^2 - 2pq}}{2q}$$

$$d_1 = \left(1 - c_1 u\right)\frac{n + pu + \sqrt{\left(n + pu\right)^2 - p\left(pu^2 + 2nu + 2q\right)}}{2nu + pu^2 + 2q}$$

$$a_1 = \left(d_1 + c_1\right)u - 1$$

$$b_1 = \frac{d_1\left(n + 2qc_1\right) + a_1\left(nc_1 + p\right)}{a_1 c_1 - d_1}.$$

We find closed-form solutions for $\Phi_u\left(t,u\right)$ and $\Psi_u\left(t,u\right)$ with the symbolic toolbox from Matlab.

A.2.2 Functions $\widetilde{\Phi}\left(t,u\right)$, $\widetilde{\Psi}\left(t,u\right)$

The following solutions are based on Lemma 2 from Filipovic (2009), for CIR process λ_t:

$$d\lambda_t = \left(b + \beta\lambda_t\right)dt + \sigma\sqrt{\lambda_t}dW_t$$

which yields

$$\Phi\left(t,u\right) = \frac{2b}{\sigma^2}\log\left(\frac{2\theta e^{\frac{(\theta-\beta)t}{2}}}{L_3\left(t\right) - L_4\left(t\right)u}\right) \tag{28}$$

$$\Psi\left(t,u\right) = -\frac{L_1\left(t\right) - L_2\left(t\right)u}{L_3\left(t\right) - L_4\left(t\right)u} \tag{29}$$

where $\theta = \sqrt{\beta^2 + 2\sigma^2}$ and

$$
\begin{aligned}
L_1(t) &= 2\left(e^{\theta t} - 1\right) \\
L_2(t) &= \theta\left(e^{\theta t} + 1\right) + \beta\left(e^{\theta t} - 1\right) \\
L_3(t) &= \theta\left(e^{\theta t} + 1\right) - \beta\left(e^{\theta t} - 1\right) \\
L_4(t) &= \sigma^2\left(e^{\theta t} - 1\right).
\end{aligned}
$$

A.3   Description of the MCMC Estimation Method

We sample a path of λ_t and the model parameters with 50000 MCMC steps, where the first 5000 samples are discarded as burn-in steps. The major blocks are given by:

Step A: sample Θ from $p(\Theta|y, \lambda, \lambda_0)$

Step B: sample λ from $p(\lambda|y, \Theta, \lambda_0)$

A.3.1   Step A: Drawing the Parameter Vector Θ

The Bayesian MCMC solution of the estimation problem is based on the joint posterior distribution $p(\lambda, \lambda_0, \Theta, |y)$. The joint posterior is given by equation (19), and we may derive from it the marginal distribution $p(\Theta|y)$ to infer the model parameter vector Θ. The parameters β_0 and β_2 can be obtained by means of a Gibbs sampler. All the other parameters from the vector Θ are sampled by means of the Metropolis-Hastings algorithm. Denote $\Theta_A = \{\beta_0, \beta_2\}$, $\Theta_B = \{\kappa^{\mathbb{P}}, \kappa^{\mathbb{Q}}, \theta^{\mathbb{P}}, \theta^{\mathbb{Q}}, \sigma, \beta_1, a_0, a_1, a_2\}$ and $\Theta = \{\Theta_A, \Theta_B\}$.

A1. Gibbs sampler for Θ_A We observe that equation (16) can be expressed as a linear function of β_0 and β_1 and

$$
\begin{aligned}
y_t &= (1 - \beta_2) \times \underbrace{\frac{\int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v\right] dv}{\frac{1}{4}\sum_{j=1}^{4T} E_t^{\mathbb{Q}}\left[e^{-\int_t^{\frac{1}{4}j} \lambda_s ds}\right] + \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v\right]\left(v - T_{I(v)}\right) dv}}_{=A(t, T, \Theta_B, \lambda_t)} \\
&\quad + \beta_0 \times \underbrace{\frac{\int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v \beta_0 e^{\beta_1 \lambda(v)}\right] dv}{\frac{1}{4}\sum_{j=1}^{4T} E_t^{\mathbb{Q}}\left[e^{-\int_t^{\frac{1}{4}j} \lambda_s ds}\right] + \int_t^T E_t^{\mathbb{Q}}\left[e^{-\int_t^v \lambda_s ds}\lambda_v\right]\left(v - T_{I(v)}\right) dv}}_{=B(t, T, \Theta_B, \lambda_t)} + e_t \\
&= (1 - \beta_2) \times A(t, T, \Theta_B, \lambda_t) + \beta_0 \times B(t, T, \Theta_B, \lambda_t) + e_t.
\end{aligned}
$$

Hence, $(1 - \beta_2)$ and β_0 are coefficients in a panel regression conditional on the other parameters, state variables and the data. We assume a truncated normal prior with $1_{\{\Theta_A \in (0,1)\}}$ as a truncation function. Specifically, we regress $A(t, T, \Theta_B, \lambda_t)$ and $B(t, T, \Theta_B, \lambda_t)$ on the observed CDS spreads y at each point in time t and for all maturities T.

We construct the standard Gibbs sampler as follows. y is a panel of 1146×5 CDS data-points, and $A(t, T, \Theta_B, \lambda_t)$ and $B(t, T, \Theta_B, \lambda_t)$ for all $t \in (1, ..., 1146)$ and $T = (1, 3, 5, 7, 10)$ form two panels of regressors of the same size as

y. Therefore, in order to sample $(1 - \beta_2)$ and β_0, we perform a Bayesian panel regression. We weight $A\left(\cdot, T, \Theta_B, \lambda\right)$ and $B\left(\cdot, T, \Theta_B, \lambda\right)$ separately for each maturity T corresponding to the particular maturity error term from the matrix $(\Sigma_e)_{ii}$ defined in equation (17). Then we stack different maturities on to each other to obtain:

$$
\underbrace{\begin{bmatrix} y^{T=1} \\ y^{T=3} \\ y^{T=5} \\ y^{T=7} \\ y^{T=10} \end{bmatrix}}_{=Y} = (1-\beta_2) \times \underbrace{\begin{bmatrix} A\left(\cdot, T=1, \Theta_B, \lambda\right)/(\Sigma_e)_{11} \\ A\left(\cdot, T=3, \Theta_B, \lambda\right)/(\Sigma_e)_{22} \\ A\left(\cdot, T=5, \Theta_B, \lambda\right)/(\Sigma_e)_{33} \\ A\left(\cdot, T=7, \Theta_B, \lambda\right)/(\Sigma_e)_{44} \\ A\left(\cdot, T=10, \Theta_B, \lambda\right)/(\Sigma_e)_{55} \end{bmatrix}}_{=\widetilde{A}} + \beta_0 \times \underbrace{\begin{bmatrix} B\left(\cdot, T=1, \Theta_B, \lambda\right)/(\Sigma_e)_{11} \\ B\left(\cdot, T=3, \Theta_B, \lambda\right)/(\Sigma_e)_{22} \\ B\left(\cdot, T=5, \Theta_B, \lambda\right)/(\Sigma_e)_{33} \\ B\left(\cdot, T=7, \Theta_B, \lambda\right)/(\Sigma_e)_{44} \\ B\left(\cdot, T=10, \Theta_B, \lambda\right)/(\Sigma_e)_{55} \end{bmatrix}}_{=\widetilde{B}} + \underbrace{\begin{bmatrix} e^{T=1} \\ e^{T=3} \\ e^{T=5} \\ e^{T=7} \\ e^{T=10} \end{bmatrix}}_{=e}. \quad (30)
$$

For brevity, denote $b = [(1 - \beta_2), \beta_0]'$ and $X = \left[\widetilde{A}, \widetilde{B}\right]$. Then the panel regression (28) can be rewritten as $Y = b \times X + e$. With conjugate priors, $b \sim N\left(\widehat{b(0)}, \Sigma_{\widehat{b}}(0)\right)$, b can sampled from a normal distribution, $b|\Sigma_e, \lambda, \Theta_B, y \sim N\left(\widehat{b}, \Sigma_{\widehat{b}}\right)$, where

$$
\begin{aligned}
\widehat{b} &= \Sigma_{\widehat{b}}\left(\left(\Sigma_{\widehat{b}}(0)\right)^{-1}\widehat{b(0)} + X'Y\right) \\
\Sigma_{\widehat{b}} &= \left(\left(\Sigma_{\widehat{b}}(0)\right)^{-1} + X'X\right)^{-1}.
\end{aligned}
$$

The conjugate truncated normal prior has the following parameters: $\widehat{b(0)} = (0.5, 0.5)'$ and $\Sigma_{\widehat{b}}(0) = 1000 \times I_2$ where I_2 is a 2×2 identity matrix. By drawing samples from $N\left(\widehat{b}, \Sigma_{\widehat{b}}\right)$ that fulfill the parametric restrictions of the truncated normal priors, we obtain samples from the desired conditional distribution.

A2. Metropolis algorithm   We use random walk proposals and for every $i \in \Theta_B$ we accept with probability:

$$
\alpha\left((\Theta_B)_i^{(g+1)}, (\Theta_B)_i^{(g)}\right) = \min\left\{1, \frac{p\left(y|\lambda, \lambda_0, (\Theta_B)_i^{(g+1)}\right) p\left(\lambda|\lambda_0, (\Theta_B)_i^{(g+1)}\right) p\left((\Theta_B)_i^{(g+1)}\right)}{p\left(y|\lambda, (\Theta_B)_i^{(g)}\right) p\left(\lambda|(\Theta_B)_i^{(g)}\right) p\left((\Theta_B)_i^{(g)}\right)}\right\}
$$

The proposal densities cancel out due to the symmetry of the random walk proposals. The variance of random walk proposals was scaled to obtain aacceptance rate in the range of $(0, 25, 0.75)$. The Metropolis algorithm consists of the following steps, for every $i \in \Theta_B$:

Step A2.1: Draw $(\Theta_B)_i^{(g+1)}$ from the proposal density $q\left((\Theta_B)_i^{(g+1)} | (\Theta_B)_i^{(g)}\right)$

Step A2.2: Accept $(\Theta_B)_i^{(g+1)}$ with probability $\alpha\left((\Theta_B)_i^{(g+1)}, (\Theta_B)_i^{(g)}\right)$.

A.3.2   Step B: Latent Factor Sampling

The latent state process cannot be inverted directly from the term structure of CDS spreads because of the observation error e_t. Therefore, the latent process λ_t is sampled using a combination of Metropolis-Hastings algorithm with a Kalman filtering method, as described in Geyer (1996). The basic idea used in this paper is to combine an approximate extended Kalman filtering method (EKF) with a Metropolis-Hastings algorithm.

The MCMC algorithm in iteration $(g+1)$ samples a new starting value $\lambda_0^{(g+1)}$, which is used by the extended Kalman filter to find the whole process $\lambda^{(g+1)}$. This new proposal for the latent factor is either accepted or rejected with probability:

$$\alpha\left(\lambda^{(g+1)}, \lambda^{(g)}\right) = \min\left\{1, \frac{p\left(y|\lambda^{(g+1)}, \lambda_0^{(g+1)}, \Theta\right) p\left(\lambda^{(g+1)}|\Theta\right) p\left(\lambda_0^{(g+1)}\right)}{p\left(y|\lambda^{(g)}, \lambda_0^{(g)}, \Theta\right) p\left(\lambda^{(g)}|\Theta\right) p\left(\lambda_0^{(g)}\right)} \times \frac{q\left(\lambda^{(g)}|\lambda^{(g+1)}\right)}{q\left(\lambda^{(g+1)}|\lambda^{(g)}\right)}\right\}$$

Thus, it can be summarized as a three-step procedure:

Step B.1: Draw $\lambda_0^{(g+1)}$ from the proposal density

Step B.2: Using $\lambda_0^{(g+1)}$ as a starting value filter $\lambda^{(g+1)}$ with an Extended Kalman Filter

Step B.3: Accept $\lambda^{(g+1)}$ with probability $\alpha\left(\lambda^{(g+1)}, \lambda^{(g)}\right)$

As a proposal density, we use the normal distribution of the following form:

$$q\left(\lambda^{(g)}|\lambda^{(g+1)}\right) \propto \prod_{t=1}^{1146} e^{-\frac{\left(\lambda_t^{(g)} - \lambda_0^{(g+1)}\right)^2}{2c_\lambda^2}} \tag{31}$$

where c_λ is a variance from the random walk proposal for parameter $\lambda_0^{(g+1)} = \lambda_0^{(g)} + c_\lambda \varepsilon_\lambda$, $\varepsilon_\lambda \sim N(0,1)$. The discretized diffusion process for the latent factor λ_t under the objective P measure, defined in equation (16), has a non-central χ^2 transition density (Cox et al. (1985)). In our implementation of the EKF, the exact non-central χ^2 transition density for the latent factor is substituted with a normal density:

$$\lambda_{t|t-1} \sim N\left(F_t \lambda_{t-1} + u_t, Q_t\right) \tag{32}$$

where, from Chen and Scott (1995), we know that F_t, u_t and Q_t are chosen in such a way that the first two moments of the approximate normal and the exact transition density are equal:

$$F_t = e^{-\kappa^{\mathbb{P}} \Delta t} \tag{33}$$

$$u_t = \left(1 - e^{-\kappa^{\mathbb{P}} \Delta t}\right) \theta^{\mathbb{P}} \tag{34}$$

$$Q_t = \sigma^2 \frac{1 - e^{-\kappa^{\mathbb{P}} \Delta t}}{\kappa^{\mathbb{P}}} \left[\left(1 - e^{-\kappa^{\mathbb{P}} \Delta t}\right) \frac{\theta^{\mathbb{P}}}{2} + e^{-\kappa^{\mathbb{P}} \Delta t} \lambda_{t-1}\right] \tag{35}$$

In order to cope with the nonlinearity of the CDS function, we will apply an extended Kalman filter. It relies on the first-order Taylor expansion of equation (16) around the predicted state $\lambda_{t|t-1}$:

$$y_t = \underbrace{cds\left(\lambda_{t|t-1}|\Theta\right) + J_t\left(\lambda_t - \lambda_{t|t-1}\right)}_{=cds(\lambda_t, \Theta)} + e_t \tag{36}$$

where

$$J_t = \left.\frac{\partial cds}{\partial \lambda}\right|_{\lambda=\lambda_{t|t-1}} \tag{37}$$

denotes the Jacobian matrix of the non-linear function $cds\left(\lambda_{t|t-1}, \Theta\right)$ and $\lambda_{t|t-1}$ is the predicted state. The

derivation of the Jacobian J_t can be found in section 7.4. For a detailed description of latent factor sampling, see Geyer (1996).

A.4 First-Order Taylor Approximation of CDS Price Function for the Extended Kalman Filter

In order to implement the extended Kalman filter, we solve the Jacobian of the CDS function with respect to λ:

$$J_t = \left. \frac{\partial cds}{\partial \lambda} \right|_{\lambda = \lambda_{t|t-1}}$$

at the predicted point $\lambda_{t|t-1}$. In our case, the Jacobian J_t will be a 5×1 vector, because λ is one-dimensional and the CDS equation (14) has to be computed at five different maturities, that is:

$$J_t = \frac{\partial}{\partial \lambda} \begin{bmatrix} cds_t^{T=1} \left(\lambda = \lambda_{t|t-1} \right) \\ cds_t^{T=3} \left(\lambda = \lambda_{t|t-1} \right) \\ cds_t^{T=5} \left(\lambda = \lambda_{t|t-1} \right) \\ cds_t^{T=7} \left(\lambda = \lambda_{t|t-1} \right) \\ cds_t^{T=10} \left(\lambda = \lambda_{t|t-1} \right) \end{bmatrix}$$

We show how to compute J_t by decomposing equation (14) into a numerator and denominator.

A.4.1 Numerator:

For the numerator of the CDS function, we need to find:

$$\frac{\partial}{\partial \lambda} L_t^{default} (T) = \int_t^T \left\{ \frac{\partial}{\partial \lambda} E_t^{\mathbb{Q}} \left[e^{-\int_t^v \lambda_s ds} \lambda_v \left(1 - \phi \left(\lambda_v \right) \right) \right] \right\} dv.$$

We know that

$$\frac{\partial}{\partial \lambda} L_t^{default} (T) = \int_0^T \left\{ (1 - \beta_2) \frac{\partial}{\partial \lambda} E_t^{\mathbb{Q}} \left[e^{-\int_0^t \lambda_s ds} \lambda_t \right] + \beta_0 \frac{\partial}{\partial \lambda} E_t^{\mathbb{Q}} \left[e^{-\int_0^t \lambda_s ds} \lambda_t e^{\beta_1 \lambda_t} \right] \right\} dt,$$

so it is enough to find the following derivatives:

$$\frac{\partial}{\partial \lambda} E_t^{\mathbb{Q}} \left[e^{-\int_t^v \lambda_s ds} \lambda_v \right] \quad = \quad \frac{\partial}{\partial \lambda} \left\{ e^{\Phi^\lambda (v-t,0) + \Psi^\lambda (v-t,0) \lambda(t)} \left[\Phi_u^\lambda (v - t, 0) + \Psi_u^\lambda (v - t, 0) \lambda_t \right] \right\} \tag{38}$$

$$= \quad \Psi^\lambda (v - t, 0) E_t^{\mathbb{Q}} \left[e^{-\int_t^v \lambda_s ds} \lambda_v \right] + \Psi_u^\lambda (v - t, 0) e^{\Phi^\lambda (v-t,0) + \Psi^\lambda (v-t,0) \lambda_t} \tag{39}$$

$$= \quad \Psi^\lambda (v - t, 0) E_t^{\mathbb{Q}} \left[e^{-\int_t^v \lambda_s ds} \lambda_v \right] + \Psi_u^\lambda (v - t, 0) E_t^{\mathbb{Q}} \left[e^{-\int_t^v \lambda_s ds} \right] \tag{40}$$

and

$$\frac{\partial}{\partial \lambda} E_t^{\mathbb{Q}} \left[e^{-\int_t^v \lambda_s ds} \lambda_v e^{\beta_1 \lambda(v)} \right] = \frac{\partial}{\partial \lambda} \left\{ e^{\Phi^\lambda (v-t,\beta_1) + \Psi^\lambda (v-t,\beta_1) \lambda(t)} \left[\Phi_u^\lambda (v - t, \beta_1) + \Psi_u^\lambda (v - t, \beta_1) \lambda_v \right] \right\}$$

$$= \Psi^\lambda (v - t, \beta_1) E_t^{\mathbb{Q}} \left[e^{-\int_t^v \lambda_s ds} \lambda_v e^{\beta_1 \lambda_t} \right] + \Psi_u^\lambda (v - t, \beta_1) e^{\Phi^\lambda (v-t,\beta_1) + \Psi^\lambda (v-t,\beta_1) \lambda(t)}$$

$$= \Psi^\lambda \left(v - t, \beta_1 \right) E_t^\mathbb{Q} \left[e^{-\int_t^v \lambda_s ds} \lambda_v e^{\beta_1 \lambda(v)} \right] + \Psi_u^\lambda \left(v - t, \beta_1 \right) E_t^\mathbb{Q} \left[e^{-\int_t^v \lambda_s ds} e^{\beta_1 \lambda(v)} \right]. \tag{41}$$

These formulae give us the numerator of the first derivative of the CDS price formula.

A.4.2 Denominator:

For the denominator of the CDS function, we need to find $L_t^{fixed}(T)$:

$$\frac{\partial}{\partial \lambda} L_t^{fixed}(T) = \frac{1}{4} \sum_{j=1}^{4T} \frac{\partial}{\partial \lambda} E_t^\mathbb{Q} \left[e^{-\int_t^{\frac{1}{4}j} \lambda_s ds} \right] + \int_t^T \left\{ \frac{\partial}{\partial \lambda} E_t^\mathbb{Q} \left[e^{-\int_t^v \lambda_s ds} \lambda_v \right] \left(v - T_{I(v)} \right) \right\} dv$$

and, with the following derivative, we obtain all the necessary functions:

$$\frac{\partial}{\partial \lambda} E_t^\mathbb{Q} \left[e^{-\int_t^v \lambda_s ds} \right] \quad = \quad \frac{\partial}{\partial \lambda} \left\{ e^{\Phi^\lambda(v-t,0) + \Psi^\lambda(v-t,0)\lambda(t)} \right\} \tag{42}$$

$$= \quad \Psi^\lambda \left(v - t, 0 \right) E_t^\mathbb{Q} \left[e^{-\int_t^v \lambda_s ds} \right]. \tag{43}$$

A.4.3 Numerator and denominator:

Once we have all the necessary components of the numerator and denominator of equation (14), we can obtain $\frac{\partial}{\partial \lambda} cds \left(\lambda, \Theta \right)$. Rewrite the CDS pricing function in terms of functions f, g and h:

$$cds \left(\lambda, \Theta \right) \quad = \quad \frac{\underbrace{\int_t^T E_t^\mathbb{Q} \left[e^{-\int_t^v \lambda_s ds} \lambda_v \left(1 - \beta_2 + \beta_0 e^{\beta_1 \lambda(v)} \right) \right] dv}_{=f(\lambda)}}{\underbrace{\frac{1}{4} \sum_{j=1}^{4T} E_t^\mathbb{Q} \left[e^{-\int_t^{\frac{1}{4}j} \lambda_s ds} \right]}_{=g(\lambda)} + \underbrace{\int_t^T E_t^\mathbb{Q} \left[e^{-\int_t^v \lambda_s ds} \lambda_v \right] \left(v - T_{I(v)} \right) dv}_{h(\lambda)}} \tag{44}$$

$$= \quad \frac{f \left(\lambda \right)}{g \left(\lambda \right) + h \left(\lambda \right)}. \tag{45}$$

Then, by using

$$\frac{\partial}{\partial \lambda} cds \left(\lambda \right) = \frac{\frac{\partial}{\partial \lambda} f \left(\lambda \right) \times \left[g \left(\lambda \right) + h \left(\lambda \right) \right] + f \left(\lambda \right) \times \left[\frac{\partial}{\partial \lambda} g \left(\lambda \right) + \frac{\partial}{\partial \lambda} h \left(\lambda \right) \right]}{\left[g \left(\lambda \right) + h \left(\lambda \right) \right]^2} \tag{46}$$

we may combine equations (36)-(40) in order to derive the Jacobian J_t.